IDIAP RESEARCH REPORT

ON THE RECENT USE OF LOCAL BINARY PATTERNS FOR FACE AUTHENTICATION

Sébastien Marcel, Yann Rodriguez and Guillaume Heusch

IDIAP–RR 06-34

JUNE 2006

TO APPEAR IN
International Journal of Image and Video Processing – Special Issue on Facial Image Processing

IDIAP Research Institute **www.idiap.ch**
Rue du Simplon 4 P.O. Box 592 1920 Martigny − Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

IDIAP Research Report 06-34

On the Recent Use of Local Binary Patterns for Face Authentication

Sébastien Marcel, Yann Rodriguez and Guillaume Heusch

June 2006

to appear in
International Journal of Image and Video Processing – Special Issue on Facial Image Processing

Abstract. This paper presents a survey on the recent use of Local Binary Patterns (LBPs) for face recognition. LBP is becoming a popular technique for face representation. It is a non-parametric kernel which summarizes the local spacial structure of an image and it is invariant to monotonic gray-scale transformations. This is a very interesting property in face recognition. This probably explains the recent success of Local Binary Patterns in face recognition. In this paper, we describe the LBP technique and different approaches proposed in the literature to represent and to recognize faces. The most representatives are considered for experimental comparison on a common face authentication task. For that purpose, the XM2VTS and BANCA databases are used according to their respective experimental protocols.

1 Introduction

Local Binary Pattern (LBP) is becoming a popular technique for face representation as well as for image representation in general. Recently, LBP has been applied to the specific problem of face recognition. The LBP is a non-parametric kernel which summarizes the local spacial structure of an image. Moreover, it is invariant to monotonic gray-scale transformations, hence the LBP representation may be less sensitive to changes in illumination. This is a very interesting property in face recognition. Indeed, one of the major problem in face recognition systems is to deal with variations in illumination. In a realistic scenario, it is very likely that the lighting conditions of the probe image does not correspond to those of the gallery image, hence there is a need to handle such variations. This probably explains the recent success of Local Binary Patterns in the face recognition community.

We propose in this paper an overview of different LBP techniques proposed for face recognition in general and we experimentally compare the most representative ones on the face authentication task. *Face authentication* (or *verification*) involves confirming or denying the identity claimed by a person (one-to-one matching). In contrast, *face identification* (or *recognition*) attempts to establish the identity of a given person out of a closed pool of N people (one-to-N matching). Both mode are generally grouped under the generic *face recognition* term. Authentication and identification share the same preprocessing and feature extraction steps and a large part of the classifier design. However, both modes target distinct applications. In authentication mode, people are supposed to cooperate with the system (the claimant wants to be accepted). The main applications are access control systems, such as computer or mobile devices log-in, building gate control, digital multimedia access. On the other hand, in identification mode, people are generally not concerned by the system and often even do not want to be identified. Potential applications includes video surveillance (public places, restricted areas) and information retrieval (police databases, video or photo album annotation/identification).

The problem of face authentication has been addressed by different researchers using various approaches. Thus, the performance of face authentication systems has steadily improved over the last few years. For a comparison of different approaches see [18]. These approaches can be divided mainly into *discriminant* approaches and *generative* approaches. A *discriminant* approach takes a binary decision (whether or not the input face is a client) and considers the whole input for this purpose. Such *holistic* approaches are using the original gray-scale face image or its projection onto a Principal Component subspace (referred to as PCA or Eigenfaces [26]) or Linear Discriminant subspace (referred to as LDA or Fisherfaces [3, 6]) as input of a discriminant classifier such as Multi-Layer Perceptrons (MLPs), Support Vector Machines (SVMs) [13] or simply a metric [15, 14]. Recently, it has been shown that *generative* approaches such as Gaussian Mixture Models (GMMs) [5] and Hidden Markov Models (HMMs) [19, 4] were more robust to automatic face localization than the above discriminant methods. A generative approach computes the likelihood of an observation (a holistic representation of the face image) or a set of observations (local observations of particular facial features) given a client model and compares it to the corresponding likelihood given an impostor model.

