Face Recognition with Local Binary Patterns

Conference Paper *in* Lecture Notes in Computer Science · May 2004

DOI: 10.1007/978-3-540-24670-1_36 · Source: DBLP

3 authors, including:

Matti Pietikäinen
University of Oulu
329 PUBLICATIONS **66,714** CITATIONS

Face Recognition with Local Binary Patterns

Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen

Machine Vision Group, Infotech Oulu
PO Box 4500, FIN-90014 University of Oulu, Finland,
{tahonen,hadid,mkp}@ee.oulu.fi, http://www.ee.oulu.fi/mvg/

Abstract. In this work, we present a novel approach to face recognition which considers both shape and texture information to represent face images. The face area is first divided into small regions from which Local Binary Pattern (LBP) histograms are extracted and concatenated into a single, spatially enhanced feature histogram efficiently representing the face image. The recognition is performed using a nearest neighbour classifier in the computed feature space with Chi square as a dissimilarity measure. Extensive experiments clearly show the superiority of the proposed scheme over all considered methods (PCA, Bayesian Intra/extrapersonal Classifier and Elastic Bunch Graph Matching) on FERET tests which include testing the robustness of the method against different facial expressions, lighting and aging of the subjects. In addition to its efficiency, the simplicity of the proposed method allows for very fast feature extraction.

1 Introduction

The availability of numerous commercial face recognition systems [1] attests to the significant progress achieved in the research field [2]. Despite these achievements, face recognition continues to be an active topic in computer vision research. This is due to the fact that current systems perform well under relatively controlled environments but tend to suffer when variations in different factors (such as pose, illumination etc.) are present. Therefore, the goal of the ongoing research is to increase the robustness of the systems against different factors. Ideally, we aim to develop a face recognition system which mimics the remarkable capabilities of human visual perception. Before attempting to reach such a goal, one needs to continuously learn the strengths and weaknesses of the proposed techniques in order to determine new directions for future improvements. To facilitate this task, the FERET database and evaluation methodology have been created [3]. The main goal of FERET is to compare different face recognition algorithms on a common and large database and evaluate their performance against different factors such as facial expression, illumination changes, aging (time between the acquisition date of the training image and the image presented to the algorithm) etc.

Among the major approaches developed for face recognition are Principal Component Analysis (PCA) [4], Linear Discriminant Analysis (LDA) [5] and

T. Pajdla and J. Matas (Eds.): ECCV 2004, LNCS 3021, pp. 469–481, 2004.

Elastic Bunch Graph Matching (EBGM) [6]. PCA is commonly referred to as the "eigenface" method. It computes a reduced set of orthogonal basis vectors or eigenfaces of the training face images. A new face image can be approximated by a weighted sum of these eigenfaces. PCA provides an optimal linear transformation from the original image space to an orthogonal eigenspace with reduced dimensionality in the sense of least mean squared reconstruction error. LDA seeks to find a linear transformation by maximising the between-class variance and minimising the within-class variance. In the EBGM algorithm, faces are represented as graphs, with nodes positioned at fiducial points and edges labelled with distance vectors. Each node contains a set of Gabor wavelet coefficients, known as a jet. Thus, the geometry of the face is encoded by the edges while the grey value distribution (texture) is encoded by the jets. The identification of a new face consists of determining among the constructed graphs, the one which maximises the graph similarity function. Another proposed approach to face recognition is the Bayesian Intra/extrapersonal Classifier (BIC) [7] which uses the Bayesian decision theory to divide the difference vectors between pairs of face images into two classes: one representing intrapersonal differences (i.e. differences in a pair of images representing the same person) and extrapersonal differences.

In this work, we introduce a new approach for face recognition which considers both shape and texture information to represent the face images. As opposed to the EBGM approach, a straightforward extraction of the face feature vector (histogram) is adopted in our algorithm. The face image is first divided into small regions from which the Local Binary Pattern (LBP) features [8,9] are extracted and concatenated into a single feature histogram efficiently representing the face image. The textures of the facial regions are locally encoded by the LBP patterns while the whole shape of the face is recovered by the construction of the face feature histogram. The idea behind using the LBP features is that the face images can be seen as composition of micro-patterns which are invariant with respect to monotonic grey scale transformations. Combining these micro-patterns, a global description of the face image is obtained.

