

# Automatic Face Recognition System Based on the SIFT Features

Ladislav Lenc<sup>a</sup>, Pavel Král<sup>a,b</sup>

*<sup>a</sup>Dept. of Computer Science & Engineering  
Faculty of Applied Sciences  
University of West Bohemia  
Plzeň, Czech Republic*

*<sup>b</sup>NTIS - New Technologies for the Information Society  
Faculty of Applied Sciences  
University of West Bohemia  
Plzeň, Czech Republic  
{llenc,pkral}@kiv.zcu.cz*

---

## Abstract

The main goal of this paper is to propose and implement an experimental fully automatic face recognition system which will be used to annotate photographs during insertion into a database. Its main strength is to successfully process photos of a great number of different individuals taken in a totally uncontrolled environment. The system is available for research purposes for free. It uses our previously proposed SIFT based Kepenekci approach for the face recognition, because it outperforms a number of efficient face recognition approaches on three large standard corpora (namely FERET, AR and LFW). The next goal is proposing a new corpus creation algorithm that extracts the faces from the database and creates a facial corpus. We show that this algorithm is beneficial in a preprocessing step of our system in order to create good quality face models. We further compare the performance of our SIFT based Kepenekci approach with the original Kepenekci method on the created corpus. This comparison proves that our approach significantly outperforms the original one. The last goal is to propose two novel supervised confidence measure methods based on *a posterior* class probability and a multi-layer perceptron to identify incorrectly recognized faces. These faces are then removed from the recognition results. We experimentally validated that the proposed confidence measures are very efficient and thus suitable

for our task.

*Keywords:* Face Recognition, Face Detection, Czech News Agency, Corpus Creation, Confidence Measure, Scale Invariant Feature Transform (SIFT), FERET, AR, LFW

---

## 1. Introduction

Automatic Face Recognition (AFR) consists in identification of a person from an image or from a video frame by a computer. This field has been intensively studied by many researchers during the past few decades. Nowadays, it can be seen as one of the most progressive biometric authentication methods. Numerous AFR methods have been proposed and the face recognition has become the key task in several applications as for instance surveillance of wanted persons, access control to restricted areas, automatic annotation of photos in photo sharing applications or in social networks, and so on.

From the viewpoint of the image nature and quality, it is possible to divide this spectrum of applications to two groups: controlled and uncontrolled scenario systems. The image quality of the controlled systems is usually high, a face pose variation and other dissimilarities (lighting conditions, face tilt and rotation, etc.) are limited. Automatic face recognition is easier because the face detection and further image preprocessing are usually not necessary and the performance of current AFR approaches is often sufficient. Most of the current AFR systems belong into this group.

However, few uncontrolled scenario systems exist. The main issue of these systems is that the face pose and position significantly differ and also the other dissimilarities within the images are common. A successful face recognition is thus much more challenging. This issue will be addressed in this paper and solved by the goals described next.

The first goal consists in proposing an algorithm to extract the faces from the photographs and to create a facial corpus as high quality as possible. The design and implementation of this algorithm is one important contribution of this paper. We will further show that the usage of this algorithm in our AFR system in the preprocessing step is beneficial.

It is necessary to identify an optimal face recognition approach to be integrated into our system. In our previous work [23] we proposed an efficient *SIFT based Kepenekci* face recognition approach and evaluated it on the

standard ORL [15] corpus. We showed that this approach gives very good results using this small well controlled face dataset. The second goal thus consists in evaluation of our SIFT based Kepenekci on the other larger standard face corpora. We have chosen three large challenging face databases: FERET [33], AR [29] and LFW [17]. Our method is then compared with a number of very efficient AFR approaches in order to identify the most suitable one for the large data and for processing of the less controlled images. This method will be integrated into our system.

Face recognition in an uncontrolled environment is erroneous and it is beneficial to identify incorrectly recognized examples. The third goal of this paper thus consists in proposing and evaluating the confidence measure methods in order to identify such examples and remove them from the recognition results or propose them to the user for manual correction. Two novel supervised confidence measures based on *a posterior* class probability and a multi-layer perceptron are proposed.

The last goal of this paper is to introduce an experimental fully automatic face recognition system. This system will be used by the Czech News Agency (ČTK)<sup>1</sup> to annotate people in photographs during insertion into the ČTK database<sup>2</sup>. Its main strength is to successfully process photos of a great number of different persons taken in a totally uncontrolled environment. The original images have high resolution but the size of the faces varies substantially. There are also significant variances in lighting, ageing, face poses and angles. Moreover, another property is presence of more objects/faces in the images. The recognition of such images is thus very challenging. The system (with the source code) is available for research purposes for free. Note that the resolution of face images is strictly related to our proposed application. However, our face recognition approach is general enough to process images of different resolutions.

This paper is organized as follows. The following section summarizes successful approaches in the face recognition domain including confidence measure methods and face recognition systems. Section 3 describes the proposed corpus creation algorithm. Section 4 details the SIFT based Kepenekci method which we use for the face recognition. The next section details the proposed confidence measure methods which are used in order to detect and

---

<sup>1</sup><http://www.ctk.eu>

<sup>2</sup><http://multimedia.ctk.cz/en/foto/>

remove incorrectly recognized faces from the recognition results. Section 6 describes the architecture of our face recognition system. The performed experiments are presented in Section 7. This section describes the corpora which are used to evaluate the system. Furthermore, we compare our AFR module with the other efficient methods on these corpora. The performance of our system on the lower quality real  $\check{C}$ TK data with the particular focus on the proposed corpus creation algorithm is shown further. In the last part of this section we evaluate our confidence measure methods. Section 8 briefly summarizes the main scientific contribution of this paper. In the last section, we conclude the results and propose some future research directions.

## 2. Related Work

One of the first successful approaches in the AFR domain is the Principal Component Analysis (PCA), so called Eigenfaces [39]. This is a statistical method that takes into account the whole image as a vector. First, the image vectors are put together and the image matrix is formed. The eigenvectors of this matrix are calculated. The face images can then be expressed as a linear combination of these vectors. Each image is represented as a set of weights for the corresponding vectors. The PCA based approaches are still popular, as shown in [34].

Another method, the Fisherfaces [5], is derived from Fisher's Linear Discriminant (FLD). Similarly to Eigenfaces, Fisherfaces project an image into another, less dimensional, space. The original dimensionality, which is given by the resolution of the images, is reduced to the number of images (distinct classes). The projections of the facial images are then compared using a suitable similarity measure. The key point is maximization of the ratio of between-class scatter and within-class scatter. However, Eigenfaces maximize the total scatter across all images. According to the authors, this approach should be insensitive to variations in lighting conditions. A recent extension of this approach, called L-Fisherfaces, is proposed in [48]. Authors experimentally showed that this method achieves higher accuracy than four other evaluated methods on the three datasets.

Independent Component Analysis (ICA) can be also successfully used in the automatic face recognition field [11]. The main principle is to find a linear combination of non-Gaussian data that reconstructs the original data. Contrary to PCA, ICA uses higher order statistics. ICA thus provides more powerful data representation. ICA approach is still worthy of attention as

shown in [10]. Authors further propose a Locality Pursuit (LP) approach, which uses locality preserving projections in the high-dimensional whitened space. This approach addresses the issue of pursuing low-dimensional features. Experimental results show that the LP approach achieves on the FERET corpus higher accuracy than ICA.