Finally, the decision to accept or reject a claim depends on a score (distance measure, MLP output or Likelihood ratio) which could be either above (accept) or under (reject) a given threshold.

This paper is organized as follows. First, we introduce Local Binary Patterns and we describe different approaches for the representation of faces with LBP. Finally, we present experimental results and we draw some conclusions.

2 Local Binary Patterns

The Local Binary Pattern (LBP) operator is a non-parametric 3x3 kernel which summarizes the local spacial structure of an image. It was first introduced by Ojala et al. [20] who showed the high discriminative power of this operator for texture classification. At a given pixel position (x_c, y_c), LBP is defined as an ordered set of binary comparisons of pixel intensities between the center pixel and its eight surrounding pixels.

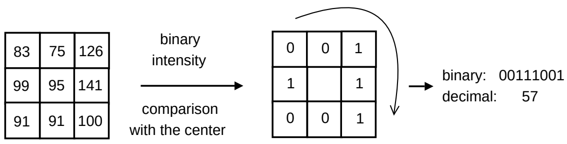

Figure 1: Calculating the original LBP code

The decimal form of the resulting 8-bit word (LBP code) can be expressed as follows (Figure 1):

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

where i_c corresponds to the grey value of the center pixel (x_c, y_c), i_n to the grey values of the 8 surrounding pixels, and function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \; . \end{cases} \tag{2}$$

Note that each bit of the LBP code has the same significance level and that two successive bit values may have a totally different meaning. Actually, The LBP code may be interpreted as a kernel structure index (Figure 2). By definition, the LBP operator is unaffected by any monotonic gray-scale transformation which preserves the pixel intensity order in a local neighborhood.

Later, Ojala et al. [21] extended their original LBP operator to a circular neighborhood of different radius size. Their $LBP_{P,R}$ notation refers to P equally spaced pixels on a circle of radius R. In [21], they also noticed that most of the texture information was contained in a small subset of LBP patterns.

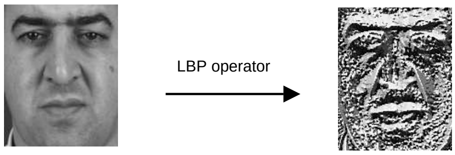

Figure 2: Original image (left) processed by the LBP operator (right).

These patterns, called uniform patterns, contain at most two bitwise 0 to 1 or 1 to 0 transitions (circular binary code). 11111111, 00000110 or 10000111 are for instance uniform patterns. They mainly represent primitive micro-features such as lines, edges, corners. $LBP_{P,R}^{u2}$ denotes the extended LBP operator ($u2$ for only uniform patterns, labelling all remaining patterns with a single label). The $LBP_{8,2}$ operator is illustrated in Figure 3.

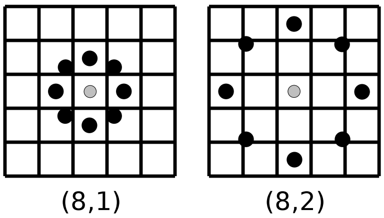

Figure 3: Examples of extended LBP operators

Recently, new variants of LBP have appeared. For instance, Jin et al. [12] remarked that LBP features miss the local structure under some certain circumstance, and thus they introduced the *Improved* Local Binary Pattern (ILBP). Huang et al. [11] pointed out that LBP can only reflect the first derivation information of images, but could not present the velocity of local variation. To solve this problem, they propose an *Extended* version of Local Binary Patterns (ELBP).

Due to its texture discriminative property and its very low computational cost, LBP is becoming very popular in pattern recognition. Recently, LBP has been applied for instance to face detection [12], face recognition [29, 1], image retrieval [25], motion detection [8] or visual inspection [27][1]. We finally point out that, approximately in the same time the original LBP operator was introduced by Ojala [20], Zabih and Woodfill [28] proposed a very similar local structure feature. This feature, called *Census Transform*, also maps the local neighborhood surrounding a pixel. With respect to LBP, the *Census Transform* only differs by the order of the bit string. Later, the *Census Transform* has been extended to become the *Modified Census Transform* (MCT) [7]. Again, one can point out the same similarity between ILBP and MCT (also published at the same time).