2 Face Description with Local Binary Patterns

The original LBP operator, introduced by Ojala *et al.* [9], is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. See Figure 1 for an illustration of the basic LBP operator.

Later the operator was extended to use neigbourhoods of different sizes [8]. Using circular neighbourhoods and bilinearly interpolating the pixel values allow any radius and number of pixels in the neighbourhood. For neighbourhoods we will use the notation (P, R) which means P sampling points on a circle of radius of R. See Figure 2 for an example of the circular (8,2) neighbourhood.

Another extension to the original operator uses so called *uniform patterns* [8]. A Local Binary Pattern is called uniform if it contains at most two bitwise

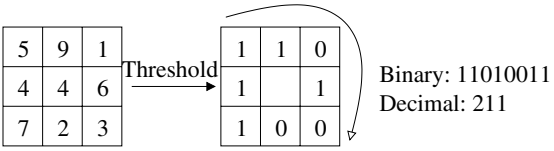

Fig. 1. The basic LBP operator.

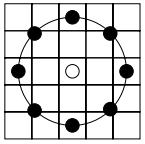

Fig. 2. The circular (8,2) neigbourhood. The pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.

transitions from 0 to 1 or vice versa when the binary string is considered circular. For example, 00000000, 00011110 and 10000011 are uniform patterns. Ojala *et al.* noticed that in their experiments with texture images, uniform patterns account for a bit less than 90 % of all patterns when using the (8,1) neighbourhood and for around 70 % in the (16,2) neighbourhood.

We use the following notation for the LBP operator: $\mathrm{LBP}^{u2}_{P,R}$. The subscript represents using the operator in a (P, R) neighbourhood. Superscript $u2$ stands for using only uniform patterns and labelling all remaining patterns with a single label.

A histogram of the labeled image $f_l(x, y)$ can be defined as

$$H_i = \sum_{x,y} I\left\{f_l(x,y) = i\right\}, i = 0, \ldots, n-1, \tag{1}$$

in which n is the number of different labels produced by the LBP operator and

$$I\{A\} = \begin{cases} 1, & \text{A is true} \\ 0, & \text{A is false.} \end{cases}$$

This histogram contains information about the distribution of the local micropatterns, such as edges, spots and flat areas, over the whole image. For efficient face representation, one should retain also spatial information. For this purpose, the image is divided into regions $R_0, R_1, \ldots R_{m-1}$ (see Figure 5 (a)) and the spatially enhanced histogram is defined as

$$H_{i,j} = \sum_{x,y} I\left\{f_l(x,y) = i\right\} I\left\{(x,y) \in R_j\right\}, i = 0, \ldots, n-1, j = 0, \ldots, m-1. \tag{2}$$

In this histogram, we effectively have a description of the face on three different levels of locality: the 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.

From the pattern classification point of view, a usual problem in face recognition is having a plethora of classes and only a few, possibly only one, training sample(s) per class. For this reason, more sophisticated classifiers are not needed but a nearest-neighbour classifier is used. Several possible dissimilarity measures have been proposed for histograms:

– Histogram intersection:

$$D(\mathbf{S}, \mathbf{M}) = \sum_i \min(S_i, M_i) \tag{3}$$

– Log-likelihood statistic:

$$L(\mathbf{S}, \mathbf{M}) = -\sum_i S_i \log M_i \tag{4}$$

– Chi square statistic (χ^2):

$$\chi^2(\mathbf{S}, \mathbf{M}) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \tag{5}$$

All of these measures can be extended to the spatially enhanced histogram by simply summing over i and j.

When the image has been divided into regions, it can be expected that some of the regions contain more useful information than others in terms of distinguishing between people. For example, eyes seem to be an important cue in human face recognition [2,10]. To take advantage of this, a weight can be set for each region based on the importance of the information it contains. For example, the weighted χ^2 statistic becomes

$$\chi^2_w(\mathbf{S}, \mathbf{M}) = \sum_{i,j} w_j \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}}, \tag{6}$$

in which w_j is the weight for region j.

