Pattern Recognition Letters 33 (2012) 431–437

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Learning local binary patterns for gender classification on real-world face images

Caifeng Shan

Philips Research, Eindhoven, The Netherlands

ARTICLE INFO

Article history:
Available online 27 May 2011

Keywords:
Gender classification
Local binary patterns
Face image analysis

ABSTRACT

Gender recognition is one of fundamental face analysis tasks. Most of the existing studies have focused on face images acquired under controlled conditions. However, real-world applications require gender classification on real-life faces, which is much more challenging due to significant appearance variations in unconstrained scenarios. In this paper, we investigate gender recognition on real-life faces using the recently built database, the Labeled Faces in the Wild (LFW). Local Binary Patterns (LBP) is employed to describe faces, and Adaboost is used to select the discriminative LBP features. We obtain the performance of 94.81% by applying Support Vector Machine (SVM) with the boosted LBP features. The public database used in this study makes future benchmark and evaluation possible.

1. Introduction

Gender recognition is a fundamental task for human beings, as many social functions critically depend on the correct gender perception. Automatic gender classification has many important applications, for example, intelligent user interface, visual surveillance, collecting demographic statistics for marketing, etc. Human faces provide important visual information for gender perception. Gender classification from face images has received much research interest in the last two decades.

In the early 1990s various neural network techniques were employed to recognize gender by frontal faces (Golomb et al., 1991; Brunelli and Poggio, 1992), for example, Golomb et al. (1991) trained a fully connected two-layer neural network, SEXNET, which achieves the recognition accuracy of 91.9% on 90 face images. Recent years have witnessed many advances (Yang et al., 2006); we summarize recent studies in Table 1. Moghaddam and Yang (2002) used raw image pixels with nonlinear SVMs for gender classification on thumbnail faces (12×21 pixels); their experiments on the FERET database (1,755 faces) demonstrated SVMs are superior to other classifiers, achieving the accuracy of 96.6%. In BenAbdelkader and Griffin (2005), local region matching and holistic features were exploited with Linear Discriminant Analysis (LDA) and SVM for gender recognition. On the 12,964 frontal faces from multiple databases (including FERET and PIE), local region-based SVM achieved the performance of 94.2%. Lapedriza et al. (2006) compared facial features from internal zone (eyes, nose, and mouth) and external zone (hair, chin, and ears). Their experiments on the FRGC database show that the external face zone contributes useful information for gender classification. Baluja and Rowley (2007) introduced an efficient gender recognition system by boosting pixel comparisons in face images. On the FERET database, their approach matches SVM with 500 comparison operations on 20×20 pixel images. Mäkinen and Raisamo (2008) systematically evaluated different face alignment and gender recognition methods on the FERET database. More recently face appearance and motion cues are combined for gender recognition in videos (Hadid and Pietikäinen, 2009).

A common problem of the above studies is that face images acquired under controlled conditions (e.g., the FERET database) are considered, which usually are frontal, occlusion-free, with clean background, consistent lighting, and limited facial expressions. However, in real-world applications, gender classification needs to be performed on real-life face images captured in unconstrained scenarios; see Fig. 1 for examples of real-life faces. As can be observed, there are significant appearance variations on real-life faces, which include facial expressions, illumination changes, head pose variations, occlusion or make-up, poor image quality, and so on. Therefore, gender recognition in real-life faces is much more challenging compared to the case for faces captured in constrained environments. Few studies in the literature have addressed this problem. Shakhnarovich et al. (2002) made an early attempt by collecting over 3,500 face images from the web. On this difficult data set, using Harr-like features, they obtained the performance of 79.0% (Adaboost) and 75.5% (SVM). Recently Gao and Ai (2009) adopted the probabilistic boosting tree with Harr-like features, and obtained the accuracy of 95.51% on 10,100 real-life faces. However, the data sets used in these studies are not public available; therefore, it is difficult for benchmark in research community. Kumar et al. (2008, 2009) recently investigated face verification on real-world images, where many binary "attribute" classifiers (including gender) were trained. They reported the performance of 81.22% on gender classification; however, as they mainly focused on face verification, they did not fully study gender

E-mail address: caifeng.shan@philips.com

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.05.016

Table 1
Overview of recent studies on gender classification from face images.

