

Local binary pattern based face recognition with automatically detected fiducial points

Ladislav Lenc^{a,b,*} and Pavel Král^{a,b}

^a*Department of Computer Science and Engineering, University of West Bohemia, Plzeň, Czech*

^b*NTIS – New Technologies for the Information Society, University of West Bohemia, Plzeň, Czech*

Abstract. This paper deals with automatic face recognition in the context of a real application for person identification developed for the Czech News Agency (ČTK). We focus on popular Local Binary Patterns (LBP) that are frequently used in this field with high recognition accuracy. One drawback of current LBP based methods is that the positions and number of the fiducial points are fixed. These points thus do not reflect the properties of a particular image whereas we believe it is beneficial to identify the most representative ones. The main contribution consists in proposing and comparing several LBP-based approaches that detect such points fully automatically. We use a set of Gabor filters for this task. Local extrema in the filter responses are detected and then used as the feature points. The number of points is further reduced by a clustering algorithm. Our approaches also differ from the other ones in the matching procedure. The proposed methods are evaluated on three standard corpora: ORL, FERET, AR face database and our ČTK dataset containing uncontrolled face images. Recognition results clearly show high quality of the proposed approaches that outperform significantly the baseline LBP approach on all corpora. The benefits of our methods are particularly evident in the case of real non-controlled images (ČTK corpus) where the accuracy is increased by more than 20% in absolute value.

Keywords: Automatic face recognition, czech news agency, gabor filter, local binary patterns, LBP

1. Introduction

Automatic identification of persons using their biometric characteristics became an important research and practice area. Numerous identification systems which use/ combine different types of biometrics have been proposed. The most important biometric information are DNA, iris, face, signature, fingerprint, voice, gait and palmprint [17]. When we compare their respective advantages and drawbacks, we must conclude that face recognition belongs to the most progressive biometric techniques today.

There are numerous applications of the automatic face recognition as for instance an access control, searching of wanted criminals, face annotation in social networks or in photo sharing applications, and

so on. An important progress in this field was done during the last decades and the number of proposed approaches is overwhelming. Face recognition from controlled images (i.e. small variances in face pose, tilt, lighting conditions, etc.) is considered to be sufficiently solved. However, the face recognition from real photographs is an open issue yet, because very sophisticated approaches are required.

Local Binary Patterns (LBP) are a very efficient approach for feature extraction. It encodes differences between pixel intensities in a local neighbourhood of a pixel. It is used in various computer vision applications. In the case of the face recognition, histograms of LBP values are used. The face is usually separated into rectangular regions and histograms are computed in each of them. The concatenation of these histogram values is used as the face representation. The main strength of this method consists in its high capability to retain important information in the images and in its low computational cost. However, most existing LBP-

*Corresponding author: Ladislav Lenc, Department of Computer Science and Engineering, University of West Bohemia, Plzeň, Czech. E-mail: llenc@kiv.zcu.cz.

based face recognition methods perform the computation of LBP histograms on a uniform and fixed grid in the image and thus do not respect the properties of the particular images.

We have solved this issue in [12] by proposing an approach to detect positions of the fiducial point candidates automatically using Gabor Wavelet Transform (GWT). K-means clustering algorithm has been used next in order to identify the final point positions. We have further compared the features separately instead of concatenating them into one feature vector as in the case of other existing LBP based methods. The validity of the proposed approach has been evaluated on two corpora: standard AT&T Database of Faces and our Czech News Agency (ČTK) corpus containing uncontrolled face images. The experimental results have shown promising potential of the proposed method.

In our previous work, the fiducial points have been detected and clustered individually for each image. Therefore, they should correspond to the most representative points for every image. However, computing the GWT for every image is a time consuming task. Therefore, we solve this issue in this work by proposing another approach that identifies fiducial points *globally* for the whole gallery. Another important extension consists in integrating two different distance metrics (other than the Chi square distance already used by Ahonen in [1]) to improve our matching procedure. We also show the performance of our matching used separately in the so called *Grid position* approach. To the best of our knowledge, no existing method used the LBP algorithm in this way.

The proposed approaches are evaluated on three standard corpora, namely ORL, FERET, AR face database and on our ČTK dataset.

The results of this work will be used for implementation of an experimental face recognition system. This system will be used by the reporters of the Czech News Agency (ČTK¹) to automatically annotate people in photographs during insertion into the photo-database.² Its main strength is to successfully process photos of a great number of different persons taken in a totally uncontrolled environment.

The following section lists important face recognition methods with a particular focus on the local binary patterns. Section 3 shortly introduces Gabor wavelets and the LBP algorithm. The proposed face recognition methods are described next. Section 4 evaluates and

compares the results of the proposed approaches. In the last section, we conclude the experimental results and propose some future research directions.

2. Related work

It is possible to divide the existing face recognition approaches into holistic and feature based ones. Holistic approaches represent the face image as a whole while the feature based methods represent a face as a set of features. The first successful holistic approach are the Eigenfaces [28], which are based on the Principal Component Analysis (PCA). The main drawback of this method is its sensitivity to different lighting conditions, pose and scale. However, the PCA based approaches are still popular, as shown for instance in [22]. The other important methods belonging into this group are Fisherfaces [4] based on Linear Discriminant Analysis (LDA) or the Independent Component Analysis (ICA) [3]. These approaches achieve good results on the data acquired in laboratory conditions. However, their accuracy decreases significantly on the low quality real data.