3 Experimental Design

The CSU Face Identification Evaluation System [11] was utilised to test the performance of the proposed algorithm. The system follows the procedure of the FERET test for semi-automatic face recognition algorithms [12] with slight modifications. The system uses the full-frontal face images from the FERET database and works as follows (see Figure 3):

1. The system preprocesses the images. The images are registered using eye coordinates and cropped with an elliptical mask to exclude non-face area from the image. After this, the grey histogram over the non-masked area is equalised.
2. If needed, the algorithm is trained using a subset of the images.

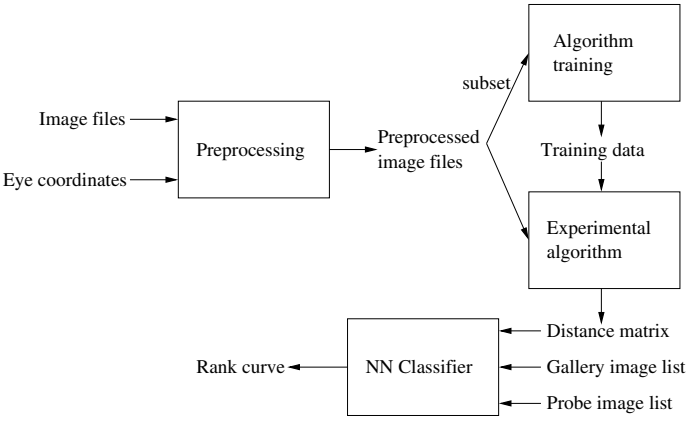

Fig. 3. The parts of the CSU face recognition system.

3. The preprocessed images are fed into the experimental algorithm which outputs a distance matrix containing the distance between each pair of images.
4. Using the distance matrix and different settings for gallery and probe image sets, the system calculates rank curves for the system. These can be calculated for prespecified gallery and probe image sets or by choosing a random permutations of one large set as probe and gallery sets and calculating the average performance. The advantage of the prior method is that it is easy to measure the performance of the algorithm under certain challenges (e.g. different lighting conditions) whereas the latter is more reliable statistically.

The CSU system uses the same gallery and probe image sets that were used in the original FERET test. Each set contains at most one image per person. These sets are:

- **fa** set, used as a gallery set, contains frontal images of 1196 people.
- **fb** set (1195 images). The subjects were asked for an alternative facial expression than in 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.

In this paper, we use two statistics produced by the permutation tool: the mean recognition rate with a 95 % confidence interval and the probability of one algorithm outperforming another [13]. The image list used by the tool[1] contains 4 images of each of the 160 subjects. One image of every subject is selected to the gallery set and another image to the probe set on each permutation. The number of permutations is 10000.

[1] list640.srt in the CSU Face Identification Evaluation System package

The CSU system comes with implementations of the PCA, LDA, Bayesian intra/extrapersonal (BIC) and Elastic Bunch Graph Matching (EBGM) face recognition algorithms. We include the results obtained with PCA, BIC[2] and EBGM here for comparison.

There are some parameters that can be chosen to optimise the performance of the proposed algorithm. The first one is choosing the LBP operator. Choosing an operator that produces a large amount of different labels makes the histogram long and thus calculating the distace matrix gets slow. Using a small number of labels makes the feature vector shorter but also means losing more information. A small radius of the operator makes the information encoded in the histogram more local. The number of labels for a neighbourhood of 8 pixels is 256 for standard LBP and 59 for LBP^{u2}. For the 16-neighbourhood the numbers are 65536 and 243, respectively. The usage of uniform patterns is motivated by the fact that most patterns in facial images are uniform: we found out that in the preprocessed FERET images, 79.3 % of all the patterns produced by the $\text{LBP}_{16,2}$ operator are uniform.

Another parameter is the division of the images into regions $R_0, \ldots, R_{m-1}$. The length of the feature vector becomes $B = mB_r$, in which m is the number of regions and B_r is the LBP histogram length. A large number of small regions produces long feature vectors causing high memory consumption and slow classification, whereas using large regions causes more spatial information to be lost. We chose to divide the image with a grid into $k * k$ equally sized rectangular regions (windows). See Figure 5 (a) for an example of a preprocessed facial image divided into 49 windows.

4 Results

To assess the performance of the three proposed distance measures, we chose to use two different LBP operators in windows of varying size. We calculated the distance matrices for each of the different settings and used the permutation tool to calculate the probabilities of the measures outperforming each other. The results are in Table 1.