Study	Data set			Approach		Result (%)
	Data	Real-Life	Public	Feature	Classifier	
2002(Moghaddam and Yang, 2002)	1,755	No	Yes	Raw pixels	SVM	96.62
2002(Shakhnarovich et al., 2002)	3,500	**Yes**	**No**	Haar-like features	Adaboost	79.0
2005(BenAbdelkader and Griffin, 2005)	12,964	No	Yes	Local-region matching	SVM	94.2
2006(Lapedriza et al., 2006)	5,326	No	Yes	Fragment-based filter banks	Boosting	91.72
2007(Baluja and Rowley, 2007)	2,409	No	Yes	Pixel comparisons	Adaboost	94.3
2008(Mäkinen and Raisamo, 2008)	500	No	Yes	Raw pixels	SVM	86.54
2009(Hadid and Pietikäinen, 2009)	4,000 videos	No	Yes	LBP features	SVM	91.0
2009(Gao and Ai, 2009)	10,100	**Yes**	**No**	Haar-like features	Probabilistic Boosting tree	95.51
Our work	7,443	**Yes**	**Yes**	Boosted LBP features	SVM	94.81

recognition on real-life faces. In this paper, we use a recently built public database, the Labeled Faces in the Wild (LFW) (Huang et al., 2007), to investigate gender classification on real-world face images. The public database used in this study enables future benchmark and evaluation.

Similar to other face analysis tasks, deriving an effective facial representation from original face images is a vital step for successful gender classification. If inadequate features are used, even the best classifier could fail to achieve accurate recognition. As an efficient non-parametric method summarizing the local structure of an image, Local Binary Patterns (LBP) has been exploited for face analysis (Ahonen et al., 2004). For example, in Sun et al. (2006), Lian and Lu (2007), LBP was exploited for gender recognition on face images acquired under controlled conditions. In the existing work, LBP histograms are extracted from local facial regions as the region-level description, where the n-bin histogram is utilized as a whole. However, not all bins in the LBP histogram are necessary to contain useful information for facial representation. In this paper, we propose to learn discriminative LBP-Histogram (LBPH) bins for gender classification. Our experiments illustrate that the selected LBPH bins provide a compact facial representation, reducing feature length greatly, while producing better recognition performance. By adopting SVM with the selected LBPH bins, we obtain the recognition performance of 94.81% on the LFW database.

The paper is structured as follows. Section 2 describes local binary patterns. In Section 3, learning LBPH bin using Adaboost is

Fig. 1. Examples of real-life faces (from the LFW database). (*top 2 rows*) Female; (*bottom 2 rows*) Male.

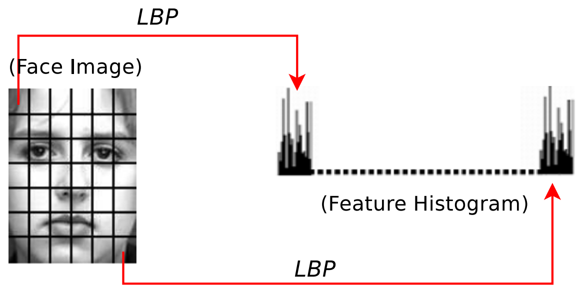

Fig. 2. A face image is divided into sub-regions from which LBP histograms are extracted and concatenated into a single, spatially enhanced feature histogram.

discussed. Section 4 presents our extensive experiments. Finally Section 5 concludes the paper.

2. Local binary patterns

The original LBP operator (Ojala et al., 2002) labels the pixels of an image by thresholding a 3×3 neighborhood of each pixel with the center value and considering the results as a binary number. Formally, given a pixel at (x_c, y_c), the resulting LBP can be expressed in the decimal form as

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^n \tag{1}$$

where n runs over the 8 neighbors of the central pixel, i_c and i_n are the gray-level values of the central pixel and the surrounding pixel, and $s(x)$ is 1 if $x \geqslant 0$ and 0 otherwise.

