

# Learning Multi-scale Block Local Binary Patterns for Face Recognition

Shengcai Liao, Xiangxin Zhu, Zhen Lei, Lun Zhang, 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  
{scliao, xxzhu, zlei, lzhang, szli}@nlpr.ia.ac.cn  
<http://www.cbsr.ia.ac.cn>

**Abstract.** In this paper, we propose a novel representation, called Multi-scale Block Local Binary Pattern (MB-LBP), and apply it to face recognition. The Local Binary Pattern (LBP) has been proved to be effective for image representation, but it is too local to be robust. In MB-LBP, the computation is done based on average values of block subregions, instead of individual pixels. In this way, MB-LBP code presents several advantages: (1) It is more robust than LBP; (2) it encodes not only microstructures but also macrostructures of image patterns, and hence provides a more complete image representation than the basic LBP operator; and (3) MB-LBP can be computed very efficiently using integral images. Furthermore, in order to reflect the uniform appearance of MB-LBP, we redefine the uniform patterns via statistical analysis. Finally, AdaBoost learning is applied to select most effective uniform MB-LBP features and construct face classifiers. Experiments on Face Recognition Grand Challenge (FRGC) ver2.0 database show that the proposed MB-LBP method significantly outperforms other LBP based face recognition algorithms.

**Keywords:** LBP, MB-LBP, Face Recognition, AdaBoost.

## 1 Introduction

Face recognition from images has been a hot research topic in computer vision for recent two decades. This is because face recognition has potential application values as well as theoretical challenges. Many appearance-based approaches have been proposed to deal with face recognition problems. Holistic subspace approach, such as PCA [14] and LDA [3] based methods, has significantly advanced face recognition techniques. Using PCA, a face subspace is constructed to represent “optimally” only the face; using LDA, a discriminant subspace is constructed to distinguish “optimally” faces of different persons. Another approach is to construct a local appearance-based feature space, using appropriate image filters, so the distributions of faces are less affected by various changes. Local features analysis (LFA) [10], Gabor wavelet-based features [16,6] are among these.

Recently, Local Binary Patterns (LBP) is introduced as a powerful local descriptor for microstructures of images [8]. The LBP operator labels the pixels of an image by thresholding the  $3 \times 3$ -neighborhood of each pixel with the center value and considering the result as a binary string or a decimal number. Recently, Ahonen et al proposed a novel approach for face recognition, which takes advantage of the Local Binary Pattern (LBP) histogram [1]. In their method, the face image is equally divided into small sub-windows from which the LBP features are extracted and concatenated to represent the local texture and global shape of face images. Weighted Chi square distance of these LBP histograms is used as a dissimilarity measure of different face images. Experimental results showed that their method outperformed other well-known approaches such as PCA, EBGm and BIC on FERET database. Zhang et al [17] propose to use AdaBoost learning to select best LBP sub-window histograms features and construct face classifiers.

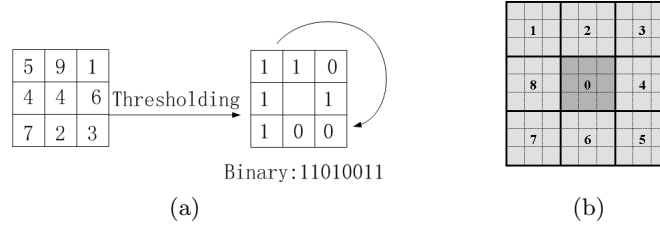
However, the original LBP operator has the following drawback in its application to face recognition. It has its small spatial support area, hence the bit-wise comparison therein made between two single pixel values is much affected by noise. Moreover, features calculated in the local  $3 \times 3$  neighborhood cannot capture larger scale structure (macrostructure) that may be dominant features of faces.

In this work, we propose a novel representation, called Multi-scale Block LBP (MB-LBP), to overcome the limitations of LBP, and apply it to face recognition. In MB-LBP, the computation is done based on average values of block subregions, instead of individual pixels. This way, MB-LBP code presents several advantages: (1) It is more robust than LBP; (2) it encodes not only microstructures but also macrostructures of image patterns, and hence provides a more complete image representation than the basic LBP operator; and (3) MB-LBP can be computed very efficiently using integral images. Considering this extension, we find that the property of the original uniform LBP patterns introduced by Ojala et al [9] can not hold to be true, so we provide a definition of statistically effective LBP code via statistical analysis. While a large number of MB-LBP features result at multiple scales and multiple locations, we apply AdaBoost learning to select most effective uniform MB-LBP features and thereby construct the final face classifier.

