Automatic Frequency Band Selection for Illumination Robust Face Recognition

Hazım Kemal Ekenel, Rainer Stiefelhagen
Institute of Anthropomatics, Karlsruhe Institute of Technology, Germany
{ekenel,rainer.stiefelhagen}@kit.edu

***Abstract*—Varying illumination conditions cause a dramatic change in facial appearance that leads to a significant drop in face recognition algorithms' performance. In this paper, to overcome this problem, we utilize an automatic frequency band selection scheme. The proposed approach is incorporated to a local appearance-based face recognition algorithm, which employs discrete cosine transform (DCT) for processing local facial regions. From the extracted DCT coefficients, the approach determines to the ones that should be used for classification. Extensive experiments conducted on the Yale face database B and the extended Yale face database B have shown that benefiting from frequency information provides robust face recognition under changing illumination conditions.**

***Keywords*-Illumination robust face recognition, local appearance representation, discrete cosine transform, frequency band selection**

I. INTRODUCTION

Face recognition under varying illumination conditions is known to be one of the most difficult problems of face recognition research [1]. The variation in the facial appearance of an individual caused by illumination changes can dominate the one due to identity differences [2]. Therefore, face recognition algorithms have to consider this issue and they should be able to recognize the person, although his or her face looks more like a different person's face illuminated the same way, than the same person's face under different illumination.

Face recognition under varying lighting has attracted significant attention and there have been many solutions proposed for this problem to provide illumination robust face recognition [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]. These solutions can be classified as: invariant features, canonical forms, and variation modeling [3]. In the first approach, features insensitive to illumination variations are searched for [2]. The second approach tries to remove the illumination variation either by an image transformation or by synthesizing a new image [3], [4], [5], [6]. Finally in the third approach, illumination variation is learned and modeled in a suitable subspace [7], [8], [9], [10], [11]. Besides these solutions, in [12] near-infrared lighting is proposed to have illumination invariant capture conditions.

In this study, we propose an automatic frequency band selection approach to handle facial appearance variations caused by changing illumination conditions. The employed face recognition algorithm is based on the one presented in [13], which uses discrete cosine transform (DCT) coefficients to represent local facial regions. The approach automatically determines which frequency band to rely on for the classification task. In the study, we first analyzed the effect of using different frequency bands for face recognition and assess their classification performance under different conditions. We observed that there is no unique frequency band that can handle all the facial appearance variations. For different conditions, different frequency bands are found to be useful. Following this observation, we develop a face recognition system that adapts itself automatically to the environmental conditions by utilizing a multi-band classification scheme, in which the frequency band that is most confident about its classification output is chosen to be the most reliable one to perform classification. Experimental results obtained on the Yale face database B and extended Yale face database B show that the proposed approach is able to handle illumination variations successfully.

The organization of the paper is as follows. In Section II, used local appearance-based face recognition algorithm is described. The effect of frequency bands on classification performance is assessed in Section III. In Section IV, automatic frequency band selection scheme is explained. Experimental results are presented and discussed in Section V. Finally, in Section VI, conclusions are given.

II. LOCAL APPEARANCE-BASED FACE RECOGNITION

A local appearance-based face recognition algorithm is used for face recognition. It is a generic face recognition approach that has been found to be robust against expression, illumination, and occlusion variations as well as real-world conditions [13]. The algorithm has been evaluated on several benchmark face databases and found to be significantly superior to other generic face recognition algorithms [13]. In addition, it achieved the best recognition rates in the CLEAR 2007 evaluations [14]. The algorithm uses discrete cosine transform for local appearance representation. There are several advantages of using the DCT. Its data independent bases make it very practical to use. There is no need to prepare a representative set of training data to compute a subspace. In addition, it provides frequency information, which is very useful for handling changes in facial appearance. It also facilitates fast feature extraction.

In the proposed approach, a detected and registered face image is divided into non-overlapping blocks of 8×8 pixels size. Afterwards, on each block, the DCT is performed. The obtained DCT coefficients are ordered using zig-zag scanning. From the ordered coefficients, according to a feature selection strategy, M of them are selected and normalized resulting in an M-dimensional local feature vector. Finally, the DCT coefficients extracted from each block are concatenated to construct the overall feature vector. Classification is done using a nearest neighbor classifier with L1 norm as the distance metric. For details of the algorithm please see [13].

III. Analysis of Frequency Bands

It is known that different frequency bands play different roles depending on the classification task [15]. Therefore, an important aspect in local appearance-based face recognition using the DCT is the selection of frequency content to be used for classification. Each DCT basis has a different response which goes from coarser to finer as the basis index increases. The bases' outputs depict how strong a specific basis pattern is observed in the corresponding block. The low frequency coefficients represent most of the input block's energy, whereas the higher frequency coefficients correspond to finer details. However, neither conserving more energy nor having finer details guarantees better discrimination. Depending on the identification task, the required frequency band may change. In order to analyze the effect of feature dimensionality and frequency content on face recognition performance simultaneously, we employed a sliding window scheme, in which windows with varying sizes were moved from the beginning to the end of the ordered DCT coefficients. The coefficients obtained this way were used to derive the local feature vectors. From each block, the same frequency band is utilized for feature extraction. A separate classification is performed for each frequency band.