[1] a more exhaustive list of applications can be found on Oulu University web site at: http://www.ee.oulu.fi/research/imag/texture/lbp/lbp.php

3 Face Representation using Local Binary Patterns

In [1], Ahonen proposed a face recognition system based on a LBP representation of the face. The individual sample image is divided into R small non-overlapping blocks (or regions) of same size. Histograms of LBP codes H^r, with $r \in \{1, 2, \ldots, R\}$ are calculated over each block and then concatenated into a single histogram representing the face image. A block histogram can be defined as:

$$H^r(i) = \sum_{x,y \in block_r} I(f(x,y) = i),\ i = 1, ..., N, \tag{3}$$

where N is the number of bins (number of different labels produced by the LBP operator), $f(x, y)$ the LBP label [2] at pixel (x, y) and I the indicator function.

This model contains information on three different levels: (1) LBP code labels for the local histograms (pixel level), (2) local histograms (region level) and (3) a concatenated histogram which builds a global description of the face image (image level). Because some regions are supposed to contain more information (such as eyes), Ahonen propose an empirical method to assign weights to each region. For classification, a nearest-neighbor classifier is used with Chi square (χ^2) dissimilarity measure, defined as follows:

$$\chi^2(\mathbf{S}, \mathbf{M}) = \sum_{r,i} \frac{(S^r(i) - M^r(i))^2}{S^r(i) + M^r(i)}, \tag{4}$$

where $\mathbf{S}$ and $\mathbf{M}$ correspond to the sample and the model histograms.

Following the work of Ahonen, Zhang et al. [29] underlined some limitations. First, the size and position of each region are fixed which limits the size of the available feature space. Second, the weighting region method is not optimal. To overcome these limitations, they propose to shift and scale a scanning window over pairs of images, extract the local LBP histograms and compute a dissimilarity measure between the corresponding local histograms. If both images are from the same identity, the dissimilarity measure are labelled as positive features, otherwise as negative features. Classification is performed with AdaBoost learning, which solves the feature selection and classifier design problem. Optimal position/size, weight and selection of the regions are then chosen by the boosting procedure. Comparative study with Ahonen's method showed similar results. Zhang et al.'s system uses however much less features (local LBP histograms).

More recently, Huang et al. [10] proposed an improved version of Zhang et al. system, based on a modified version of the boosting procedure called *JSBoost*. LBP has been also combined with Kernel Fisher Discriminant Analysis (KDA) [30]. This technique uses the Chi square measure to calculate the inner product in the KDA but no improvement nor extension of the LBP kernel itself is proposed.

---

[2] Note that $LBP(x, y)$, the LBP operator value, may not be equal to $f(x, y)$ which is the label assigned to the LBP operator value. With the $LBP^{u2}_{P,R}$ operator, for instance, all non-uniform patterns are labelled with a single label.

Another approach [31] proposes to apply the LBP to Gabor filtered images and to use histogram intersection instead of Chi square as a similarity measure. Although this method obtains good performance on a face recognition task, it requires to compute a large number of histograms over 40 Gabor magnitude images from a single original image.

Even more recently, alternative approaches have been also proposed. Heusch et al. [9] suggested to use the LBP directly as an illumination normalization technique and then to keep standard face recognition techniques such as LDA [15] or HMM [4]. This method, called INORM LBP, obtained the better results using HMM than using LDA. In [22], Rodriguez et al. proposed to use a generative approach. This method, called LBP/MAP, considers local histograms as probability distributions and computes a log-likelihood ratio instead of a Chi square similarity. A generic face model is represented by collection of LBP-histograms. Then, a client-specific model is obtained by an adaptation technique from this generic model under a probabilistic framework.

4 Experiments and Results

In this section, we provide comparative experiments with several systems introduced in Section 3 on two face authentication benchmark databases, namely XM2VTS and BANCA, which we briefly describe in this section.