The rest of this paper is organized as follows: In Section 2, we introduce the MB-LBP representation. In Section 3, the new concept of uniform patterns are provided via statistical analysis. In Section 4, A dissimilarity measure is defined to discriminate intra/extrapersonal face images, and then we apply the AdaBoost learning for MB-LBP feature selection and classifier construction. The experiment results are given in Section 5 with the FRGC ver2.0 data sets [11]. Finally, we summarize this paper in Section 6.

## 2 Multi-scale Block Local Binary Patterns

The original LBP operator labels the pixels of an image by thresholding the  $3 \times 3$ -neighborhood of each pixel with the center value and considering the result as

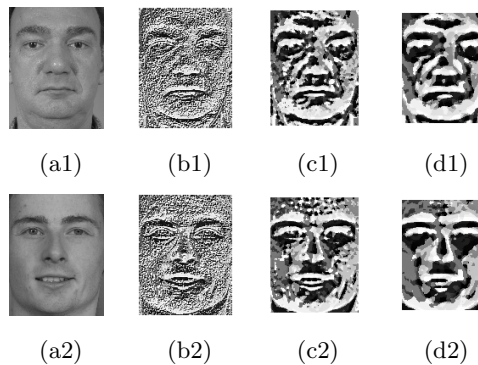


**Fig. 1.** (a) The basic LBP operator. (b) The  $9 \times 9$  MB-LBP operator. In each sub-region, average sum of image intensity is computed. These average sums are then thresholded by that of the center block. MB-LBP is then obtained.

a binary string or a decimal number. Then the histogram of the labels can be used as a texture descriptor. An illustration of the basic LBP operator is shown in Fig. 1(a). Multi-scale LBP [9] is an extension to the basic LBP, with respect to neighborhoods of different sizes.

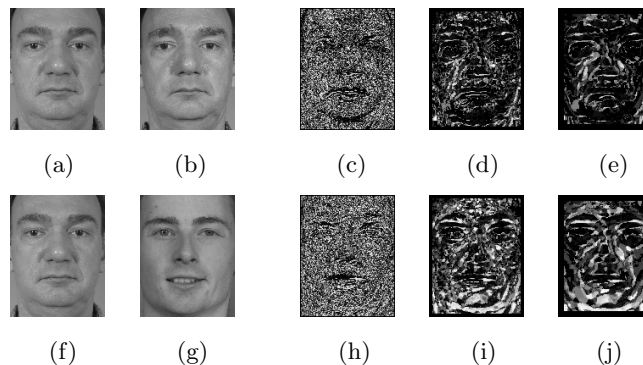
In MB-LBP, the comparison operator between single pixels in LBP is simply replaced with comparison between average gray-values of sub-regions (*cf.* Fig. 1(b)). Each sub-region is a square block containing neighboring pixels (or just one pixel particularly). The whole filter is composed of 9 blocks. We take the size  $s$  of the filter as a parameter, and  $s \times s$  denoting the scale of the MB-LBP operator (particularly,  $3 \times 3$  MB-LBP is in fact the original LBP). Note that the scalar values of averages over blocks can be computed very efficiently [13] from the summed-area table [4] or integral image [15]. For this reason, MB-LBP feature extraction can also be very fast: it only incurs a little more cost than the original  $3 \times 3$  LBP operator.

Fig. 2 gives examples of MB-LBP filtered face images by  $3 \times 3$ ,  $9 \times 9$  and  $15 \times 15$  blocks. From this example we can see what influence parameter  $s$  would make. For a small scale, local, micro patterns of a face structure is well represented, which may beneficial for discriminating local details of faces. On the



**Fig. 2.** MB-LBP filtered images of two different faces. (a) original images; (b) filtered by  $3 \times 3$  MB-LBP (c) filtered by  $9 \times 9$  MB-LBP; (d) filtered by  $15 \times 15$  MB-LBP.

other hand, using average values over the regions, the large scale filters reduce noise, and makes the representation more robust; and large scale information provides complementary information to small scale details. But much discriminative information is also dropped. Normally, filters of various scales should be carefully selected and then fused to achieve better performance.



