Discriminative Metric Learning with Convolutional Feature Descriptors for Age-Invariant Face Recognition and Verification

Wataru OHYAMA

Department of Information Systems, Faculty of Engineering, Saitama Institute of Technology

ohyama@sit.ac.jp

畳み込み特徴に対する識別計量学習を用いた経年不変顔認識

大山　　航

埼玉工業大学工学部情報システム学科

Abstract

Aging includes internal and external factors that cause variation in appearance of face and, consequently, it is a difficult problem to handle in person identification and verification using face images. In this paper, we propose a method for face recognition and verification that is robust against variation of facial appearance caused by aging. Our proposed method uses discriminative metric learning over convolutional feature descriptors extracted from frontal face images. The results of an experiments for performance evaluation on the FG-Net and CACD face aging datasets empirically clarify that the proposed method is effective for improving the performance of person identification and verification in the scenario where input face images contain appearance variation due to aging.

Key Words: face recognition, face verification, age invariant, discriminative metric learning, convolutional feature descriptors

Introduction

Recently, the social demands for protecting privacy and personal information have increased. To protect privacy and personal information from unauthorized attacks, authentication based on biometric information has gained much attention because it has inherent robustness against theft and falsification. Of the various modalities in biometric authentication systems, face recognition and authentication have the following advantages:

(1) A forgery is expected to be confirmed using an image which is saved previously.
(2) Unauthorized access is also expected to be prevented because unauthorized person would not like to be taken a picture of their face.

In the authentication viewpoint, face images have the disadvantage where its performance is easily affected by variation in appearance caused by changes of pose, illumination, and facial

expression (PIE). Traditionally, several studies for face recognition and verification that handled appearance variations due to change in PIE have been reported. Although these researches proposed some promising techniques, methods for the recognition and verification of faces images that contain appearance variation caused by aging are not mature. The age-invariant face recognition or verification systems are quite important to reduce frequency of updating registered face information and find persons who are missing for long time using surveillance cameras.

Developing a face recognition method which is robust against appearance variation caused by aging involves challenging problems. First, aging of the facial appearance is affected by not only internal genetic factors but several external factors that come from the living environment of an individual. Furthermore, whereas developing a face recognition system requires a large dataset of face images for classifier training, acquiring face images including several ages over the long term is not so easy.

Another problem for a dataset is that the image quality of old face images is usually degraded. Because spreading digital cameras to consumers started only approximately past 20 years, the digitization of old photographs requires the scanning of low-quality pictures. Old pictures printed on a paper contains a particle-like texture. When we scan these pictures, the texture also appears in the image. An age-invariant face recognition method is required to manage both facial appearance changes caused by subjects aging and low image quality.

The authors propose a face recognition method that is robust against facial appearance variation caused by aging. The proposed method matches two input face images and evaluates the similarity between the images in Euclidean space, which is robust against aging. Euclidean space is obtained by discriminative metric learning (DML) in which a nonlinear projection function decreases the distance between samples extracted from face images of the same person and increases the distance between other ones. In the proposed method, the projection function is obtained using a neural network.

The main contributions of this research are as follows:

1) A new method that uses a neural network-based DML combined with a convolutional feature extractor is proposed.
2) For DML, neural network-based optimization is implemented.
3) A comprehensive performance evaluation using two major open datasets, FG-Net and CACD, is performed to confirm the effectiveness of the proposed method.

While a similar concept of DML has been proposed by Hu et al.[1], their method was applied to conventional hand-crafted feature descriptors. In this paper we introduce the convolutional feature extractor and some modifications to their concept.

2. Previous Work

Face recognition has been the most active research topic since eigenfaces[2] were reported. Recently, the face recognition performance has been dramatically improved by the introduction of the convolutional neural network (CNN). Some methods, such as Facenet[3] and DeepID[4], demonstrate very high performance that outperforms even human skills.

Face recognition is considered as a special case of object recognition; however, it is distinguished from other object recognition problems in terms of appearance variations, such as PIE and aging. Many methods managing PIE variation have been proposed and have achieved high recognition performance. However, face recognition performance against aging variation is still required for improvement because of a lack of datasets and the low quality of images.

In our proposed method, we introduce DML and show that its effectiveness for age-invariant face recognition.