Two experiments have been conducted on the *Face Recognition Grand Challenge* (FRGC) database [16]. One, *FRGC1*, with face images collected under controlled studio settings and one, *FRGC4*, with face images collected under uncontrolled conditions, such as in hallways or outdoors. For both cases, we selected 120 individuals from the database, who have at least ten images both in fall 2003 and spring 2004 recordings. We used the images from fall 2003 for training and the ones from spring 2004 for testing. Sample images from the data sets are shown in Figure 1.

The results of the experiments can be seen from Figure 2 and Figure 3. In the figures, the x-axes show how many DCT coefficients are removed from the beginning, while y-axes show the local feature dimension. The number of possible shifts depends on the dimensionality of the local feature vector. For example, when the local feature dimension is two, there are 63 possible shifts, and when 63-dimensional local feature is used only two shifts are possible. The upper diagonals in the figures are padded with zeros, since at that region there exists no local feature dimension and shift combination. Dark red color indicates high correct recognition rates, whereas dark blue color corresponds to low correct recognition rates. It can be observed from Figure 2 that, in the case of controlled settings, low dimensional feature vectors that contain low frequency content are beneficial for face recognition. On the other hand, as can be seen from Figure 3, higher dimensional feature vectors that contain higher frequency content are required to achieve high correct recognition rates under uncontrolled conditions. This outcome justifies that different frequency bands are useful for handling different conditions.

Figure 1. Sample images from the FRGC database. First row, controlled settings. Second row, uncontrolled conditions

IV. Automatic Frequency Band Selection

As shown in the previous section, different types of variations can be handled by different frequency bands. Therefore, an automatic frequency band selection scheme is employed for the local appearance-based face recognition

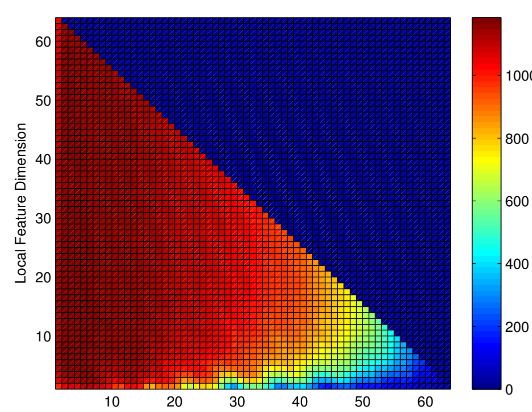

Figure 2. Comparison of different frequency bands on *FRGC1*

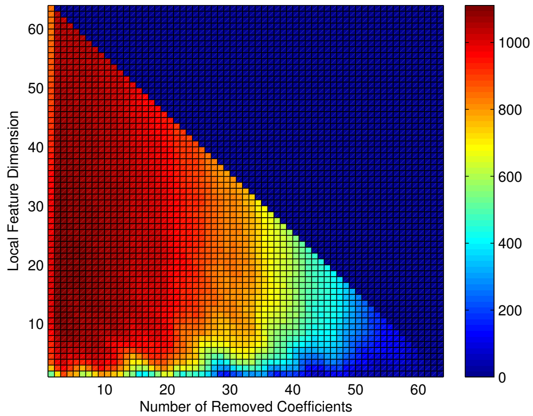

Figure 3. Comparison of different frequency bands on *FRGC4*

Figure 4. Sample images from the Yale face database. First row: Sample training images. Other rows, from top to bottom: Sample images from subsets 2, 3, 4 and 5, respectively

algorithm. In the utilized scheme, the classification is done using multiple frequency bands, that is, by selecting different DCT coefficients with a sliding window of size M and performing classification using the feature vectors extracted from each frequency band. In order to save processing time, we used a fixed window size of $M = 10$. In addition, we only scan the DCT coefficients from lower diagonal, since most of the DCT coefficients from upper diagonal are zeros. The frequency band that provides the maximum separation between the closest two identity candidates is chosen to be the most reliable band, and its decision is used as the classification output. The separation between the closest two candidates are measured by calculating the ratio between the closest and the second closest identity candidates' distance values, which is shown to be a robust measure to assess the classification outputs' reliability [17]. Note that this approach can be seen as using *max rule* [18] to combine multiple classification outputs. Other classifier combination approaches, such as *sum rule* or *product rule* [18], can also be used for this task. However, *max rule* provides a faster combination scheme and it is found to perform as well as the *sum rule* and *product rule* on a validation set.

V. Experiments

The algorithm's robustness against varying illumination conditions is tested extensively using the Yale face database B [7] and the extended Yale face database B [10]. The extended Yale face database B contains 38 subjects under 64 different illumination conditions. The first subset that has close to frontal illumination is used for training. For testing, subsets 2, 3, 4 and 5 are used. These subsets contain 12, 12, 14, 19 images per person, and the experiments with them are labeled *ExtYale2*, *ExtYale3*, *ExtYale4*, and *ExtYale5*, respectively. With increasing subset number, the illumination variations become stronger. The Yale face database B [7] is also used in the experiments, which is a subset of the extended Yale face database B and contains only 10 subjects. The extended Yale face database B is recently released, and there are many studies already conducted on the Yale face database B. Therefore, in order to be able to compare the obtained results with the ones in the literature, both of them are used. Sample images can be seen in Figure 4.

