Comparative Study of Local Descriptor for Face Description

J.Vikram [1], M.Gobi [2]

[1] Computer Science, Teacher, Agaram Public School.

Abstract- **This paper presents a novel and effective facial image representation based on Local Binary Pattern (LBP) texture features. The face image is split into several regions from which the LBP feature distributions are abstracted and linked into an augmented feature vector to be used as face descriptor. The accomplishment of the suggested method is determined in the face recognition problem under different attempts.**

1. INTRODUCTION

Finding good descriptors for the appearance of local facial regions is an open issue. Ideally, these descriptors should be easy to compute and have high extra-class variance (i.e., between different persons in the case of face recognition) and low intraclass variance, which means that the descriptor should be robust with respect to aging of the subjects, alternating illumination and other factors.

Using local photometric features [1] for object recognition in the more general context has become a widely accepted approach. In that setting, the typical approach is to detect interest points or interest regions in images, perform normalization with respect to affine transformations, and describe the normalized interest regions using local descriptors. This bag-of-keypoints approach is not suited for face description as such since it does not retain information on the spatial setting of the detected local regions but it does bear certain similarities to local feature-based face description.

One of the first face descriptors based on information extracted from local regions is the eigen features method proposed by Pentland et al. [2]. This is a hybrid approach in which the features are obtained by performing PCA to local face regions independently. In Local Feature Analysis [3], kernels of local spatial support are used to extract information about local facial components. Elastic Bunch Graph Matching (EBGM) [4] describes faces using Gabor filter responses in certain facial landmarks and a graph describing the spatial relations of these landmarks. The validity of the component-based approach is also attested by the study conducted by Heisele et al. in which a component-based face recognition system clearly outperformed global approaches on a test database containing faces rotated in depth [5].

AUTOMATIC face analysis which includes, e.g., face detection, face recognition, and facial expression recognition has become a very active topic in computer vision research [6]. A key issue in face analysis is finding efficient descriptors for face appearance. Different holistic methods such as Principal Component Analysis (PCA) [7], Linear Discriminant Analysis (LDA) [8], and the more recent 2D

PCA [9] have been studied widely but lately also local descriptors have gained attention due to their robustness to challenges such as pose and illumination changes. This paper presents a novel descriptor based on local binary pattern texture features extracted from local facial regions.

The texture analysis community has developed a variety of different descriptors for the appearance of image patches. However, face recognition problem has not been associated to that progress in texture analysis field as it has not been investigated from such point of view. Recently, we investigated the representation of face images by means of local binary pattern features, yielding in outstanding results that were published in the ECCV 2004 conference [10]. After this, several research groups have adopted our approach. In this paper, we provide a more detailed analysis of the proposed representation, present additional results and discuss further extensions.

2. LOCAL BINARY PATTERN

The LBP operator [11] is one of the best performing texture descriptors and it has been widely used in various applications. It has proven to be highly discriminative and its key advantages, namely, its invariance to monotonic gray-level changes and computational efficiency, make it suitable for demanding image analysis tasks

The idea of using LBP for face description is motivated by the fact that faces can be seen as a composition of micropatterns which are well described by such operator.

The LBP operator was originally designed for texture description. The operator assigns a label to every pixel of an image by thresholding the 3 _ 3-neighborhood of each pixel with the center pixel value and considering the result as a binary number. Then, the histogram of the labels can be used as a texture descriptor. See Fig. 1 for an illustration of the basic LBP operator.

To be able to deal with textures at different scales, the LBP operator was later extended to use neighborhoods of different sizes [12]. Defining the local neighborhood as a set of sampling points evenly spaced on a circle centered at the pixel to be labeled allows any radius and number of sampling

points. Bilinear interpolation is used when a sampling point does not fall in the center of a pixel. In the following, the notation (P, R) will be used for pixel neighborhoods which means P sampling points on a circle of radius of R. See Fig. 2 for an example of circular neighborhoods.

