

CONFIDENCE MEASURE FOR AUTOMATIC FACE RECOGNITION

Ladislav Lenc and Pavel Král

Department of Computer Science and Engineering, University of West Bohemia, Plzeň, Czech Republic

Keywords: Automatic face recognition, Confidence measure, Gabor wavelets, Czech news agency.

Abstract: This paper deals with the use of confidence measure for Automatic Face Recognition (AFR). AFR is realized by the adapted Kepenekci face recognition approach based on the Gabor wavelet transform. This work is motivated by the fact that obtained recognition rate on the real-world corpus is only about 50% which is not sufficient for our application, a system for automatic labelling of the photographs in a large database. The main goal of this work is thus the proposition of the post-processing of the classification result in order to remove the “incorrectly” classified face images. We show that the use of confidence measure to filter out incorrectly recognized faces is beneficial. Two confidence measures are proposed and evaluated on the Czech News Agency (ČTK) corpus. Experimental results confirm the benefit of the use of confidence measure for the automatic face recognition task.

1 INTRODUCTION

Automatic Face Recognition (AFR) consists of automatic identification of a person from a digital image or from a video frame by a computer. A huge amount of methods for face recognition was proposed in the last two decades. As many researchers agree (Wiskott et al., 1999; Bolme, 2003; Shen and Bai, 2006), some of the most successful methods are based on Gabor wavelet transform.

Particularly, the method proposed by Kepenekci in (Kepenekci, 2001) gives a very good face recognition accuracy and outperforms the majority of the other approaches. In our previous work, we thus adapted the Kepenekci method in order to increase significantly the recognition accuracy and robustness of the algorithm and decrease the computation time. We experimentally confirmed that the proposed modifications significantly improve recognition rate and robustness of the algorithm, while the computation time doesn't change. When the parameters are set correctly, the use of more training examples increases the recognition rate on the ORL and Czech News Agency (ČTK) corpora by 17% and by 35% in absolute value, respectively. The final obtained recognition rate on the ORL database is 100%, however for the ČTK corpus it is only 50%.

The results of this work will be used by the ČTK. ČTK owns a large database (about 2 millions) of photographs. A significant number of photos is manually

annotated (i.e. the photo identity is known). However, other photos are unlabelled; the identities are thus unknown. The main task of our application consists in automatic labelling of the unlabelled photos. This application must also handle cases when a new photograph is added into the database (automatic labelling of this picture). The system must also guarantee the cases when the image is not well aligned, when its pose varies and when the lighting conditions differ. Note that about ten labelled images of each person are available.

The above presented recognition accuracy is not sufficient for our application. The main goal of this work thus consists of the proposition of the post-processing of the classification result in order to remove the “incorrectly” classified face images. We proposed the confidence measure technique for this task. Note that to the best of our knowledge, there is very little existing work on confidence measure in the automatic face recognition domain. Alternatively, a number of studies have been published for other domains, and particularly for automatic speech recognition.

Section 2 presents a short review of face recognition approaches. Section 3 describes confidence measure methods that we used in order to remove the incorrectly recognized examples. Section 4 deals with the evaluation of the results. In the last section, we discuss the research results and we propose some future research directions.

2 RELATED WORK

Early face recognition approaches are based on normalized error measures between significant face points, as proposed for instance in (Bledsoe, 1966). The main drawback of such methods consists in the necessity of manual labelling of the images. Another method that uses similar measurements was designed by Kanade (Kanade, 1977). In this case, the labelling of important face points is automatic.

One of the first successful approaches is Principal Component Analysis (PCA), so called Eigenfaces (Turk and Pentland, 1991). Eigenfaces is a statistical method that takes into account the whole image as a vector. Image vectors are put together and create a matrix. Eigenvectors of this matrix are calculated. Face images can then be expressed as a linear combination of these vectors. Each image is represented as a set of weights for corresponding vectors. This method performs very well when images are well aligned and have approximately the same pose. Changing lighting conditions, pose variations, scale variations and other dissimilarities between images decrease the recognition rate rapidly (Sirovich and Kirby, 1987).

Further methods using Neural networks or Hidden Markov Models (HMMs) are introduced. One of the best performing methods based on neural networks is presented in (Lawrence et al., 1997), while HMMs are successfully used in (Nefian and Hayes, 1998).

In the last couple of years, several successful approaches based on Gabor wavelets were introduced (Shen and Bai, 2006). The first method for face recognition based on Gabor wavelets was proposed by Lades (Lades et al., 1993). He used Gabor wavelets for image preprocessing. Some approaches (Shen, 2005) also combine the preprocessing with Gabor wavelets with well established methods such as Eigenfaces, Fisherfaces etc.

One successful approach described in (Wiskott et al., 1999; Bolme, 2003) is Elastic Bunch Graph Matching (EBGM). Gabor wavelet convolutions (Jets) in specified positions (Landmarks) are used for image representation. Another Gabor wavelet based method was proposed in (Kepenekci, 2001). This method differs from the previous one, that it determines the landmarks automatically and their number is thus not constant.

