Pattern Recognition Letters 33 (2012) 182–190

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Ear recognition based on local information fusion

Li Yuan *, Zhi chun Mu

School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 3 August 2010
Available online 20 October 2011
Communicate by M.S. Nixon

Keywords:
Ear recognition
Partial occlusion
Neighborhood preserving embedding
Sub-classifier fusion

ABSTRACT

Ears have rich structural features that are almost invariant with increasing age and facial expression variations. Therefore ear recognition has become an effective and appealing approach to non-contact biometric recognition. This paper gives an up-to date review of research works on ear recognition. Current 2D ear recognition approaches achieve good performance in constrained environments. However the recognition performance degrades severely under pose, lighting and occlusion. This paper proposes a 2D ear recognition approach based on local information fusion to deal with ear recognition under partial occlusion. Firstly, the whole 2D image is separated to sub-windows. Then, Neighborhood Preserving Embedding is used for feature extraction on each sub-window, and we select the most discriminative sub-windows according to the recognition rate. Each sub-window corresponds to a sub-classifier. Thirdly, a sub-classifier fusion approach is used for recognition with partially occluded images. Experimental results on the USTB ear dataset and UND dataset have illustrated that using only few sub-windows we can represent the most meaningful region of the ear, and the multi-classifier model gets higher recognition rate than using the whole image for recognition.

1. Introduction

As an emerging biometrics technology, ear recognition is attracting more and more attention in biometrics recognition. Human ears offer some distinct advantages over other biometric modalities: they have a wealthy of structural features that are permanent with increasing age from about 8–70 years old, and they are not affected by the expression variations (Burge and Burger, 2000). Ear image is smaller under the same resolution, which can be favorable in some situations, such as the audio-visual person authentication using speech and ear images for mobile phone usage. According to the evaluations in Choraś (2006), the ear is a kind of highly accepted biometrics, and subjects to be identified feel more comfortable with ear images enrollment compared to face images enrollment. Ear recognition is user-friendly and can be used in non-intrusive recognition and surveillance scenarios.

Ears have played a significant role in forensic science for many years (Nixon et al., 2010), especially in the United States, where an ear classification system based on manual measurements has been developed by Iannarelli, and has been in use for more than 40 years (Iannarelli, 1989). The United States Immigration and Naturalization Service (INS) has a form giving specifications for the photograph that indicate that the right ear should be visible (INS Form M-378 (6-92)). During crime scene investigation or airplane crashes, earmarks are often used for identification (Alberink and Ruifrok, 2007; Choraś, 2007). The history of using ear images or ear prints shows their potential value for human identification applications such as access control, security monitoring and video surveillance (Hurley et al., 2008).

An ear recognition system usually involves ear detection, feature extraction and ear recognition/verification. As the first stage of the ear recognition system, real-time ear detection and tracking is a key component for the whole system. It mainly focuses on detecting and tracking human ear from the input video images of a scene and then returning the location and extent of each ear in the image if one or more ears are present. The next step is to represent the ear by appropriate features and design effective classifier. Most of the present ear recognition papers are focused on this step. But in real scenarios, the performance of ear recognition will be affected by illumination variation, pose variation and partial occlusion.

In this paper, we deal with ear recognition under partial occlusion. The main contribution of this paper is to propose a fusion method for component based ear recognition based on our previous work (Yuan et al., 2010). In the previous work, we proposed a multi-classifier fusion scheme for ear recognition with partially occluded images. Top discriminative sub-classifiers ware combined on the decision level for ear recognition. This paper has made the following further improvements compared with our previous work: (1) we add the ear detection step to form a more complete ear recognition system, and give more comparisons with other up-to-date references on ear recognition under partial occlusion; (2) recent advances in ear recognition have been reviewed in this

* Corresponding author. Tel.: +86 10 62334995.
E-mail address: yuanli64@hotmail.com (L. Yuan).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.09.041

Table 1
Summary on ear image database.