Another extension to the original operator is the definition of so-called uniform patterns [13]. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not. In the computation of the LBP histogram, uniform patterns are used so that the histogram has a separate bin for every uniform pattern and all nonuniform patterns are assigned to a single bin. Ojala et al. noticed that in their experiments with texture images, uniform patterns account for a bit less than 90 percent of all patterns when using the (8,1) neighborhood and for around 70 percent in the (16,2) neighborhood. We have found that 90.6 percent of the patterns in the (8,1) neighborhood and 85.2 percent of the patterns in the (8,2) neighborhood are uniform in case of preprocessed FERET facial images.

We use the following notation for the LBP operator: 〖LB〗_(P,R)^Pu2. The subscript represents using the operator in a (P,R) neighborhood. Superscript u2 stands for using only uniform patterns.

2.2 DESCRIPTION

In this work, the LBP method presented in the previous section is used for face description. The procedure consists of using the texture descriptor to build several local descriptions of the face and combining them into a global description. Instead of striving for a holistic description, this approach was motivated by two reasons: the local feature-based or hybrid approaches to face recognition have been gaining interest lately [3], [5], [13], which is understandable given the limitations of the holistic representations. These local feature-based and hybrid methods seem to be more robust against variations in pose or illumination than holistic methods.
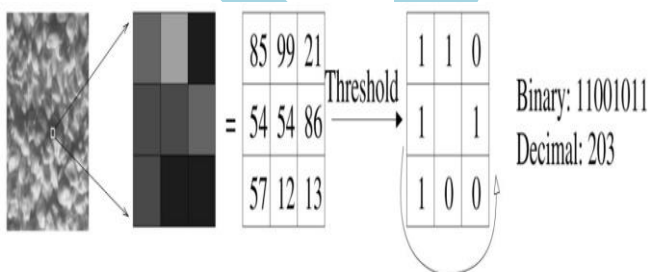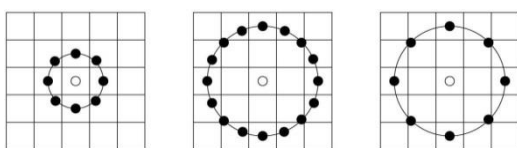

Fig. 1. The basic LBP Operator

Fig.2. The circular (8, 1), (16, 2), and (8, 2) neighborhoods. The pixel values are bilinearly interpolated whenever the sampling point is not in the center of the pixel.

Another reason for selecting the local feature-based approach is that trying to build a holistic description of a face using texture methods is not reasonable since texture descriptors tend to average over the image area. This is a desirable property for ordinary textures, because texture description should usually be invariant to translation or even rotation of the texture and, especially, for small repetitive textures, the small-scale relationships determine the appearance of the texture and, thus, the largescale relations do not contain useful information. For faces, however, the situation is different: retaining the information about spatial relations is important.

This reasoning leads to the basic methodology of this work. The facial image is divided into local regions and texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face. See Fig. 3 for an example of a facial image divided into rectangular regions.

The basic histogram can be extended into a spatially enhanced histogram which encodes both the appearance and the spatial relations of facial regions. As the m facial regions R0, R1, . . .Rm-1 have been determined, a histogram is computed independently within each of the m regions. The resulting m histograms are combined yielding the spatially enhanced histogram. The spatially enhanced histogram has size m x n, where n is the length of a single LBP histogram. In the spatially enhanced histogram, we effectively have a description of the face on three different levels of locality: The LBP labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level, and the regional histograms are concatenated to build a global description of the face.