**Fig. 3.** Differential images. (a)(b) are the original intra-personal images, (c)(d)(e) are the corresponding differential images generated by  $3 \times 3$ ,  $9 \times 9$  and  $15 \times 15$  MB-LBP operators. And (f)(g) are the original extra-personal images, (h)(i)(j) are the corresponding differential images generated by  $3 \times 3$ ,  $9 \times 9$  and  $15 \times 15$  MB-LBP operators.

Fig. 3 demonstrates MB-LBP features for intra-personal and extra-personal difference images (brighter pixels indicate greater difference). Using differential images, We also have another way to demonstrate the discriminative power of MB-LBP. We see that when scale of the MB-LBP filter become larger, thought both intra-personal variance and extra-personal variance decrease, it is more clear that with larger scale MB-LBP, intra-personal differences are smaller than that of the extra-personal.

### 3 Statistically Effective MB-LBP (SEMB-LBP)

Using the “uniform” subset of LBP code improves the performance of LBP based methods. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular [9]. It is observed that there are a limited number of transitions or discontinuities in the circular presentation of the  $3 \times 3$  texture patterns; according to Ojala, this uniform patterns are fundamental properties of local image texture, they provide a vast majority amount all patterns: 90% for (8,2) type LBP, 70% for (16,2) type [9].

However, the original definition of uniform LBP patterns based on the transition of pixel values can not be used for MB-LBP with blocks containing more

than a single pixel. The reason is obvious: the same properties of circular continuities in  $3 \times 3$  patterns can not hold to be true when parameter  $s$  becomes larger. Using average gray-values instead of single pixels, MB-LBP reflects the statistical properties of local sub-region relationship. The larger parameter  $s$  becomes, the harder for the circular presentation to be continuous.

So we need to redefine uniform patterns. Since the term uniform refers to the uniform appearance of the local binary pattern, we can define our uniform MB-LBP via statistical analysis. Here, we present the concept of Statistically Effective MB-LBP (SEMB-LBP), based on the percentage in distributions, instead of the number of 0-1 and 1-0 transitions as in the uniform LBP.

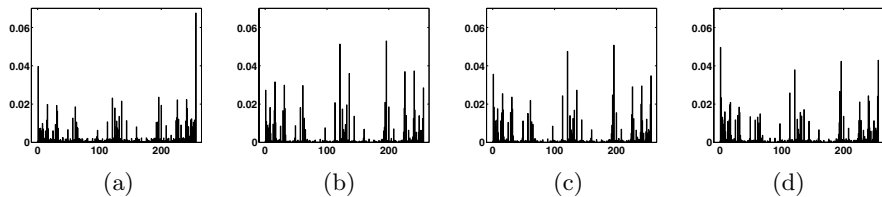
Denote  $f_s(x, y)$  as an MB-LBP feature of scale  $s$  at location  $(x, y)$  computed from original images. Then a histogram of the MB-LBP feature  $f_s(\cdot, \cdot)$  over a certain image  $I(x, y)$  can be defined as:

$$H_s(\ell) = 1_{[f_s(x,y)=\ell]}, \quad \ell = 0, \dots, L - 1 \quad (1)$$

where  $1_{(S)}$  is the indicator of the set  $S$ , and  $\ell$  is the label of the MB-LBP code. Because all MB-LBP code are 8-bit binary string, so there are total  $L = 2^8 = 256$  labels. Thus the histogram has 256 bins.

This histogram contains information about the distribution of the MB-LBP features over the whole image. The histograms, with the large still training set of the FRGC ver2.0, for the scales  $3 \times 3$ ,  $9 \times 9$ ,  $15 \times 15$ , and  $21 \times 21$  are shown in Fig. 4. We sort the bins of a histogram according to its percentage in the histogram. A statistical analysis shows the following:

- For the  $3 \times 3$  LBP operator, the top 58 bins correspond to the so-called uniform patterns;
- However, for MB-LBP filters with block containing more than one pixel, the top 58 bins are not the same as those for the  $3 \times 3$  LBP filters.



**Fig. 4.** Histograms of MB-LBP labels on the large still training set of the FRGC ver2.0. (a)  $3 \times 3$ , (b)  $9 \times 9$ , (c)  $15 \times 15$ , and (d)  $21 \times 21$ .