Another efficient AFR approach is the Elastic Bunch Graph Matching (EBGM) [6]. This approach uses features constructed by the Gabor wavelet transform. Initially, a set of manually labelled landmarks is presented to the algorithm. These landmarks are used as examples to determine the landmark positions in novel images. The Gabor wavelet convolutions (so called Jets) are computed in the landmark positions and are used for a face representation. A “bunch graph” is created from these examples. Each node in the graph contains a set of Jets for one landmark across all of the images. The similarity of faces is determined from the landmark positions and jet values.

Kepekci proposes in [20] an algorithm that outperforms the classical EBGM and that addresses the main issue of Elastic Bunch Graph Matching, manual labelling of the landmarks. Landmark positions are obtained dynamically by the Gabor filter responses.

Other successful approaches [1] use so-called, Local Binary Patterns (LBP) for facial feature extraction. The LBP operator thresholds a local region in the image by a value of the central pixel. It labels the pixels either 0 or 1 if the value is lower or higher than the threshold. Then, a histogram of the labels is computed and used as a descriptor. The original method used a  $3 \times 3$  neighbourhood which was later extended to use neighbourhoods of various sizes. In the application for face recognition, a facial image is first divided into rectangular regions. The LBP descriptor is constructed in each region and the results are put together to create one vector representing the face. The face representations are compared using a nearest neighbour rule.

A modification of the original LBP approach called Dynamic Threshold Local Binary Pattern (DTLBP) is proposed in [25]. It takes into consideration the mean value of the neighbouring pixels and also the maximum contrast between the neighbouring points. This variation is less sensitive to the noise than the original LBP method. Another extension of the original method is Local Ternary Patterns (LTP) proposed in [38]. It uses three states to capture the differences between the center pixel and the neighbouring ones. Similarly to the DTLBP the LTP is less sensitive to the noise. The so called Local Derivative Patterns (LDP) are proposed in [47]. In contrast

with the original LBP, it uses features of higher order. It thus can represent more information than the original LBP.

Yang et al. show in [45] the accuracy of the LBP based approaches on the low resolution images. The experimental results are compared with some subspace algorithms on the four down-sampled benchmark face datasets (FERET, Extended YaleB, CMU PIE and AR [29]). The image resolution vary from  $32 \times 32$  to  $60 \times 60$  pixels.

Speeded-Up Robust Features (SURF) [3] is a useful method for key-point detection and creation of a descriptor. An integral image [40] is utilized to speed-up the key-point detection process. The detector is based on the Hessian matrix<sup>3</sup>. Therefore, it is called the “Fast-Hessian” detector. Box filters are used as an approximation of second order Gaussian derivatives. The box filters are then up-scaled and applied to the original image. This method is invariant to face rotation. To ensure rotation invariance, one orientation is assigned to each key-point. The computation is based on the circular neighbourhood of the key-points.

Another efficient AFR method is based on the Scale Invariant Feature Transform (SIFT) [2]. The SIFT algorithm is used to create a set of descriptors (features) for each image. These features have the ability to detect and describe local image properties and the algorithm ensures invariance to image scaling, translation, rotation and also lighting conditions.

Another approach using SIFT is called Fixed-key-point-SIFT (FSIFT) [21]. As opposed to the previous method, SIFT descriptors are fixed in predefined locations determined in the training step.

Another interesting face recognition approach based on linear regression is proposed by Naseem et al. [32]. The authors reported very good recognition scores on two different subsets of the ORL and FERET corpora. Lu et al. proposed in [28] a novel Discriminative multi-manifold analysis method which learn discriminative features from image patches. Conversely to our approach, this approach uses exactly one image sample for training. Experimental results on three face databases (AR, FERET and FG-NET) showed the efficacy of the proposed approach. Wang et al. [42] describe another successful approach based on the linear regression which also uses one face example for training. This approach is evaluated on two down-sampled bench-

---

<sup>3</sup>The Hessian matrix is a square matrix of the second order partial derivatives of a given function.

mark datasets (AR and FERET) and the results are compared with the other state-of-the-art methods.

Probably the most important task of the face recognition consists in identification of the important face points (landmarks). A number of recent studies is focused on this research topic as shown for instance in the comparative study [9]. These approaches are often evaluated on the Annotated Facial Landmarks in the Wild (AFLW). dataset. The landmarking is recently very popular, because it can be also used in many other tasks as for instance, face coding, face expression, gesture understanding, gaze detection, face tracking, etc.

For further information about the face recognition, please refer to the review [4].

### 2.1. Confidence Measures

Confidence Measure (CM) is used as a post-processing of the recognition to determine whether a recognition result is correct or not. The incorrectly recognized samples should be removed from the recognition set or another processing (e.g. manual correction) can be further realized. This technique is mainly used in the automatic speech processing field [19, 43] and mostly based on the *a posterior* class probability. However, it can be successfully used in another research areas as shown in [37] for genome maps construction, in [16] for stereo vision or in [31] for handwriting sentence recognition.

Another related confidence measure approach is proposed by Proedrou et al. in the pattern recognition task [36]. The authors use a classifier based on the nearest neighbours. Their confidence measure is based on the algorithmic theory of randomness and on transductive learning.

Unfortunately, only few works about the confidence measure in the face recognition domain exist. Li and Wechsler propose a face recognition system which integrates a confidence measure [24] in order to reject unknown individuals or to detect incorrectly recognized faces. Their confidence measure is, as in the previous case, based on the theory of randomness. The proposed approaches are validated on the FERET database.

Eickeler et al. propose and evaluate in [12] five other CMs also in the face recognition task. They use a pseudo 2-D Hidden Markov Model classifier with features created by the Discrete Cosine Transform (DCT). Three proposed confidence measures are based on *a posterior* probabilities and two others on ranking of results. Authors experimentally show that *a posterior* class probability gives better results for the recognition error detection task.

## 2.2. Face Recognition Systems

As already shown above, numerous papers presented in the face recognition domain concentrate only on the recognition task itself. In a few cases, the process of face localization (or detection) is considered. Because the faces in the original images are usually sufficiently well aligned, this task is often reduced to define the face position more accurately. Moreover, to the best of our knowledge, relatively few works about whole face recognition systems exist.

One example of a system that addresses the issue of imprecisely localized faces is proposed in [30]. The system compensates the face position and also solves partial occlusion and different facial expressions. Only one training example per person is used.

A complete face recognition system is described in [41]. The training images are well aligned (acquired in controlled conditions) whereas the recognized images are real-world photos. The system is based on the Sparse Representation and Classification (SRC) [18] algorithm. It achieves very good results on the FERET database.

Another face recognition system is presented in [8]. The authors localize the face in the images and then compute the facial features. Their face recognition algorithm is based on the EBGm, but the fiducial points are detected completely automatically. The system is evaluated on the FERET corpus. The authors show that their system has recognition scores comparable to the elastic bunch graph matching.