It should be noted that when using the histogram-based methods, despite the examples in Fig. 3, the regions R0, R1, . . .Rm-1 do not need to be rectangular. Neither do they need to be of the same size or shape and they do not necessarily have to cover the whole image. For example, they could be circular regions located at the fiducial points like in the EBGM method. It is also possible to have partially overlapping regions. If recognition of faces rotated in depth is considered, it may be useful to follow the procedure of Heisele et al. [5] and automatically detect each region in the image instead of first detecting the face and then using a fixed division into regions.6

The idea of a spatially enhanced histogram can be exploited further when defining the distance measure. An indigenous property of the proposed face description method is that each element in the enhanced histogram corresponds to a certain small area of the face. Based on the psychophysical findings, which indicate that some facial features (such as eyes) play

---

more important roles in human face recognition than other features [14], it can be expected that, in this method, some of the facial regions contribute more than others in terms of extra personal variance. Utilizing this assumption, the regions can be weighted based on the importance of the information they contain. The following notation is used for the LBP operator: LBPP,Ru2. The subscript represents using the operator in a (P,R) neighborhood. Superscript u2 stands for using only uniform patterns and labeling all remaining patterns with a single label. After the LBP labeled image $f_l(x,y)$ has been obtained, the LBP histogram can be defined as

$$H_i = \sum_{x,y} I\{f_l(x,y) = i\}, i = 0,\dots,n-1,$$

in which n is the number of different labels produced by the LBP operator, and I{A} is 1 if A is true and 0 if A is false.

When the image patches whose histograms are to be compared have different sizes, the histograms must be normalized to get a coherent description:

$$N_i = \frac{H_i}{\sum_{j=0}^{n-1} H_j}.$$

3. FACE DESCRIPTION WITH – LBP

Our approach is assessed on the face recognition problem using the Colorado State University Face Identification Evaluation System [15] with images from the FERET [16] database. PCA [5], Bayesian Intra/Extrapersonal Classifier (BIC) [17], and EBGM were used as control algorithms.

To ensure the reproducibility of the tests, the publicly available CSU face identification evaluation system [15] was utilized to test the performance of the proposed algorithm. The system uses the FERET face images and follows the procedure of the FERET test for semiautomatic face recognition algorithms [18] with slight modifications.

The FERET database consists of a total of 14,051 gray-scale images representing 1,199 individuals. The images contain variations in lighting, facial expressions, pose angle, etc. In this work, only frontal faces are considered. These facial images can be divided into five sets as follows:

- fa set, used as a gallery set, contains frontal images of 1,196 people.
- fb set (1,195 images). The subjects were asked for an alternative facial expression than in the fa photograph.
- fc set (194 images). The photos were taken under different lighting conditions.
- dup I set (722 images). The photos were taken later in time.
- dup II set (234 images). This is a subset of the dup I set containing those images that were taken at least a year after the corresponding gallery image.

Along with recognition rates at rank 1, two statistical measures are used to compare the performance of the algorithms: the mean recognition rate with a 95 percent confidence interval and the probability of one algorithm outperforming another. The probability of one algorithm outperforming another is denoted by P(R(alg1)>R(alg2)). These statistics are computed by permuting the gallery and probe sets, see [15] for details. The CSU system comes with implementations of the PCA, LDA, BIC, and EBGM face recognition algorithms. The results obtained with PCA, BIC, and EBGM are included here for comparison.
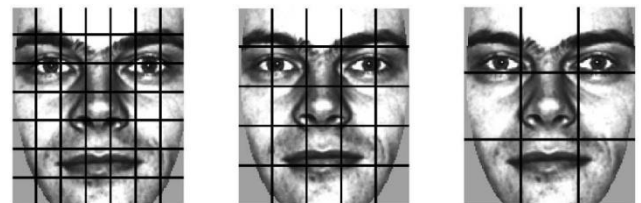

Fig.3. A facial image divided into 7x7, 5x5, and 3x3 rectangular regions

(a)   (b)

Fig.4. (a) A facial image divided into 7x7 windows. (b) The weights set for the weighted x2 dissimilarity measure. Black squares indicate weight 0.0, dark gray 1.0, light gray 2.0, and white 4.0

here are some parameters that can be chosen to optimize the performance of the LBP-based algorithm. These include choosing the type of the LBP operator, division of the images into regions $R_0, \dots, R_{m-1}$, selecting the distance measure for the nearest neighbor classifier, and finding the weights w_j for the weighted x^2 statistic (1). The extensive experiments to find the parameters for the proposed method are detailed in [10].