From the statistical hypothesis testing point of view, it cannot be said that any of the metrics would be the best one with a high (>0.95) probability. However, histogram intersection and χ^2 measures are clearly better than log-likelihood when the average number of labels per histogram bin is low but log-likelihood performs better when this number increases. The log-likelihood measure has been preferred for texture images [8] but because of its poor performance on small windows in our experiments it is not appealing for face recognition. The χ^2 measure performs slightly better than histogram intersection so we chose to use it despite the simplicity of the histogram intersection.

When looking for the optimal window size and LBP operator we noticed that the LBP representation is quite robust with respect to the selection of the

[2] Two decision rules can be used with the BIC classifier: Maximum A Posteriori (MAP) or Maximum Likelihood (ML). We include here the results obtained with MAP.

Table 1. The performance of the histogram intersection, log-likelihood and χ^2 dissimilarity measures using different window sizes and LBP operators.

Operator	Window size	P(HI > LL)	P(χ^2 > HI)	P(χ^2 > LL)
$LBP^{u2}_{8,1}$	18x21	1.000	0.714	1.000
$LBP^{u2}_{8,1}$	21x25	1.000	0.609	1.000
$LBP^{u2}_{8,1}$	26x30	0.309	0.806	0.587
$LBP^{u2}_{16,2}$	18x21	1.000	0.850	1.000
$LBP^{u2}_{16,2}$	21x25	1.000	0.874	1.000
$LBP^{u2}_{16,2}$	26x30	1.000	0.918	1.000
$LBP^{u2}_{16,2}$	32x37	1.000	0.933	1.000
$LBP^{u2}_{16,2}$	43x50	0.085	0.897	0.418

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. For example, changing from $LBP^{u2}_{16,2}$ in 18*21-sized windows to $LBP^{u2}_{8,2}$ in 21*25-sized windows drops the histogram length from 11907 to 2124, while the mean recognition rate reduces from 76.9 % to 73.8 %.

The mean recognition rates for the $LBP^{u2}_{16,2}$, $LBP^{u2}_{8,2}$ and $LBP^{u2}_{8,1}$ as a function of the window size are plotted in Figure 4. The original 130*150 pixel image was divided into $k * k$ windows, $k = 4, 5, \ldots, 11, 13, 16$ resulting in window sizes from 32*37 to 8*9. The five smallest windows were not tested using the $LBP^{u2}_{16,2}$ operator because of the high dimension of the feature vector that would have been produced. As expected, a larger window size induces a decreased recognition rate because of the loss of spatial information. The $LBP^{u2}_{8,2}$ operator in 18*21 pixel windows was selected since it is a good trade-off between recognition performance and feature vector length.

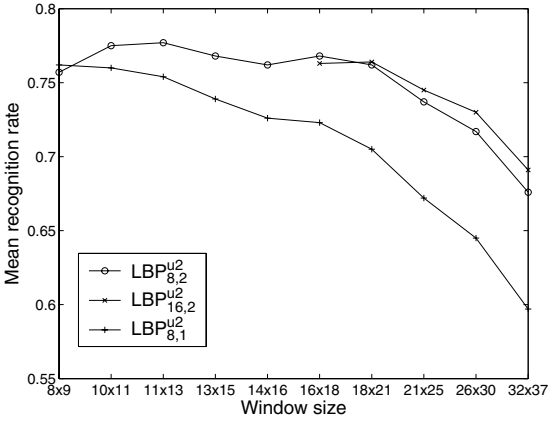

Fig. 4. The mean recognition rate for three LBP operators as a function of the window size.

To find the weights w_j for the weighted χ^2 statistic (Equation 6), the following procedure was adopted: a training set was classified using only one of the 18*21 windows at a time. The recognition rates of corresponding windows on the left and right half of the face were averaged. Then the windows whose rate lay below the 0.2 percentile of the rates got weight 0 and windows whose rate lay above the 0.8 and 0.9 percentile got weights 2.0 and 4.0, respectively. The other windows got weight 1.0.

The CSU system comes with two training sets, the standard FERET training set and the CSU training set. As shown in Table 2, these sets are basically subsets of the *fa*, *fb* and *dup I* sets. Since illumination changes pose a major challenge to most face recognition algorithms and none of the images in the *fc* set were included in the standard training sets, we defined a third training set, called the subfc training set, which contains half of the *fc* set (subjects 1013–1109).

