Neurocomputing 144 (2014) 484–492

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Distance metric learning for recognizing low-resolution iris images

Jing Liu [a,*], Zhenan Sun [b], Tieniu Tan [b]

[a] Department of Automation, University of Science and Technology of China, Hefei 230026, PR China
[b] Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 9 February 2014
Received in revised form
17 April 2014
Accepted 2 May 2014
Communicated by X. Li
Available online 27 May 2014

Keywords:
Mahalanobis distance
Metric learning
Low-resolution iris recognition

ABSTRACT

Low-resolution (LR) iris images are inevitable, especially in the iris recognition systems under less constrained imaging conditions which are desirable to extend the applicability of iris biometrics. It is a challenging problem to match LR probe iris images with high-resolution (HR) ones captured at enrollment stage. This paper presents a heterogeneous metric learning algorithm which can favorably improve the accuracy of LR iris recognition. The basic idea of the method is to learn an appropriate distance metric to transform the heterogenous (LR vs. HR) iris matching results towards the desirable homogeneous (HR vs. HR) ones and then further enhance the separability between intra-class and inter-class matching samples. This learning procedure not only utilizes label and local information, but also fully exploits the sample correspondence and the ideal application scenario as the specific prior information. Two steps are included in the proposed method. Firstly, the ideal pairwise similarities are defined on the training set to faithfully achieve the basic idea above. Secondly, the Mahalanobis distance is learnt by minimizing the divergence between the matching results measured by the target Mahalanobis distance and the ideally defined matching results. Extensive experiments show that the proposed metric learning solution consistently outperforms state-of-the-art metric learning methods and can further enhance the performance of existing LR iris recognition approaches.

1. Introduction

Biometrics gains its popularity in both public and personal security applications in recent years [2]. Iris recognition [3] as one of the most reliable biometric identification methods has received great attention. Extremely high accuracy can be guaranteed when high-quality images are available in iris recognition. Hence, almost all of real-world iris recognition applications work in constrained environments and require the inconvenient cooperation of users. For example, the Unique Identification Authority of India (UIDAI) project has enrolled high-quality iris images of half a billion subjects using iris cameras with close operating range. Such constraints prevent wider applications of iris recognition in a user-friendly manner. Currently, massive facial images containing iris regions are available on internet, and surveillance cameras are widely deployed in commercial and law enforcement applications. And some advanced systems such as iris recognition at a distance and iris on the move [4] have been developed by Sarnoff and Aoptix trying to recognize iris images from 2 to 3 meters away. They all challenge the traditional iris recognition systems, and encourage to propose effective recognition algorithms working in less constrained environments. However, the captured images usually contain low-resolution (LR) iris regions due to such factors as the limited resolution of iris sensors and the large distance between subjects and iris sensors. For instance, the resolution of iris region is only about 100 pixels in iris images captured at 11 feet away by common acquisition systems. Since the pixel resolution is limited, some texture details are inevitably lost, as shown in Fig. 1. It degrades recognition accuracy and explicitly requires improvement of the existing recognition algorithms [3,5]. Apart from this, enrolled iris images are usually captured in controlled conditions with high resolution (HR). Therefore, it is more difficult to alleviate the degraded performance caused by the heterogeneous matching of LR probe and HR gallery iris images.

1.1. Related work

The heterogeneous matching of LR and HR iris images has received increasing attention recently [6–8]. A straightforward idea is to conduct super-resolution (SR) and reconstruct iris texture details from LR images at pixel, feature or code level. Higher resolution probe iris images or enhanced feature codes can be obtained after SR and then are fed into the traditional recognition procedure. In [6], a number of LR iris image frames are firstly

* Corresponding author.
E-mail address: liujing0@mail.ustc.edu.cn (J. Liu).