This issue is partly solved by the feature based approaches. The most important representatives are shortly described next. The first successful approach is Elastic Bunch Graph Matching (EBGM) [30]. It is based on Gabor Wavelet Transform. It is later used with many modifications as shown for example in [10], which solves the main weakness of the original EBGM and extracts the fiducial points automatically. Another successful approach is based on Scale Invariant Feature Transform (SIFT) [15]. The main advantage of this approach is that the SIFT features are invariant to rotation, scale and lighting conditions. Speeded-Up Robust Features (SURF) is further used for face recognition as shown for instance in [6]. It obtains comparable results as the SIFT with significantly lower computational costs.

One interesting approach proposed by Tran et al. [27] extracts several biometric features such as total face, eyes, nose, mouth, etc. from the face image. These features are combined to obtain a rich information source for the face recognition. The authors then employ the Bees Algorithm to determine the optimal feature weights to have an accurate face representation in order to obtain the highest possible recognition accuracy. The experiments on the CASIA and ORL face corpora show that this method achieves consistently high recognition scores.

¹<http://www.ctk.eu>.

²<http://multimedia.ctk.cz/en/foto/>.

Recently, Local Binary Patterns (LBP) are successfully used for face representation and recognition. The LBP operator was first used for texture representation as presented in [20]. It is computed from the neighbourhood of a pixel and uses the intensity of the central pixel as a threshold. The pixels are marked either 0 or 1 if the value is lower or greater than this threshold. These binary values are concatenated into one binary string and its decimal value is then used as a descriptor of the pixel.

The first application of LBP for face recognition is proposed by Ahonen et al. [1]. The face is divided into rectangular regions. In each region a histogram of the LBP values is computed. All histograms are then concatenated into one vector which is used for the face representation. A histogram intersection method or Chi square distance is used for vector comparison. A weighted LBP modification is also proposed in this work. It gives more importance to the regions around the eyes and the central part of the face. The reported recognition rate on the FERET dataset [21] reaches 93% for the original method and 97% for the weighted LBP method. The strength of this method is its high ability to extract important information while the computational costs are low.

Another extension of the original method are Local Ternary Patterns (LTP) proposed in [26]. It uses three states to capture the differences between the center pixel and the neighbouring ones. It is thus less sensitive to noise. The authors report that this algorithm outperforms the LBP on FRGC experiment.³

A modification of the original LBP approach called *Dynamic Threshold Local Binary Pattern (DTLBP)* is proposed in [14]. It takes into consideration the mean value of the neighbouring pixels and also the maximum contrast between the neighbouring points. It is stated there that this variation is less sensitive to noise than the original LBP method. The recognition rates reported on the Yale B database outperform both LBP and LTP.

The so called *Local Derivative Patterns (LDP)* are proposed in [37]. The difference from the original LBP is that it uses features of higher order. It thus can represent more information than the original LBP. It is proved that LDP of second and third order outperforms LBP on the FERET database. The difference is evident for the more complicated image sets (fc, dup1, dup2) while the results for fb probe are similar to the LBP.

An important idea which is proposed already by Ojala in [20] are so called *Uniform Local Binary Patterns*. The pattern is called uniform if it contains at most two transitions from 0 to 1 or from 1 to 0. It was proved that approximately 90% of the patterns in facial images are uniform. The histogram can then be shortened from 256 intervals (bins) to 59, where the 59th bin is reserved for the non-uniform patterns. This improvement influences mainly the memory requirements and computational costs. The improvement of the recognition rate is minor.

An interesting method which uses uniform patterns is proposed in [34]. The authors state that the histogram bin containing non-uniform patterns dominates among other bins and gives thus too much importance to this bin. Therefore, they propose to assign such patterns to the closest uniform pattern. The closest pattern is selected using the Hamming distance. Authors state that this method slightly outperforms the original LBP based method on the FRGC challenge.

A novel original LBP extension called *Patch Based Descriptor* is proposed in [31]. The authors show that this approach improves the accuracy of the original LBP method in both multi-option identification and same/not-same classification on the LFW corpus [8].

Some methods also combine other preprocessing tools with the LBP. In [25], Gabor features and LBP are combined. The Gabor features as well as the LBP features are extracted and transformed using PCA. The features are then combined and used as face representation. This method outperforms the LBP on all sets of the FERET database. The difference is higher for the more complicated image sets (fc, dup1, dup2). An important drawback of this method is much higher size of the resulting descriptors.

In [33], Discrete Cosine Transform (DCT) is combined with Local Gabor Binary Patterns Histograms (LGBPH). The DCT is used to obtain low frequency features while the LGBPH captures higher frequencies and thus more local features. The experiments show that this approach outperforms the LGBPH on the ORL database.

For additional information about the LBP based methods, please see the surveys [7,19].

All the reviewed methods perform very well but a straight comparison is difficult because the experimental set-up usually differs. We can conclude that the more complicated methods bring better recognition accuracy. However, it is often accompanied with larger features and higher computational costs. It is worth mentioning that in all above described LBP methods,

³<http://www.nist.gov/itl/iad/ig/frgc.cfm>.

the images are divided into rectangular regions and histograms are computed in each region. All histograms from one image are then concatenated to create the face representation.