4.1 Databases and Experimental Setup

The XM2VTS database [17] contains synchronized video and speech data from 295 subjects, recorded during four sessions taken at one month intervals. The subjects were divided into a set of 200 training clients, 25 evaluation impostors and 70 test impostors. We performed the experiments following the *Lausanne Protocol Configuration I*.

The BANCA database [2] was designed to test multi-modal identity verification with various acquisition devices under several scenarios (controlled, degraded and adverse). In the experiments described here we used the face images from the English corpora, containing 52 subjects. Each subject participated in 12 recording sessions in different conditions and with different cameras. Each of these sessions contains two video recordings: one true client access and one impostor attack. Five frontal face images were extracted from each video recording.

Whereas XM2VTS database contains face images in well controlled conditions (uniform blue background), BANCA is a much more challenging database with face images recorded in uncontrolled environment (complex background, difficult lightning conditions). See Figure 4 for example images of each database.

To assess verification performance, the Half Total Error Rate (HTER) is generally used:

$$\mathrm{HTER}(\theta) = \frac{\mathrm{FAR}(\theta) + \mathrm{FRR}(\theta)}{2} \ . \tag{5}$$

where FAR if the false alarm rate, FRR the false rejection rate and θ the decision threshold.

(a) XM2VTS (controlled conditions): uniform background and lighting

(b) BANCA English (uncontrolled conditions): complex background and lighting variability

Figure 4: Comparison of XM2VTS (1) and BANCA (2) image conditions.

To correspond to a realistic situation, θ is chosen *a priori* on the validation set at Equal Error Rate (EER).

For experiments on XM2VTS database, we use all available training client images to build the generic model. For BANCA experiments, the generic model was trained with the additional set of images, referred to as *world data* (independent of the subjects in the client database).

4.2 Results and Discussion

For comparison purpose, we implemented the systems of Ahonen [1], Zhang [29], briefly described in Section 3, but also INORM LBP/HMM [9] and LBP/MAP [22], as well as two standard state-of-the-art methods. The first one is a combination of Linear Discriminant Analysis with Normalized Correlation (LDA/NC) based on a holistic representation of the face [23]. The second one is a generative approach based on a modified version of the Discrete Cosine Transform and Gaussian Mixture Models (DCTmod2/GMM) with a local description of the face [4].

For both XM2VTS and BANCA databases, face images are extracted to a size of 84×68 (rows $\times$ columns), according to the provided ground-truth eye positions. The cropped faces are then processed with the $LBP_{8,2}^{u2}$ operator ($N = 59$ labels). The resulting 80×64 *LBP* face images do not need any further lighting normalization, due to the illumination invariant property of LBP operators. In a block by block basis, the face images are decomposed in 8×8 blocks ($R = 80$ blocks). Histograms of LBP codes are then computed over each block r.

Table 1 reports comparative results for Ahonen, Zhang, LBP/JSBoost, INORM LBP/HMM and LBP/MAP systems, as well as for state-of-the-art methods LDA/NC and DCTmod2/GMM.

First, we remark that several LBP methods obtain state-of-the-art results. Secondly, we notice that compared to the two other methods which use a LBP representation of the face, LBP/MAP performs clearly better. However, it must be noted that these methods (Ahonen and Zhang) have been originally designed for face identification problem. We finally point out that as reported in [29] for identification, Ahonen and Zhang methods give similar results

Table 1: HTER performance comparison for two state-of-the-art methods (LDA/NC and DCTmod2/GMM) and LBP systems, for the XM2VTS database and BANCA database.

Models	XM2VTS	BANCA				
	LP1	Mc	Ud	Ua	P	G
LDA/NC	2.97	3.75	14.5	20.1	15.5	-
DCTmod2/GMM	1.67	6.2	23.7	17.6	18.6	11.0
LBP/JSBoost [10]	-	-	-	-	-	10.7
LBP Ahonen	3.40	8.3	14.3	23.1	20.8	10.4
LBP Zhang	3.94	-	-	-	-	-
INORM LBP/HMM	1.37	2.4	9.9	15.0	11.7	**1.2**
LBP/MAP	1.42	7.3	10.7	22.6	19.2	5.0

on the XM2VTS database. Therefore, we don't provide the results of Zhang method on the BANCA database.

