Face Detection Based on Multi-Block LBP Representation

Lun Zhang, Rufeng Chu, Shiming Xiang, Shengcai Liao, and Stan Z. Li

Center for Biometrics and Security Research & National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences
95 Zhongguancun Donglu Beijing 100080, China

Abstract. Effective and real-time face detection has been made possible by using the method of rectangle Haar-like features with AdaBoost learning since Viola and Jones' work [12]. In this paper, we present the use of a new set of distinctive rectangle features, called Multi-block Local Binary Patterns (MB-LBP), for face detection. The MB-LBP encodes rectangular regions' intensities by local binary pattern operator, and the resulting binary patterns can describe diverse local structures of images. Based on the MB-LBP features, a boosting-based learning method is developed to achieve the goal of face detection. To deal with the non-metric feature value of MB-LBP features, the boosting algorithm uses multi-branch regression tree as its weak classifiers. The experiments show the weak classifiers based on MB-LBP are more discriminative than Haar-like features and original LBP features. Given the same number of features, the proposed face detector illustrates 15% higher correct rate at a given false alarm rate of 0.001 than haar-like feature and 8% higher than original LBP feature. This indicates that MB-LBP features can capture more information about the image structure and show more distinctive performance than traditional haar-like features, which simply measure the differences between rectangles. Another advantage of MB-LBP feature is its smaller feature set, this makes much less training time.

1 Introduction

Face detection has a wide range of applications such as automatic face recognition, human-machine interaction, surveillance, etc. In recent years, there has been a substantial progress on detection schemes based on appearance of faces. These methods treat face detection as a two-class (face/non-face) classification problem. Due to the variations in facial appearance, lighting, expressions, and other factors [11], face/non-face classifiers with good performance should be very complex. The most effective method for constructing face/non-face classifiers is learning based approach. For example, neural network-based methods [10], support vector machines [9], etc.

Recently, the boosting-based detector proposed by Viola and Jones [12] is regarded as a breakthrough in face detection research. Real-time performance is achieved by learning a sequence of simple Haar-like rectangle features. The Haar-like features encode differences in average intensities between two rectangular regions, and they can be calculated rapidly through integral image [12]. The complete Haar-like feature set is large and contains a mass of redundant information. Boosting algorithm is introduced to

select a small number of distinctive rectangle features and construct a powerful classifier. Moreover, the use of cascade structure [12] further speeds up the computations. Li et al. extended that work to multi-view faces using an extended set of Haar features and an improved boosting algorithm [5]. However, these Haar-like rectangle features seem too simple, and the detector often contains thousands of rectangle features for considerable performance. The large number of selected features leads to high computation costs both in training and test phases. Especially, in later stages of the cascade, weak classifiers based on these features become too weak to improve the classifier's performance [7]. Many other features are also proposed to represent facial images, including rotated Haar-like features [6], census transform [3], sparse features [4], etc.

In this paper, we present a new distinctive feature, called Multi-block Local Binary Pattern (MB-LBP) feature, to represent facial image. The basic idea of MB-LBP is to encode rectangular regions by local binary pattern operator [8]. The MB-LBP features can also be calculated rapidly through integral image, while these features capture more information about the image structure than Haar-like features and show more distinctive performance. Comparing with original Local Binary Pattern calculated in a local 3×3 neighborhood between pixels, the MB-LBP features can capture large scale structure that may be the dominant features of image structures. We directly use the output of LBP operator as the feature value. But a problem is that this value is just a symbol for representing the binary string. For this non-metric feature value, multi-branch regression tree is designed as weak classifiers. We implement Gentle adaboost for feature selection and classifier construction. Then a cascade detector is built. Another advantage of MB-LBP is that the number of exhaustive set of MB-LBP features is much smaller than Haar-like features (about 1/20 of Haar-like feature for a sub-window of size 20×20). Boosting-based method use Adaboost algorithm to select a significant feature set from the large complete feature set. This process often spends much time even several weeks. The small feature set of MB-LBP can make this procedure more simple.