To reflect the uniform appearance of MB-LBP, we define the statistically effective MB-LBP (SEMB-LBP) set of scale  $s$  as follows:

$$SEMB-LBP_s = \{\ell | Rank[H_s(\ell)] < N\} \quad (2)$$

where  $Rank[H_s(\ell)]$  is the index of  $H_s(\ell)$  after descending sorting, and  $N$  is the number of uniform patterns. For the original uniform local binary patterns,  $N$

= 58. However, in our definition,  $N$  can be assigned arbitrarily from 1 to 256. Large value of  $N$  will cause the feature dimensions very huge, while small one loses feature variety. Consequently, we adopt  $N = 63$  for a trade-off. Labeling all remaining patterns with a single label, we use the following notation to represent SEMB-LBP:

$$u_s(x, y) = \begin{cases} Index_s[f_s(x, y)], & \text{if } f_s(x, y) \in SEMB-LBP_s \\ N, & \text{otherwise} \end{cases} \quad (3)$$

where  $Index_s[f_s(x, y)]$  is the index of  $f_s(x, y)$  in the set  $SEMB-LBP_s$  (started from 0).

#### 4 AdaBoost Learning

The above SEMB-LBP provides an over-complete representation. The only question remained is how to use them to construct a powerful classifier. Because those excessive measures contain much redundant information, a further processing is needed to remove the redundancy and build effective classifiers. In this paper we use Gentle AdaBoost algorithm [5] to select the most effective SEMB-LBP feature.

Boosting can be viewed as a stage-wise approximation to an additive logistic regression model using Bernoulli log-likelihood as a criterion [5]. Developed by Friedman et al, Gentle AdaBoost modifies the population version of the Real AdaBoost procedure [12], using Newton stepping rather than exact optimization at each step. Empirical evidence suggests that Gentle AdaBoost is a more conservative algorithm that has similar performance to both the Real AdaBoost and LogitBoost algorithms, and often outperforms them both, especially when stability is an issue.

Face recognition is a multi-class problem whereas the above AdaBoost learning is for two classes. To dispense the need for a training process for faces of a newly added person, we use a large training set describing intra-personal or extra-personal variations [7], and train a “universal” two-class classifier. An ideal intra-personal difference should be an image with all pixel values being zero, whereas an extra-personal difference image should generally have much larger pixel values. However, instead of deriving the intra-personal or extra-personal variations using difference images as in [7], the training examples to our learning algorithm is the set of differences between each pair of local histograms at the corresponding locations. The positive examples are derived from pairs of intra-personal differences and the negative from pairs of extra-personal differences.

In this work, a weak classifier is learned based on a dissimilarity between two corresponding histogram bins. Once the SEMB-LBP set are defined for each scale, histogram of these patterns are computed for calculating the dissimilarity: First, a sequence of  $m$  subwindows  $R_0, R_1, \dots, R_{m-1}$  of varying sizes and locations are obtained from an image; Second, a histogram is computed for each SEMB-LBP code  $i$  over a subwindow  $R_j$  as

$$H_{s,j}(i) = 1_{[u_s(x,y)=i]} \cdot 1_{[(x,y) \in R_j]}, \quad i = 0, \dots, N, j = 0, \dots, m-1 \quad (4)$$

The corresponding bin difference is defined as

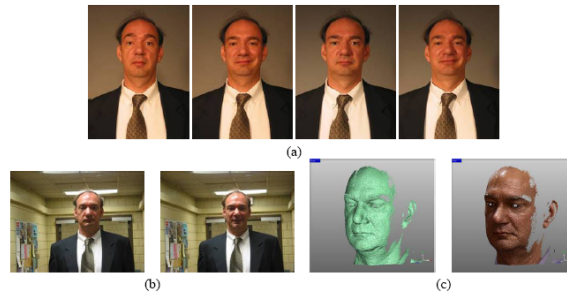
$$D(H_{s,j}^1(i), H_{s,j}^2(i)) = |H_{s,j}^1(i) - H_{s,j}^2(i)| \quad i = 0, \dots, N \quad (5)$$

The best current weak classifier is the one for which the weighted intra-personal bin differences (over the training set) are minimized while that of the extra-personal are maximized.

With the two-class scheme, the face matching procedure will work in the following way: It takes a probe face image and a gallery face image as the input, and computes a difference-based feature vector from the two images, and then it calculates a similarity score for the feature vector using the learned AdaBoost classifier. Finally, A decision is made based on the score, to classify the feature vector into the positive class (coming from the same person) or the negative class (different persons).

