Proceedings of the Fifth IIEEJ International Workshop
on Image Electronics and Visual Computing 2017
Da Nang, Vietnam, February 28 - March 3, 2017

FACE RECOGNITION FOR VIDEOS USING LOCAL BINARY PATTERNS AND SUPPORT VECTOR MACHINES

Chonnikan Chuensuwan†
†Graduate School of Engineering
Tokai University

Tomoko Ozeki††
††Dept. of Human and Information Science
Tokai University

ABSTRACT

Face recognition is one of the most important research topics in computer vision because of its many potential applications in criminal identification from surveillance cameras, person identification and so on. In this research we compare local features, holistic local binary pattern histogram (hLBPH), enhanced LBP histogram (eLBPH) and holistic LBP image (hLBPI), by using the Japanese Female Facial Expression (JAFFE) database. We apply the best feature to face recognition from video database, using grayscale images of five subjects as the training set and five videos as the test set. In this step, we examine two classifiers, Euclidean distance and support vector machines (SVM). Finally, the experimental results are compared with each other.

1. INTRODUCTION

Face recognition is used to verify a person in a photo or an image using a stored database of faces. It is still a hot research topic in the field of pattern recognition, image analysis and artificial intelligence. In addition, face recognition is a challenging field since the real face images are formed via the interactions of various factors on several conditions such as changes in illumination, pose variations, occlusions and facial expressions [1]. Especially, face recognition in videos is susceptible to such uncontrollable conditions [2].

In this paper, our purpose is to quantitatively investigate the effectiveness of local feature, LBP and the classifier, SVM for face recognition in videos. In the previous studies, LBP is applied to face recognition in still images and videos by using simple classifier, k-NN with some similarity measures. We also evaluate the dependence of parameters in spatially extended LBP on the face recognition accuracy by testing on still images and video database.

2. LOCAL BINARY PATTERN (LBP)

LBP involves dividing up the face image to regions. The neighboring pixels are compared to the central grayscale pixel value as a threshold and converted to a binary signal. Hence a binary number representing each pixel will be formed as in Fig. 1. The histograms of 256 bins will be built from the whole image (hLBPH) [3],[4].

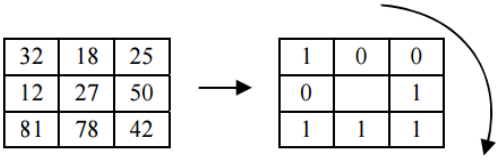

Figure 1: Example of LBP operators. Binary sequence: 10011110 Decimal: 158

The enhanced local binary pattern histogram (eLBPH) was proposed by Ahonen et al [5]. To extract the eLBPH feature of a face, a facial image is divided into d regions $R_0, R_1, \cdots, R_{d-1}$ and each regional LBPH is individually calculated, then these resulting d regional LBPHs are concatenated into a spatially enhanced LBPH (eLBPH) in the same order for all images [4].

A holistic LBP image (hLBPI) is given by the transformation of all pixels of an image into LBP codes [4]. The hLBPI of a facial image preserves spatial relations which are lost in hLBPH.

In this research, we apply hLBPH, eLBPH and hLBPI as local features to face recognition on the JAFFE database [6] and compare their performance. We use the best local feature to recognize faces in videos.

3. SUPPORT VECTOR MACHINE (SVM)

Support vector machine (SVM) is one of supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis [7]. Furthermore, suppose each of the given samples belongs to either of the two category. The main goal is to decide which category the new sample belongs to.

In the case of SVM, the samples will be considered as p-dimension vectors. The task is to separate such points with a (p-1)-dimensional hyperplane. And this can be defined as a linear classifier. There are many hyperplanes that