The rest of this paper is organized as follows. Section 2 introduces the MB-LBP features. In section 3, the AdaBoost learning for feature selection and classifier construction are proposed. The cascade detector is also described in this section. The experiment results are given in Section 4. Section 5 concludes this paper.
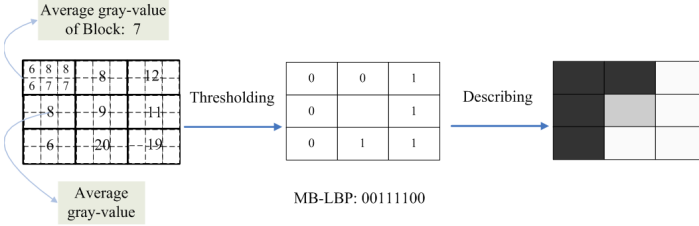
2   Multi-Block Local Binary Pattern Features

Traditional Haar-like rectangle feature measures the difference between the average intensities of rectangular regions (See Fig.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 an Jones [12] 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 [12]. However, the Haar-like features seem too simple and show some limits [7].

In this paper, we propose a new distinctive rectangle features, called Multi-block Local Binary Pattern (MB-LBP) feature. The basic idea of MB-LBP is that the
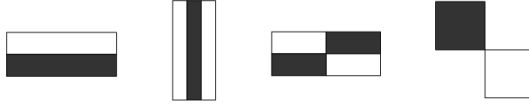
Fig. 1. Traditional Haar-like features.These features measure the differences between rectangular regions' average intensities.

Fig. 2. Multi-block LBP feature for image representation. As shown in the figure, the MB-LBP features encode rectangular regions' intensities by local binary pattern. The resulting binary patterns can describe diverse image structures. Compared with original Local Binary Pattern calculated in a local 3×3 neighborhood between pixels, MB-LBP can capture large scale structure.

simple difference rule in Haar-like features is changed into encoding rectangular regions by local binary pattern operator. The original LBP, introduced by Ojala [8], is defined for each pixel by thresholding the 3×3 neighborhood pixel value with the center pixel value. To encode the rectangles, the MB-LBP operator is defined by comparing the central rectangle's average intensity g_c with those of its neighborhood rectangles $\{g_0, ..., g_8\}$. In this way, it can give us a binary sequence. An output value of the MB-LBP operator can be obtained as follows:

$$MB - LBP = \sum_{i=1}^{8} s(g_i - g_c)2^i \qquad (1)$$

where g_c is the average intensity of the center rectangle, g_i $(i = 0, \cdots, 8)$ are those of its neighborhood rectangles,

$$s(x) = \begin{cases} 1, if\ x > 0 \\ 0, if\ x < 0 \end{cases}$$

A more detailed description of such MB-LBP operator can be found in Fig. 2. We directly use the resulting binary patterns as the feature value of MB-LBP features. Such binary patterns can detect diverse image structures such as edges, lines, spots, flat areas and corners [8], at different scale and location. Comparing with original Local Binary Pattern calculated in a local 3×3 neighborhood between pixels, MB-LBP can capture large scale structures that may be the dominant features of images. Totally, we can get 256 kinds of binary patterns, some of them can be found in Fig. 3. In section 4.1, we conduct an experiment to evaluate the MB-LBP features. The experimental results
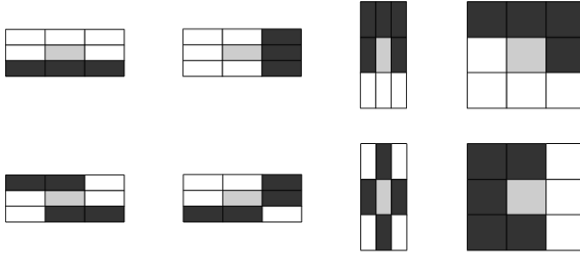
Fig. 3. A randomly chosen subset of the MB-LBP features

show the MB-LBP features are more distinctive than Haar-like features and original LBP features.