For additional information about the face recognition and the face recognition systems, please refer to the survey [49]. The authors also mention some commercial face recognition systems. Unfortunately, neither the system architecture nor the approaches used are usually reported. Moreover, these systems are not evaluated on the standard face datasets and it is thus impossible to compare them with our system.

## 3. Corpus Creation Algorithm

As already stated, our system will be used by the ČTK for automatic labelling of individuals in new photographs before insertion into the database. Currently annotated pictures contain information about the person's identity. Unfortunately, the photos contain not only the face itself. They may be composed of more people, some background objects, etc. The faces usually





Figure 1: Example of one correctly detected face (left) and one incorrectly detected face by the Viola-Jones face detector

differ in size, pose and angle and are also often partially occluded. Moreover, some labels are erroneous due to annotator errors.

For a correct estimation of face models, a face corpus as good as possible is necessary. We propose an algorithm in order to detect and extract faces from pictures and to create a face corpus automatically. This algorithm is composed of four main steps:

1. face detection,
2. eye detection,
3. face rotation,
4. corpus cleaning

Note that this task is very specific and application dependent. Therefore, to the best of our knowledge there is no similar previous work available.

### *3.1. Face Detection*

We use an implementation of the Viola-Jones algorithm [40] which uses a boosted cascade of simple classifiers. This algorithm was chosen because it is one of the most successful face detection methods.

Figure 1 shows one example of a correctly and one example of an incorrectly detected face.

### *3.2. Eye Detection*

Unfortunately, a certain number of incorrectly detected faces occur in the output. Verification of the detected faces is thus indispensable. In order to avoid manual processing, we propose utilizing eye detection for this task. The eyes are detected by the Viola-Jones algorithm (as in the face detection task). Only those face images with two successfully detected eyes are kept while the others are discarded.

Figure 2 shows one example of the eye detection task where two, one and no eye was detected by the algorithm. This step eliminates a great



Figure 2: Examples of faces with 2, 1 and 0 detected eyes (from left to right)



Figure 3: Example of face detection (left), eyes detection (middle) and face rotation according to the eyes (right)

number of incorrectly detected faces. Moreover, the successful detection of both eyes ensures we obtain more or less frontal images, while profiles should be removed.

### 3.3. Face Rotation

If both eyes are detected successfully, the face is rotated so that the eyes are on a horizontal line according to the detected positions. The face is further placed in the image centre. This transformation cannot completely resolve the issue of variations in pose and tilt. Nevertheless, face variations are significantly reduced.

Figure 3 shows the tasks of face and eye detection and face rotation according to the eyes.

### 3.4. Corpus Cleaning

We obtained a large number of face images for each individual. However, these numbers differ significantly. In our previous work [23], we proved that more training examples play a crucial role in the correct estimation of face models. Nevertheless, we also showed that this number should be balanced and the quality of faces should be as high as possible.

Therefore, we perform a step called “*corpus cleaning*” in order to choose the same number of the most representative face images. We manually verified on a randomly chosen small face sub-set that the majority of face images

is correct and the erroneous examples differ substantially from this representative set. Our algorithm used to choose only the representative faces is based on this observation.

Let  $S$  be the set of all face images extracted by the three steps described previously. Let  $S_i$  be the set of  $n$  face images  $I_1, \dots, I_n \in S_i$  representing one individual. Each face image  $I_j$  is represented by the feature vector  $F_j$  computed by the SIFT algorithm. The pseudo-code of the proposed algorithm is given below.

```

 $N$  = face image number/individual
 $K$  = required face image number/individual
for all  $S_i \in S$  do
  while  $N > K$  do
    for all  $F_j \in F$  do
      compute the face model  $M$  using the feature set  $F \setminus F_j$ 
      compute the similarity  $FS_j$  between  $F_j$  and  $M$  by the Equation 2
    end for
    compute an average value  $av_{FS}$ 
    compute a standard deviation  $sd_{FS}$ 
    compute a similarity threshold  $ST = av_{FS} + \frac{sd_{FS}}{\sigma}$ 
    if  $\exists I_j : FS_j < ST$  then
      for all  $FS_k < ST$  do
        remove the face image  $I_k$  from the face corpus
         $N \leftarrow N - 1$ 
      end for
    else
      remove the face image  $I_{min}$  with minimal similarity  $FS_{min}$  from the
      face corpus
       $N \leftarrow N - 1$ 
    end if
  end while
end for

```

The similarity  $FS_j$  of image  $I_j$  to face model  $M$  created from the remaining images  $I \setminus I_j$  is computed. The average  $av_{FS}$  of the similarity values and consequently the standard deviation  $sd_{FS}$  is computed. The similarity threshold  $ST$  is computed using these two values. Note that the optimal

value of the constant  $\sigma$  is not possible to compute analytically and is thus obtained by the development corpus. Images  $I_j$  with lower similarity  $FS_j$  than the threshold  $ST$  are discarded. The above described task is realized iteratively. If there is no similarity value lower than the threshold  $ST$  given and there are still more images than the required number  $K$ , the images with the minimal similarities  $FS_{min}$  are further discarded. Finally the number of images/person is equal to  $K$ , which is the required face image number defined at the beginning of the algorithm.

## 4. Face Recognition

For the face recognition task, we use the SIFT based Kepenekci method [23] which combines the efficient SIFT algorithm with the adapted Kepenekci matching. This algorithm was chosen, because as proven previously, it significantly outperforms the other approaches particularly for lower quality real data. Note that the above mentioned modifications do not influence the applicability of the original Kepenekci method. Moreover, this approach is not limited by the image resolution.

### 4.1. Scale Invariant Feature Transform

The Scale Invariant Feature Transform algorithm basically has four steps: extrema detection, removal of key-points with low contrast, orientation assignment and descriptor calculation [21].

#### 4.1.1. Extrema Detection

The Difference of Gaussian (DoG) filter is applied to the input image. The image is gradually down-sampled and the filtering is performed at several scales. Figure 4 demonstrates the process of the creation of DoG filters at different scales [27]. Filtering at several scales ensures scale invariance. Each pixel is then compared with its neighbours. Neighbours on its level as well as on the two neighbouring (lower and higher) levels are examined. If the pixel is the maximum or minimum of all the neighbouring pixels, it is considered to be a potential key-point.

#### 4.1.2. Low Contrast Key-point Removal

The detected key-points are further examined to choose the “best” candidates. For the resulting set of key-points their stability is determined. Locations with low contrast and unstable locations along edges are discarded.

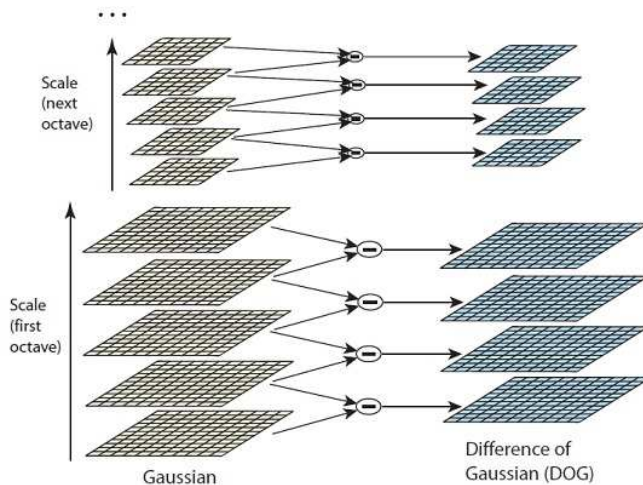


Figure 4: Difference of Gaussian filters at different scales [27]

#### 4.1.3. Orientation Assignment

The orientation of each key-point is computed. The computation is based upon gradient orientations in the neighbourhood of the pixel. The values are weighted by the magnitudes of the gradient.