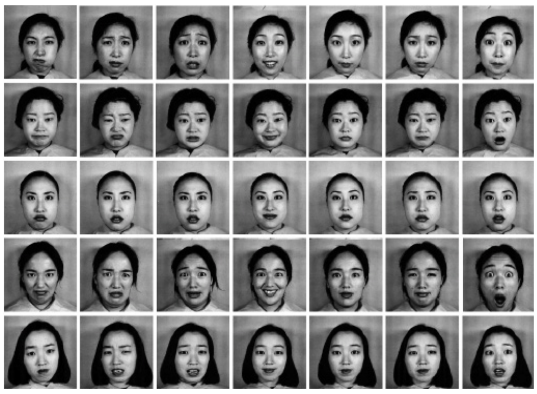

Figure 2: JAFFE database

can classify the data. One reasonable choice is that the best hyperplane is the one that maximizes the sum of the distance to the closest training samples and it is called a margin. For linear separable case, the training algorithm can always find such a hyperplane known as maximum-margin hyperplane. The linear classifier is defined as a maximum margin classifier. For the non-linear case, the training samples are mapped onto a high-dimensional space by applying the kernel trick as suggested by Vapnik [8]:

$$h(x) = \sum_{i=1}^{L} y_i \alpha_i k_i(x_i, x) + b \qquad (1)$$

where $h(x)$ is a decision boundary, $k(,)$ is a kernel function, and L is the number of training samples. In this space, the decision boundary is linear. The most commonly used kernel functions are polynomials, exponential and sigmoidal functions [7].

4. EXPERIMENT AND RESULTS

To evaluate the performance of three types of LBP and two classifiers, we performed two experiments. First we tested on static images to find the best LBP method and second we tested on a video database to find the best classifier. All of our experiments were tested on MATLAB R2016a. These experiments were running on Window10 Pro 64-bit Operating system with Intel(R) Xeon(R)CPU E3-1241 v3 @3.50GHz 16GB RAM.

As a database we have used the Japanese Female Facial Expression (JAFFE) database [6], which contains 213 images of seven facial expressions (six basic facial expressions + one neutral) posed by ten Japanese female models. Each image has been rated on six emotion adjectives as show in Figure2.

We have also used a video database that was separated into two parts. The first is a training set, that was taken at the same place and different illumination, 206 images of five subjects. The second is testing videos of five subjects that were taken from the deferent places and different illumination. Figure3 shows the training set database.

Figure 3: The training set of original database

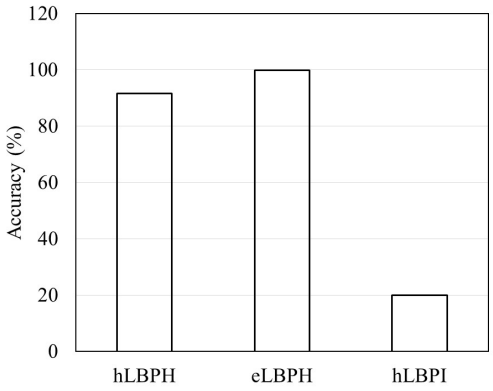

Figure 4: Comparison between hLBPH, 5×5 eLBPH and hLBPI in JAFEE database.

4.1. RESULTS ON STILL IMAGES

In this experiment, 213 gray scale images in JAFFE database are utilized to calculate the recognition accuracies of hLBPH, eLBPH and hLBPI by using the Euclidean distance. Figure 4 shows the result that eLBPH (5×5) gave the best accuracy of 99.86%. For eLBPH we further divided an image for 6 patterns of $2 \times 1, 3 \times 1, 1 \times 3, 2 \times 2, 3 \times 3, 5 \times 5$, which are shown in Fig.5. Figure 6 shows the recognition accuracies of eLBPHs obtained by setting the ratio between training and testing set as 80:20,60:40 and 20:80. We can conclude that the division by 5×5 gave the best accuracy. However, it took the longest computation time In [9] they performed an experiment of face recognition on JAFFE database with MB-LBP, PCA and k-NN. The result is 98.57% accuracy.

For JAFFE database [6], we also used SVM as a classifier and SVM performed the result shown in Fig.7. From this result we conclude that SVM gives better accuracy.

