Automatic Face Corpus Creation

Ladislav Lenc[1], Pavel Král[1]
[1]*Department of Computer Science and Engineering, University of West Bohemia, Plzeň, Czech Republic*
{*llenc, pkral*}*@kiv.zcu.cz*

Keywords: Automatic Face Recognition, Czech News Agency, Scale Invariant Feature Transform, Corpus Creation

Abstract: This paper deals with the automatic real-world face corpus creation. The main contribution consists in proposition and evaluation of the automatic face corpus creation algorithm. Next, we statistically analysed the structure of the created face corpus when the automatic algorithm is used. We further compared the face recognition accuracy of our previously developed face recognition approach on this corpus while using different size/quality datasets. We have shown that the manual verification of the corpus is not necessary. Therefore, we concluded that our proposed algorithm is suitable for the further use by the Czech News Agency, our commercial partner.

1 INTRODUCTION

In this paper, we are focusing on the automatic labeling of people in the database with a huge number of the real-world photographs. Certain portion of the pictures is labeled (information about the person identity available). The rest is unlabeled. In this case, the recognition is tightly connected with the face detection and extraction steps because the pictures do not contain only the faces.

Automatic corpus creation methods have been developed and evaluated particularly in the speech processing domain (Chen and Nie, 2000; Tomás et al., 2001). Unfortunately, to the best of our knowledge there is only little work on the automatic corpus creation in the face recognition field. All the well known face databases have been created manually. However, manual labeling is a very time-consuming and expensive task.

The main goal of this paper thus consists in creation of a huge real-world face database. The creation process must be as automatic as possible. The labeled examples will be used for this task. The main contribution of this work is proposition and evaluation of the automatic face corpus creation algorithm. Another contribution of this paper is the statistical analysis of the results of the face corpus creation process when the automatic creation algorithm is used. The newly created face corpus will be used to evaluate some face recognition approaches, which represents the next contribution of this paper. We also compare the face recognition accuracy while using different size/quality datasets. The results of this work will be used by the Czech News Agency (ČTK).

For the face recognition, we use the adapted Scale Invariant Feature Transform (SIFT)-based Kepenekci method (Lenc and Král, 2012), which has shown very good recognition accuracy on standard datasets (e.g. ORL). It is based on the SIFT algorithm proposed by David Lowe in (Lowe, 2004). The SIFT algorithm is used for feature extraction and the matching scheme proposed by Kepenekci (Kepenekci, 2001) is used for face comparison.

The following section describes the proposed corpus creation algorithm and the structure of the created face corpus. The next section contains the face recognition results on the created corpus. Finally, the Section 4 summarizes the results and gives some ideas for the future research.

2 AUTOMATIC CORPUS CREATION

We used the ČTK photo-database for all experiments. Every picture contains one face of a known person (with the label). Unfortunately, the photos contain not only the face itself. They may be composed of more people, some background objects, etc.

2.1 Proposed Algorithm

Therefore, we propose an algorithm in order to detect and extract the faces from the pictures and to create a face corpus as most automatically as possible:

1. face detection,
2. identification and deletion of the incorrectly detected faces,
3. eyes detection and head rotation according to the eyes

2.1.1 Face detection

We use the OpenCV library `http://opencv.willowgarage.com/wiki/` for the face detection task. It implements the Viola-Jones algorithm (Viola and Jones, 2001) which is one of the most successful face

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

detection algorithms. Figure 1 shows one example of the correctly and one example of the incorrectly detected face.

2.1.2 Incorrectly detected faces identification and deletion

Unfortunately, a certain number of the incorrectly detected faces occurs in the output. Therefore, a verification of the detected faces is indispensable. In order to avoid the manual processing, we propose a classifier for this task. It is used to classify the pictures into two classes: F (faces) and NF (non faces). We assume that the color distribution of these two classes differ significantly. Therefore, we compute a histogram for every picture and we use histogram values as a feature vector.

We utilize a neural network of the type Multi-Layer Perceptron (MLP) as an implementation of the classifier due to its simplicity and our know-how in this field. It is trained on manually selected sub-set of faces (50 examples) and non-faces (50 examples). The MLP topology has 3 layers: 256 input nodes (each input corresponds to one intensity value in the grayscale picture), 10 neurons in the hidden layer and two output nodes: $F \times NF$ classes. In order to evaluate the accuracy of this classifier, we chosen randomly 100 examples from each class and we verify them manually. The face $\times$ non-face classifier's accuracy is about 87% which is enough for our further experiments.

2.1.3 Eyes detection and head rotation according to the eyes

As concluded in many previous studies, the face quality influence significantly the face recognition accuracy. Our database is composed of the real-world photograph where the faces are usually not taken from the front, but from the different angles. It is thus beneficial to transform the faces in order to be as similar as possible, i.e. front-view faces. We simplify this transformation by the face rotation according to the eyes.