We also report the results from Huang et al. [10]. Recently, Huang et al. proposed an improved version of Zhang et al. system, based on a modified version of the boosting procedure called *JSBoost*, and provided results on BANCA. We then denote this method LBP/JSBoost. Unfortunately, the authors only gave results with protocol G. From the results, we notice that LBP-based generative methods (INORM LBP/HMM and LBP/MAP) perform better that the two other LBP-based methods for all conditions. On protocol G, where more client training data is available, LBP/MAP clearly outperforms the improved version of Zhang system (LBP/JSBoost).

However, according to the results the best system is INORM LBP/HMM, that is when LBP is used as a pre-processing step and when an additional face recognition technique is used. Indeed, all LBP-based face recognition techniques perform histogram comparison. Therefore, we believe there might be a large potential for performance improvement by using more appropriate generative models of Local Binary Patterns.

5 Conclusion and Future Work

In this paper, we presented a survey on some recent use of Local Binary Patterns (LBPs) for face recognition. LBP is a non-parametric kernel which summarizes the local spacial structure of an image and is invariant to monotonic gray-scale transformations. This is a very interesting property in face recognition. We described the LBP technique as well as several different approaches proposed in the literature to represent and to recognize faces. We selected the most representatives to perform an experimental comparison on a face authentication task. The XM2VTS and BANCA databases were used according to their respective experimental protocols.

For comparison purpose, we implemented the several LBP systems as well as two standard state-of-the-art methods. The first one is a combination of Linear Discriminant Analysis with Normalized Correlation based on a holistic representation of the face. The second one is a generative approach based on a modified version of the Discrete Cosine Transform and Gaussian Mixture Models (DCTmod2/GMM) with a local description of the face.

Results have shown that LBP based methods obtained state-of-the-art results and than some of them were even outperforming the state-of-the-art. Another interesting conclusion from the results suggested to combine Local Binary Patterns and generative models. We believe this might be a novel research direction to investigate.

References

[1] T. Ahonen, A. Hadid and M. Pietikäinen, “Face recognition with local binary patterns”, *European Conference on Computer Vision*, Prague, 469–481, 2004.

[2] E. Bailly-Baillière, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz and J.P. Thiran, “The BANCA database and evaluation protocol”, *International Conference on Audio- and Video-Based Biometric Person Authentication*, Guilford, UK, 2003.

[3] P. Belhumeur, J.P. Hespanha and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection”, *European Conference on Computer Vision*, Cambridge, United Kingdom, 45–58, 1996.

[4] F. Cardinaux, C. Sanderson and S. Bengio, “Face verification using adapted generative models”, *IEEE Conference on Automatic Face and Gesture Recognition*, 2004.

[5] F. Cardinaux, C. Sanderson and S. Marcel, “Comparison of MLP and GMM classifiers for face verification on XM2VTS”, *International Conference on Audio- and Video-Based Biometric Person Authentication*, Guilford, UK, 911–920, 2003.

[6] P.A. Devijver and J. Kittler, “Pattern Recognition: A Statistical Approach”, *Prentice-Hall, Englewood Cliffs, N.J.*, 1982.

[7] B. Fröba and A. Ernst, “Face detection with the modified census transform”, *IEEE Conference on Automatic Face and Gesture Recognition*, 2004.

[8] M. Heikkilä, M. Pietikäinen and J. Heikkilä, “A texture-based method for detecting moving objects”, *British Machine Vision Conference*, London, UK. Volume 1. 187–196, 2004

[9] G. Heusch, Y. Rodriguez and S. Marcel, “Local Binary Patterns as an Image Preprocessing for Face Authentication”, *IEEE International Conference on Automatic Face and Gesture Recognition*, 9–14, 2006.