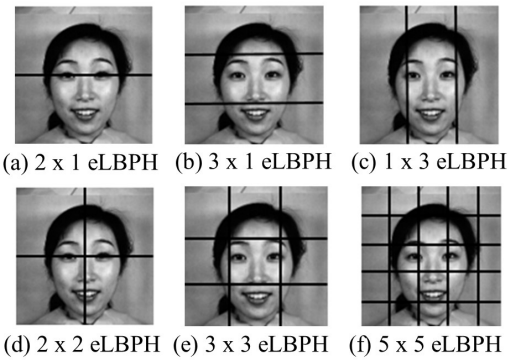

Figure 5: Division of an image for eLBP histogram.

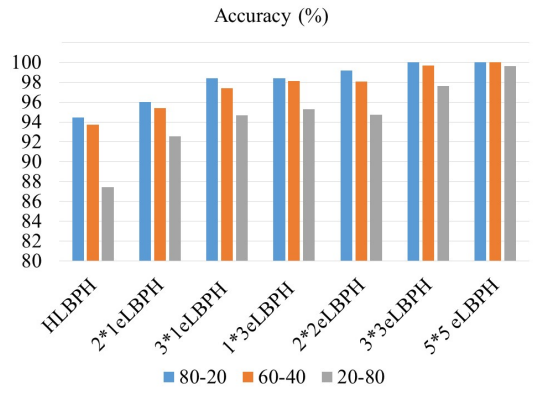

Figure 6: Results on JAFFE. Local features, hLBPH and eLBPHs are compared by using the Euclidean distance.

For our database which is obtained in relatively unconstrained environment, 206 gray scale images are utilized to calculate the recognition accuracies of hLBPH and eLBPHs. In this database the 3×1 performed the best accuracy. In the case of training and test ratio of 80:20 and 60:40, the recognition accuracy is 100% and for 20:80, the accuracy is 99.5%.

4.2. RESULTS ON VIDEOS

In this part, we consider 7×7 eLBPH to prove if 5×5 eLBPH provides the best accuracy for this database. The experiments on videos consist of two stages. Firstly, we generated eLBP histograms from 206 training images and trained the SVM for classification. Then we cut 81 frames from the videos of five subjects to use as a test set. For each frame we detected the face and cropped the face part for face recognition by using MATLAB cascade detector. We use the trained SVM for classification.

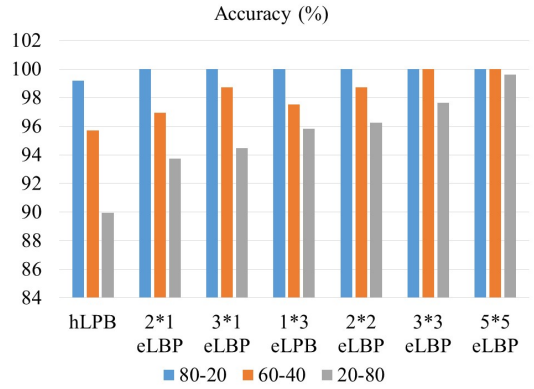

Figure 7: Results on JAFFE. Local features, hLBPH and eLBPHs are compared by using SVM.

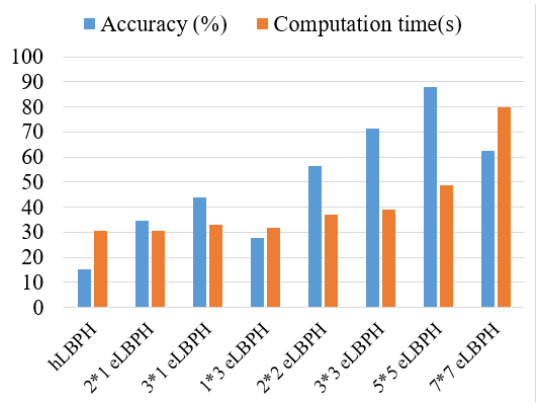

Figure 8: Results on videos. Local features, hLBPH and eLBPHs are compared by using SVM.