Table 2. Number of images in common between different training and testing sets.

Training set	fa	fb	fc	dup I	dup II	Total number of images
FERET standard	270	270	0	184	0	736
CSU standard	396	0	0	99	0	501
subfc	97	0	97	0	0	194

The permutation tool was used to compare the weights computed from the different training sets. The weights obtained using the FERET standard set gave an average recognition rate of 0.80, the CSU standard set 0.78 and the subfc set 0.81. The pairwise comparison showed that the weights obtained with the subfc set are likely to be better than the others (P(subfc > FERET)=0.66 and P(subfc > CSU)=0.88).

The weights computed using the subfc set are illustrated in Figure 5 (b). The weights were selected without utilising an actual optimisation procedure and thus they are probably not optimal. Despite that, in comparison with the nonweighted method, we got an improvement both in the processing time (see Table 3) and recognition rate (P(weighted > nonweighted)=0.976).

The image set which was used to determine the weights overlaps with the *fc* set. To avoid biased results, we preserved the other half of the *fc* set (subjects

(a) (b)

Fig. 5. (a) An example of a facial image divided into 7x7 windows. (b) The weights set for weighted χ^2 dissimilarity measure. Black squares indicate weight 0.0, dark grey 1.0, light grey 2.0 and white 4.0.

Table 3. Processing times of weighted and nonweighted LBP on a 1800 MHz AMD Athlon running Linux. Note that processing FERET images (last column) includes heavy disk operations, most notably writing the distance matrix of about 400 MB to disk.

Type of LBP	**Feature ext.** (ms / image)	**Distance calc.** (μs / pair)	**Processing FERET images** (s)
Weighted	3.49	46.6	1046
Nonweighted	4.14	58.6	1285

1110-1206) as a validation set. Introducing the weights increased the recognition rate for the training set from 0.49 to 0.81 and for the validation set from 0.52 to 0.77. The improvement is slightly higher for the training set, but the significant improvement for the validation set implies that the calculated weights generalize well outside the training set.

The final recognition results for the proposed method are in shown Table 4 and the rank curves are plotted in Figures 6 (a)–(d). LBP clearly outperforms the control algorithms in all the FERET test sets and in the statistical test. It should be noted that the CSU implementations of the algorithms whose results we 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 [11]. The results of the original FERET test can be found in [12].

Table 4. The recognition rates of the LBP and comparison algorithms for the FERET probe sets and the mean recognition rate of the permutation test with a 95 % confidence interval.

Method	**fb**	**fc**	**dup I**	**dup II**	**lower**	**mean**	**upper**
LBP, weighted	0.97	0.79	0.66	0.64	0.76	0.81	0.85
LBP, nonweighted	0.93	0.51	0.61	0.50	0.71	0.76	0.81
PCA, MahCosine	0.85	0.65	0.44	0.22	0.66	0.72	0.78
Bayesian, MAP	0.82	0.37	0.52	0.32	0.67	0.72	0.78
EBGM_Optimal	0.90	0.42	0.46	0.24	0.61	0.66	0.71

Additionally, to gain knowledge about the robustness of our method against slight variations of pose angle and alignment we tested our approach on the ORL face databse (Olivetti Research Laboratory, Cambridge) [14]. The database contains 10 different images of 40 distinct subjects (individuals). Some images were taken at different times for some people. There are variations in facial expression (open/closed eyes, smiling/non-smiling.), facial details (glasses/no glasses) and scale (variation of up to about 10 %). All the images were taken against a dark homogenous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. The images are grey scale with a resolution of 92*112. Randomly selecting 5 images for the gallery set and the other 5 for the probe set, the preliminary experiments result in 0.98 of average recognition rate and 0.012 of standard