When looking for the optimal window size and LBP operator it was noticed that the LBP representation is quite robust with respect to the selection of parameters. Changes in the parameters may cause big differences in the length of the feature vector, but the overall performance is not necessarily affected significantly [10]. Here, the $LBP_{8,2}^{u2}$ operator in 18 x 21 pixel windows was selected since it is a good trade-off between recognition performance and feature vector length. When comparing different distance measures, the x^2 measure was found to perform better than histogram intersection or log-likelihood distance. Therefore, the x^2 measure was chosen to be used.

To find the weights w_j for the weighted x^2 statistic (1), a simple procedure was adopted in which a training set was classified using only one of the 18 x 21 windows at a time and the windows were assigned a weight based on the recognition rate.

The obtained weights are illustrated in Fig. 4b. The weights were selected without utilizing an actual optimization procedure and thus they are probably not optimal. Despite that, in comparison with the nonweighted method, an improvement both in the processing time and recognition rate (P(R(weighted) >R(nonweighted)) ¼ 0:976) was obtained.

To gain better understanding on whether the obtained recognition results are due to general idea of computing texture features from local facial regions or due to the discriminatory power of the local binary pattern operator, we compared LBP to three other texture descriptors, namely, the gray-level difference histogram, homogeneous texture descriptor [19], and an improved version of the texton histogram [20]. The details of these experiments can be found in [21].

The recognition rates obtained with different descriptors are shown in Table 1. It should be noted that no weighting for local regions was used in this experiment. The results show that the tested methods work well with the easiest fb probe set, which means that they are robust with respect to variations of facial expressions, whereas the results with the fc probe set show that other methods than LBP do not survive changes in illumination. The LBP and texton give the best results in the dup I and dup II test sets.

We believe that the main explanation for the better performance of the local binary pattern operator over other texture descriptors is its tolerance to monotonic gray-scale changes. Additional advantages are the computational efficiency of the LBP operator and that no gray-scale normalization is needed prior to applying the LBP operator to the face image.

4. RESULT

The final recognition results for the proposed method are shown in Table 2 and the rank curves are plotted in Fig. 5. LBP yields clearly higher recognition rates than the control algorithms in all the FERET test sets and in the statistical test. The results on the fc and dup II sets show that especially with weighting, the LBP-based description is robust to challenges caused by lighting changes or aging of the subjects but further research is still needed to achieve even better performance.

It should be noted that the CSU implementations of the algorithms whose results are included here do not achieve the same figures as in the original FERET test due to some modifications in the experimental setup as mentioned in [15]. The results of the original FERET test can be found in [18].

Robustness of the Method to Face Localization Error
Real-world face recognition systems need to perform face detection prior to face recognition. Automatic face localization may not be completely accurate so it is desirable that face recognition works under small localization errors.

The proposed face recognition method calculates histograms over the local regions so a small change in the position of the face relative to the grid causes changes in the labels only on the borders of the local regions. Therefore, it can be expected that the proposed method is not sensitive to small changes in the face localization and that using larger local regions increases the robustness to errors.

The effect of localization errors to recognition rate of the proposed method compared to PCA MahCosine was systematically tested as follows: The training images for PCA and gallery (fa) images were normalized to size 128 x 128 using provided eye coordinates. The fb set was used as probes. The probes were also normalized to size 128 _ 128 but a random vector ($\Delta X, \Delta Y$) was added to the face location, where ΔX and ΔY are uncorrelated and normally distributed with mean 0 and standard deviation σ. Ten experiments were conducted with each probe totaling 11,950 queries for each tested σ value.

The recognition rates of the LBP-based method using window sizes 21 x 21 and 32 x 32 and PCA Mach Cosine as a function of the standard deviation of the simulated localization offset are plotted in Fig. 6. It can be seen that when no error or only a small error is present, LBP with small local regions works well but as the localization error increases, using larger local regions produces better recognition rate. Most interestingly, the recognition rate of the local region-based methods drops significantly slower than that of PCA.