The obtained correct recognition rates are presented in Table I. The results attained without automatic frequency selection are from the local appearance-based face recognition algorithm that uses only the low frequency band as in [19]. As can be noticed, the algorithm can already handle changing illumination conditions very well. Especially, when the illumination variation is not too strong, as in the case of *ExtYale2* and *ExtYale3*, it can achieve 100% correct classification rate. It can be observed that automatic frequency band selection contributes to the performance significantly when the illumination variation is strong. This shows that by using the appropriate frequency band, the algorithm is able to adapt itself automatically to the changing illumination conditions. Compared to the obtained results in the recent studies on the same databases, such as the ones in [6], [10], the proposed approach performs as well as or even better. Note that in [6], [10] prior illumination-related information is utilized. On the other hand, we did not use any illumination-specific information in the proposed approach. Therefore, it is more general. Moreover, the algorithm is very fast and able to work real-time.

VI. Conclusions

In this paper, we analyzed the effect of frequency bands on face recognition performance. Confirming the psychological studies on face perception, we found that different frequency bands are useful for handling different types of variations. We developed a classifier reliability based frequency band selection scheme to automatically determine to the frequency

Experiment	No Auto Freq. Sel.	Auto. Freq. Sel.
Yale2	100%	100%
Yale3	100%	100%
Yale4	95.6%	100%
Yale5	96.8%	100%
ExtYale2	100%	100%
ExtYale3	100%	100%
ExtYale4	93.1%	98.7%
ExtYale5	93.1%	99.0%

Table I
RESULTS OF AUTOMATIC FEATURE SELECTION EXPERIMENTS

band to be used for classification. Experimental results on the extended Yale face database B have shown that the proposed algorithm is robust against changing illumination conditions.

ACKNOWLEDGMENTS

This study is partially funded by the German Research Foundation (DFG) under SFB 588 —Humanoid Robots—, and by OSEO, French State agency for innovation, as part of the Quaero Programme.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[2] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. on PAMI*, vol. 19, no. 7, pp. 721–732, 1997.

[3] T. Sim and T. Kanade, "Combining models and exemplars for face recognition: An illumination example," in *Proc. of Workshop on Models versus Exemplars in Computer Vision*, 2001.

[4] A. Shashua and T. Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Trans. on PAMI*, vol. 23, no. 2, pp. 129–139, 2001.

[5] H. Wang, S. Z. Li, Y. Wang, and W. Zhang, "Illumination modeling and normalization for face recognition," in *Proc. of IEEE Intl. Workshop on Analysis and Modeling of Faces and Gestures*, 2003.

[6] T. Chen, Y. Wotao, S. Z. Xiang, D. Comaniciu, and T. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. on PAMI*, vol. 28, no. 9, pp. 1519–1524, 2006.

[7] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on PAMI*, vol. 23, no. 6, pp. 643–660, 2001.

[8] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear image analysis for face recognition," in *Proc. of Intl. Conf. on Pattern Recognition*, 2002, pp. 511–514.

[9] R. Gross, I. Matthews, and S. Baker, "Appearance-based face recognition and light-fields," *IEEE Trans. on PAMI*, vol. 26, no. 4, pp. 449–465, 2004.

[10] K. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. on PAMI*, vol. 27, no. 5, pp. 684–698, 2005.

[11] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs, "Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination invariant face recognition," *IEEE Trans. on PAMI*, vol. 29, no. 2, pp. 230–245, 2007.

[12] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition under near-infrared images," *IEEE Trans. on PAMI*, vol. 29, no. 4, pp. 627–639, 2007.

[13] H. K. Ekenel, "A robust face recognition algorithm for real-world applications," Ph.D. dissertation, Universität Karlsruhe (TH), Germany, 2009.

[14] R. Stiefelhagen and et al., "The CLEAR 2007 evaluation," in *Proc. of the Intl. Evaluation Workshops CLEAR 2007 and RT 2007*, 2007, pp. 3–34.

[15] J. Sergent, "Microgenesis of face perception," in *Aspects of Face Processing*, H. D. Ellis, M. A. Jeeves, F. Newcombe, and A. Young, Eds. Martinus Nijhoff, 1986, Chapter, pp. 17–33.

[16] P. J. Phillips and et al., "Overview of the Face Recognition Grand Challenge," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2005, pp. 947–954.

[17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[18] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. on PAMI*, vol. 20, no. 3, pp. 226–239, 1998.

[19] H. K. Ekenel and R. Stiefelhagen, "Analysis of local appearance-based face recognition: Effects of feature selection and feature normalization," in *Proc. of Computer Vision and Pattern Recognition Workshop*, 2006.