Face recognition results should be further improved by image pre or post processing. A successful pre-processing is for example image enhancing or denoising (see [18,23]). Another one can be a head pose estimation as shown in [2], which is used for face harmonization. It is also beneficial to use confidence measures [13] in the post-processing step to detect and remove incorrectly recognized samples from the recognized results.

3. Automatically detected fiducial points for LBP based face recognition

To the best of our knowledge, the feature positions in the current LBP based approaches are based on the predefined grid. We assume that the best features must be created in the most representative face points which are related to the particular image properties. We propose to solve this issue by employing Gabor wavelet transform to identify such positions automatically. Because of a high quantity of the detected points a clustering technique is applied further to reduce this number. The clustering is particularly beneficial because: a) it merges the multiple points detected in a small area and thus suppresses redundant information in the face model; b) it reduces the total number of the features and thus decreases significantly the computational costs.

Next, we shortly introduce the Gabor wavelets and the original LBP algorithm to facilitate the reading of the paper.

3.1. Gabor wavelets

Gabor filter is a sinusoid modulated by a Gaussian. A basic form of a two dimensional Gabor filter is shown in Eq. (1).

$$g(x, y) = \exp\left(-\frac{\hat{x} + \gamma^2 \hat{y}^2}{2\sigma^2}\right) \exp\left(j\left(2\pi\frac{\hat{x}}{\lambda} + \psi\right)\right) \quad (1)$$

where $\hat{x} = x \cos \theta + y \sin \theta$, $\hat{y} = -x \sin \theta + y \cos \theta$, λ is the wavelength of the cosine factor, θ represents the

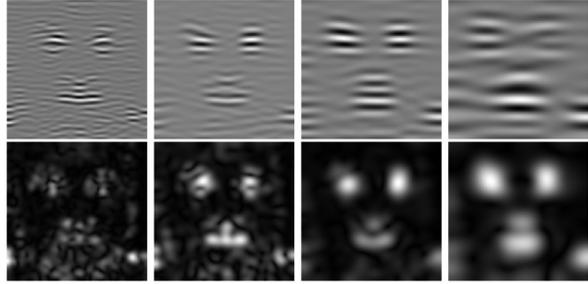


Fig. 1. Examples of an image filtered by four different Gabor filters. The real parts of the filter responses are used in the first four images, while the magnitudes in the last four ones. The original image is depicted in Fig. 4.

orientation of the filter and ψ is a phase offset, σ and γ are parameters of the Gaussian envelope, σ is the standard deviation of the Gaussian and γ defines the ellipticity (aspect ratio) of the function.

To keep only the real part of the complex value, it is possible to rewrite the equation as follows:

$$g(x, y) = \exp\left(-\frac{\hat{x} + \gamma^2 \hat{y}^2}{2\sigma^2}\right) \cos\left(2\pi\frac{\hat{x}}{\lambda} + \psi\right) \quad (2)$$

The imaginary part is then computed replacing the sine function in Eq. (2) by the cosine function. These values are then used to compute the magnitude and the phase of the filter response.

The Gabor wavelets are often used in image analysis because of their great ability to capture important information in images. They are sensitive to edges, ridges and valleys in the image [10] which usually correspond to the important parts of the face such as eyes or mouth. This approach is also used in some specific applications for medical image processing as shown for example in [24].

Figure 1 illustrates four Gabor filters with different frequencies applied to an image (original image in Fig. 4). The first four images use only the real part of the filter response whereas the other four ones use the magnitude.

These figures (see Fig. 1) show clearly that the most important face points belong in all cases (as supposed) to the eyes, nose and mouth and differ significantly from the predefined ones used in the other LBP based approaches (usually based on the grid).

3.2. Local binary patterns

The LBP operator encodes the intensity changes in a small neighbourhood of a pixel. Originally, a 3×3

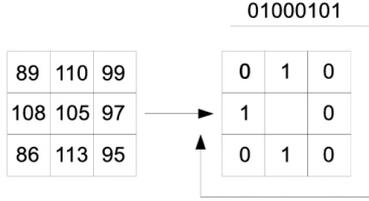


Fig. 2. An example of the feature computing by the original LBP operator.

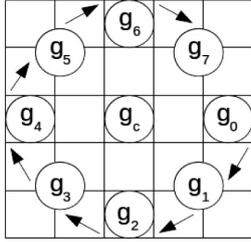


Fig. 3. Computation of the $LBP_{8,2}$ operator.

square centred at the given pixel is used. The algorithm assigns either 0 or 1 value to the 8 neighbouring pixels by Eq. (3).

$$val = \begin{cases} 0 & \text{if } g_N < g_C \\ 1 & \text{if } g_N \geq g_C \end{cases} \quad (3)$$

where val is the binary value assigned to the neighbouring pixel, g_N denotes the gray-level value of the neighbouring pixel and g_C is the gray-level value of the central pixel. The resulting values are then concatenated into an 8 bit binary number. Its decimal representation is used for further computation. This approach is illustrated in Fig. 2 and we use it further as our baseline.