3 CONFIDENCE MEASURE FOR AUTOMATIC FACE RECOGNITION

As in many other works (Lleida and Rose, 1996; Jiang, 2005), our first confidence measure for automatic face recognition is an estimate of the *a posteriori* class probability. The output of our classifier is $P(F|C)$, where C is the recognized face class and F represents the face features. The likelihoods $P(F|C)$ are normalized to compute the *a posteriori* class probabilities as follows:

$$P(C|F) = \frac{P(F|C).P(C)}{\sum_{I \in \mathcal{FIM}} P(F|I).P(I)} \quad (1)$$

\mathcal{FIM} represents the set of all faces and $P(C)$ denotes the *prior* probability of the face class C .

In the first version of our algorithm, called *Absolute confidence value* method, only the faces \hat{C} so that

$$\hat{C} = \operatorname{argmax}_C (P(C|F)) \quad (2)$$

$$P(\hat{C}|F) > T \quad (3)$$

are considered as recognized correctly.

In the second version of our approach, called *Relative confidence value* method, the difference between the *best* hypothesis and the *second best* one is computed by the following equation:

$$P\Delta = P(\hat{C}|F) - \max_{C \neq \hat{C}} (P(C|F)) \quad (4)$$

Only the faces with $P\Delta > T$ are accepted. This second approach aims at identifying the faces that "dominate" all the other candidates, which is not always well captured by the first method.

T is in the both cases an acceptance threshold and its optimal value is found experimentally.

4 EXPERIMENTS

4.1 ČTK Corpus

All experiments are evaluated on the previously created ČTK corpus. This corpus is composed of the images of individuals in uncontrolled environment that were randomly selected from the large ČTK database. All images were taken during a long time period (20 years or more). They were automatically resized to the size 92×120 pixel and transformed to grayscale. The resulting corpus contains pictures of 70 individuals, at least 8 images for each person. Note that orientation, lighting conditions and background of images

differ significantly. A correct face recognition using this dataset is thus very difficult.

Figure 1 shows one example from this corpus. This corpus is available for free for research purpose upon request to the authors.



Figure 1: Examples of the ČTK face corpus.

4.2 Experimental Setup

One modification of the Kepenekci algorithm consists of the use more training examples instead of one training image per person during the face modeling step. Therefore, our confidence measure methods are evaluated with the different number of the training examples (between 1 and 7). Only the most important “border” cases (with 1 and 7 training examples) are reported in this paper.

We used the threshold values $T \in [0; 1]$ when 0 is the case without any confidence measure (all recognized faces are accepted). The cases when the face recognition rate remains constant are not reported in any figure.

ČTK requires a guarantee that a high face recognition rate (it means 95% or higher) is achieved. However, it is not necessary to identify all individuals.

4.3 Absolute Confidence Value

Figure 2 plots the face recognition rate and number of the classified faces when one training example is used. The result obtained without any confidence measure (or equivalently for $T = 0$) is only 15%. However, when the confidence measure is used, the best obtained score is 100%. This result is constrained by the fact, that only few examples is accepted (about 2%) which is unfortunately not acceptable by the target application (see Section 1).

Figure 3 shows the face recognition rate and number of the classified faces when all seven training examples are used. The result obtained without any confidence measure is 52%. This value confirms the good performance of our modifications of the Kepenekci approach. However this score is still not high enough

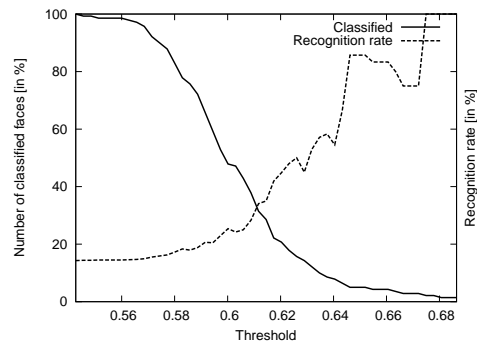


Figure 2: Absolute Confidence Value method: Face recognition rates and numbers of the classified faces when one training example is used ($T \in [0; 1]$).

for our target application. When obtained recognition rate is 95%, we filter out about 90% faces.

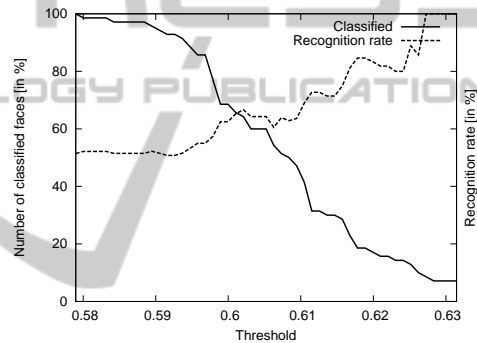


Figure 3: Absolute Confidence Value method: Face recognition rates and numbers of the identified faces when all seven training examples are used ($T \in [0; 1]$).

