ARTICLE IN PRESS
JID: NEUCOM [m5G;April 8, 2017;21:7]
Neurocomputing 000 (2017) 1–14
Contents lists available at ScienceDirect
Neurocomputing
journal homepage: www.elsevier.com/locate/neucom

Ear recognition: More than a survey

Žiga Emeršič[a,*], Vitomir Štruc[b], Peter Peer[a]

[a] Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, Ljubljana 1000, Slovenia
[b] Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, Ljubljana 1000, Slovenia

ARTICLE INFO

Article history:
Received 7 March 2016
Revised 20 July 2016
Accepted 23 August 2016
Available online xxx

Keywords:
Biometry
Dataset
In-the-wild
Unconstrained image
Descriptor-based method
Open-source toolbox
Ear recognition

ABSTRACT

Automatic identity recognition from ear images represents an active field of research within the biometric community. The ability to capture ear images from a distance and in a covert manner makes the technology an appealing choice for surveillance and security applications as well as other application domains. Significant contributions have been made in the field over recent years, but open research problems still remain and hinder a wider (commercial) deployment of the technology. This paper presents an overview of the field of automatic ear recognition (from 2D images) and focuses specifically on the most recent, descriptor-based methods proposed in this area. Open challenges are discussed and potential research directions are outlined with the goal of providing the reader with a point of reference for issues worth examining in the future. In addition to a comprehensive review on ear recognition technology, the paper also introduces a new, fully unconstrained dataset of ear images gathered from the web and a toolbox implementing several state-of-the-art techniques for ear recognition. The dataset and toolbox are meant to address some of the open issues in the field and are made publicly available to the research community.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Ear images used in automatic ear recognition systems can typically be extracted from profile head shots or video footage. The acquisition procedure is *contactless* and *nonintrusive* and also *does not depend on the cooperativeness* of the person one is trying to recognize. In this regard ear recognition technology shares similarities with other image-based biometric modalities. Another appealing property of ear biometrics is its *distinctiveness* [1]. Recent studies even empirically validated existing conjectures that certain features of the ear are distinct for identical twins [2]. This fact has significant implications for security related applications and puts ear images on par with epigenetic biometric modalities, such as the iris. Ear images can also serve as *supplements* for other biometric modalities in automatic recognition systems and provide identity cues when other information is unreliable or even unavailable. In surveillance applications, for example, where face recognition technology may struggle with profile faces, the ear can serve as a source of information on the identity of people in the surveillance footage. The importance and potential value of ear recognition technology for multi-modal biometric systems is also evidenced by the number of research studies on this topic, e.g. [3–7].

Today, ear recognition represents an active research area, for which new techniques are developed on a regular basis and several datasets needed for training and testing of the technology are publicly available, e.g., [8,9]. Nevertheless, despite the research efforts directed at ear biometrics, to the best of our knowledge, there is only one commercial system currently available on the marked that exploits ear biometrics for recognition, i.e., the Helix from Descartes Biometrics [10]. We conjecture that the limited availability of the commercial ear recognition technology is a consequence of the open challenges that by today have still not been appropriately addressed. This paper is an attempt to meet some of these challenges and provide the community with a point of reference as well as with new research tools that can be used to further advance the field.

1.1. Contributions and paper organization

Prior surveys related to ear recognition, such as [11–16], provide well written and well structured reviews of the field. In this paper we contribute to these surveys by discussing recent 2D ear recognition techniques proposed until the end of 2015. We pay special attention to descriptor-based approaches that are currently considered state-of-the-art in 2D ear recognition. We present comparative experiments with the new dataset and toolbox to establish an

* Corresponding author.
E-mail addresses: ziga.emersic@fri.uni-lj.si (Ž. Emeršič), vitomir.struc@fe.uni-lj.si (V. Štruc), peter.peer@fri.uni-lj.si (P. Peer).

http://dx.doi.org/10.1016/j.neucom.2016.08.139
0925-2312/© 2017 Elsevier B.V. All rights reserved.

independent ranking of the state-of-the-art techniques and show that there is significant room for improvement and that ear recognition is far from being solved.

We make the following contributions in this paper:

- **Survey:** We present a comprehensive survey on ear recognition, which is meant to provide researchers in this field with a recent and up-to-date overview of the state-of-technology. We introduce a taxonomy of the existing 2D ear recognition approaches, discuss the characteristics of the technology and review the existing state-of-the-art. Most importantly, we elaborate on the open problems and challenges faced by the technology.
- **Dataset:** We make a new dataset of ear images available to the research community. The dataset, named Annotated Web Ears (AWE), contains images collected from the web and is to the best of our knowledge the first dataset for ear recognition gathered “in the wild”. The images of the AWE dataset contain a high degree of variability and present a challenging problem to the existing technology, as shown in the experimental section.
- **Toolbox:** We introduce an open source (Matlab) toolbox, i.e., the AWE toolbox, for research in ear recognition. The toolbox implements a number of state-of-the-art feature extraction techniques as well as other important steps in the processing pipeline of ear recognition systems. It contains tools for generating performance metrics and graphs and allows for transparent and reproducible research in ear recognition. The toolbox is available from: http://awe.fri.uni-lj.si.
- **Reproducible evaluation:** We conduct a comparative evaluation of several state-of-the-art methods on a number of popular ear datasets using consistent experimental protocols, which enables direct comparisons of the state-of-the-art in ear recognition. All experiments are conducted with our AWE toolbox making all presented results reproducible.