The original LBP operator was extended to use circular neighbourhoods of various sizes and also with different number of points. The LBP is then denoted as $LBP_{P,R}$ where P is the number of points and R is the radius of the neighbourhood. The points are evenly placed on a circle with a given radius and they are compared with the central pixel in the same way as in the original LBP operator (see Fig. 3). The binary values are encoded, as in the previous case, according to Eq. (3). A bilinear interpolation is then used to compute the values in the points that are not placed in the pixel centres.

Figure 4 depicts the original image and the LBP image after applying the original LBP , $LBP_{8,1}$ and $LBP_{8,2}$ operators.



Fig. 4. An example of the original image and the LBP image after applying the original LBP , $LBP_{8,1}$ and $LBP_{8,2}$ operators.

3.3. Key-point position determination

The determination of the feature positions is a key part of our method. The approach is inspired by the work of Kepenekci [10] who used automatically detected feature positions to create Gabor wavelet features. All Gabor filters are applied to the facial image and the filter responses are scanned using a square sliding window W . The window centre (x_0, y_0) is considered to be a fiducial point iff:

$$R_j(x_0, y_0) = \max_{(x,y) \in W} R_j(x, y) \quad (4)$$

$$R_j(x_0, y_0) > \frac{1}{w_i * h_i} \sum_{x=1}^{w_i} \sum_{y=1}^{h_i} R_j(x, y) \quad (5)$$

where $j = 1, \dots, K$ (K is the number of Gabor filters), R_j represents an image filtered by the j -th Gabor filter and w_i and h_i are image width and height respectively.

Three different methods for feature position determination are described next.

3.3.1. Grid Position

The first method, *Grid Position*, is used in order to show the impact of our matching technique used separately. We extract the features using a predefined grid in exactly the same way as in the Ahonen's method [1]. No Gabor based feature detection technique is used in this case.

3.3.2. Face Specific Position

The second method, called *Face Specific Position*, determines the feature points individually for each face image. The images are filtered by 40 Gabor filters similarly as presented in [5] and the feature points are obtained as described previously. K-means clustering algorithm is applied to this point-set in order to reduce the point quantity and to identify the most representative ones. Unfortunately, an optimal cluster number is *a priori* not known. Therefore, this value must be set experimentally. Several different values will be tested in the experimental section in order to identify the best



Fig. 5. Example of original image, fiducial points detected by Gabor wavelets and clustered points (64 clusters) using *Face Specific Position* approach (left to right).

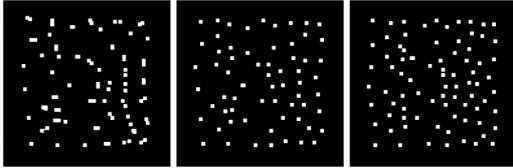


Fig. 6. Fiducial points selected by the *Global Position* method; number of clusters 48, 64 and 80 (left to right).

one. The resulting points are used to create the face representation. The fiducial point positions determined this way thus differ for each image and should represent every face as accurate as possible (see Fig. 5). This figure further shows that the most important fiducial points mostly correspond to the eyes, nose and mouth.

However, the main issue of this approach are high computational costs of the Gabor wavelets. This issue will be solved by the second approach.

3.3.3. Global Position

The last proposed method, called *Global Position*, detects the feature points using a certain sub-set of the image gallery. This sub-set is composed of 100 randomly chosen face samples. The detected points are grouped and the k-mean clustering is applied on this set. The cluster centres are then used to create the representation of all faces. The feature positions are thus constant for all images. Figure 6 shows the fiducial points detected by this method with different cluster numbers. These fiducial points do not no longer strictly correspond to the eyes, nose and mouth. However, they differ significantly from the points placed on a grid and we assume that they belong still to the most representative ones.

3.4. Feature construction

The feature vectors are created in the detected fiducial points. Each feature is composed of its coordinates and the histogram of the LBP values (see Section 3.2) in the surroundings of the point. We use a square neighbourhood centred in the feature point to calculate the

histogram. The size of the neighbourhood is referred in the following text as the *LBP window*. The histogram has 256 bins and is normalized according to the L1 norm.

If more than one image of one person is available we use the so called *Composed face model* as already described in our previous work [11]. It means that the features obtained from all training examples are put together to create the face model. This approach is used because it allows to create a more robust face representation.

3.5. Face comparison

Contrary to the other LBP based approaches we do not concatenate the feature histograms into one feature vector, however we compare the feature vectors separately. Three following distance metrics frequently used for histogram comparison are used to compute the distance of feature vectors f and r .

3.5.1. Histogram intersection

The first evaluated metric is the Histogram Intersection (HI). It was first used for an image retrieval task. The advantage of this method is its simplicity and the fast computation.

$$HI(f, r) = 1 - \sum_i \min(f_i, r_i) \quad (6)$$

where i is the number of histogram bins. We use this form of HI to be able to interpret it as a distance where 0 value means the identical histograms. The histograms are normalized using the L1 norm.

3.5.2. Chi square statistic

Chi square statistic is the second successful metric that can be used for histogram comparison:

$$\chi^2(f, r) = \sum_i \frac{(f_i - r_i)^2}{f_i + r_i} \quad (7)$$

3.5.3. Bhattacharyya distance

The last metric is the Bhattacharyya distance:

$$D_B(f, r) = \sqrt{1 - \frac{1}{\sqrt{f \cdot r} N^2} (BC(f, r))} \quad (8)$$

where BC is the Bhattacharyya coefficient defined as:

$$BC(f, r) = \sum_i \sqrt{f_i r_i} \quad (9)$$

and \bar{f} and \bar{r} are the average values of the feature vectors f and r respectively.