Database	Acquisition equipment	Dataset	Image type	Size	Pose variation	Other variation
UND (http://www.nd.edu/~cvrl/)	Minolta Vivid 910 camera	Collection E	2D	114 Subjects, 464 images	–	–
		Collection F	3D + 2D	302 Subjects, 942 3D + corresponding 2D ear images	–	Time -lapse. Now a subset of collection J2
		Collection J2	3D + 2D	415 Subjects, 1800 3D + corresponding 2D ear images	24 Subjects with images taken at four different viewpoints	Time-lapse, Illumination variation
UCR (Chen and Bhanu, 2009)	Minolta Vivid 300 camera	–	3D + 2D	155 Subjects, 902 shots	Frontal, left, and right within ±35°	Partial occlusion of 18 subjects
USTB (http://www.ustb.edu.cn/resb/)	CCD camera	Dataset1	2D	60 Subjects, 3 ear images per subject	–	–
		Dataset2	2D	77 Subjects, 4 ear images per subject	Frontal, left, and right within ±35°	Illumination variation
		Dataset3	2D	79 Subjects, 40 profile images per subject	Frontal, left (+), and right (−) rotation of ±5°, ±10°, ±15°, ±20°, ±25°, ±30°, ±35°, ±40°, ±45°	Partial occlusion of 24subjects
		Dataset4	2D	500 Subjects, 85 profile images per subject	Five poses (frontal, tilt ±30°, yaw ±30°), 17 images at 10° interval	–

paper in Section 2, including commonly used ear image databases, ear recognition in 3D, ear recognition in 2D and some challenging problems in ear recognition; (3) score level fusion is applied for ear recognition/verification. Extensive performance evaluations have been carried out on larger ear image databases: USTB dataset3 (79 subjects) and UND ear dataset (150 subjects). In ear recognition with un-occluded ear images, rank-1 recognition rate for identification scenario and EER for verification scenario are analyzed for both USTB and UND datasets.

This paper is organized as follows: Section 2 presents the related research work on ear recognition. In Section 3 the sub-classifier fusion approach is proposed for recognition with partially occluded images. Then we conclude the paper in Section 4 with experimental results on related ear image database.

2. Related work

Current ear recognition approaches have exploited how to use 2D ear image and 3D ear model for human identification. At present, although 3D ear recognition performs well in illumination variation or pose variation, it needs expensive computation and special equipments (Zeng et al., 2009), most of the resent works on ear recognition are focused on 2D images because using 2D images is more consistent with deployment in surveillance or other planar image scenarios (Arbab-Zavar and Nixon, 2011). In this section, we will start with intruding the frequently used ear image databases, and then review recent advances in ear recognition in 3D and 2D respectively, and propose some challenging problems in ear recognition.

2.1. Ear image databases

At present, the ear images used in most works mainly come from three specific databases: UND database collected by University of Norte Dame (http://www.nd.edu/~cvrl/CVRL/CVRL_Home_Page.html), UCR database collect by University of California at Riverside (Chen and Bhanu, 2009) and USTB database collected by University of Science and Technology Beijing (http://www.ustb.edu.cn/resb/), more details about these databases are shown in Table 1.

2.2. Ear recognition in 3D

Ear recognition in 3D gets good performance in illumination variation or orientation variation. Chen and Bhanu (2007) used local surface patch (LSP) to represent the ear and a modified Iterative Closest Point (ICP) algorithm to improve the matching. The root mean square (RMS) registration error was used as matching criterion. In their later work (Chen and Bhanu, 2009), due to the high dimensionality of LSP feature space, FastMap algorithm was used for dimension reduction. By searching the nearest neighbors in low dimensions, the similarity between a model-test pair was computed using the LSP features. The similarities for all model-test pairs were ranked using SVM to generate a short list of candidate models for verification. The verification was performed by aligning a model with the test object with ICP algorithm. On the UND collection *F*, the rank-1 recognition rate was 96.7%, and EER was 1.8% for verification. Yan and Bowyer (2007) presented an automatic ear biometric system using 2D and 3D information. After the ear pit detection, an active contour algorithm using both color and depth information was applied to outline the visible ear region. The recognition subsystem used an ICP-based approach for 3D shape matching. On UND collection J2, this method achieved a rank-one recognition rate of 97.8% for identification and an EER of 1.2% for verification. However the computation involved in ICP algorithm is considerably intensive, which makes the recognition sluggish. Besides, ICP algorithm is easy to fall into local minimum.

The above research is based on the range images, which need to be taken by a special 3D scanner. The data acquisition process requires the user to maintain a relatively still pose for several seconds. In non-intrusion recognition scenario, surveillance video or cameras are more popular to get ear images, which makes it necessary to use 2D ear recognition from application point of view. So 3D ear recognition based on 2D images becomes another option for ear recognition, which includes two problems: 3D ear reconstruction and 3D ear recognition. So far only few publications have reported the relevant works on 3D ear reconstruction based on 2D image. Liu et al. (2006) developed a SFM (Shape From Motion) algorithm based 3D ear reconstruction method, which was semi-automatic because the feature points are selected in a way of human–computer interaction. Cadavid and Abdel-Mottaleb (2007, 2008) proposed a SFS (structure from shading) algorithm