4.4 Relative Confidence Value

Figure 4 shows the face recognition accuracy and number of the classified faces in relation to the threshold T (one training example is used). This figure shows that this method performs better than the previous one. However, the number (about 5% with the 95% recognition rate) of the accepted examples is still too low.

Figure 5 plots the face recognition accuracy and number of the classified faces in relation to the threshold T when seven training examples are used. When obtained recognition rate is 95%, we accept about 35% faces. This score is high enough for our application.

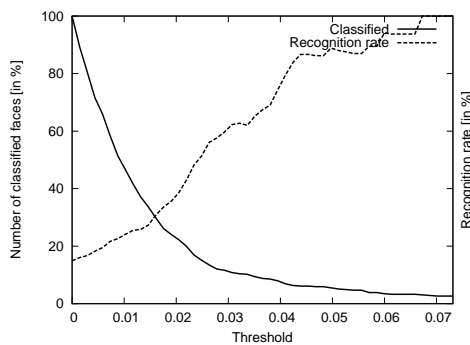


Figure 4: Relative Confidence Value method: Face recognition rates and numbers of the classified faces when one training example is used ($T \in [0; 1]$).

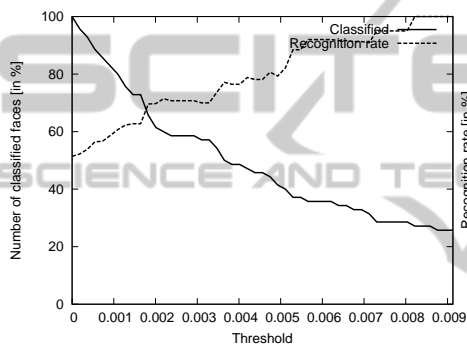


Figure 5: Relative Confidence Value method: Face recognition rates and numbers of the classified faces when all seven training examples are used ($T \in [0; 1]$).

5 CONCLUSIONS AND PERSPECTIVES

The main contribution of this work is the proposition and evaluation of two confidence measure techniques as a post-processing of the automatic face recognition task. We suggest using a confidence measure due to the relatively low face recognition rate on the ČTK corpus. The experiments show that around 10% of the images is classified with nearly 100% accuracy by the *Absolute confidence value* approach. We further show, that the second proposed approach, *Relative confidence value* method, is more suitable for the practical use. Around 30% of images is classified with an accuracy close to 100% by this method.

The perspectives of this work are numerous, including evaluation of the methods on the larger real-world corpora (the faces are not well aligned and the lighting conditions differ), the development of more sophisticated confidence measures or in adding the pre-processing step, i.e. eye localization and the sub-

sequent image rotation and aligning.

ACKNOWLEDGEMENTS

This work has been partly supported by ČTK and by the UWB grant SGS-2010-028 Advanced Computer and Information Systems. We also would like to thank ČTK for providing the photographic data.

REFERENCES

- Bledsoe, W. W. (1966). Man-machine facial recognition. Technical report, Panoramic Research Inc., Palo Alto, CA.
- Bolme, D. S. (2003). *Elastic Bunch Graph Matching*. PhD thesis, Colorado State University.
- Jiang, H. (2005). Confidence measures for speech recognition: a survey. *Speech Communication*, 45(4):455–470.
- Kanade, T. (1977). *Computer recognition of human faces*. Birkhauser Verlag.
- Kepenekci, B. (2001). *Face Recognition Using Gabor Wavelet Transform*. PhD thesis, The Middle East Technical University.
- Lades, M., Vorbrüggen, J. C., Buhmann, J., Lange, J., and von der Malsburg, C. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions On Computers*.
- Lawrence, S., Giles, S., Tsoi, A., and Back, A. (1997). Face recognition: A convolutional neural network approach. *IEEE Trans. on Neural Networks*.
- Lleida, E. and Rose, R. C. (1996). Likelihood Ratio Decoding and Confidence Measures for Continuous Speech Recognition. In *ICSLP'96*, volume 1, pages 478–481, Philadelphia, USA.
- Nefian, A. V. and Hayes, M. H. (1998). Hidden markov models for face recognition. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*.
- Shen, L. (2005). *Recognizing Faces - An Approach Based on Gabor Wavelets*. PhD thesis, University of Nottingham.
- Shen, L. and Bai, L. (2006). A review on gabor wavelets for face recognition. *Pattern Analysis & Applications*.
- Sirovich, L. and Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America*, 4.
- Turk, M. A. and Pentland, A. P. (1991). Face recognition using eigenfaces. In *IEEE Computer Society Conference on In Computer Vision and Pattern Recognition*. Computer Vision and Pattern Recognition.
- Wiskott, L., Fellous, J.-M., Krüger, N., and von der Malsburg, C. (1999). Face recognition by elastic bunch graph matching. *Intelligent Biometric Techniques in Fingerprint and Face Recognition*.