3. DML with convolutional feature descriptors

In this section, we provide an outline of the proposed recognition method for aging faces and detail of each process. The novelties of the method are a method that combines convolutional feature descriptors and DML, and its implementation using a deep neural network.

Fig. 1 shows the outline of the proposed method. The process of face image recognition or verification using the proposed method consists of the following four steps:

1) As preprocessing, the face region is extracted by a tight bounding box and image quality enhancement is applied.
2) A convolutional feature descriptor is extracted from input and registered face images using a pre-trained CNN.
3) The extracted convolutional feature descriptors are projected into low-dimensional Euclidean space using nonlinear projection function obtained by DML.
4) The dissimilarity between input and registered face images is calculated as the Euclidean distance between them in the transformed space.

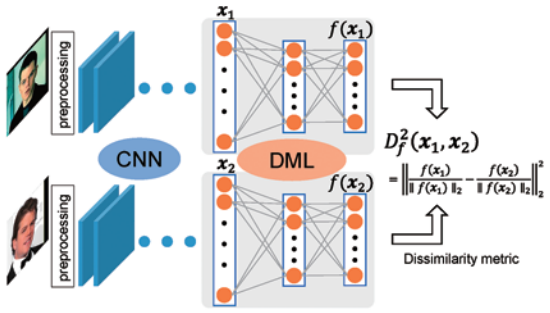

Fig. 1 Outline of the proposed age-invariant face recognition method. The method consists of preprocessing, extraction of convolutional feature descriptors, projection by a nonlinear function obtained by DML, and calculation of dissimilarity.

4. Evaluation experiments

To evaluate the effectiveness of the proposed DML method, we conducted evaluation experiments using two public facial image datasets that include appearance variation caused by aging.

The datasets used for performance evaluation in conventional research that proposes age-invariant face recognition sometimes contains bias regarding the race or age of a person. Consequently, evaluation and comparison using multiple datasets that have different characteristics are requested to perform a fair assessment. From this point of view, the authors conducted evaluation experiments, which evaluated and compared the performance of the proposed method with that of conventional methods using two major public datasets that contain facial appearance variation caused by aging.

We evaluated the performance improvement by introducing DML using two major open datasets: FG-Net and Cross-Age Celebrity Dataset (CACD)[5]. FG-Net contains 1,002 images from 82 subjects and 6-18 images per subject. The age variation of subjects is 0 to 69 years old. Each image is provided with annotations that consist of vertical and horizontal orientations and image quality. The CACD dataset consists of 163,446 images from 2,000 subjects. The age variation of subjects is 12 to 62 years old. All images were taken between 2004 and 2013. The CACD dataset also provides the verification subset (CACD-VS) for the evaluation of verification performance. CACD-VS consists of 4,000 pairs of face images of which 2,000 pairs are of the same subjects and 2,000 pairs are of different subjects.

5. Results and discussion

Following the experimental procedure in Li et al.[6], we evaluated the performance of face recognition based on leave-one-person-out (LOPO), where all images of one subject were excluded from the dataset as test subjects and remaining images were used for training.

Fig. 2 shows the rank performance of face recognition with and without DML. Face recognition without DML denotes the condition in which the pre-trained VGG16 network was naively applied on the FG-Net dataset. We observed that the Rank-1 recognition accuracy significantly improved from 83.93% to 91.32%.

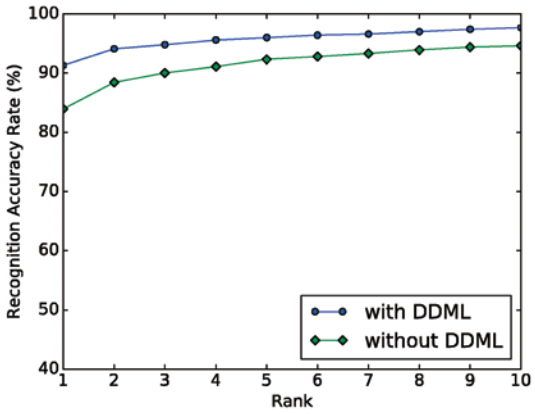

Fig. 2 Improvement of recognition accuracy by introducing DML for FG-Net.