#### 4.1.4. Descriptor Calculation

The final step consists in the creation of descriptors. The computation involves the  $16 \times 16$  neighbourhood of the pixel. Gradient magnitudes and orientations are computed in each point of the neighbourhood. Their values are weighted by a Gaussian. For each sub-region of size  $4 \times 4$  (16 regions), orientation histograms are created. Finally, a vector containing 128 ( $16 \times 8$ ) values is created.

#### 4.2. SIFT based Kepenekci Method

The original Kepenekci method achieves very good accuracy. The reported recognition rates on the ORL and FERET datasets are 95.25% and 96.3%, respectively.

In our previous work [23] we modified this approach by replacing the features obtained using Gabor wavelet filters with the SIFT features. The original Kepenekci matching was simplified by removing the “similarity” threshold which excludes some feature vectors from the matching procedure. We

proved that this modified approach significantly outperforms the original Kepenekci method from the viewpoint of recognition accuracy particularly on lower quality real data. Moreover, because of the lower number of image representation features, the computation cost of the proposed method is lower than the original one. This fact is very important for practical usage.

#### 4.2.1. Adapted Kepenekci Matching

This approach combines two methods of matching and uses the weighted sum of the two values as a result as follows.

Let  $T$  be a test image and  $G$  a gallery image. For each feature vector  $t$  of face  $T$  we determine a set of relevant vectors  $g$  of face  $G$ . Vector  $g$  is relevant iff:

$$\sqrt{(x_t - x_g)^2 + (y_t - y_g)^2} < distanceThreshold \quad (1)$$

where  $x$  and  $y$  are the coordinates of the feature vector points.

If no relevant vector to vector  $t$  is identified, vector  $t$  is excluded from the comparison procedure. The overall similarity of two faces  $OS$  is computed as the average of similarities  $S$  between each pair of corresponding vectors as:

$$OS_{T,G} = mean \{S(t, g), t \in T, g \in G\} \quad (2)$$

Then, the face with the most similar vector to each of the test face vectors is determined. The  $C_i$  value informs how many times gallery face  $G_i$  was the closest one to some of the vectors of test face  $T$ . The similarity is computed as  $C_i/N_i$  where  $N_i$  is the total number of feature vectors in  $G_i$ . The weighted sum of these two similarities  $FS$  is used for similarity measure:

$$FS_{T,G} = \alpha OS_{T,G} + \beta \frac{C_G}{N_G} \quad (3)$$

The face is recognized as follows:

$$F\hat{S}_{T,G} = \arg \max_G (FS_{T,G}) \quad (4)$$

Note that the cosine similarity is used for vector comparison.

## 5. Confidence Measures

The Confidence Measure (CM) is used in the last step of the pipeline in order to identify and remove incorrectly recognized faces from the resulting set. As in many other papers [19, 13, 43], our confidence measure is based on the estimation of the *a posterior* class probability.

Let the output of the classifier be  $P(F|C)$ , where  $C$  is the recognized face class and  $F$  represents the face features. The values  $P(F|C)$  are normalized to compute *a posterior* class probabilities as follows:

$$P(C|F) = \frac{P(F|C).P(C)}{\sum_{I \in \mathcal{FLM}} P(F|I).P(I)} \quad (5)$$

$\mathcal{FLM}$  represents the set of all individuals and  $P(C)$  denotes the *prior* probability of the individual’s (face) class  $C$ .

The two confidence measure methods already presented in [22] are used. In the first approach, called *absolute confidence value*, only faces  $\hat{C}$  complying with

$$\hat{C} = \arg \max_C (P(C|F)) \quad (6)$$

$$P(\hat{C}|F) > T \quad (7)$$

are considered as being recognized correctly.

The second approach, called *relative confidence value*, computes the difference between the *best* score and the *second best* one by the following equation:

$$P\Delta = P(\hat{C}|F) - \max_{C \neq \hat{C}} (P(C|F)) \quad (8)$$

Only the faces with  $P\Delta > T$  are accepted. This approach aims to identify the “dominant” faces among all the other candidates.  $T$  is the acceptance threshold and its optimal value is adjusted experimentally.

### 5.1. Supervised Confidence Measure

We further propose another confidence measure approach which uses the scores  $R$  obtained by the methods presented previously. We use a Multi-layer Perceptron (MLP) to model *a posterior* probability  $P(H|R)$ . The variable  $H$  has only two values and determines whether the face image was classified correctly or not.

Three MLP configurations are built and evaluated:

1. supervised *absolute confidence value* method,
2. supervised *relative confidence value* method,
3. combination of methods 1 and 2.

The MLP topology will be described in detail in the experimental section. Note, that this confidence measure is proposed in order to improve the scores of the methods described above.

## 6. System Architecture

The presented system has (as shown in Figure 5) a modular architecture. It is composed of five modules (see the rectangles) connected by dependencies (see the oriented edges). The input image and the recognition results are represented by parallelograms. The storage of the face representation is shown by the *Face Gallery* sign.

The first module  $M1$  deals with face extraction. This module converts a color image into its grey-scale representation, then it performs face detection. The detected face is further extracted from the image in the next step. This module also detects the eyes in the detected face region and transforms and resizes the face.

The second module  $M2$  is used to create the face representation. It detects the SIFT key-points and creates a set of SIFT descriptors for a representation of the face image.

The next module  $M3$  is used to select the most representative face vectors in order to create a precise face model  $M$ . The algorithm implemented within this module is described in Section 3.4.

The fourth module  $M4$  deals with face recognition. A recognized face is compared to the face models stored in the *Face Gallery* and the most similar model is chosen as the recognized one.

The last *confidence measure* module  $M5$  is dedicated to identifying whether the recognition result is correct or not. This unique step is particularly important, because when the user knows that the recognition is probably not correct, he can manually correct the recognition result.

Note that the modules  $M1$  and  $M2$  are used in both face representation (or modelling) and face recognition tasks. However, module  $M3$  is used only for face representation and modules  $M4$  and  $M5$  are used only for recognition. The last remark is that every module should be used separately in order to create another face processing system.