The eyes are detected using a Viola-Jones algorithm (as for the face detection task). If both eyes are detected successfully the face is rotated and positioned so that the eyes are in the horizontal line for all images. Finally, we perform a brief manual control (on 100 examples) in order to evaluate the accuracy of our eyes detection task. The resulting eyes detection accuracy is thus about 94%. Figure 2 shows one example of the eyes detection task where two, one and no eye was detected by the algorithm.

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

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

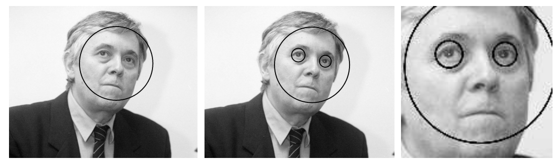

Figure 3: Example of the face detection (left), eyes detection (middle) and the face rotation according to the eyes (right)

2.2 Resulting Corpus

Figure 4 shows the structure of the automatically created face corpus by the above described algorithm. We can summarize the important information as follows:

- in 1478 pictures no face was detected by the Viola Jones method,
- another 3158 images was marked as no faces by the MLP face $\times$ non-face classifier,
- another 6623 face-pictures was not rotated according to the eyes because no eye or one eye detected,
- in the 4562 faces two eyes were detected and these faces were rotated according to the eyes,

Note, that the whole corpus contains 15821 labeled images.

3 FACE RECOGNITION

The experiments have been motivated by the fact, that our previously proposed face recognition approaches work

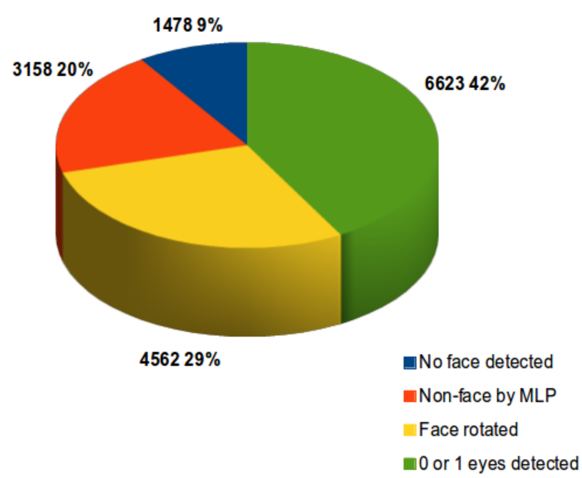

Figure 4: Structure of the created ČTK face corpus

very well (recognition accuracy close to 100% as already shown in (Lenc and Král, 2012)) on the standard ORL dataset and we would like to evaluate our best approach on the real-world data (i.e. previously created corpus).

In all following experiments, an adapted SIFT-based Kepenekci face recognition approach (Lenc and Král, 2012) and a cross-validation procedure will be used.

3.1 Face datasets

The face recognition accuracy is significantly influenced by the three parameters:

1. quality of the face corpus,
2. number of the recognized faces,
3. number of the training face examples per person.

Therefore, we create some different face sub-sets (see Table 1) in order to show the accuracy of the face recognition when these parameters differ.

The first part of this table shows the number of individuals when the face detection step is not verified, while the second part reports the number of individuals when the verification of the face detection was used. The second column represents the number of individuals with a successful eyes detection, while the number of individuals with incorrectly detected eyes is reported in the third column.

Table 1: Face dataset sizes in relation to: a) the number of the face examples per person; b) verification of the face detection; c) result of the eye detection

Example no. per person	Correctly detected eyes	Incorrectly detected eyes
1. Face detection without any verification		
	Number of individuals	
10	121	101
8	229	238
6	395	458
4	594	781
2. Face detection with the MLP verification		
	Number of individuals	
10	34	25
8	91	106
6	244	262
4	468	568

3.2 Face recognition results

Table 2 shows the face recognition accuracy on the different datasets. We can conclude some interesting knowledge from this table:

- eyes detection is very important for the face recognition (poor face recognition accuracies of the sub-sets with the incorrect eyes detection),
- the face detection using the Viola Jones method is good enough and the further verification by the MLP classifier is not necessary (better recognition accuracy of the MLP verified set could be caused by the smaller number of recognized examples, only 34 individuals in the smallest set),
- face recognition accuracy is significantly lower when the number of the recognized individuals increases and the number of training examples decreases,
- the best recognition accuracy (about 71%) is obtained on the sub-set with the smallest number of recognized individuals and with the highest number of training examples.