The rest of the paper is structured as follows. In Section 2, we present the background and basic terminology related to ear recognition. In Section 3, existing ear datasets are discussed and compared on the basis of some common criteria of interest. Section 4 introduces our new AWE dataset and the accompanying AWE toolbox. Comparative experiments and results with the new toolbox are presented in Section 5. In Section 6 open problems and promising future research directions are examined. The paper concludes with some final comments in Section 7.

2. Ear recognition essentials

2.1. Ear structure

The human ear develops early during pregnancy and is already fully formed by the time of birth. Due to its role as the human hearing organ, the ear has a characteristic structure that is (for the most part) shared across the population. The appearance of the outer ear is defined by the shapes of the tragus, the antitragus, the helix, the antihelix, the incisura, the lope and other important structural parts as shown in Fig. 2. These anatomical cartilage formations differ in shape, appearance and relative positions from person to person and can, therefore, be exploited for identity recognition.

In general, the left and right ears of a person are similar to such an extent that makes matching the right ear to the left ear (and vice versa) with automatic techniques perform significantly better than chance. Yan and Bowyer [17], for example, reported a recognition performance of around 90% in cross-ear matching experiments. They observed that for most people the left and right ears are at least close to bilateral symmetric, though the shape of the two ears is different for some [17]. Similar findings were also reported by Abaza and Ross in [18].

The literature suggests that the size of the ear changes through time [15,16,19,20]. Longitudinal studies from India [21] and Europe [20,22,23] have found that the length of the ear increases significantly with age for men and women, while the width remains relatively constant. How ear growth affects the performance of automatic recognition system is currently still an open research question. The main problem here is the lack of appropriate datasets captured over a long enough period of time that could help provide final and conclusive answers. Some initial studies appeared recently on this topic, but only featured images captured less than a year apart [24].

2.2. Chronological development of ear recognition

The chronological development of ear recognition techniques can be divided into a *manual (pre-automatic)* and *automatic era*. During the pre-automatic era several studies and empirical observations were published pointing to the potential of ears for identity recognition [25–28]. One of the biggest contributions to the field during this era was made by Iannarelli in 1989 [19], when he published a long-term study on the potential of ear recognition. Iannarelli’s seminal work included more than 10,000 ears and addressed various aspects of recognition, such as ear similarity of siblings, twins and triplets, relations between the appearance of the ears of parents and children as well as racial variations of ear appearance [11].

The 1990s marked the beginning of automatic ear recognition. Various methods were developed during this time and were introduced in the literature. In 1996, for example, Burge and Burger [29] used adjacency graphs computed from Voronoi diagrams of the ears curve segments for ear description and in 1999 Moreno et al. [30] presented the first fully automated ear recognition procedure exploiting geometric characteristics of the ear and a compression network. In 2000 Hurley et al. [31] described an approach for ear recognition that relied on the Force Field Transform, which proved highly successful for this task. A year later, in 2001, the forensic ear identification project (FEARID) project was launched, marking the first large-scale project in the field of ear recognition [32].

With the beginning of the new millennium, automatic ear recognition techniques started to gain traction with the biometric community and new techniques were introduced more frequently. In 2002 Victor et al. [33] applied principal component analysis (PCA) on ear images and reported promising results. In 2005 the scale invariant feature transform (SIFT) [34] was used for the first time with ear images, raising the bar for the performance of the existing recognition techniques. In 2006 a method based on non-negative matrix factorization (NMF) was developed by Yuan et al. [35] and applied to occluded and non-occluded ear images with competitive results. In 2007 a method based on the 2D wavelet transform was introduced by Nosrati et al. [36], followed by a technique based on log-Gabor wavelets in the same year [37]. More recently, in 2011, local binary patterns (LBP) were used for ear-image description in [38], while later binarized statistical image features (BSIF) and local phase quantization (LPQ) features also proved successful for this task [39–41]. A graphical representation of the main milestones[1] in the development of ear recognition technology (briefly discussed above) is shown in Fig. 1.

2.3. Ear recognition approaches

Techniques for automatic identity recognition from ear images can in general be divided into techniques operating on either 2D

[1] In the opinion of the authors.

Download English Version:

https://daneshyari.com/en/article/4947236

Download Persian Version:

https://daneshyari.com/article/4947236

Daneshyari.com