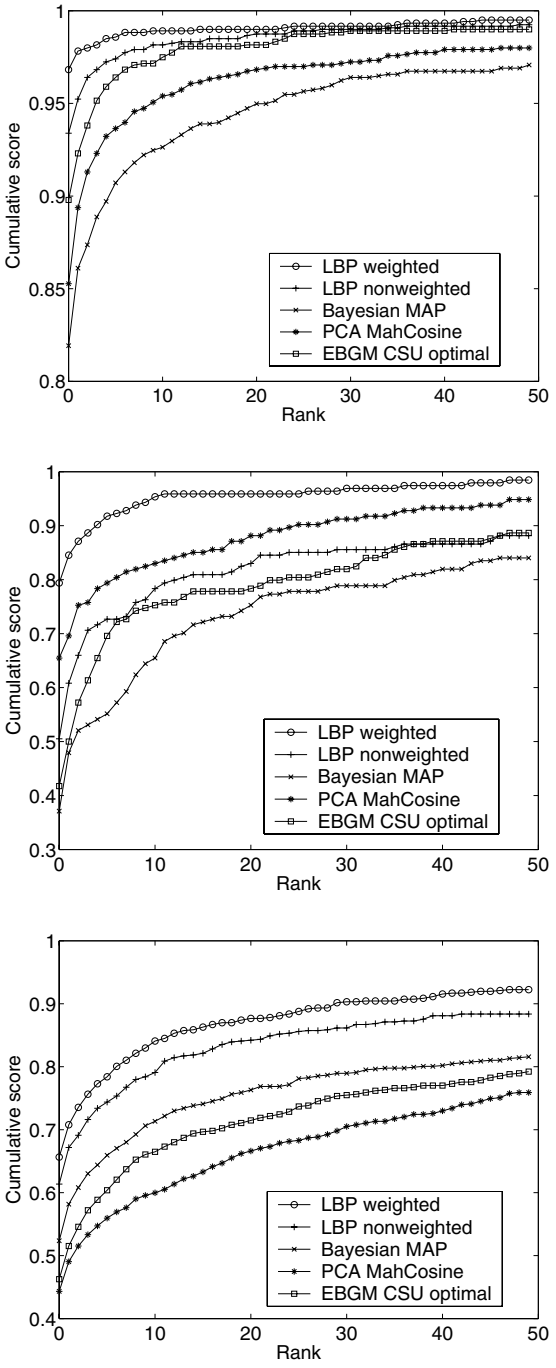

Fig. 6. (a), (b), (c) Rank curves for the *fb*, *fc* and *dup1* probe sets (from top to down).

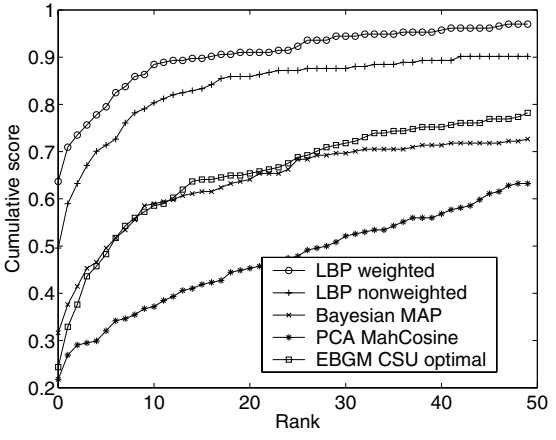

Fig. 6. (d) Rank curve for the *dup2* probe set.

deviation of 100 random permutations using $\mathrm{LBP}^{u2}_{16,2}$, a windows size of 30*37 and χ^2 as a dissimilarity measure. Window weights were not used. Note that no registration or preprocessing was made on the images. The good results indicate that our approach is also relatively robust with respect to alignment. However, because of the lack of a standardised protocol for evaluating and comparing systems on the ORL database, it is to difficult to include here a fair comparison with other approaches that have been tested using ORL.

5 Discussion and Conclusion

Face images can be seen as a composition of micro-patterns which can be well described by LBP. We exploited this observation and proposed a simple and efficient representation for face recognition. In our approach, a face image is first divided into several blocks (facial regions) from which we extract local binary patterns and construct a global feature histogram that represents both the statistics of the facial micro-patterns and their spatial locations. Then, face recognition is performed using a nearest neighbour classifier in the computed feature space with χ^2 as a dissimilarity measure. The proposed face representation can be easily extracted in a single scan through the image, without any complex analysis as in the EBGM algorithm.