Ojala et al. (2002) later made two extensions of the original operator. Firstly, the operator was extended to use neighborhood of different sizes, to capture dominant features at different scales. Using circular neighborhoods and bilinearly interpolating the pixel values allow any radius and number of pixels in the neighborhood. The notation (P,R) denotes a neighborhood of P equally spaced sampling points on a circle of radius of R. Secondly, they proposed to use a small subset of the 2^P patterns, produced by the operator LBP(P,R), to describe the texture of images. These patterns, called *uniform patterns*, contain at most two bitwise transitions from 0 to 1 or vice versa when considered as a circular binary string. For example, 00000000, 001110000 and 11100001 are uniform patterns. It was observed that most of the texture information was contained in the uniform patterns. Labeling the patterns which have more than 2 transitions with a single label yields an LBP operator, denoted LBP$(P,R,u2)$, which produces much less patterns without losing too much information.

After labeling an image with a LBP operator, a histogram of the labeled image can be used as texture descriptor. Each face image can be seen as a composition of micro-patterns which can be effectively described by LBP. In the existing studies (Ahonen et al., 2004), to consider the shape information, face images are divided into non-overlapping sub-regions (as shown in Fig. 2); the LBP histograms extracted from sub-regions are concatenated into a single, spatially enhanced feature histogram The extracted feature histogram describes the local texture and global shape of face images.

The limitations of the above LBP-based facial representation are that dividing the face into a grid of sub-regions is somewhat arbitrary, as sub-regions are not necessary well aligned with facial features, and that the resulting facial representation suffers from fixed size and position of sub-regions. In Zhang et al. (2004), Sun et al. (2006), Adaboost was used to learn the discriminative sub-regions (in term of LBP histogram) from a large pool of sub-regions generated by shifting and scaling a sub-window over face images. In these studies, the Chi square distance between corresponding LBP histograms of the sample image and the template is used to construct the weak classifier.

3. Learning LBP-histogram bins

In the existing work, the LBP histograms are always extracted from local regions, and used as a whole for the regional description. However, not all bins in the LBP histogram are discriminative for facial representation. Here we propose to learn discriminative LBP-Histogram (LBPH) bins for better gender classification.

Adaboost (Freund and Schapire, 1997; Schapire and Singer, 1999) provides a simple yet effective approach for stagewise learning of a nonlinear classification function. Here we adopt Adaboost to learn the discriminative LBPH bins. Adaboost learns a small number of weak classifiers whose performance is just better than random guessing, and boosts them iteratively into a strong classifier of higher accuracy. The process of Adaboost maintains a distribution on the training samples. At each iteration, a weak classifier which minimizes the weighted error rate is selected, and the distribution is updated to increase the weights of the misclassified samples and reduce the importance of the others. Similar to Viola and Jones (2001), the weak classifier $h_j(x)$ consists of a feature f_j which corresponds to a single LBPH bin, a threshold θ_j and a parity p_j indicating the direction of the inequality sign:

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) \leqslant p_j \theta_j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

4. Experiments

We conduct experiments on the LFW database (Huang et al., 2007). LFW is a database for studying the problem of unconstrained face recognition, which contains 13,233 color face photographs of 5,749 subjects collected from the web. All the faces were detected by the Viola-Jones face detector (Viola and Jones, 2004), and the images were centered using detected faces and scaled to the size of 250×250 pixels. We manually labeled the ground truth regarding gender for each face. The faces that are not (near) frontal, as well as those for which it is difficult to establish the ground truth, were not considered (see Fig. 3 for some examples). In our experiments, we chose 7,443 face images (2,943 females and 4,500 males); see Fig. 1 for some examples. All experimental results were obtained using the 5-fold cross-validation. We partitioned the data set into five subsets of similar size, keeping

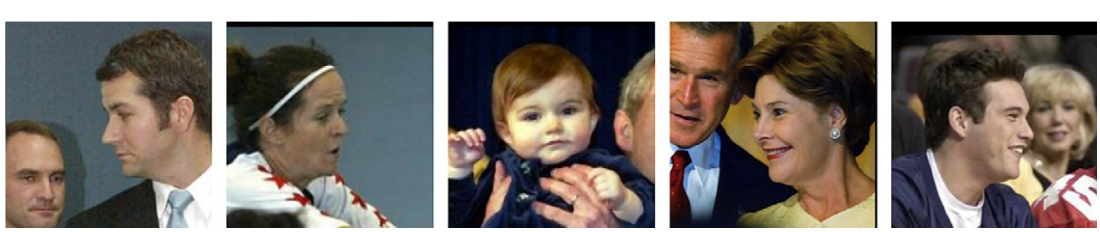

Fig. 3. Example images that are not considered.

