中央研究院
資訊科學研究所
Institute of Information Science, Academia Sinica • Taipei, Taiwan, ROC

TR-IIS-14-003

Review and Implementation of High-Dimensional Local Binary Patterns and Its Application to Face Recognition

Bor-Chun Chen, Chu-Song Chen, Winston Hsu

July. 24, 2014 || Technical Report No. TR-IIS-14-003
http://www.iis.sinica.edu.tw/page/library/TechReport/tr2014/tr14.html

Review and Implementation of High-Dimensional Local Binary Patterns and Its Application to Face Recognition

Bor-Chun Chen
Academia Sinica, Taiwan
bcsiriuschen@gmail.com

Chu-Song Chen
Academia Sinica, Taiwan
song@iis.sinica.edu.tw

Winston Hsu
National Taiwan University
whsu@ntu.edu.tw

ABSTRACT

High-dimensional local binary patterns [5] have been proved to be a useful feature for face recognition, which provides near-human performance in a widely used face verification benchmark. In this report, we first review the technical aspect of this promising feature, and then we provide our implementation details of the feature. Finally, we show some experimental results using this feature on two public datasets, LFW and CACD.

Keywords

Local binary patterns, high-dimensional features, face recognition, LFW dataset, CACD dataset.

1. INTRODUCTION

Face recognition has always been an important problem in the field of computer vision, and many researchers try to develop different features for the task. In particular, local binary patterns (LBP) [7], a descriptor originally developed for texture classification, have been proved to be effective for face recognition [1], and therefore many different variations of LBP was proposed to use as feature for describing face images. Recently, Chen [5] propose a variation of LBP called high-dimensional local binary patterns (HD-LBP) and achieve near-human performance on face verification task. They combine LBP descriptors extracted from different landmarks of face at different scales together as a high-dimensional feature, and then they propose a method called rotated sparse regression to reduce the dimension of the features while maintaining the verification accuracy. In this report, we focus on the HD-LBP feature proposed in [5]. In Section 2, we provide the technical aspect of HD-LBP. We first review the original operator for extracting LBP descriptors, and then we describe how to use LBP descriptors for face recognition. Finally, we describe the details for extracting HD-LBP. Figure 1 shows the pipeline for extracting HD-LBP feature. In Section 3, we describe our implementation details of HD-LBP including the face detection and facial landmark detection algorithms, face alignment method, and parameters settings for extracting the LBP features. Some experimental results using HD-LBP on two public large-scale face datasets, cross-age celebrity dataset (CACD) [4] and Labeled Face in the Wild (LFW) [6], will be shown in Section 4, and last, Section 5 draws conclusions.

2. REVIEWS OF HIGH-DIMENSIONAL LOCAL BINARY PATTERNS

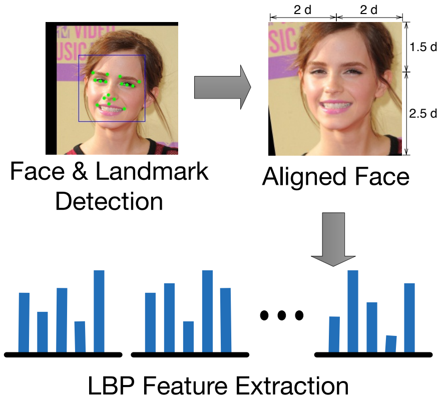

Figure 1: Pipeline for extracting HD-LBP features. For each input image, face detection and facial landmark detection is applied to detect the face and landmark positions. Transformation are then applied to align the face image. Finally, LBP features are extracted from different landmark positions at different scales and concatenated together to form the final HD-LBP feature.

2.1 Local Binary Patterns

Local binary patterns were originally used for texture classification. A LBP operator computes a value for each pixel in the image based on its relationship with neighborhood pixels so that it can capture the local texture of the image. Figure 2 shows a simple LBP operator. To compute a LBP of one pixel, the values of the neighborhood pixels are found, and then the values are binarized via thresholding with the value of the current pixel. Finally, the binary values are concatenated together as the LBP of the current pixel. LBPs are usually aggregated into histogram for further processing. In order to increase the discriminability of the LBP, a concept called uniform pattern is proposed in [7]. A LBP is called uniform if the binary pattern contains at most two bitwise transitions. For example, the pattern 01000000 contains 2 transitions and so it is uniform; while the pattern 10001001 contains 4 transitions, so it is non-uniform. Uniform patterns account for most of the patterns in face images and therefore they are more important for face recognition. There are totally 58 different uniform patterns. When aggregating LBPs into histogram, all non-uniform patterns are assigned into the same bin, so the final dimension of the uniform LBP histogram is 59. More details of LBP and uniform LBP can be found in [7].