The distance of two faces is then for all previous metrics defined as:

$$\text{sim}(F, R) = \sum_{f_i} \min_{r_j \in N(f_i)} (D(f_i, r_j)) \quad (10)$$

where $D \in \{HI, \chi^2, D_B\}$ symbol represents the metric used, $N(f_i)$ is the neighbourhood of the feature f_i defined by the *distanceThreshold* that specifies the maximum distance within that the features are compared. It means that for each feature of the face F we find the closest one within the neighbourhood $N(f_i)$ from the face R . The distance of the two faces is computed as a sum of these minimum distances.

The recognized face \hat{F} is given by the following equation:

$$\hat{F} = \arg \min_R (\text{sim}(F, R)) \quad (11)$$

4. Experiments

This section first describes the four datasets that we used for experimental evaluation of our methods. Then, the experiments performed on each of the databases are presented.

4.1. Corpora

This section briefly summarizes the face databases used for evaluation of our approaches.

4.1.1. AT&T database of faces

This database [9] was formerly known as the ORL database. It was created at the AT&T Laboratories.⁴ The database contains the faces of 40 people, 10 pictures for each person are available. Each image contains one face with a black homogeneous background. They may vary due to different times of acquisition, head size and pose and lighting conditions. The size of the pictures is 92×112 pixels.

Figure 7 shows three example images from this database.

4.1.2. FERET dataset

FERET dataset [21] contains 14,051 images of 1,199 individuals. The resolution of the images is 256×384 pixels. The images are divided into the following categories according to the face pose: frontal,

Table 1

Image numbers in the main frontal probe sets of the FERET dataset

| Type | Images no. |
|------|------------|
| fa | 1,196 |
| fb | 1,195 |
| fc | 194 |
| dup1 | 722 |
| dup2 | 234 |

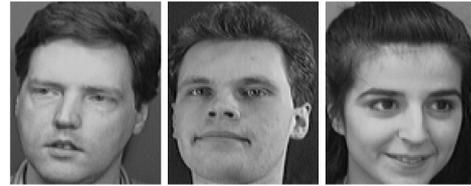


Fig. 7. Three example images from the AT&T database of faces.



Fig. 8. Three example images of one person from the FERET face corpus (fa, fb and fc – left to right).

quarter-left, quarter-right, half-left, half-right, full-left and full-right, and they are stored in the *.tiff* format. The images are also grouped into several probe sets. The main probe sets of the frontal images are summarized in Table 1. There are usually only a few seconds between the capture of the gallery-probe pairs in the f^* sets. The *fa* set is used for training (one image per person available). The individuals in the *fb* set differ in facial expressions, while the images in the *fc* set differ in illumination conditions. The images in the *dup1* probe set were obtained over a three year period and the *dup2* set is a sub-set of the *dup1*.

Figure 8 shows three example images from the FERET database.

4.1.3. AR face database

AR Face Database⁵ [16] was created at the Univerzitat Autònoma de Barcelona. This database contains more than 4,000 colour images of 126 individuals. The images are stored in a *raw* format and their size is 768×576 pixels. The individuals are captured under sig-

⁴<http://www.cl.cam.ac.uk/research/dtg/attarchive/face/database.html>.

⁵<http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>.



Fig. 9. Three example images from the AR face database.



Fig. 10. Three example images from the ČTK face database.

nificantly different lighting conditions and with varying expressions. Another characteristic is a possible presence of glasses or scarf.

Figure 9 shows three example images from this dataset.

4.1.4. Czech news agency (ČTK) database

This database was created automatically from real-world photographs owned by the Czech News Agency and contains gray-scale images of 638 people of the size 128×128 pixels. All images were taken over a long time period (20 years or more) and have significant variations in pose and lighting conditions. Up to 10 images for each person are available. The testing part contains one image for each person whereas the remaining part is used for training. Note that only the testing part was checked manually.

Figure 10 shows three example images from this corpus. The corpus is available freely for the research purposes upon request to the authors.

4.2. Experimental setup

We use the approach proposed by Ahonen [1] with the $LBP_{8,2}$ operator and the grid composed of 64 cells as a baseline in all following experiments. We also add the results of some recent state-of-the-art (sota) approaches to further compare with the proposed methods. The window size used for feature point detection is set to 25 in all experiments. This value was set experimentally in our previous work. These points are used as an input of the clustering algorithm. We chose the cluster numbers as multiples of 16 (i.e. 48, 64, 80, 96 and 112) in order to allow a straightforward comparison with the baseline with 64 features. The LBP window is set to 13 points for all experiments.