Fig. 4. The pre-processing process on face images. (*left*) original image; (*middle*) aligned image; (*right*) cropped face.

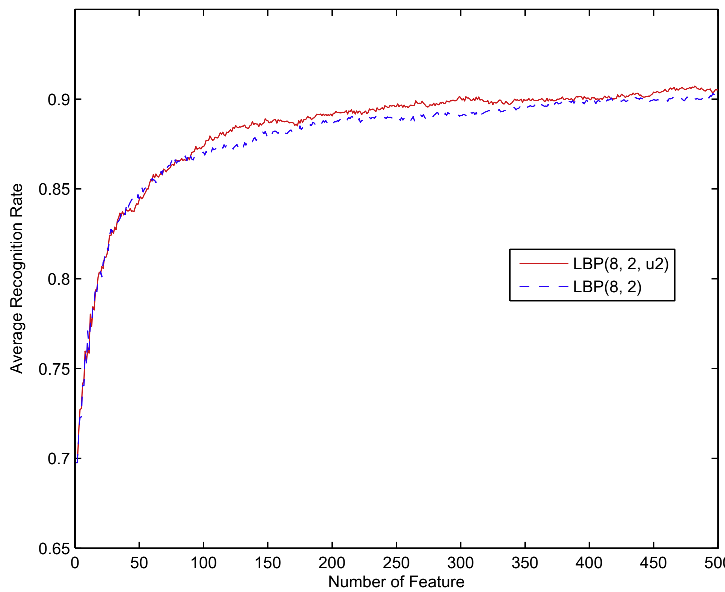

Fig. 5. Classification performance of the boosted strong classifiers, as a function of the number of feature selected: LBP$(8,2,u2)$ vs LBP$(8,2)$.

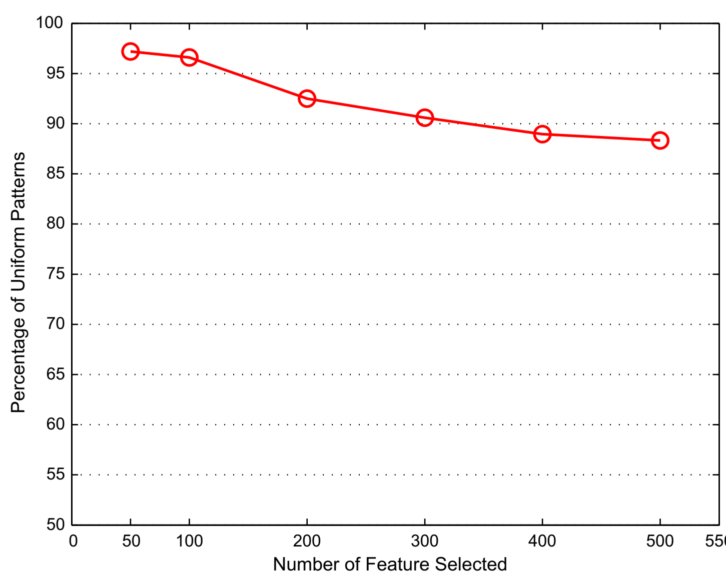

Fig. 6. The percentage of uniform patterns in the selected LBP $(8,2)$ patterns for gender classification.

the same ratio between female and male. The images of a particular subject appear only in one subset. As illustrated in Fig. 4, all images were aligned with commercial face alignment software (Wolf et al., 2009); the grayscale faces of 127×91 pixels were cropped from aligned images for use.

4.1. Experiments: limited sub-regions

In Ahonen et al. (2004), face images were divided into 42 sub-regions, and the 59-label LBP$(8,2,u2)$ operator was adopted to extract LBP features. We started our experiments with this parameter