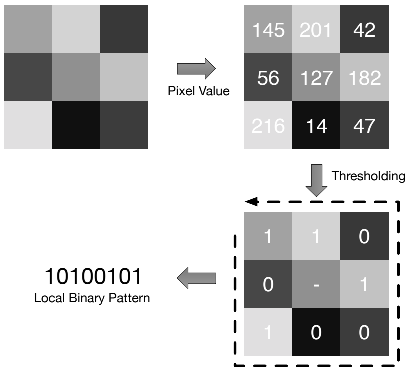

Figure 2: Illustration of LBP operation. The LBP value of each pixel is decided by the values of its neighborhood pixels.

2.2 Local Binary Patterns for Face Recognition

Ahonen et al. [1] find that LBP achieves salient results when applied to face images. After they detect and align the face image, they divide the whole face into small grids. From each grid, they compute a LBP histogram. And finally they concatenate all histograms together into a feature vector to represent the face image. They use weighted chi-square distance to compute the distance between faces. Details of LBP for face recognition can be found in [1].

2.3 High-Dimensional Local Binary Patterns

Recently, Chen et al. proposed high-dimensional LBP features for face recognition. Compared to original LBP features used in [1] which has 2,891 dimensions ($59 \times 7 \times 7$), feature used in [5] is higher than 100K dimensions. They proposed two ways to increase the dimension of the features, including multiple landmarks and multiple scales. In [5], they extract LBP histograms from up to 27 facial landmarks. From each landmarks, sixteen grids with 10×10 pixels are used for computing the LBP histograms. Up to five scales are used to compute LBP histograms in different resolutions. Therefore the total dimensions are up to 127,440 ($59 \times 16 \times 27 \times 5$). After they extracted the high-dimensional LBP features, they use PCA to reduce the redundant features and shorten the features dimension and apply supervised learning methods such as LDA [3] to train a verification classifier. To reduce the space and time required for testing, they proposed a method called rotated sparse regression. A sparse matrix is learned and used for approximating the PCA and supervised learning with linear projection. Here we only focus on the feature extraction part in [5], other details of the method can be found in their paper.

3. IMPLEMENTATION DETAILS

3.1 Face Detection

We apply the Viola-Jones face detector [10] implemented in openCV for the task. Each image is first converted to gray scale and enhanced with histogram equalization before applying the detection. Default scale factor is set to 1.1, minimal neighbor is set to 4, and minimal size is set to 30×30. After detecting the faces in the image, we crop the largest face in the image with a border. The size of the cropped image is two times the size of the face and it is then rescaled to 250×250 pixels which is similar to images in the LFW dataset [6]. When applying face detections to the CACD dataset [4] (275,803 images) download from the Internet, average time for detecting a single image is about 0.37 seconds on a single machine with Intel Xeon CPU E5420.

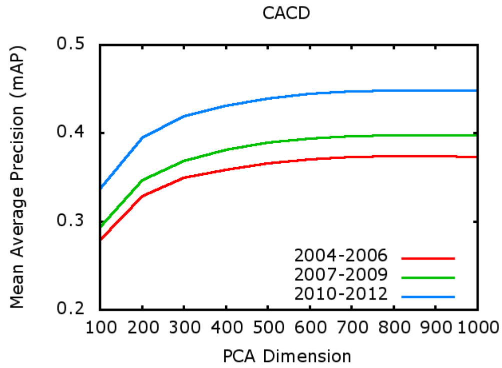

Figure 3: Performance of HD-LBP on CACD with different PCA dimensions.

3.2 Facial Landmark Detection

After face detection, we use IntraFace library [11] to detect facial landmarks. We then compute the eye locations by finding the mid-points of eye corners. The eye locations are later used to align the face image. We use 0.5 as our confident threshold and the indices of the landmarks used in IntraFace include: $\{0, 4, 5, 9, 19, 22, 25, 28, 13, 14, 18, 31, 34, 37\}$. Landmark locations on the face can be found in Figure 1.

3.3 Face Alignment

We use the eye locations to align the face image. Images are first rotated so that the two eyes are horizontal even. We then compute the distance between two eyes, and use the obtained distance d to crop a rectangular region of the rotated face. Default values are $2 \times d, 2 \times d, 1.5 \times d, 2.5 \times d$ for the left, right, top, and bottom borderline lengths respectively. Figure 1 shows a example of aligned face.