Figure 8 shows the accuracies of hLBPH and eLBPHs in different division. The highest accuracy of 88.0% is obtained by the 5×5 eLBPH. We compare our results with previous researches, where the results were obtained from video database with many uncontrollable environments. In [10] they performed an experiment on a video surveillance that also had many uncontrollable environments. They also used SVM as a classifier and AdaBoost as face region detection. The feature was a holistic feature given by the naive pixel values of the image. The result is about 70% accuracy. They have a very large training set about 1900 images and the test set is 1476 images. In [11] they have a training set about 130 images and test images are separated into two parts. The first part includes 1950 images that were taken in daytime. The second part includes 780 images that were taken in nighttime. For daytime, the accuracy is 37.2% on LBP method. For nighttime, the accuracy is 11.8%. From [10] and [11], we conclude that the number of training set

and illumination conditions in daytime and nighttime have significant effects on the accuracy.

5. CONCLUSION

In this paper we performed experiments to evaluate the performance of various LBPs and SVM for face recognition task in the still image database and the video database. JAFFE database is used as testing the performance of LBP methods. The best LBP is eLBPH. We examined the dependence of the accuracy on the division of images in eLBPH. Our results show that the finer division results in the better accuracy. We also compare the accuracy between two types of classifiers, the Euclidean distance and the SVM. SVM performed better accuracy.

For video database, we have examined the hLBPH and eLBPH with various divisions as features of the video frames by using support vector machine. However, as suggested by the results of hLBPI, it is expected that the accuracy saturates at the certain number of divisions and that it becomes even worse as the number of divisions increases. So we performed an experiment of 7×7 eLBPH to prove that 5×5 eLBPH provided the best accuracy for video database in our experiments.

From our experiments, it is suggested that the finer division results in the better accuracy and that the support vector machine gives better performance compared to the simple k-NN with the similarity measure which was used in the previous studies. For still images, although the best accuracy is given by 5×5 eLBPH, the 3×3 eLBPH gives almost the same accuracy with less computation. However 7×7 gives lower accuracy with heavy computation time when compared with 5×5 and 3×3. Our quantitative estimation of parameters in LBP suggests the availability of suboptimal division of eLBPH in real applications for still images and videos.

For the original video database, it has many controlled environments such as pose variations, illumination and video resolution. In the future work, database should be changed into more realistic video from surveillance camera that includes various poses. However, it is also expected that the face recognition performance in the video suffers from the heavy computation time.

6. REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, ”Face Recognition: A Literature Survey,” *ACM Computing Surveys*, Vol. 35, No. 4, pp. 399-458, 2003.

[2] H. Wang, Y. Wang, Y. Cao, “Video-based Face Recognition: A Survey,” *International Journal of Computer, Electrical,Automation, Control and Information Engenieering*, Vol. 3, No. 12, pp. 273-283, 2009.

[3] T. Ojala, M. Pietikainen and T. Maenpaa, “Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[4] B. Yang and S. Chen, “A Comparative Study on Local Binary Pattern (LBP) Based Face Recognition: LBP Histogram Versus LBP Image, ” *Neurocomputing*, Vol. 120, pp.365–379, 2013.

[5] T. Ahonen, A. Hadid and M. Pietikainen, ”Face Description with Local Binary Patterns: Application to Face Recognition,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041 (2006).

[6] M. J. Lyons, S. Akemastu, M. Kamachi and J. Gyoba, “Coding Facial Expressions with Gabor Wavelets,” *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200–205 (1998).

[7] C. J. Burges: ”A tutorial on support vector machines for pattern recognition, ” Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121-167, Jun. 1998.

[8] V. N. Vapnik, “Statistical Learning Theory,” Wiley-Interscience, ISBN: 978-0-471-03003-4, 1998.

[9] G. N. Girish, N. C. L. Shrinivasa and P. K. Das, ”Face Recognition Using MB-LBP and PCA: A Comparative Study,” Computer Communication and Informatics (ICCCI), 2014 International Conference on, Coimbatore, pp. 1-6 (2014).

[10] C. T. Yew and S. A. Suandi, “A Study on Face Recognition in Video Surveillance System Using Multi-class Support Vector Machines,” *TENCON 2011 - 2011 IEEE Region 10 Conference*, Bali, pp. 25–29, 2011.

[11] H.-T. Nguyen and A. Caplier “Elliptical Local Binary Patterns for Face Recognition ” Springer-Verlag Berlin Heidelberg, ACCV 2012 Workshops, Part I, LNCS 7728, pp. 8596 (2013).