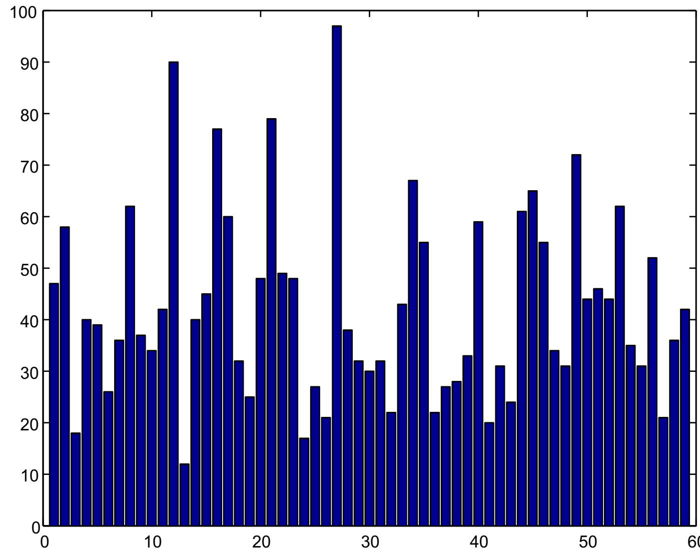

Fig. 7. The distribution of the top 500 LBPH bins selected.

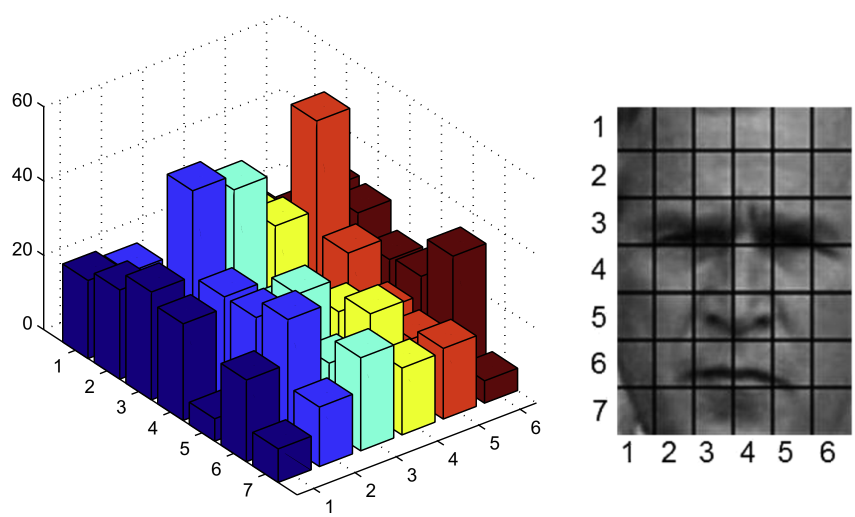

Fig. 8. Spatial distribution of the selected LBPH bins (an example face image divided in sub-regions is shown in the right side for illustration).

setting; thus each face image was described by a LBP histogram of 2,478 (42 × 59) bins. We adopted Adaboost to learn discriminative LBPH bins and boost a strong classifier. We plot in Fig. 5 the recognition performance of the boosted strong classifier as a function of the number of features selected. With the 500 selected LBPH bins, the boosted strong classifier achieves the recognition rate of 90.7%.

Uniform patterns – It was observed that most of the texture information was contained in the *uniform patterns* (Ojala et al., 2002), so uniform patterns was widely used to reduce the length of LBP histograms. Here we verify the validity of uniform patterns with machine learning. By using the LBP(8,2) operator, we represent each face image by a LBP histogram of 10,752 (42 × 256) bins, and then adopt Adaboost to learn the discriminative LBPH bins. We plot in Fig. 5 the recognition performance of the boosted strong classifier. We can see that the strong classifier of *LBP*(8,2) performs similarly to that of LBP(8,2,$u2$), which illustrates that the non-uniform patterns do not provide additional discriminative information for gender classification. To further verify this, we have a closer look at the learned LBPH bins of LBP(8,2), and show the percentage of uniform patterns in Fig. 6. It is observed that most of patterns selected are uniform patterns, e.g., 97.2% for the top 50 selected features. Therefore, we experimentally verify that most of discriminative information for gender classification is contained in the uniform patterns.

Feature distribution – We plot in Fig. 7 the distribution of the top 500 selected features among the 59 labels of LBP(8,2,$u2$). As can be observed, the selected features come from all the 59 labels, but some labels have more contributions to discriminative features. Fig. 8 shows the spatial distribution of the top 500 features selected. It is observed that, for gender classification, discriminative LBPH bins features mainly distribute in the regions around/above eyes.