Another advantage of MB-LBP is that the number of exhaustive set of MB-LBP features (rectangles at various scales, locations and aspect ratios) is much smaller than Haar-like features. Given a sub-window size of 20×20, there are totally 2049 MB-LBP features, this amount is about 1/20 of Haar-like features (45891). People usually select significant features from the whole feature set by Adaboost algorithm, and construct a binary classifier. Owing to the large feature set of haar-like feature, the training process usually spends too much time. The fewer number of MB-LBP feature set makes the implementation of feature selection significantly easy.

It is should be emphasized that the value of MB-LBP features is non-metric. The output of LBP operator is just a symbol for representing the binary string. In the next section, we will describe how to design the weak classifiers based on MB-LBP features, and apply the Adaboost algorithm to select significant features and construct classifier.

3   Feature Selection and Classifier Construction

Although the feature set of MB-LBP feature is much smaller than Haar-like features, it also contains much redundant information. The AdaBoost algorithm is used to select significant features and construct a binary classifier. Here, AdaBoost is adopted to solve the following three fundamental problems in one boosting procedure: (1) learning effective features from the large feature set, (2) constructing weak classifiers, each of which is based on one of the selected features, (3) boosting the weak classifiers into a stronger classifier.

3.1   AdaBoost Learning

We choose to use the version of boosting called gentle adaboost [2] due to it is simple to be implemented and numerically robust. Given a set of training examples as $(x_1, y_1), ..., (x_N, y_N)$, where $y_i \in \{+1, -1\}$ is the class label of the example $x_i \in R^n$. Boosting learning provides a sequential procedure to fit additive models of the form $F(x) = \sum_{m=1}^{M} f_m(x)$. Here $f_m(x)$ are often called weak learners, and $F(x)$ is called a strong learner. Gentle adaboost uses adaptive Newton steps for minimizing the cost

Table 1. Algorithm of Gentle AdaBoost

1. Start with weight $w_i = \frac{1}{N}, i = 1, 2, ..., N, F(x) = 0$
2. Repeat for m = 1, ... ,M
  (a) Fit the regression function by weighted least squares
    fitting of Y to X.
  (b) Update $F(x) \leftarrow F(x) + f_m(x)$
  (c) Update $w_i \leftarrow w_i e^{-y_i f_m(x_i)}$ and normalization
3. Output the Classifier $F(x) = sign[\sum_{m=1}^{M} f_m(x)]$

function: $J = E[e^{-yF(x)}]$, which corresponds to minimizing a weighted squared error at each step.

In each step, the weak classifier $f_m(x)$ is chosen so as to minimize the weighted squared error:

$$J_{wse} = \sum_{i=1}^{N} w_i(y_i - f_m(x_i))^2 \tag{2}$$

3.2 Weak Classifiers

It is common to define the weak learners $f_m(x)$ to be the optimal threshold classification function [12], which is often called a stump. However, it is indicated in Section 2 that the value of MB-LBP features is non-metric. Hence it is impossible to use threshold-based function as weak learner.

Here we describe how the weak classifiers are designed. For each MB-LBP feature, we adopt multi-branch tree as weak classifiers. The multi-branch tree totally has 256 branches, and each branch corresponds to a certain discrete value of MB-LBP features. The weak classifier can be defined as:

$$f_m(x) = \begin{cases} a_0, & x^k = 0 \\ & \dots \\ a_j, & x^k = j \\ & \dots \\ a_{255}, & x^k = 255 \end{cases} \tag{3}$$

Where x^k denotes the k-th element of the feature vector x, and a_j, $j = 0, \cdots, 255$, are regression parameters to be learned. These weak learners are often called decision or regression trees. We can find the best tree-based weak classifier (the parameter k, a_j with minimized weighted squared error as Equ.(2)) just as we would learn a node in a regression tree.The minimization of Equ.(2))gives the following parameters:

$$a_j = \frac{\sum_i w_i y_i \delta(x_i^k = j)}{\sum_i w_i \delta(x_i^k = j)} \tag{4}$$

As each weak learner depends on a single feature, one feature is selected at each step. In the test phase, given a MB-LBP feature, we can get the corresponding regression

value fast by such multi-branch tree. This function is similar to the lookup table (LUT) weak classifier for Haar-like features [1], the difference is that the LUT classifier gives a partition of real-value domain.