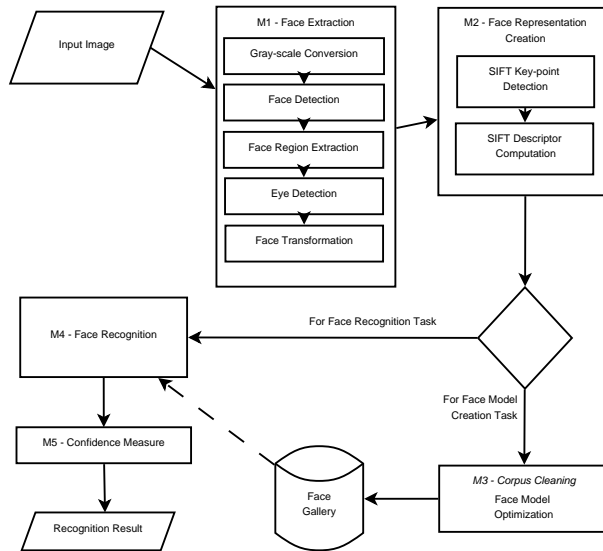


Figure 5: System Architecture

The time complexity of the whole system is  $O(N^3)$  for training and  $O(N^2)$  for recognition. The first value is the time complexity of the module  $M3$ , i.e. module with the maximal complexity over all modules in the training task. The second value is the time complexity of the modules  $M2$  and  $M4$ , i.e. also the modules with the maximal complexity over all modules in the recognition pipeline.

## 7. Experiments

In our previous work [23], we evaluated and compared the recognition results of our proposed approach on the ORL dataset and on a sub-set of the ČTK corpus of comparable size. The number of recognized individuals for the ORL and ČTK datasets was 40 and 37, respectively. We used 9 face examples for training and the remaining one for testing. The achieved recognition rate was 100% on the ORL dataset and 72.7% on the ČTK dataset. Both these scores outperform the original Kepenekci method.

We would like to extend our previous work and evaluate our approach on huge and real corpora. In the first experiment, we compare our approach with the other most efficient methods on a large sub-set of the FERET corpus. The second and third experiments show the performance of our approach on two other large standard corpora (AR Face Database and Labeled Faces in

the Wild) composed of the significantly varying face images. The remaining experiments show the results of our approach on a real ČTK dataset. These experiments are motivated by the need to evaluate our face recognition system under real conditions. It is worth mentioning that the image resolutions for each experiment differ.

### 7.1. FERET Dataset

The FERET dataset [33] contains 14,051 images of 1,199 individuals. The images were collected between December 1993 and August 1996. The resolution of the images is  $256 \times 384$  pixels. The images are divided into the following categories according to the face pose: frontal, 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 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*. Note that only one image per person/set is available. Figure 6 shows three example images from the FERET database.

Table 1: Images 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

### 7.2. AR Face Database

Another face dataset, called the AR Face Database<sup>4</sup> [29], was created at the Univerzitat Autònoma de Barcelona. This database contains more than

---

<sup>4</sup><http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>



Figure 6: Three example images of one person from the FERET face corpus (fa, fb and fc - *left to right*).



Figure 7: Four example images from the AR face database

4,000 colour images of 126 individuals. The images are stored in a *raw* format and their size is  $768 \times 576$  pixels. The individuals are captured under significantly different lighting conditions and with varying expressions. Another characteristic is a possible presence of glasses or scarf. Figure 7 shows four example images from this dataset.

### 7.3. Labeled Faces in the Wild (LFW)

The Labeled Faces in the Wild (LFW)<sup>5</sup> [17] is a new dataset created in 2007 in order to allow testing of face recognition methods in unconstrained environment. It contains more than 13,000 images of 5,749 people in a resolution  $250 \times 250$  pixels which were collected from the web. The images are stored in a *jpg* format. 1,680 people have at least two distinct images and they are used usually for face recognition experiments. The face images were detected by the Viola-Jones algorithm [40].

The database contains four different sub-sets for testing. The first one is an original dataset. The other three sets are composed of images obtained using different methods of alignment. The proposed testing protocol concentrates mainly on the face verification task. It means to decide whether two images belong to the same person or not. Figure 8 shows three example

---

<sup>5</sup><http://vis-www.cs.umass.edu/lfw/>



Figure 8: Three example images from the LFW database



Figure 9: Example images of one person from the ČTK face corpus. The top row shows the original images and the bottom row the extracted faces.

images from the LFW face database.

#### 7.4. Czech News Agency (ČTK) Database

This corpus was created as a result of the proposed corpus creation algorithm presented in Section 3. It is composed of images of individuals in an uncontrolled environment that were randomly selected from the large ČTK database of high resolution photographs. All images were taken over a long time period (20 years or more). The resulting corpus contains gray-scale images of 638 individuals of size  $128 \times 128$  pixels. Two versions of this corpus are used in following experiments. The first one has at least 10 images / person and represents the output of the proposed algorithm without the corpus cleaning step, while the second version (cleaned) contains at most 10 images for each person. Note that the orientation, lighting conditions and background of the images differ significantly. A correct face recognition on this dataset is thus very difficult.

Figure 9 shows examples of one face from this corpus. This corpus is available for free for research purposes at <http://home.zcu.cz/~pkral/sw/> or upon request to the authors.

### 7.5. Recognition Results on the FERET Database

Table 2 compares and evaluates our approach with the other very efficient methods on a large sub-set of the FERET corpus. The *fa* set is used for training, while the *fb* set is for testing.

The first reported method is an implementation of the EBGM algorithm by Bolme [6]. The second algorithm is proposed by Ahonen in [1]. This approach is based on Local binary patterns. Wagner et al. propose in [41] another efficient approach, based on the Sparse Representation and Classification (SRC) algorithm, whose score is reported next. The results of the novel approach based on linear regression proposed by Naseem et al. [32] are shown in the following line of this table. Recognition accuracy of the novel Discriminative Multi-Manifold Analysis (DMMA) method (Lu) [28] is presented next. The sixth method is introduced by Kepenekci in [20] and is based on the Gabor wavelets. The last approach is our implementation of the SIFT based Kepenekci method with the adapted Kepenekci matching.

The table shows that the recognition rates of all approaches are very good and close to one another. This experiment also shows that our SIFT based Kepenekci method slightly outperforms the other approaches. Regarding this fact and our previous experiments we decided to integrate this method into our AFR system.

The following experiments are undertaken in order to show the performance of this method on real data. We would also like to determine the optimal parameters of this approach for our task.

Table 2: Recognition results of various very efficient AFR methods on the FERET dataset

No.	Method	Recognition rate [%]
1.	EBGM (Bolme)	96.4
2.	Local binary patterns (Ahonen)	97.0
3.	SRC algorithm (Wagner)	95.2
4.	Linear regression (Naseem)	93.5
5.	DMMA (Lu)	93.0
6.	Kepenekci method (Kepenekci)	96.3
7.	<b>SIFT based Kepenekci method</b>	<b>97.3</b>

### 7.6. Recognition Results on the AR Database

Table 3 compares and evaluates our approach with the other very efficient methods on the AR face corpus composed of the images recorded in significantly varying conditions. This experiment was realized in order to show the performance of our method on another corpus recorded in relatively real conditions.

Unfortunately, the other studies often use different experimental set-ups on this database. Subsets with less individuals are usually used to evaluate the proposed methods. Therefore, we report also the number of the recognized individuals (*subj.*) of the other studies. In order to show the superior recognition abilities of the proposed SIFT based Kepenekci approach, we recognize 120 individuals.

The recognition rates of the first four reported methods presented by Wright et al. in [44] show that the SVM approach is the best performing one on this dataset. The results of the second three methods are chosen from the comparative study of Yeh et al. [46]. The authors show that the LBP method outperforms significantly two other methods, the SVM and the SRC on the AR corpus. The following two methods proposed by Geng et al. [14] are based on the SIFT features. The last line of this table shows the results of our proposed SIFT based Kepenekci approach. This table reports that our proposed method gives better recognition accuracy than all previous approaches.