http://dx.doi.org/10.1016/j.neucom.2014.05.016
0925-2312/

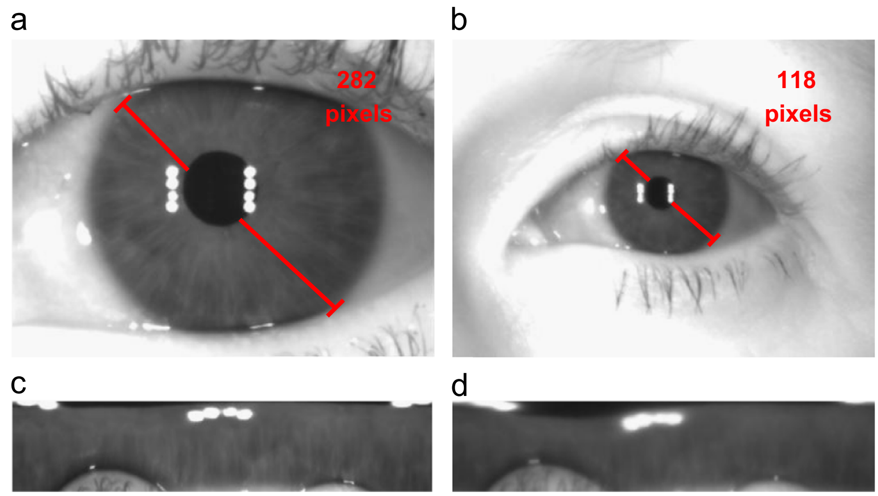

Fig. 1. Comparison of high-resolution (HR) and low-resolution (LR) iris images of the same subject (0162635L in Q-FIRE [1]) captured at 5 and 11 feet away. (a) A HR iris image (the diameter of iris ring is 282 pixels). (b) A LR iris image (the diameter of iris ring is 118 pixels). (c) The normalized version of the HR iris image. (d) The normalized version of the LR iris image.

localized and normalized, and then they are interpolated and fused into the final HR iris image according to the image qualities. Nguyen et al. [7] and Liu et al. [8] believe that a higher processing level of SR gets more relevant to recognition performance, thus their algorithms are performed on Gabor filtered images and iris binary codes respectively. These existing methods are generally found to be effective for LR iris recognition, however they require multiple perfectly aligned LR iris images which are not always available in real-world applications. Even worse, they cannot explicitly guarantee the improvement of recognition performance, because they [7,8] only focus on the relevance to recognition items and are not optimized with respect to machine perspective.

The basic idea behind the solutions above is actually to find a transformation between iris samples with different resolutions and use it to recover the informative patterns of LR iris images or feature codes. As we know, iris recognition decisions are made according to the similarities between iris images, i.e., the Hamming distances [3] between iris codes. Therefore, compared with finding the transformations between LR and HR samples for satisfied visual effects, it is more effective to learn an appropriate transformation of distance metric by optimizing objective functions for better separability of iris samples. Essentially, it refers to the task of metric learning (ML) [9] which is a well exploited research area in the literature. It aims at learning "a distance metric for the input space of data from a given collection of pair of similar/dissimilar points that preserves the distance relation among the training data" [10]. Depending on whether the label information is available, metric learning methods can be divided into two categories, i.e., unsupervised and supervised. Unsupervised metric learning methods [11,12] focus on the underlying manifold structures, and attempt to preserve the geometric relationship of the input data points. Supervised methods [9,13] utilize the label information, and try their best to achieve compactness (i.e., keeping all the data points in the same class close together) and separability (i.e., ensuring the data points from different classes far apart) of data points simultaneously for better classification performance [14]. Among them, learning a Mahalanobis distance [9,13] is one of the most representative approaches. Formally, the Mahalanobis distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as

$$d_{\mathbf{A}}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{A} (\mathbf{x}_i - \mathbf{x}_j)}, \quad (1)$$

where $\mathbf{A}$ is a symmetric positive semi-definite (p.s.d) matrix and actually the optimization objective of Mahalanobis distance learning. Since the Mahalanobis distance generalizes the traditional Euclidean distance by scaling and rotating the feature spaces, it can provide strong generalization ability. Furthermore, some advanced metric learning methods [14] realize the problem of multi-modal data distributions and exploit both globality and locality information for geometry preservation. In this way, two identically labeled data points in different collections cannot be pulled together but are still separated based on the learnt metric distances.

The promising applications of metric learning in heterogeneous biometric recognition are shown by [15,16]. Appropriately determined distance metrics are used in face and gait recognition. Biswas et al. [15] find two mappings for HR and LR samples, and use them to calculate the latent HR vs. HR distance according to the given HR vsp. LR distance. Iterative optimization procedures are conducted to maximize the correspondence between LR and HR face images of the same subject. In [16], they also learn two transformation functions for probe and gallery samples by a generalized eigne-decomposition to find a unified representation between query and registered samples. The above methods only concern a single aspect in solving heterogeneous biometric recognition, i.e., [15] only ensures the correspondence between each pair of HR and LR samples while [16] only focuses on the compactness of data points according to the local geometry. However, it is promising and necessary to propose a comprehensive solution which systematically studies all the specialities of heterogeneous biometric recognition.

1.2. Contributions

In this paper, we propose a novel Mahalanobis distance learning algorithm to improve classification accuracy on multi-modal data sets and apply it as the means of recognizing low-resolution probe iris images from high-resolution gallery templates. Among different kinds of solutions to LR iris recognition, learning an appropriate transformation of distance metric is more effective in terms of recognition accuracy improvement, since the distance metric is obtained by optimizing the objective function for better separability of iris matching samples. Compared with common metric learning approaches, the proposed algorithm realizes the problem of multi-modal data distributions and makes contributions to alleviate it by exploiting the potential information among training data set. Specifically, in LR iris recognition, the multi-modal data distributions are caused by the matchings of both HR