Multi-scale LBP – By varying the sampling radius R, LBP of different resolutions can be obtained. We also investigate multiscale LBP for gender classification. We applied the LBP(8,R,$u2$) ($R = 1,\ldots,8$) to extract multiscale LBP features, resulting a LBP histogram of 19,824 (42 × 59 × 8) bins for each face image. We then run Adaboost to learn discriminative LBPH bins from the multiscale feature

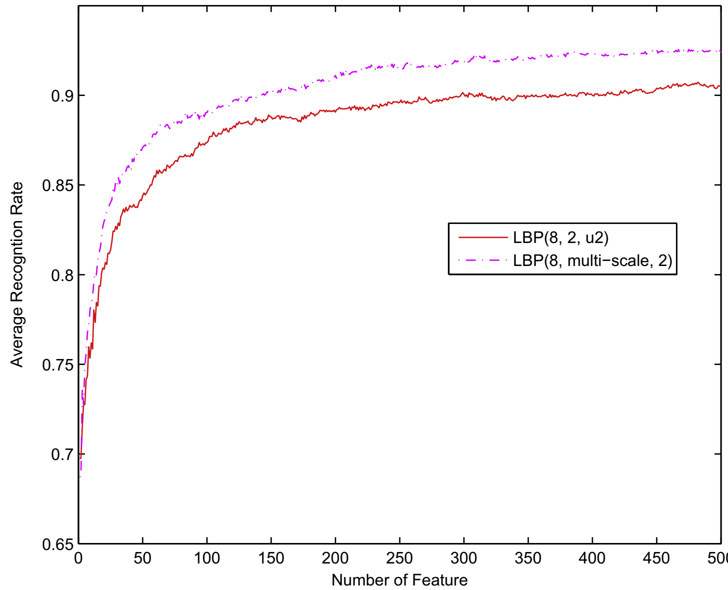

Fig. 9. Gender recognition performance of the boosted strong classifiers, as a function of the number of feature selected: LBP$(8,2,u2)$ vs LBP$(8, multiscale, u2)$.

pool. We plot in Fig. 9 the recognition performance of the boosted strong classifier. As can be observed, the strong classifier of multiscale LBP$(8,R,u2)$ $(R = 1,\ldots,8)$ produces consistently better performance than that of single scale LBP$(8,2,u2)$, providing the recognition rate of 92.6% with the 500 selected LBPH bins. Thus the multiscale LBP brings more discriminative information for gender classification. Fig. 10 shows the scale distribution of the selected LBPH bins, and we can see that discriminative LBPH bins distribute at all scales, especially scales $R = 1, 2, 3, 4, 8$. Our experimental results suggest that the multiscale LBP features bring more discriminative information for gender recognition. This reinforces the observation in facial expression recognition experiments (Shan and Gritti, 2008).

4.2. Experiments: More sub-regions

In the above experiments, only 42 equally divided sub-regions were considered for feature selection. By shifting and scaling a sub-window over face images, we can get many more sub-regions, which potentially contain more complete and discriminative information about face images. We shifted the sub-window with the shifting step of 12 pixels vertically and 10 pixels horizontally. The sub-window was scaled as 12, 18, or 24 pixels (height) and 10, 15, or 20 pixels (width) respectively. In total 725 sub-regions were obtained. By using multiscale LBP$(8,R,u2)$ $(R = 1,\ldots,8)$, a histogram of 342,200 $(725 \times 59 \times 8)$ bins was extracted from each face image.

To improve computation efficiency, we adopted a coarse to fine feature selection scheme: We first run Adaboost to select LBPH bins from each single scale LBP$(8,R,u2)$, then applied Adaboost to the selected LBPH bins at different scales to obtain the final feature selection results. We plot in Fig. 11 the recognition performance of the boosted strong classifiers as a function of the number of features selected. We can see that the boosted strong classifier of multiscale LBP provides better performance than that of each single scale, achieving the recognition rate of 94.40%.

SVM Classification – We further adopted SVM (RBF kernel) to perform gender classification using the selected LBPH bins, and obtained the best recognition rate of 94.81%. As a baseline to compare against, we also applied SVM with raw image pixels, which delivers the best performance on face images acquired in controlled environments (Moghaddam and Yang, 2002). For computational simplicity, face images of 127×91 pixels were down-scaled to 64×46 pixels, thus each image represented by a vector of 2,944 dimensions. We summarize the results of SVM with raw pixels and standard LBP features in Table 2. It is observed that, with the raw pixel intensities, SVM can achieve the recognition rate of 91.27%. This is very promising, although it is inferior to 96.62% reported in Moghaddam and Yang (2002) on the FERET database. This indicates that the faces in the LFW database are not difficult for gender classification given good face alignment. It is seen in Table 2 that LBP features produce better performance than raw image pixels. With the 500 selected LBPH bins, Adaboost achieves better performance than SVM using the standard LBP (2,478 bins). Overall the best performance of 94.81% is obtained by applying SVM with the boosted LBP features.