[10] X. Huang, S.Z. Li and Y. Wang, “Jensen-shannon boosting learning for object recognition”, *IEEE International Conference on Computer Vision and Pattern Recognition*, San Diego, USA, 2005.

[11] X. Huang, S. Li, and Y. Wang, "Shape localization based on statistical method using extended local binary pattern", *International Conference on Image and Graphics*, Hong Kong, China, 184–187, 2004.

[12] H. Jin, Q. Liu, H. Lu and X. Tong, "Face detection using improved LBP under bayesian framework", *International Conference on Image and Graphics*, Hong Kong, China. 306–309, 2004.

[13] K. Jonsson, J. Matas, J. Kittler, and Y. Li, "Learning support vectors for face verification and recognition", *International Conference on Automatic Face and Gesture Recognition*, 208–213, 2000.

[14] J. Kittler, R. Ghaderi, T. Windeatt and G. Matas, "Face verification via ECOC", *British Machine Vision Conference*, 593–602, 2001.

[15] Y. Li, J. Kittler and J. Matas, "On matching scores of LDA-based face verification", *British Machine Vision Conference*, 2000.

[16] S. Lucey and T. Chen, "A GMM parts based face representation for improved verification through relevance adaptation", *IEEE International Conference on Computer Vision and Pattern Recognition*, Washington D.C., USA. 2004.

[17] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre, "XM2VTSDB: The Extended M2VTS Database", *International Conference on Audio and Video-based Biometric Person Authentication*, 1999.

[18] K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostyn, S. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, N. Poh, Y. Rodriguez and al.: "Face authentication competition on the BANCA database", *International Conference on Biometric Authentication*, Hong Kong, 2004.

[19] A. Nefian and M. Hayes, "Face recognition using an embedded HMM", *IEEE Conference on Audio and Video-based Biometric Person Authentication*, 19–24, 1999.

[20] T. Ojala and M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", *Pattern Recognition*, Volume 29, 51–59, 1996.

[21] T. Ojala and M. Pietikäinen and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with loval binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24, 971–987, 2002.

[22] Y. Rodriguez and S. Marcel, "Face Authentication Using Adapted Local Binary Pattern Histograms", *European Conference on Computer Vision*, *to appear*, 2006.

[23] M. Sadeghi, J. Kittler, A. Kostin and K. Messer, "A comparative study of automatic face verification algorithms on the banca database", *International Conference on Audio- and Video-Based Biometric Person Authentication*, Guilford, UK, 35–43, 2003.

[24] C. Sanderson and K. Paliwal, “Fast features for face authentication under illumination direction changes”, *Pattern Recognition Letters*, 2409–2419, 2003.

[25] V. Takala, T. Ahonen and M. Pietikäinen, “Block-based methods for image retrieval using local binary patterns”, *Scandinavian Conference on Image Analysis*, Joensuu, Finland, 882–891, 2005.

[26] M. Turk and A. Pentland, “Eigenface for recognition” *Journal of Cognitive Neuroscience*, Volume 3, 70–86, 1991.

[27] M. Turtinen, M. Pietikäinen and O. Silven, “Visual characterization of paper using isomap and local binary patterns”, *Conference on Machine Vision Applications*, Tsukuba Science City, Japan, 210–213, 2005.

[28] R. Zabih and J. Woodfill, “A non-parametric approach to visual correspondence”, *IEEE Transactions on Pattern Analysis and Machine intelligence*, 1996.

[29] G. Zhang, X. Huang, S. Li, Y. Wang and X. Wu, “Boosting local binary pattern (LBP)-based face recognition”, *Chinese Conference on Biometric Recognition*, Guangzhou, China, 179–186, 2004.

[30] J. Zhao, H. Wang, H. Ren and S.-C. Kee, “LBP Discriminant Analysis for Face Verification”, *IEEE Workshop on Face Recognition Grand Challenge Experiments*, Volume 3, 2005.

[31] Z. Wenchao, S. Shiguang, G. Wen, C. Xilin and Z. Hongming, “Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition”, *IEEE International Conference on Computer Vision*, Volume 1, 786–791, 2005.