### 7.7. Recognition Results on the LFW Database

As already stated, this face corpus is used for face verification. Therefore, it is not possible to compare our proposed approach directly with the other ones. We implemented three best performing methods described in the literature. These approaches are used to recognize 1,680 different people from this corpus. Then, we use our proposed approach with the similar experimental set-up.

Table 4 reports the recognition results of this experiment. This table shows that the proposed method outperforms significantly the other approaches. The recognition rate improvement is 1.79% in the absolute value.

### 7.8. Parameter Optimization for the ČTK Corpus

The images in the ČTK corpus differ significantly from the images in the FERET database. Therefore, an optimization of our method for this corpus is beneficial.

Table 3: Recognition results of various very efficient AFR methods on the AR face database

No.	Method	Recognition rate [%]
1.	Nearest neighbour (Wright, 100 subj.)	89.7
2.	Nearest subspace (Wright, 100 subj.)	90.3
3.	SRC (Wright, 100 subj.)	94.7
4.	SVM (Wright, 100 subj.)	95.7
5.	SVM (Yeh, 120 subj.)	87.5
6.	LBP (Yeh, 120 subj.)	94.0
7.	SRC (Yeh, 120 subj.)	89.7
8.	Original SIFT (Geng, 75 subj.)	93.9
9.	PD-SIFT (Geng, 75 subj.)	95.2
10.	<b>SIFT based Kepenekci method (Lenc, 120 subj.)</b>	<b>95.8</b>

The most important parameter is the *distance threshold* (see Equation 1) which significantly influences recognition accuracy.

This threshold defines the surrounding region to a given key-point. Only the vectors in this region are used for comparison in the recognition step. The optimal value of this threshold is strictly related to the size and quality of the recognized images. Unfortunately, it is not possible to analytically set its optimal value. Therefore, this value will be set experimentally on a development corpus as demonstrated in Figure 10. The evaluated values are in the interval  $]0; 50]$ . However, for better clarity of this figure, the reported interval is reduced by 10 from both sides.

The highest obtained recognition rate is 61.91% when the threshold value is set at 35. Therefore, this value will be used in the following experiments.

### 7.9. Recognition Results on the ČTK Corpus

The first experiment demonstrates the performance of the proposed corpus creation algorithm presented in Section 3 on the ČTK corpus with a particular focus on corpus cleaning. The recognition rates in two cases, without and with the corpus cleaning step, are compared.

From the viewpoint of the application of our system, we would like to identify whether a correctly recognized face belongs to the  $N$  best recognized

Table 4: Recognition results of various very efficient AFR methods on the LFW face corpus

No.	Method	Recognition rate [%]
1.	Gabor wavelets (Kepenekci)	95.96
2.	LBP (Ahonen)	96.25
3.	SIFT (Aly)	95.83
4.	<b>SIFT based Kepenekci method (Lenc)</b>	<b>98.04</b>

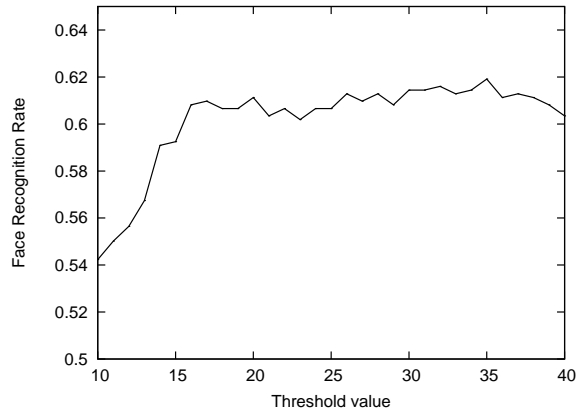


Figure 10: Optimal distance threshold value estimation for the ČTK corpus

ones. If this is the case, we can propose to a user choosing the correctly recognized face, for example, from a drop-down list.

Therefore, three different evaluation metrics are described next. The *Correct* metric is a classical recognition rate computed as  $\frac{\text{correct}}{\text{all}}$  where *correct* represents the number of correctly recognized faces and *all* is the number of all recognized faces. The second metric, called *Correct in 5 most similar*, considers a correct recognition result when a correct face belongs to the set of the *five* most similar faces. The last metric is similar to the previous one, while a correct result is considered when a correct face belongs to the set of the *ten* most similar (*best*) recognized faces.

The first section of Table 5 compares the recognition rates on the ČTK corpus without and with the corpus cleaning step, i.e. on the raw and on the cleaned version of the ČTK corpus. The table shows that corpus cleaning



Table 5: Comparison of the recognition results on the raw and the cleaned version of the ČTK corpus using three different evaluation metrics

	Raw corpus	Clean corpus
Evaluation metric	Recognition rate [%]	
<b>1. SIFT based Kepenekci method</b>		
Correct	30.72	61.91
Correct in 5 most similar	36.83	70.06
Correct in 10 most similar	40.28	74.61
<b>2. Original Kepenekci method</b>		
Correct	13.32	48.75
Correct in 5 most similar	17.24	57.21
Correct in 10 most similar	19.44	60.50

is beneficial for the correct estimation of the face models. The improvement in recognition accuracy is greater than 30% for all three metrics. This table also shows that the error rate of the *Correct* metric is about 38% and that about 9% and 13% of errors belong to the 5 and 10 best recognized faces, respectively.

Note that we assume the use of the confidence measures in order to automatically detect incorrectly recognized faces. In these cases, users should manually correct these incorrectly recognized examples.

In the next experiment, we compare the performance of the SIFT-based Kepenekci method with the original Kepenekci approach. The results of the original Kepenekci approach are shown in the second section of Table 5. This table clearly shows that the SIFT-based Kepenekci method significantly outperforms the original Kepenekci approach. The error rate reduction is about 26% in relative value when corpus cleaning is used (i.e. on the cleaned corpus).

#### 7.10. Recognition Results with Confidence Measures

As in many other articles in the confidence measure field [26], we will use the Receiver Operating Characteristic (ROC) curve [7] to evaluate our methods. This curve clearly shows the relationship between the True Positive

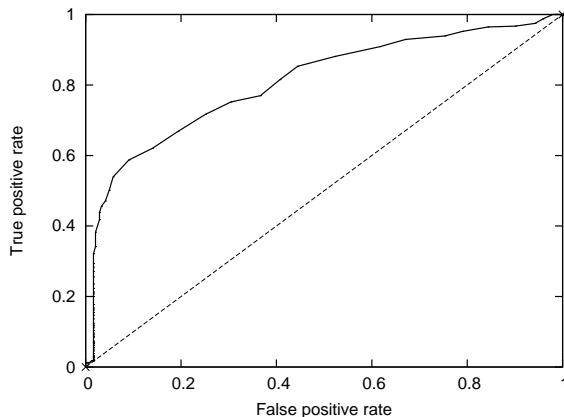


Figure 11: ROC curve of the *absolute confidence value* method

(TP) and the False Positive (FP) rate for different values of the *acceptance* threshold.