4   Experiments

In this section, we conduct two experiments to evaluate proposed method. (1) Comparing MB-LBP features with Haar-like features and original LBP features. (2) Evaluating the proposed detector on CMU+MIT face database.

A total of 10,000 face images were collected from various sources, covering out-of-plane and in-plan rotation in the range of $[-30°, 30°]$. For each aligned face example, four synthesized face examples were generated by following random transformation: mirroring, random shifting to +1/-1 pixel, in-plane rotation within 15 degrees and scaling within 20% variations. The face examples were then cropped and re-scaled to 20×20 pixels. Totally, we get a set of 40,000 face examples. More than 20,000 large images which do not contain faces are used for collecting non-face samples.

4.1   Feature Comparison

In this subsection, we compare the performance of MB-LBP feature with Haar-like rectangle features and conventional LBP features. In the experiments, we use 26,000 face samples and randomly divide them to two equally parts, one for training the other for testing. The non-face samples are randomly collected from large images which do not contain faces. Our training set contains 13,000 face samples and 13,000 non-face samples, and the testing set contains 13,000 face samples and 50,000 non-face samples.

Based on Adaboost learning framework, three boosting classifiers are trained. Each of them contains selected 50 Haar-like features, conventional LBP features and MB-LBP features, respectively. Then they are evaluated on the test set. Fig. 4(a) shows the curves of the error rate (average of false alarm rate and false rejection rate) as a function of the number of the selected features in the training procedure. We can see the curve
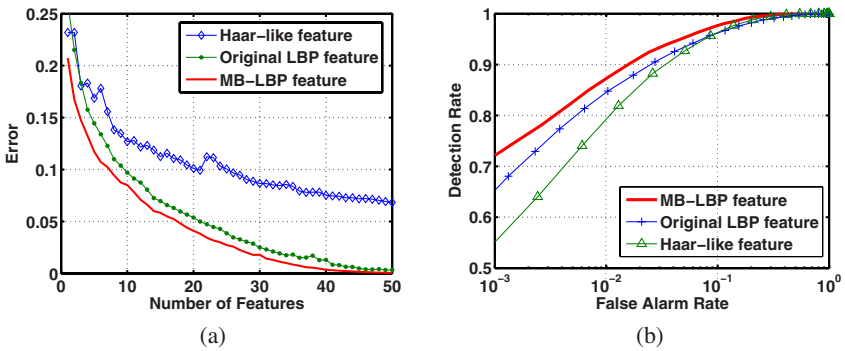
Fig. 4. Comparative results with MB-LBP features, Haar-like features and original LBP features. (a) The curves show the error rate as a function of the selected features in training process. (b) The ROC curves show the classification performance of the three classifiers on the test set.

corresponding to MB-LBP features has the lowest error rate. It indicates that the weak classifiers based on MB-LBP features are more discriminative. The ROC curves of the three classifiers on the test set can be found in Fig. 4(b). It is shown that in the given false alarm rate at 0.001, classifier based on MB-LBP features shows 15% higher correct rate than haar-like feature and 8% higher than original LBP feature. All the above shows the distinctive of MB-LBP features. It is mainly because the MB-LBP features can capture more information about the image structures.

4.2   Experimental Results on CMU+MIT Face Set

We trained a cascade face detector based on MB-LBP features and tested it on the MIT+CMU database which is widely used to evaluate the performance of face detection algorithm. This set consists of 130 images with 507 labeled frontal faces. For training the face detector, all collected 40,000 face samples are used, the bootstrap strategy is also used to re-collect non-face samples. Our trained detector has 9 layers including 470 MB-LBP features.Comparing with the Viola's cascade detector [12] which has 32 layers and 4297 features, our MB-LBP feature is much more efficient. From the results, we can see that our method get considerable performance with fewer features. The processing time of our detector for a 320x240 image is less than 0.1s on a P4 3.0GHz PC.