Regarding support vectors, with raw pixels, the learned SVMs utilized 51–53% of the total number of training samples (in each trial of cross-validation, the number varies slightly), while SVMs

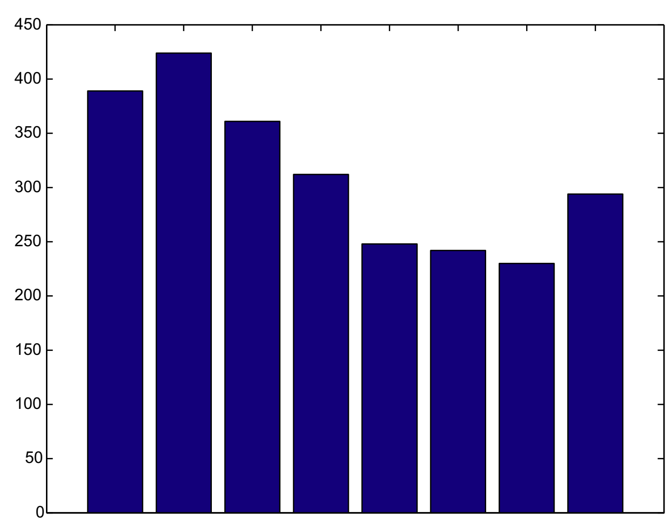

Fig. 10. Scale distribution of the selected LBPH bins.

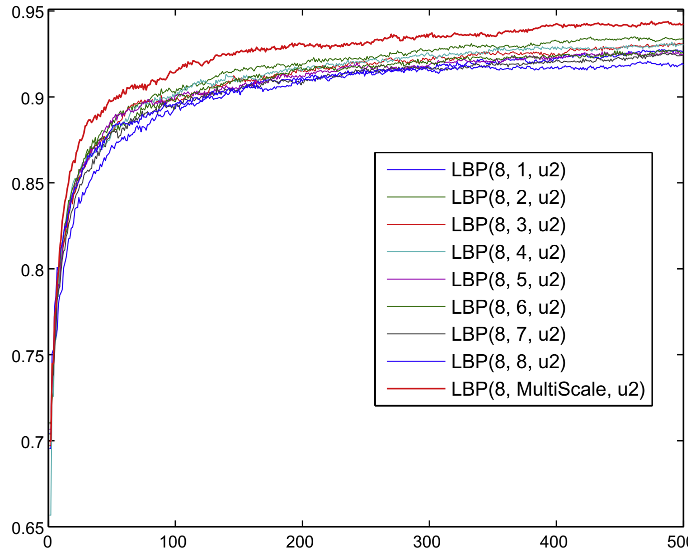

Fig. 11. Classification performance of boosted strong classifiers, as a function of the number of feature selected.

Table 2
Experimental results of gender classification.

Approach			Recognition rates (%)		
Feature	Dimension	Classifier	Female	Male	Overall
Raw pixels	2,944	SVM	86.89	94.13	91.27 ± 1.67
Standard LBP	2,478	SVM	89.78	95.73	93.38 ± 1.50
Boosted LBP	500	Adaboost	91.98	95.98	94.40 ± 0.86
Boosted LBP	500	SVM	92.02	96.64	94.81 ± 1.10

with LBP features employ 58–61%. In contrast, for boosted LBP based SVM, the numbers of support vectors were 40–44% of the number of training samples. As observed in Table 2, the boosted LBP features also produce smaller standard variation. We see in Table 2 there is notable bias towards males in all experiments, as observed in existing studies (Shakhnarovich et al., 2002). This might be due to the unbalanced training data.

5. Conclusions

In this paper, we investigate gender classification on real-life faces acquired in unconstrained conditions, a challenging but relatively understudied problem. We learn discriminative LBP-Histogram bins as compact facial representation for gender classification. By adopting SVM with the selected LBPH bins, we obtain the classification rate of 94.81% on the LFW database.