Figure 11 shows the results of the *absolute confidence value* method, while the results of the *relative confidence value* approach are given in Figure 12. These figures show that both confidence measures are suitable for our task in order to identify incorrectly recognized faces. Moreover, the *relative confidence value* method significantly outperforms the *absolute confidence value* approach.

In the last part of this experiment, we would like to evaluate the results of the proposed supervised confidence measure methods which use an MLP. The best MLP topology uses three layers (in all cases): one or two input neurons, 10 neurons in the hidden layer and two outputs (*correctly* and *incorrectly* recognized face). One input is used in the case when the CM methods are used separately and two input neurons are used when we combine both confidence measures. The MLP topology was found empirically on a small development corpus which contains 120 examples (i.e. 120 confidence values).

Table 6 shows the results of these CMs. As in many other works, F-measure (F-mes) [35] is used as an evaluation metric. Because of the importance of evaluating the quality of the classification as well as the identification, Precision (Prec) and Recall (Rec) are also reported in this table. This table clearly shows that the second supervised *relative confidence value* method significantly outperforms the first supervised *absolute confidence value* approach. However, the first method brings further relevant

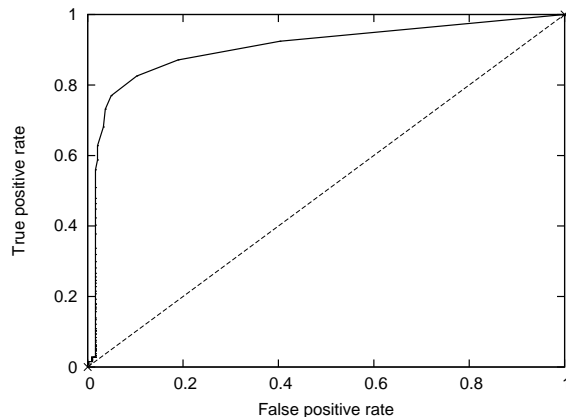


Figure 12: ROC curve of the *relative confidence value* method

information regarding the second approach. This fact is confirmed by the result of the third combined approach, which gives the best recognition score from the whole.

Table 6: Performance of the three supervised confidence measures [in %]

No	Confidence Measure	Prec	Rec	F-mes
1.	supervised <i>absolute confidence value</i> method	89.8	68.1	77.5
2.	supervised <i>relative confidence value</i> method	94.1	82.5	87.9
3.	combination of methods 1 and 2	95.6	87.9	91.6

## 8. Contributions

The main scientific contribution of this work can be summarized below:

- We proposed and implemented a new face corpus creation algorithm. This algorithm allows to create a face corpus from annotated pictures which may be composed of more people, some background objects, etc. The algorithm was evaluated on the lower quality real ČTK data.
- We created a new facial corpus from the ČTK data designed particularly for the evaluation of automatic face recognition in real condi-

tions. This corpus is available for research purposes for free at <http://home.zcu.cz/~pkral/sw/> or upon request to the authors.

- We evaluated a novel face recognition method, called the SIFT based Kepenekci approach, which is based on the SIFT features and adapted Kepenekci matching on four large corpora: FERET, AR, LFW and ČTK. This algorithm significantly outperforms the other approaches particularly on the lower quality real data.
- We evaluated two confidence measure techniques based on *a posterior* class probability. Then, we proposed and evaluated two supervised confidence measures which use an MLP as a classifier. These CMs extend both previously proposed ones.
- We have proposed, implemented and evaluated an experimental fully automatic face recognition system. This system slightly outperforms the other efficient approaches from the viewpoint of face recognition accuracy and it is also sufficiently robust on lower quality real data.
- The last and most important contribution is that we have made this system (with the source code) publicly available for research purposes for free at <http://home.zcu.cz/~pkral/sw/> or upon request to the authors.

## 9. Conclusions and Future Work

In this paper, we proposed an experimental fully automatic face recognition system for the Czech News Agency. The system is used for the labelling of people in photographs during insertion into the ČTK database.

We compared our SIFT based Kepenekci face recognition approach with a number of the most efficient AFR methods on three large corpora: FERET, AR and LFW. We showed that this approach outperforms all other methods especially on the lower quality real data. Therefore, we decided to integrate this method into our face recognition system.

Then, we proposed a corpus creation algorithm in order to extract the faces from the database and to create a facial corpus. We showed that it is necessary to use this algorithm in our AFR system in the preprocessing step in order to create sufficiently precise face models. We also compared the performance of our SIFT based Kepenekci face recognition approach with

the original Kepenekci method on the created ČTK corpus. The experiments show that our approach significantly outperforms the original one and that the error rate reduction is about 26% in a relative value.

The last goal of this paper consisted in proposing and evaluating confidence measure methods in order to identify the incorrectly recognized examples and remove them from the recognition results. Two novel supervised confidence measures based on the *a posterior* class probability and a multi-layer perceptron are proposed. We experimentally proved that both proposed CMs are suitable for our task and that they bring complementary information for the detection of incorrectly recognized examples. Therefore, the best F-measure is obtained by the combined MLP approach.

We envisage two perspectives for our further research in the near future. The recognition rate of our face recognition system on the large FERET dataset is about 97%. The recognition score on the real ČTK data is quite good, however, it is still far to be perfect. We assume that our face recognition module is accurate enough whereas the detection and extraction module needs significant improvement. Therefore, we would like to propose a more suitable face detection and extraction approach. For example, the simple rotation according to the eyes will be replaced by more sophisticated image transformations.

Another perspective consists in proposing better confidence measures in the post-processing step. The novel confidence measures will be based on the main properties of the face model. We further assume combining them with the CMs proposed in this paper.

## Acknowledgements

This work was partly supported by the UWB grant SGS-2013-029 Advanced Computer and Information Systems and by the European Regional Development Fund (ERDF), project “NTIS - New Technologies for Information Society”, European Centre of Excellence, CZ.1.05/1.1.00/02.0090. We would also like to thank the Czech News Agency for their support and for providing the photo data.

- [1] Ahonen, T., Hadid, A., Pietikinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J. (eds.) Computer Vision - ECCV 2004, Lecture Notes in Computer Science, vol. 3021, pp. 469–481. Springer Berlin Heidelberg (2004), [http://dx.doi.org/10.1007/978-3-540-24670-1\\_36](http://dx.doi.org/10.1007/978-3-540-24670-1_36)