We implemented the proposed approach and compared it against well-known methods such as PCA, EBGM and BIC. To achieve a fair comparison, we considered the FERET face database and protocol, which are a *de facto* standard in face recognition research. In addition, we adopted normalisation steps and implementation of the different algorithms (PCA, EBGM and BIC) from the CSU face identification evaluation system. Reporting our results in such a way does not only make the comparative study fair but also offers the research community new performances to which they are invited to compare their results.

The experimental results clearly show that the LBP-based method outperforms other approaches on all probe sets (*fb*, *fc*, *dup I* and *dup II*). For instance, our method achieved a recognition rate of 97% in the case of recognising faces under different facial expressions (*fb* set), while the best performance among the tested methods did not exceed 90%. Under different lighting conditions (*fc* set), the LBP-based approach has also achieved the best performance with a recognition rate of 79% against 65%, 37% and 42% for PCA, BIC and EBGM, respectively. The relatively poor results on the *fc* set confirm that illumination change is still a challenge to face recognition. Additionally, recognising duplicate faces (when the photos are taken later in time) is another challenge, although our proposed method performed better than the others.

To assess the performance of the LBP-based method on different datasets, we also considered the ORL face database. The experiments not only confirmed the validity of out approach, but also demonstrated its relative robustness against changes in alignment.

Analyzing the different parameters in extracting the face representation, we noticed a relative insensitivity to the choice of the LBP operator and region size. This is an interesting result since the other considered approaches are more sensitive to their free parameters. This means that only simple calculations are needed for the LBP description while some other methods use exhaustive training to find their optimal parameters.

In deriving the face representation, we divided the face image into several regions. We used only rectangular regions each of the same size but other divisions are also possible as regions of different sizes and shapes could be used. To improve our system, we analyzed the importance of each region. This is motivated by the psychophysical findings which indicate that some facial features (such as eyes) play more important roles in face recognition than other features (such as the nose). Thus we calculated and assigned weights from 0 to 4 to the regions (See Figure 5 (b)). Although this kind of simple approach was adopted to compute the weights, improvements were still obtained. We are currently investigating approaches for dividing the image into regions and finding more optimal weights for them.

Although we clearly showed the simplicity of LBP-based face representation extraction and its robustness with respect to facial expression, aging, illumination and alignment, some improvements are still possible. For instance, one drawback of our approach lies in the length of the feature vector which is used for face representation. Indeed, using a feature vector length of 2301 slows down the recognition speed especially, for very large face databases. A possible direction is to apply a dimensionality reduction to the face feature vectors. However, due to the good results we have obtained, we expect that the methodology presented here is applicable to several other object recognition tasks as well.

Acknowledgements. This research was supported in part by the Academy of Finland.

References

1. Phillips, P., Grother, P., Micheals, R.J., Blackburn, D.M., Tabassi, E., Bone, J.M.: Face recognition vendor test 2002 results. Technical report (2003)
2. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.J.: Face recognition: a literature survey. Technical Report CAR-TR-948, Center for Automation Research, University of Maryland (2002)
3. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. Image and Vision Computing **16** (1998) 295–306
4. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience **3** (1991) 71–86
5. Etemad, K., Chellappa, R.: Discriminant analysis for recognition of human face images. Journal of the Optical Society of America **14** (1997) 1724–1733
6. Wiskott, L., Fellous, J.M., Kuiger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Transaction on Pattern Analysis and Machine Intelligence **19** (1997) 775–779
7. Moghaddam, B., Nastar, C., Pentland, A.: A bayesian similarity measure for direct image matching. In: 13th International Conference on Pattern Recognition. (1996) II: 350–358
8. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 971–987
9. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition **29** (1996) 51–59
10. Gong, S., McKenna, S.J., Psarrou, A.: Dynamic Vision, From Images to Face Recognition. Imperial College Press, London (2000)
11. Bolme, D.S., Beveridge, J.R., Teixeira, M., Draper, B.A.: The CSU face identification evaluation system: Its purpose, features and structure. In: Third International Conference on Computer Vision Systems. (2003) 304–311
12. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The FERET evaluation methodology for face recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 1090–1104
13. Beveridge, J.R., She, K., Draper, B.A., Givens, G.H.: A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2001) I: 535–542
14. Samaria, F.S., Harter, A.C.: Parameterisation of a stochastic model for human face identification. In: IEEE Workshop on Applications of Computer Vision. (1994) 138–142