3.4 HD-LBP feature extraction

We rescale the aligned face into five different resolutions including: 300, 212, 150, 106, 75 pixels and extract uniform LBP from each of the 16 landmarks (14 plus two eye locations) at each scale. For each landmark location, we extract 4×4 grids with each grid containing 10×10 pixels. From each grid, we extract a 59-dimensional uniform LBP feature. All features are then concatenated to form a high-dimensional LBP feature. Since we use 16 landmark locations and five scales, the final feature dimension is 75,520 ($59 \times 4 \times 4 \times 16 \times 5$). We use PCA to reduce the feature dimension, and employ cosine similarity to compute the similarity between two faces. The code of our implementation of the HD-LBP can be found at http://bcsiriuschen.github.io/High-Dimensional-LBP/.

4. EXPERIMENTAL RESULTS

We conduct experiments on two public datasets, CACD and LFW, using our implementation of HD-LBP.

4.1 Experimental Results on CACD

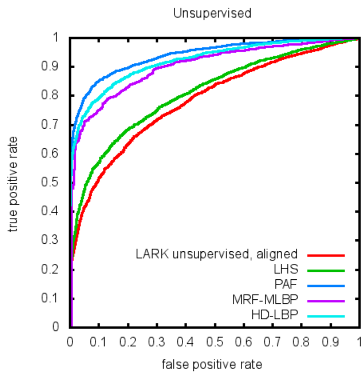

Figure 4: ROC curve of HD-LBP on LFW compared with other state-of-the-art methods under unsupervised protocol.

Table 1: The AUC of different methods on LFW dataset. The HD-LBP can achieve very competitive results compared to other state-of-the-art methods.

Method	AUC
LARK [8]	78.30%
LHS [9]	81.07%
PAF [12]	94.05%
MRF-MLBP [2]	89.94%
HD-LBP	92.11%

CACD contains 163,446 images of 2000 celebrities. We follow the protocol in [4] and use images taken in 2013 as query and images taken in 2004-2012 as database for face retrieval. We use images of 600 celebrities to compute PCA subspace as suggested in [4] and reduce the features extracted from each facial landmarks to 100-1000 dimensions. The performance for each of the three subsets is shown in Figure 3. We can see that the performance has only marginal improvement when the dimension is higher than 500 for all three subsets, and the 2010-2012 subset has the best performance while the 2004-2006 dataset has the worst performance. These findings are consistent with the results in [4].

4.2 Experimental Results on LFW

LFW contains 13,233 images from 5,749 people. We follow the unsupervised protocol of the dataset and use 6,000 face pairs with 3,000 positive pairs and 3,000 negative pairs to evaluate face verification performance of HD-LBP. We reduce the feature dimension to 400 using PCA as suggested in [5]. The ROC curves compared with other unsupervised methods are shown in Figure 4 and AUC results are shown in Table 1. We can see that the implemented HD-LBP can achieve very competitive results compared to other state-of-the-art methods. Only PAF [12] achieves better performance than HD-LBP, but it uses a more complicate 3D-alignment method to estimate the pose. More results on LFW can be found in the LFW website. Note that in [5] they did not provide the ROC curve and AUC under the unsupervised setting, instead, they provide a recognition accuracy of 84.08% using high-dimensional LBP. To compared with their result, we use SVM to find a threshold for classifying image pairs into positive or negative pair. The average accuracy over ten fold is 84.85%.

5. CONCLUSION

In this report, we review a state-of-the art face recognition feature proposed in [5], called high-dimensional local binary patterns. We provide our implementation details for the algorithm, and conduct experiments using our implementation on two public datasets, CACD and LFW. We also release the source code of our implementation, and hope it can help the future research to the related directions.

6. REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041, 2006.

[2] S. R. Arashloo and J. Kittler. Efficient processing of mrfs for unconstrained-pose face recognition. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–8. IEEE, 2013.

[3] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):711–720, 1997.

[4] B.-C. Chen, C.-S. Chen, and W. H. Hsu. Cross-age reference coding for age-invariant face recognition and retrieval. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2014.

[5] D. Chen, X. Cao, F. Wen, and J. Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3025–3032. IEEE, 2013.

[6] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[7] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.

[8] H. J. Seo and P. Milanfar. Face verification using the lark representation. *Information Forensics and Security, IEEE Transactions on*, 6(4):1275–1286, 2011.

[9] G. Sharma, S. ul Hussain, and F. Jurie. Local higher-order statistics (lhs) for texture categorization and facial analysis. In *Computer Vision–ECCV 2012*, pages 1–12. Springer, 2012.

[10] P. Viola and M. J. Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.

[11] X. Xiong and F. De la Torre. Supervised descent method and its applications to face alignment. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[12] D. Yi, Z. Lei, and S. Z. Li. Towards pose robust face recognition. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3539–3545. IEEE, 2013.