## 5 Experiments

The proposed method is tested on the Face Recognition Grand Challenge (FRGC) ver2.0 database [11]. The Face Recognition Grand Challenge is a large face recognition evaluation data set, which contains over 50,000 images of 3D scans and high resolution still images. Fig. 5 shows a set of images in FRGC for one subject section. The FRGC ver2.0 contains a large still training set for training still face recognition algorithms. It consists of 12,776 images from 222 subjects. A large validation set is also provided for FRGC experiments, which contains 466 subjects from 4,007 subject sessions.



**Fig. 5.** FRGC images from one subject session. (a) Four controlled still images, (b) two uncontrolled stills, and (c) 3D shape channel and texture channel pasted on the corresponding shape channel.

The proposed method is evaluated on two 2D experiments of FRGC: Experiment 1 and Experiment 2. Experiment 1 is designed to measure face recognition performance on frontal facial images taken under controlled illumination. In this experiment, only one single controlled still image is contained in one biometric sample of the target or query sets. Experiment 2 measures the effect of multiple

still images on performance. In Experiment 2, each biometric sample contains four controlled images of a person taken in a subject session. There are 16 matching scores between one target and one query sample. These scores are averaged to give the final result.

There are total 16028 still images both in target and query sets of experiment 1 and 2. In the testing phase, it will generate a large similarity matrix of 16028 by 16028. Furthermore, three masks are defined over this similarity matrix, thus three performance results will be obtained, corresponding to Mask I, II, and III. In mask I all samples are within semesters, in mask II they are within a year, while in mask III the samples are between semesters, i.e., they are of increasing difficulty.

In our experiments, All images are cropped to 144 pixels high by 112 pixels wide, according to the provided eyes positions. A boosting classifier is trained on the large still training set. The final strong classifier contains 2346 weak classifiers, e.g. 2346 bins of various sub-region SEMB-LBP histograms, while achieving zero error rate on the training set. The corresponding MB-LBP filter size of the first 5 learned weak classifiers are  $s = 21, 33, 27, 3, 9$ . We find that a middle level of scale  $s$  has a better discriminative power.

To compare the performances of LBP based methods, we also evaluate Ahonen et al's LBP method [1] and Zhang et al's Boosting LBP algorithm [17]. We use  $LBP_{8,2}^{u2}$  operator as Ahonen's, and test it directly on FRGC experiment 1 and 2 since their method needs no training. For Zhang's approach, we train an AdaBoost classifier on the large still training set in the same way of our MB-LBP. The final strong classifier of Zhang's contains 3072 weak classifiers, yet it can not achieve zero error rate on the training set.

The comparison results are given in Fig. 6, and Fig. 7 describes the verification performances on a receiver operator characteristic(ROC) curve (mask III). From the results we can see that MB-LBP outperforms the other two algorithms on all experiments. The comparison proves that MB-LBP is a more robust representation than the basic LBP. Meanwhile, because MB-LBP provides a more complete image representation that encodes both microstructure and macrostructure, it can achieve zero error rate on the large still training set, while generate well on the validation face images. Furthermore, we can also find that MB-LBP performs well on Mask III, which means that it is robust with time elapse.

|              | Experiment 1 |              |              | Experiment 2 |              |              |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|              | Mask I       | Mask II      | Mask III     | Mask I       | Mask II      | Mask III     |
| MB-LBP       | <b>98.07</b> | <b>97.04</b> | <b>96.05</b> | <b>99.78</b> | <b>99.58</b> | <b>99.45</b> |
| Zhang's LBP  | 84.17        | 80.35        | 76.67        | 97.67        | 96.73        | 95.84        |
| Ahonen's LBP | 82.72        | 78.65        | 74.78        | 93.62        | 90.54        | 87.46        |