Acknowledgement

This work was supported by the Visual Context Modeling (ViCoMo) project.

References

Golomb, B.A., Lawrence, D.T., Sejnowski, T.J., 1991. Sexnet: a neural network identifies sex from human faces. In: Adv. Neural Inform. Process. Systems (NIPS), pp. 572–577.
Brunelli, R., Poggio, T., 1992. Hyperbf networks for gender classification. In: DRAPA Image Understanding Workshop, pp. 311–314.
Yang, Z., Li, M., Ai, H., 2006. An experimental study on automatic face gender classification. In: Internat. Conf. on Pattern Recognition (ICPR), pp. 1099–1102.
Moghaddam, B., Yang, M., 2002. Learning gender with support faces. IEEE Trans. Pattern Anal. Machine Intell. 24 (5), 707–711.
BenAbdelkader, C., Griffin, P., 2005. A local region-based approach to gender classification from face images. In: IEEE Conf. on Computer Vision and Pattern Recognition Workshop, pp. 52–52.
Lapedriza, A., Marin-Jimenez, M.J., Vitria, J., 2006. Gender recognition in non controlled environments, In: Internat. Conf. on Pattern Recognition (ICPR), pp. 834–837.
Baluja, S., Rowley, H.A., 2007. Boosting sex identification performance. Internat. J. Computer Vision 71 (1), 111–119.
Mäkinen, E., Raisamo, R., 2008. Evaluation of gender classification methods with automatically detected and aligned faces. IEEE Trans. Pattern Anal. Machine Intell. 30 (3), 541–547.
Hadid, A., Pietikäinen, M., 2009. Combining appearance and motion for face and gender recognition from videos. Pattern Recognition 42 (11), 2818–2827.
Shakhnarovich, G., Viola, P.A., Moghaddam, B., 2002. A unified learning framework for real time face detection and classification. In: IEEE Internat. Conf. on Automatic Face & Gesture Recognition (FG), pp. 14–21.
Gao, W., Ai, H., 2009. Face gender classification on consumer images in a multiethnic environment. In: Internat. Conf. on Biometrics (ICB), pp. 169–178.
Kumar, N., Belhumeur, P.N., Nayar, S.K., 2008. Facetracer: A search engine for large collections of images with faces. In: European Conf. on Computer Vision (ECCV), pp. 340–353.
Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K., 2009. Attribute and simile classifier for face verification. In: IEEE Internat. Conf. on Computer Vision (ICCV).
Huang, G., Ramesh, M., Berg, T., Learned-Miller, E., 2007. Labeled faces in the wild: A database for studying face recognition in unconstrained environments, Tech. Rep. 07-49, University of Massachusetts, Amherst.
Ahonen, T., Hadid, A., Pietikäinen, M., 2004. Face recognition with local binary patterns. In: European Conf. on Computer Vision (ECCV), pp. 469–481.
Sun, N., Zheng, W., Sun, C., Zou, C., Zhao, L., 2006. Gender classification based on boosting local binary pattern. In: Internat. Symp. on Neural Networks.
Lian, H., Lu, B., 2007. Multi-view gender classification using multi-resolution local binary patterns and support vector machines. Internat. J. Neural Systems 17 (6), 479–487.
Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Machine Intell. 24 (7), 971–987.
Zhang, G., Huang, X., Li, S.Z., Wang, Y., Wu, X., 2004. Boosting local binary pattern (lbp)-based face recognition. In: Chinese Conf. on Biometric Recognition, pp. 179–186.
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. System Sci. 55 (1), 119–139.
Schapire, R.E., Singer, Y., 1999. Improved boosting algorithms using confidence-rated predictions. Maching Learning 37 (3), 297–336.
Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 511–518.
Viola, P., Jones, M., 2004. Robust real-time face detection. Internat. J. Comput. Vision 57 (2), 137–154.
Wolf, L., Hassner, T., Taigman, Y., 2009. Similarity scores based on background samples. In: Asian Conf. on Comput. Vision (ACCV).
Shan, C., Gritti, T., 2008. Learning discriminative lbp-histogram bins for facial expression recognition. In: British Machine Vision Conf. (BMVC), Leeds, UK.