Table 2. Experimental results on MIT+CMU set

False Alarms	6	10	21	31	57	78	136	167	293	422
Ours	80.1%	-	85.6%	-	90.7%	-	91.9%	-	93.5%	-
Viola	-	78.3%	-	85.2%	-	90.1%	-	91.8%	-	93.7%
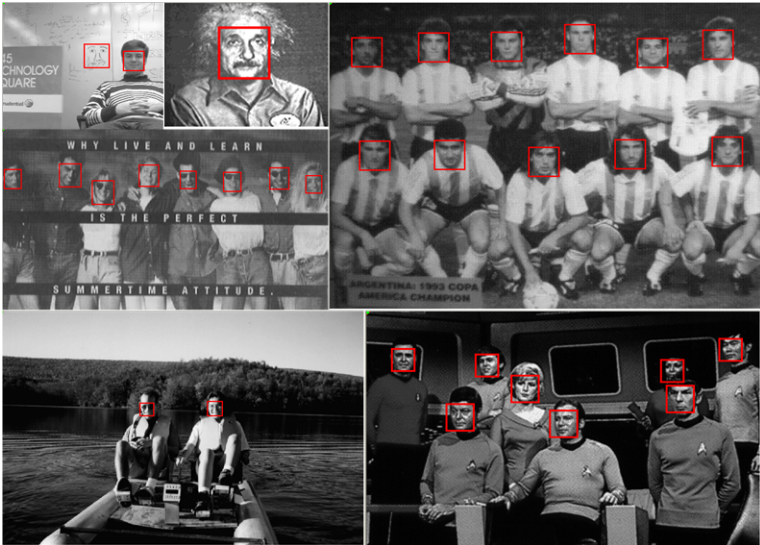
Fig. 5. Some detection results on MIT+CMU set

5   Conclusions

In this paper, we proposed multi-block local binary pattern(MB-LBP) features as descriptor for face detection. A boosting-based detector is implemented. Aims at the non-metric feature value of MB-LBP features, multi-branch regression tree is adopted to construct the weak classifiers. First, these features can capture more information about image structure than traditional Haar-like features and show more distinctive performance. Second, fewer feature number of the completed feature set makes the training process easier. In our experiment, it is shown that at the given false alarm rate 0.001, MB-LBP shows 15% higher correct rate than Haar-like feature and 8% higher than original LBP feature. Moreover, our face detector gets considerable performance on CMU+MIT database with fewer features.

Acknowledgements

References

1. Wu, B., Ai, H.Z., Huang, C., Lao, S.H.: Fast rotation invariant multi-view face detection based on real adaboost. In: FG (2004)
2. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: A statistical view of boosting. Annals of Statistics (2000)
3. Froba, B., Ernst, A.: Face detection with the modified census transform. In: AFGR (2004)
4. Huang, C., Ai, H., Li, Y., Lao, S.: Learning sparse features in granular space for multi-view face detection. In: IEEE International conference on Automatic Face and Gesture Recognition, April 2006. IEEE Computer Society Press, Los Alamitos (2006)
5. Li, S.Z., Zhu, L., Zhang, Z.Q.: Statistical learning of multi-view face detection. In: Tistarelli, M., Bigun, J., Jain, A.K. (eds.) ECCV 2002. LNCS, vol. 2359. Springer, Heidelberg (2002)
6. Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. In: ICIP (2002)
7. Mita, T., Kaneko, T., Hori, O.: Joint haar-like features for face detection. In: ICCV (2005)
8. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition (January 1996)
9. Osuna, E., Freund, R., Girosi, F.: Training support vector machines: an application to face detection. In: CVPR (1997)
10. Rowley, H.A., Baluja, S., Kanade, T.: Neural network-based face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence (1998)
11. Simard, P.Y., Cun, Y.A.L., Denker, J.S., Victorri, B.: Transformation invariance in pattern recognition - tangent distance and tangent propagation. Neural Networks: Tricks of the Trade (1998)
12. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Press, Los Alamitos (2001)