Table 2

Results of the proposed approaches on the AT&T database

| Recognition rate (%) | | | | | | |
|--|---------------|----------|-------|--------|----------|-------|
| Ahonen (baseline) | | | | 56.17 | | |
| Grid (to show the matching importance) | | | | 65.67 | | |
| Approach | Face specific | | | Global | | |
| | distance | χ^2 | HI | D_B | χ^2 | HI |
| cluster # | | | | | | |
| No | 57.75 | 71.36 | 64.97 | – | – | – |
| 48 | 50.28 | 64.86 | 51.22 | 51.25 | 72.67 | 72.17 |
| 64 | 54.56 | 69.28 | 58.97 | 53.72 | 74.06 | 74.11 |
| 80 | 56.31 | 70.67 | 61.78 | 55.58 | 75.44 | 75.44 |
| 96 | 56.53 | 71.61 | 63.39 | 56.64 | 75.47 | 75.64 |
| 112 | 57.64 | 72.22 | 65.14 | 56.97 | 76.28 | 76.53 |

4.3. Results on the AT&T database of faces

We already showed in [12] that with the common setup (9 images for training and the remaining one for testing) the recognition results are very close to 100%. The method comparison is thus very difficult. Therefore, we use another setup with only one sample for training and 9 remaining ones for testing and a ten-fold cross-validation procedure. To the best of our knowledge, no other work uses similar scenario on this face dataset. Therefore, we compare the proposed approaches only with our implementation of the Ahonen's method using the $LBP_{8,2}$ operator.

The recognition results on this database are shown in Table 2. This table shows that all proposed approaches outperform the baseline method. It is also evident that Global position approach brings better results than the Face specific one. This fact can be explained by the small variances among the face images. We can also conclude that the novel matching is important because it improves (used separately) the recognition score over the baseline. These results further confirm our assumption that the automatically detected fiducial points play an important role for the face recognition and that the clustering is also beneficial.

Note that it was not possible to evaluate Global Position approach without any clustering because of the enormous feature number in such case (see “–” symbols in the table).

4.4. Results on the FERET dataset

This dataset is probably the most often used one in the face recognition field. We used the fa set for training while the fb set for testing. The faces were cropped according to the eye positions provided with the database and resized to 130×150 pixels.

Table 3
Results of the proposed approaches on the FERET database

| Recognition rate (%) | | | | | | |
|--|---------------|-----------|-------|----------|-----------|-------|
| Ahonen (baseline) | | 93.89 | | | | |
| SRC (Wagner [29]) | | 95.20 | | | | |
| LGBPH (Yao [35]) | | 97.00 | | | | |
| Grid (to show the matching importance) | | 98.49 | | | | |
| Approach | Face specific | | | Global | | |
| distance | χ^2 | <i>HI</i> | D_B | χ^2 | <i>HI</i> | D_B |
| cluster # | | | | | | |
| No | 87.20 | 92.47 | 75.23 | – | – | – |
| 48 | 69.96 | 77.07 | 40.67 | 95.90 | 98.58 | 98.49 |
| 64 | 79.00 | 86.69 | 54.64 | 94.14 | 98.66 | 98.49 |
| 80 | 84.94 | 89.96 | 66.19 | 94.14 | 98.83 | 98.58 |
| 96 | 86.44 | 92.22 | 71.46 | 94.48 | 98.91 | 98.66 |
| 112 | 87.70 | 94.39 | 74.90 | 95.82 | 98.81 | 98.83 |

Table 4
Results of the proposed approaches on the AR database

| Recognition rate (%) | | | | | | |
|--|---------------|-----------|-------|----------|-----------|-------|
| Ahonen (baseline) | | 87.71 | | | | |
| SRC (Wright [32]) | | 94.70 | | | | |
| SVM (Yeh [36]) | | 87.50 | | | | |
| Grid (to show the matching importance) | | 97.29 | | | | |
| Approach | Face specific | | | Global | | |
| distance | χ^2 | <i>HI</i> | D_B | χ^2 | <i>HI</i> | D_B |
| cluster # | | | | | | |
| No | 70.71 | 98.14 | 97.43 | – | – | – |
| 48 | 60.86 | 96.00 | 94.00 | 80.86 | 98.86 | 98.71 |
| 64 | 61.57 | 97.00 | 96.71 | 71.29 | 99.14 | 99.14 |
| 80 | 64.43 | 97.43 | 97.00 | 73.71 | 99.43 | 99.29 |
| 96 | 66.29 | 98.14 | 97.43 | 70.57 | 99.57 | 99.29 |
| 112 | 68.43 | 98.14 | 97.86 | 70.71 | 99.57 | 99.43 |

Table 3 shows the results of our methods in comparison with the baseline and with two other state-of-the-art algorithms. This table also demonstrates that the proposed approaches outperform the baseline Ahonen’s method. This experiment further confirms better performance of the Global position method on the corpora with relatively small variances. The highest recognition rate is obtained with the Histogram intersection distance and 96 feature vectors (clusters). This table further shows that feature positions play smaller role than in the previous experiment related to the proposed Grid position approach.

4.5. Results on the AR database

This database is used because it contains images with important variances. We use a common scenario where 7 images are used for training and 7 for testing. The image size is 120×165 pixels.