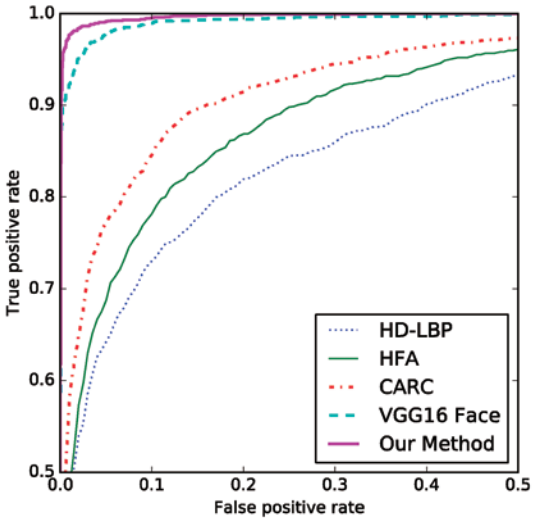

Fig. 3 Comparison of verification performance using the ROC curve among the proposed and conventional methods.

The DML is also applicable for verification tasks. We evaluated the verification performance using the CACD-VS dataset, which consists of 4,000 pairs of face images. We adopted 10-fold cross validation for performance evaluation. Fig. 3 shows the receiver operating characteristic (ROC) curves of each verification method. LF-CNN was excluded because no information was provided in the paper[7]. From the ROC curves, we observed that the verification performance effectively improved by introducing the DML.

Fig. 4 shows some examples which are improved by introducing DML. While each pair in Fig. 4 belongs to a same subject, it was verified as belonging to different ones before introducing DML. These failures are corrected by introducing DML and each pair is verified as belonging to the same subject. We can observe that image pairs having a large age gap are handled properly by the proposed method. Similarly, examples in Fig. 5 were originally verified as belonging to the same subjects but corrected by the proposed method as that they are from different subjects. We can see that the pairs of different persons in similar appearance age are correctly verified.

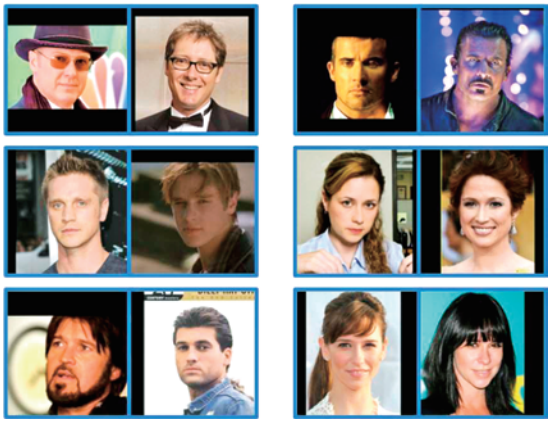

Fig. 4 Examples which are recovered by introducing DML (True Positive).

Fig. 5 Examples which are recovered by introducing DML (True Negative).

6. Conclusions

In this paper, we proposed DML for age-invariant face recognition and verification. The proposed DML method used a fully connected neural network to obtain an adaptive nonlinear

projection function for samples that included facial appearance change caused by aging.

The effectiveness of the proposed method was empirically evaluated using two major public datasets that contained aging. We confirmed that both recognition and verification performance were effectively improved by introducing DML. Particularly, in the face recognition experiment using the FG-Net database, the proposed method outperformed LF-CNN, which is one of state-of-the-art methods in this field.

By contrast, the verification performance of the proposed method was quite competitive but slightly lower than that of LF-CNN. We need to introduce fine-tuning and careful parameter setting for further performance improvement.

References

1) Hu, J., Lu, J., Tan, Y.P.: Discriminative deep metric learning for face verification in the wild. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. pp.1875–1882 (2014)
2) Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognitive Neuroscience **3**(1), 71–86 (Jan 1991)
3) Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: 2016 IEEE International Conference on Computer Vision and Pattern Recognition. pp.815–823. IEEE Computer Society (2015)
4) Sun, Y., Liang, D., Wang, X., Tang, X.: Deep ID3: Face recognition with very deep neural networks. Computing Research Repository (2015)
5) Chen, B.C., Chen, C.S., Hsu, W.H.: Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset. IEEE Transactions on Multimedia **17**(6), 804–815 (June 2015)
6) Li, Z., Park, U., Jain, A.K.: A discriminative model for age invariant face recognition. IEEE transactions on information forensics and security **6**(3), 1028–1037 (2011)
7) Wen, Y., Li, Z., Qiao, Y.: Latent factor guided convolutional neural networks for age-invariant face recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. pp.4893–4901 (2016)