- [2] Aly, M.: Face recognition using sift features (2006)
- [3] Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110(3), 346–359 (Jun 2008), <http://dx.doi.org/10.1016/j.cviu.2007.09.014>
- [4] Beham, M.P., Roomi, S.M.M.: A review of face recognition methods. *International Journal of Pattern Recognition and Artificial Intelligence* 27(4) (2013)
- [5] Belhumeur, P.N., Hespanha, J.a.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1997)
- [6] Bolme, D.S.: Elastic Bunch Graph Matching. Ph.D. thesis, Colorado State University (2003)
- [7] Brown, C.D., Davis, H.T.: Receiver operating characteristics curves and related decision measures: A tutorial. *Chemometrics and Intelligent Laboratory Systems* 80(1), 24–38 (2006)
- [8] Campadelli, P., Lanzarotti, R.: A face recognition system based on local feature characterization. In: *Advanced Studies in Biometrics*, pp. 147–152. Springer (2005)
- [9] Çeliktutan, O., Ulukaya, S., Sankur, B.: A comparative study of face landmarking techniques. *EURASIP Journal on Image and Video Processing* 2013(1), 13 (2013)
- [10] Deng, W., Liu, Y., Hu, J., Guo, J.: The small sample size problem of ICA: A comparative study and analysis. *Pattern Recognition* 45(12), 4438–4450 (2012)
- [11] Draper, B.A., Baek, K., Bartlett, M.S., Beveridge, J.R.: Recognizing faces with PCA and ICA. In: *COMPUTER VISION AND IMAGE UNDERSTANDING, SPECIAL ISSUE ON FACE RECOGNITION*. pp. 115–137 (2003)
- [12] Eickeler, S., Jabs, M., Rigoll, G.: Comparison of confidence measures for face recognition. In: *FG*. pp. 257–263. IEEE Computer Society (2000), <http://dblp.uni-trier.de/db/conf/fgr/fg2000.html#EickelerJR00>

- [13] Eickeler, S., Jabs, M., Rigoll, G.: Comparison of confidence measures for face recognition. In: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. pp. 257–262. IEEE (2000)
- [14] Geng, C., Jiang, X.: Face recognition using SIFT features. In: Image Processing (ICIP), 2009 16th IEEE International Conference on. pp. 3313–3316. IEEE, Cairo, Egypt (November 2009)
- [15] Gross, R.: Face databases. In: S.Li, A. (ed.) Handbook of Face Recognition. Springer, New York (February 2005)
- [16] Hu, X., Mordohai, P.: A quantitative evaluation of confidence measures for stereo vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(11), 2121–2133 (2012)
- [17] Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst (October 2007)
- [18] Huang, K., Aviyente, S.: Sparse representation for signal classification. Advances in neural information processing systems 19, 609 (2007)
- [19] Jiang, H.: Confidence measures for speech recognition: A survey. Speech Communication 45(4), 455–470 (2005)
- [20] Kepenekci, B.: Face Recognition Using Gabor Wavelet Transform. Ph.D. thesis, The Middle East Technical University (2001)
- [21] Krizaj, J., Struc, V., Pavesic, N.: Adaptation of sift features for robust face recognition (2010)
- [22] Lenc, L., Král, P.: Confidence measure for automatic face recognition. In: International Conference on Knowledge Discovery and Information Retrieval. Paris, France (October 26-29 2011)
- [23] Lenc, L., Král, P.: Face recognition under real-world conditions. In: International Conference on Agents and Artificial Intelligence. Barcelona, Spain (February 14-18 2013)

- [24] Li, F., Wechsler, H.: Open world face recognition with credibility and confidence measures. In: *Audio-and Video-Based Biometric Person Authentication*. pp. 462–469. Springer (2003)
- [25] Li, W., Fu, P., Zhou, L.: Face recognition method based on dynamic threshold local binary pattern. In: *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*. pp. 20–24. ACM (2012)
- [26] Lleida, E., Rose, R.C.: Likelihood ratio decoding and confidence measures for continuous speech recognition. In: *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*. vol. 1, pp. 478–481. IEEE (1996)
- [27] Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 2 (2004)
- [28] Lu, J., Tan, Y.P., Wang, G.: Discriminative multimanifold analysis for face recognition from a single training sample per person. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35(1), 39–51 (2013)
- [29] Martinez, A., Benavente, R.: *The AR Face Database*. Tech. rep., Universitat Autònoma de Barcelona (1998)
- [30] Martinez, A.M.: Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class (2002)
- [31] Marukatat, S., Artières, T., Gallinari, P., Dorizzi, B.: Rejection measures for handwriting sentence recognition. In: *Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on*. pp. 24–29. IEEE (2002)
- [32] Naseem, I., Togneri, R., Bennamoun, M.: Linear regression for face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32(11), 2106–2112 (2010)
- [33] Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing* 16(5), 295–306 (1998)



- [34] Poon, B., Amin, M.A., Yan, H.: Performance evaluation and comparison of pca based human face recognition methods for distorted images. *International Journal of Machine Learning and Cybernetics* 2(4), 245–259 (2011)
- [35] Powers, D.: Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation. *Journal of Machine Learning Technologies* 2(1), 37–63 (2011)
- [36] Proedrou, K., Nouretdinov, I., Vovk, V., Gammerman, A.: Transductive confidence machines for pattern recognition. In: *ECML'02*. pp. 381–390 (2002)
- [37] Servin, B., de Givry, S., Faraut, T.: Statistical confidence measures for genome maps: application to the validation of genome assemblies. *Bioinformatics* 26(24), 3035–3042 (2010)
- [38] Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Image Processing, IEEE Transactions on* 19(6), 1635–1650 (2010)
- [39] Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. In: *IEEE Computer Society Conference on In Computer Vision and Pattern Recognition*. *Computer Vision and Pattern Recognition* (1991)
- [40] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Conference on Computer Vision and Pattern Recognition* (2001)
- [41] Wagner, A., Wright, J., Ganesh, A., Zhou, Z., Mobahi, H., Ma, Y.: Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34(2), 372–386 (2012)
- [42] Wang, B., Li, W., Li, Z., Liao, Q.: Adaptive linear regression for single-sample face recognition. *Neurocomputing* 115, 186–191 (2013), cited By (since 1996)1
- [43] Wessel, F., Schluter, R., Macherey, K., Ney, H.: Confidence measures for large vocabulary continuous speech recognition. *Speech and Audio Processing, IEEE Transactions on* 9(3), 288–298 (2001)

- [44] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31(2), 210–227 (2009)
- [45] Yang, B., Chen, S.: A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image. *Neurocomputing* 120, 365–379 (2013), cited By (since 1996)3
- [46] Yeh, C.H., Shih, P.R., Lin, Y.T., Liu, K.T., Chang, H.M., Ouhyoung, M.: A comparison of three methods of face recognition for home photos. In: *SIGGRAPH'09: Posters*. p. 30. ACM (2009)
- [47] Zhang, B., Gao, Y., Zhao, S., Liu, J.: Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *Image Processing, IEEE Transactions on* 19(2), 533–544 (2010)
- [48] Zhang, C.Y., Ruan, Q.Q.: Face recognition using l-fisherfaces. *Journal of information science and engineering* 26(4), 1525–1537 (2010)
- [49] Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. *Acm Computing Surveys (CSUR)* 35(4), 399–458 (2003)

**Ladislav Lenc** is a lecturer/researcher at the University of West Bohemia at Dept. of Computer Science & Engineering (Czech Republic). His research interests are face recognition and computer vision. He received his Ph.D. degree in 2014 at Dept. of Computer Science & Engineering at the University of West Bohemia.

**Pavel Král** is a lecturer/researcher at the University of West Bohemia at Dept. of Computer Science & Engineering (Czech Republic). His research domains are natural language processing and pattern recognition. He received his Ph.D. degree in 2007 at Dept. of Computer Science & Engineering at the University of West Bohemia and at Henri Poincaré University in Nancy (France).