Table 4 compares the results with the baseline algorithm and with two others state-of-the-art approaches. As previously, this table demonstrates that the pro-

Table 5
Results of the proposed approaches on ČTK dataset

| Recognition rate (%) | | | | | | |
|--|---------------|-----------|-------|----------|-----------|-------|
| Ahonen (baseline) | | 39.81 | | | | |
| Grid (to show the matching importance) | | 51.72 | | | | |
| Approach | Face specific | | | Global | | |
| distance | χ^2 | <i>HI</i> | D_B | χ^2 | <i>HI</i> | D_B |
| cluster # | | | | | | |
| No | 8.15 | 61.91 | 61.91 | – | – | – |
| 48 | 7.37 | 57.05 | 54.55 | 9.56 | 53.92 | 55.80 |
| 64 | 7.68 | 59.87 | 58.93 | 7.37 | 55.33 | 56.58 |
| 80 | 7.21 | 59.87 | 59.72 | 6.27 | 58.78 | 59.40 |
| 96 | 7.99 | 61.44 | 62.38 | 3.76 | 59.72 | 58.62 |
| 112 | 7.99 | 62.38 | 61.91 | 3.61 | 59.09 | 59.72 |

posed approaches outperform the baseline and the sota methods. This experiment also shows a slightly better performance of the Global position method. Note that this difference is significantly smaller than in the previous cases. This is probably caused by the greater variances in this corpus. These results further confirm the superior performance of the simplest Histogram intersection metric. The best result is obtained with Global position approach and Histogram intersection metric.

4.6. Results on the ČTK database

The last experiment is performed on the ČTK database, which contains low quality images automatically extracted from real-world photographs. As already mentioned, we use one image for testing and the remaining part (up to 9 images) for training of the face models.

Table 5 shows the results of our methods compared with the baseline Ahonen’s algorithm. Note that we cannot compare our results with other sota approaches because no other methods are evaluated on this dataset. This table shows that the Face specific position approach reaches significantly better results than the Global position method. This difference is probably caused by the variances among the images that are significantly higher than in the other datasets. This table further shows that the Histogram intersection and Bhattacharyya distances perform significantly better than the Chi square statistic. The best results are obtained with about 100 clusters.

5. Conclusions and perspectives

In this work, we studied the importance of the fiducial point positions for automatic face recognition. One approach proposed previously and two new ones have

been described and compared. The first approach, *Grid Position*, which uses points in a regular grid is proposed in order to show the performance of our matching algorithm. The second approach, called *Face Specific Position*, locates the feature positions individually for each face, while the third one, *Global Position*, defines the feature positions based on a representative sub-set of the gallery.

In the last two approaches, Gabor wavelets are used for key-point detection and then, these points are clustered by the k-means algorithm. Another important contribution is a proposal of a novel matching method based on the comparison of the feature vectors separately. Three different distance measures are used for this task.

We considered all approaches in terms of their respective theoretical advantages and drawbacks. They are further evaluated experimentally on three standard corpora: AT&T, FERET, AR face databases and on our ČTK dataset containing real-world face images. It was clearly demonstrated that the fiducial point positions are very important and that well defined fiducial points improve significantly the recognition accuracy. All proposed approaches outperformed significantly the baseline LBP approach on all corpora.

We further demonstrated that the Global position approach brings better results on the controlled corpora (i.e. AT & T, FERET and AR) with relatively small variances. However, it is beneficial to use the Face specific position method in the case of the face recognition of real face images (i.e. ČTK dataset) with significant changes in pose, tilt and illumination conditions. We can conclude that the sufficient feature number is about 100. In the case of easier datasets also lower numbers can be sufficient. It was also demonstrated that the proposed matching algorithm outperforms the baseline used with the identical features (see Grid position approach). The importance of our methods is particularly evident in the case of real-world photographs from the ČTK corpus where the accuracy is increased by more than 20% in absolute value. From the viewpoint of the distance measure, we can conclude that the simplest Histogram intersection brings often the best results, Bhattacharyya distance is still comparable and that both methods outperform significantly the Chi square statistic.

The first perspective is applying a weighting to the automatically detected key-points. Another improvement consists in using only uniform patterns or in defining different neighbourhood for the LBP operator. We would like further significantly adapt and extend the proposed approaches into medical field for example for cancer detection.

Acknowledgments

This work has been supported by the project LO1506 of the Czech Ministry of Education, Youth and Sports. We also would like to thank ČTK for support and for providing the photographic data.

References

- [1] T. Ahonen, A. Hadid and M. Pietikäinen, Face recognition with local binary patterns, in: *Computer Vision-Eccv 2004*, Springer (2004), 469–481.
- [2] E. Arcoverde Neto, R. Duarte, R. Barreto, J. Magalhaes, C. Bastos, T. Ren and G. Cavalcanti, Enhanced real-time head pose estimation system for mobile device, *Integrated Computer-Aided Engineering* **21**(3) (2014), 281–293.
- [3] M.S. Bartlett, J.R. Movellan and T.J. Sejnowski, Face recognition by independent component analysis, *Neural Networks, IEEE Transactions on* **13**(6) (2002), 1450–1464.
- [4] 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) (1997), 711–720.
- [5] D.S. Bolme, Elastic bunch graph matching, Ph. D. thesis, Colorado State University, 2003.
- [6] P. Dreuw, P. Steingrube, H. Hanselmann, H. Ney and G. Aachen, Surf-face: Face recognition under viewpoint consistency constraints, in: *BMVC*, (2009), 1–11.
- [7] D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, Local binary patterns and its application to facial image analysis: A survey, *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* **41**(6) (2011), 765–781.
- [8] 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.
- [9] A.K. Jain and S.Z. Li, *Handbook of Face Recognition*, Springer, 2005.
- [10] B. Kepenekci, Face recognition using gabor wavelet transform, Ph. D. thesis, Citeseer, 2001.
- [11] L. Lenc and P. Král, Gabor wavelets for automatic face recognition, in: *38th International Conference on Current Trends in Theory and Practice of Computer Science Špindleruv Mlýn, Czech Republic (January 2012)*, 2012.
- [12] L. Lenc and P. Král, Automatically detected feature positions for LBP based face recognition, in: *International Conference on Artificial Intelligence Applications and Innovations (AIAI 2014)*, Springer, Rhodes, Greece (September 2014), 246–255.
- [13] L. Lenc and P. Král, Two-step supervised confidence measure for automatic face recognition, in: *24th IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2014)*, IEEE, Reims, France (21–24 September 2014), 1–6.
- [14] W. Li, P. Fu and L. Zhou, Face recognition method based on dynamic threshold local binary pattern, in: *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, ACM (2012), 20–24.
- [15] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* **2** (2004).
- [16] A. Martinez and R. Benavente, The AR face database, *Universidad Autónoma de Barcelona*, Technical report, 1998.