**Fig. 6.** Verification performance (FAR=0.1%) on FRGC Experiment 1 and 2



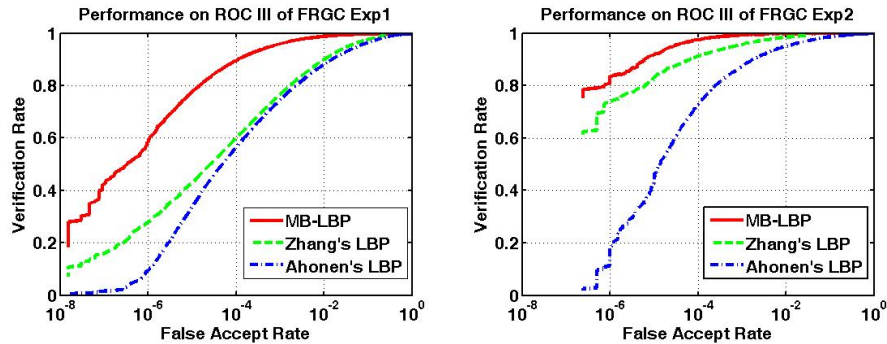


Fig. 7. ROC performance on FRGC Experiment 1 and 2

## 6 Summary and Conclusions

In this paper, we present a Multi-scale Block Local Binary Pattern (MB-LBP) based operator for robust image representation. The Local Binary Pattern (LBP) has been proved to be effective for image representation, but it is too local to be robust. The Multi-scale Block Local Binary Patterns (MB-LBP) use sub-region average gray-values for comparison instead of single pixels. Feature extraction for MB-LBP is very fast using integral images. Considering this extension, uniform patterns may not remain the same as those defined by Ojala et al [9]. To reflect the uniform appearance of the MB-LBP, we define our Statistically Effective MB-LBP (SEMB-LBP) via statistical analysis. Simply using absolute difference between the same bin of two histograms as dissimilarity measure, we finally apply AdaBoost learning to select the most effective weak classifiers and construct a powerful classifier. Experiments on FRGC ver2.0 show that MB-LBP significantly outperforms other LBP based method.

Since MB-LBP can be viewed as a certain way of combination using 8 ordinal rectangle features, our future work may be combining eight or more ordinal features [2] in a circular instead of rectangle features. While ordinal filters are sufficient and robust for comparison of image regions, we expect that this kind of combination will be more powerful for image representation.

**Acknowledgements.** This work was supported by the following funding resources: National Natural Science Foundation Project #60518002, National Science and Technology Supporting Platform Project #2006BAK08B06, National 863 Program Projects #2006AA01Z192 and #2006AA01Z193, Chinese Academy of Sciences 100-people project, and the AuthenMetric Collaboration Foundation.

## References

1. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Proceedings of the European Conference on Computer Vision, Prague, Czech, pp. 469–481 (2004)

2. Balas, B., Sinha, P.: Toward dissociated dipoles: Image representation via non-local comparisons. In: CBCL Paper #229/AI Memo #2003-018, MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, USA (August 2003)
3. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 711–720 (1997)
4. Crow, F.: Summed-area tables for texture mapping. In: SIGGRAPH, vol. 18(3), pp. 207–212 (1984)
5. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting. Technical report, Department of Statistics, Sequoia Hall, Stanford University (July 1998)
6. Liu, C., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing* 11(4), 467–476 (2002)
7. Moghaddam, B., Nastar, C., Pentland, A.: A Bayesian similarity measure for direct image matching. Media Lab Tech Report No.393, MIT (August 1996)
8. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 29(1), 51–59 (1996)
9. Ojala, T., Pietikainen, M., Maenpaa, M.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002)
10. Penev, P., Atick, J.: Local feature analysis: A general statistical theory for object representation. *Neural Systems* 7(3), 477–500 (1996)
11. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press, Los Alamitos (2005)
12. Schapire, R.E., Singer, Y.: Improved boosting algorithms using confidence-rated predictions. In: *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pp. 80–91 (1998)
13. Simard, P.Y., Bottou, L., Haffner, P., Cun, Y.L.: Boxlets: a fast convolution algorithm for signal processing and neural networks. In: Kearns, M., Solla, S., Cohn, D. (eds.) *Advances in Neural Information Processing Systems*, vol. 11, pp. 571–577. MIT Press, Cambridge (1998)
14. Turk, M.A., Pentland, A.P.: Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3(1), 71–86 (1991)
15. Viola, P., Jones, M.: Robust real time object detection. In: *IEEE ICCV Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada, July 13, 2001 (2001)
16. Wiskott, L., Fellous, J., Kruger, N., Malsburg, C.v.d.: Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 775–779 (1997)
17. Zhang, G., Huang, X., Li, S.Z., Wang, Y., Wu, X.: Boosting local binary pattern (LBP)-based face recognition. In: Li, S.Z., Lai, J.-H., Tan, T., Feng, G.-C., Wang, Y. (eds.) *SINOBIOMETRICS 2004*. LNCS, vol. 3338, pp. 180–187. Springer, Heidelberg (2004)