Coding to get pixels value from matlab

```
impixel(I)
P = impixel(I,c,r)
P = impixel(X,map)
P = impixel(X,map,c,r)
[c,r,P] = impixel(___)
P = impixel(x,y,I,xi,yi)
[xi,yi,P] = impixel(x,y,I,xi,yi)
```

Read a truecolor image into the workspace.
```
RGB = imread('peppers.png');
```

Determine the column c and row r indices of the pixels to extract.
```
c = [1 12 146 410];
r = [1 104 156 129];
```

Return the data at the selected pixel locations.
```
pixels = impixel(RGB,c,r)

pixels =

    62    29    64
    62    34    63
   166    54    60
    59    28    47
```
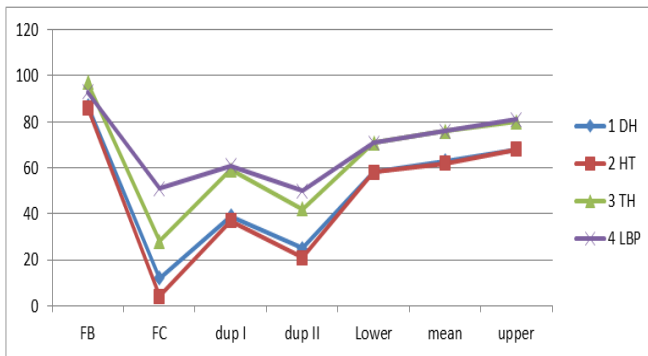
TABLE 1

The Recognition Rates Obtained Using Texture Descriptors of Local Facial Regions

Method	fb	fc	dup I	dup II	lower	mean	upper
Difference histogram	0.87	0.12	0.39	0.25	0.58	0.63	0.68
Homogeneous texture	0.86	0.04	0.37	0.21	0.58	0.62	0.68
Texton Histogram	0.97	0.28	0.59	0.42	0.71	0.76	0.80
LBP (nonweighted)	0.93	0.51	0.61	0.50	0.71	0.76	0.81

the first four columns show the recognition rates for the FERET test sets and the last three columns contain the mean recognition rate of the permutation test with a 95 percent confidence interval.

Graph of the Recognition Rates Obtained Using Texture Descriptors of Local Facial Regions

5. CONCLUSION

In this paper, a novel and efficient facial representation is proposed. It is based on dividing a facial image into small regions and computing a description of each region using local binary patterns. These descriptors are then combined into a spatially enhanced histogram or feature vector. The texture description of a single region describes the appearance of the region and the combination of all region descriptions encodes the global geometry of the face.

The LBP operator has been widely used in different applications such as texture classification, image retrieval, etc. Before our study, it was not obvious to imagine that such texture operator might be useful in representing also facial images. Our results clearly show that facial images can be seen as a composition of micropatterns such as flat areas, spots, lines, and edges which can be well described by LBP.

In this paper, the proposed methodology is assessed with the face recognition task. However, a similar method has yielded in outstanding performance in face detection [22] and facial expression recognition [23], [24]. We also believe that the developed approach is not limited to these few examples as it can be easily generalized to other types of object detection and recognition tasks.

We are also included the more advanced methods for dividing the facial image into local regions and finding the weights for these regions. The AdaBoost method presented in [25] serves as a good basis for this research. Another important topic is looking for image preprocessing methods and descriptors that are more robust against image transformations that change the appearance of the surface texture such as image blurring caused by imaging device being slightly out-of-focus.

REFERENCES

(1) K.Mikolajczyk and C.Schmid, "A Performance Evaluation of Local Descriptors," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1615-1630, Oct. 2005.

(2) A.Pentland, B.Moghaddam, and T.Starner, "View-Based and Modular Eigenspaces for Face Recognition," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, pp. 84-91, 1994.