- [17] A. Meraoumia, S. Chitroub and A. Bouridane, 2d and 3d palmprint information, pca and hmm for an improved person recognition performance, *Integrated Computer-Aided Engineering* **20**(3) (2013), 303–319.
- [18] R.G. Mesquita, C.A. Mello and L. Almeida, A new thresholding algorithm for document images based on the perception of objects by distance, *Integrated Computer-Aided Engineering* **21**(2) (2014), 133–146.
- [19] L. Nanni, A. Lumini and S. Brahmam, Survey on lbp based texture descriptors for image classification, *Expert Systems with Applications* **39**(3) (2012), 3634–3641.
- [20] T. Ojala, M. Pietikäinen and D. Harwood, A comparative study of texture measures with classification based on featured distributions, *Pattern Recognition* **29**(1) (1996), 51–59.
- [21] P.J. Phillips, H. Wechsler, J. Huang and P. Rauss, The FERET database and evaluation procedure for face recognition algorithms, *Image and Vision Computing* **16**(5) (1998), 295–306.
- [22] B. Poon, M.A. Amin and H. Yan, Performance evaluation and comparison of PCA based human face recognition methods for distorted images, *International Journal of Machine Learning and Cybernetics* **2**(4) (2011), 245–259.
- [23] G. Pérez, A. Conci, A.B. Moreno and J.A. Hernandez-Tamames, Rician noise attenuation in the Wavelet packet transformed domain for brain MRI, *Integrated Computer-Aided Engineering* **21**(2) (2014), 163–175.
- [24] M. Rizzi, M. D’Aloia and B. Castagnolo, Computer aided detection of microcalcifications in digital mammograms adopting a wavelet decomposition, *Integrated Computer-Aided Engineering* **16**(2) (2009), 91–103.
- [25] X. Tan and B. Triggs, Fusing gabor and lbp feature sets for kernel-based face recognition, in: *Analysis and Modeling of Faces and Gestures*, Springer (2007), 235–249.
- [26] X. Tan and B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, *Image Processing, IEEE Transactions on* **19**(6) (2010), 1635–1650.
- [27] Q. Tran, P. Kantartzis and P. Liatsis, Improving fusion with optimal weight selection in face recognition, *Integrated Computer-Aided Engineering* **19**(3) (2012), 229–237.
- [28] M.A. Turk and A.P. Pentland, Face recognition using eigenfaces, in: *Computer Vision and Pattern Recognition, 1991 Proceedings CVPR’91, IEEE Computer Society Conference on*, IEEE (1991), 586–591.
- [29] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi and Y. Ma, Toward a practical face recognition system: Robust alignment and illumination by sparse representation, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **34**(2) (2012), 372–386.
- [30] L. Wiskott, J.-M. Fellous, N. Kuiger and C. Von Der Malsburg, Face recognition by elastic bunch graph matching, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **19**(7) (1997), 775–779.
- [31] L. Wolf, T. Hassner, Y. Taigman et al., Descriptor based methods in the wild, in: *Workshop on Faces In’Real-Life’Images: Detection, Alignment, and Recognition*, (2008).
- [32] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry and Y. Ma, Robust face recognition via sparse representation, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **31**(2) (2009), 210–227.
- [33] Z. Xie, Single sample face recognition based on dct and local gabor binary pattern histogram, in: *Intelligent Computing Theories*, Springer (2013), 435–442.
- [34] H. Yang and Y. Wang, A lbp-based face recognition method with hamming distance constraint, in: *Image and Graphics, 2007 ICIG 2007 Fourth International Conference on*, IEEE (2007), 645–649.
- [35] B. Yao, H. Ai, Y. Ijiri and S. Lao, Domain-partitioning rank-boost for face recognition, in: *Image Processing, 2007 ICIP 2007 IEEE International Conference on*, IEEE **1** (2007), 1–129.
- [36] C.-H. Yeh, P.-R. Shih, Y.-T. Lin, K.-T. Liu, H.-M. Chang and M. Ouhyoung, A comparison of three methods of face recognition for home photos, in: *SIGGRAPH’09: Posters*, ACM (2009), 30.
- [37] B. Zhang, Y. Gao, S. Zhao and J. Liu, Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor, *Image Processing, IEEE Transactions on* **19**(2) (2010), 533–544.