Table 2: Face recognition accuracy on the different datasets in [%]

Example no. per person	Correctly detected eyes	Incorrectly detected eyes
1. Face detection without any verification		
	Face recognition accuracy in [%]	
10	56.94	19.01
8	44.60	14.45
6	34.89	12.23
4	24.28	7.74
2. Face detection with the MLP verification		
	Face recognition accuracy in [%]	
10	71.18	47.20
8	52.75	22.76
6	35.79	16.17
4	24.84	9.99

In our previous experiments (Lenc and Král, 2012) we show that the number of training examples is very important for the face recognition. Unfortunately, we do not know, whether the second parameter influences the face recognition in the same manner or it is less important. Therefore, we realized the following experiment that shows the face recognition accuracy for constant number of training examples (equal to 3). In this experiment, only the cases with the correctly detected eyes and the rotated faces are considered.

Table 3 shows the face recognition accuracy of this experiment. We can conclude that the number of the training examples influences our face recognition method much more than the number of the recognized individuals. In all cases, the recognition accuracy was comparable except the case with the lowest count of individuals (slightly higher recognition accuracy).

Table 3: Face recognition accuracy on the different datasets with 4 face examples per person in [%]

Number of individuals	Face recognition accuracy in [%]
1. Face detection without any verification	
121	24.79
229	25.67
395	26.27
594	24.28
2. Face detection with the MLP verification	
34	32.35
91	27.75
244	23.77
468	24.84

In the previously described experiments, only part of the results of the proposed face detection algorithm was manually controlled. Therefore, the face recognition errors occur due to two reasons:

1. error of the face classification method,
2. error of the corpus creation algorithm (face detection and extraction).

In the last experiment, we would like to identify the corpus creation errors in order to evaluate our classifier as accurate as possible. Table 4 reports the results of this experiment. This experiment shows that manual verification improve the recognition accuracy only slightly, improvement about 3% and 5% for 34 and 91 individuals, respectively.

Table 4: Comparison of the face recognition accuracy of the automatic and manual verified face datasets

Number of individuals	Number of examples/individual	Face recognition accuracy in [%]
1. Face detection and extraction without manual verification		
34	10	71.18
91	8	52.75
2. Face detection and extraction with manual verification		
34	10	74.12
91	8	57.28

4 CONCLUSIONS AND PERSPECTIVES

In this work, we proposed and evaluated a new face corpus creation algorithm with the objective to create a real-world face corpus for testing of our previously developed face recognition approaches. We also statistically analysed the results of the face corpus creation process when the proposed algorithm was used. This analysis shows that 4562 images was processed successfully by the proposed algorithm. We further compared the face recognition accuracy while using different size/quality datasets. We demonstrated that the eyes detection is very important step in the face corpus creation process. We also proved, that the number of the training examples influences the face recognition much more than the number of the recognized individuals. In the last experiment, we showed that manual verification of the corpus improve the recognition accuracy only slightly. Therefore, we conclude that our proposed algorithm is suitable for the further use by the Czech News Agency, our commercial partner.

As mentioned previously, the face recognition accuracy is influenced significantly by the quality of the face database. Therefore, the first perspective consists in harmonization of the faces by using better face transformations than the simple rotation according to the eyes. Another perspective consists in the use of confidence measures as the post-processing step (Lenc and Král, 2011). The confidence measure technique will be used to detect and remove incorrectly detected (and rotated) face examples from the resulting corpus. The confidence measure will be also used after the face recognition task.

ACKNOWLEDGEMENTS

This work has been partly supported by the UWB grant SGS-2010-028 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 also would like to thank Czech New Agency (ČTK) for support and for providing the photographic data.

REFERENCES

Chen, J. and Nie, J.-Y. (2000). Automatic construction of parallel English-Chinese corpus for cross-language information retrieval. In *The sixth conference on Applied natural language processing*. Association for Computational Linguistics.

Kepenekci, B. (2001). *Face Recognition Using Gabor Wavelet Transform*. PhD thesis, The Middle East Technical University.

Lenc, L. and Král, P. (2011). Confidence measure for automatic face recognition. In *International Conference on Knowledge Discovery and Information Retrieval*, Paris, France.

Lenc, L. and Král, P. (2012). Novel matching methods for automatic face recognition using SIFT. In *8th AIAI (Artificial Intelligence Applications and Innovations) Confence*, Halkidiki, Greece.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2.

Tomás, J., Fabregat, M. del Val, J., Casacuberta, F., Picó, D., Sanchís, A., and Vidal, E. (2001). Automatic development of spanish-catalan corpora for machine translation. In *Second International Workshop on Spanish Language Processing and Language Technologies*, Jaén, Spain.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Conference on Computer Vision and Pattern Recognition*.