(3) P.S.Penev and J.J.Atick, "Local Feature Analysis: A General Statistical Theory for Object Representation," Network-Computation in Neural Systems, vol. 7, no. 3, pp. 477-500, Aug. 1996.

(4) L.Wiskott, J.-M.Fellous, N.Kuiger, and C.von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, pp. 775-779, 1997.

(5) B.Heisele, P.Ho, J.Wu, and T. Poggio, "Face Recognition: Component- Based versus Global Approaches," Compter Vision and Image Understanding, vol. 91, nos. 1-2, pp. 6-21, 2003.

(6) Handbook of Face Recognition, S.Z. Li and A.K. Jain, eds. Springer, 2005.

(7) M.Turk and A.Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

(8) K.Etemad and R.Chellappa, "Discriminant Analysis for Recognition of Human Face Images," J. Optical Soc. Am., vol. 14, pp. 1724-1733, 1997.

(9) J.Yang, D.Zhang, A.F.Frangi, and J. Yang, "Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131 137, Jan 2004.

(10) T.Ahonen, A.Hadid, and M.Pietika¨inen, "Face Recognition with Local Binary Patterns," Proc. Eighth European Conf. Computer Vision, pp. 469-481, 2004.

(11) T.Ojala, M.Pietikainen, and D.Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.

(12) T.Ojala, M.Pietikainen, and T.Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture

Classification with Local Binary Patterns," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, July 2002.

(13) R.Gottumukkal and V.K.Asari, "An Improved Face Recognition Technique Based on Modular PCA Approach," Pattern Recognition Letters, vol. 25, pp. 429-436, Mar. 2004.

(14) W.Zhao, R.Chellappa, P.J.Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, Dec.2003.

(15) J.R.Beveridge, D.Bolme, B.A.Draper, and M.Teixeira, "The CSU Face Identification Evaluation System: Its Purpose, Features, and Structure," Machine Vision and Applications, vol. 16, no. 2, pp. 128-138, Feb. 2005.

(16) P.J.Phillips, H.Wechsler, J.Huang, and P.Rauss, "The FERET Database and Evaluation Procedure for Face Recognition Algorithms," Image and Vision Computing, vol. 16, no. 10, pp. 295-306, Apr. 1998.

(17) B.Moghaddam, C.Nastar, and A.Pentland, "A Bayesian Similarity Measure for Direct Image Matching," Proc. 13th Int'l Conf. Pattern Recognition, pp. II: 350-358, 1996.

(18) P.J.Phillips, H. Moon, S.A.Rizvi, and P.J.Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1090-1104, Oct. 2000.

(19) B.S.Manjunath, J.-R.Ohm, V.V.Vasudevan, and A. Yamada, "Color and Texture Descriptors," IEEE Trans. Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 703-715, June 2001.

(20) M.Varma and A.Zisserman, "Texture Classification: Are Filter Banks Necessary?" Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, pp. 691-698, 2003.

(21) T.Ahonen, M.Pietika¨inen, A.Hadid, and T.Ma¨enpa¨a¨, "Face Recognition Based on Appearance of Local Regions," Proc. 17th Int'l Conf. Pattern Recognition, pp. III: 153-156, 2004.

(22) A.Hadid, M.Pietika¨inen, and T.Ahonen, "A Discriminative Feature Space for Detecting and Recognizing Faces," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 797-804, 2004.

(23) X. Feng, M.Pietika¨inen, and A.Hadid, "Facial Expression Recognition with Local Binary Patterns and Linear Programming," Pattern Recognition and Image Analysis, vol. 15, no. 2, pp. 546-548, 2005.

(24) C.Shan, S.Gong, and P.W.McOwan, "Robust Facial Expression Recognition Using Local Binary Patterns," Proc. IEEE Int'l Conf. Image Processing, pp. II: 914-917, 2005.

(25) G.Zhang, X.Huang, S.Z.Li, Y.Wang, and X. Wu, "Boosting Local Binary Pattern (LBP)-Based Face Recognition," Proc. Advances in Biometric Person Authentication, pp. 179-186, 2004.