Pattern Recognition Letters 68 (2015) 183–189

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Synthesis of large scale hand-shape databases for biometric applications☆

Aythami Morales[a,*], Miguel A. Ferrer[b], Raffaele Cappelli[c], Davide Maltoni[c], Julian Fierrez[a], Javier Ortega-Garcia[a]

[a] ATVS-Universidad Autonoma de Madrid, C\Francisco Tomas y Valiente, Madrid 28049, Spain
[b] IDeTIC-Universidad de Las Palmas de Gran Canaria, Campus de Tafira, Las Palmas E35017, Spain
[c] Biometric System Laboratory, DISI-University of Bologna, Cesena (FC), Italy

ARTICLE INFO

Article history:
Received 13 July 2015
Available online 9 October 2015

Keywords:
Biometric synthesis
Biometric recognition
Hand-shape

ABSTRACT

This work proposes and analyzes a novel methodology for hand-shape image synthesis. The hand-shape is a popular biometric trait with a high convenience of use and non-intrusive acquisition. The proposed algorithm allows to generate realistic images with natural intra-person and inter-person variability. The method is based on the Active Shape Model algorithm which has been modified in order to add the biometric information typical of new synthetic identities. The generated images are evaluated using three public databases and two hand-shape recognition systems. The results show the suitability of the synthetic data for biometric recognition works. In addition, two novel applications have been proposed to provide new insights in hand-shape biometric recognition including: improvement of machine learning classification based on synthetic training sets and scalability analysis of hand-shape biometrics when the population of the database is increased by two orders of magnitude with respect to existing databases.

1. Introduction

Biometric recognition technologies play an increasingly important role in security applications. The reliable authentication of people is crucial in nowadays digital society and biometric recognition provides the tools for improving the security of systems and services. In last decades many technologies have been proposed based on different human physical or behavioral traits: fingerprint, face, iris, hand-shape, palmprint, signature, gait, voice, keystroke, among many others.

The research community has invested time and efforts to improve the performance of biometric recognition systems. Technology evaluations require large scale datasets and their acquisition is time consuming and problematic [23]. The privacy and legal concerns related with biometric data make difficult the promotion of large acquisition campaigns and most of the research work in the literature comprise experiments with few hundreds of users. Initially, the synthesis of biometric data emerges as a solution to the lack of large scale biometric databases. Additionally, synthetic databases have several applications such as performance evaluation [3,21], security assessment [17], modeling [24] or improving enrollment [13].

The synthesis of biometric samples involves challenging tasks related with the natural variability modeling. The generation process involves modeling the patterns related with the biometric traits as well as the intra-person and inter-person variability inherent to real data. Synthetic databases are important for both research and industry because of the lack of large-scale databases derived from the difficulties of data collection; and the privacy/legal concerns related with the use of biometric information from real people.

Researchers have proposed models to generate the most popular biometric traits as fingerprint [4,31], face [2], iris [5,16,28], palmprint [9,22,11] or signature [14,15,25,32]. However, the synthesis of databases for hand-shape biometric applications remains unexplored. The hand-shape biometric recognition systems become popular in 90′s because of: (i) its convenience and non-intrusive acquisition; (ii) low-cost hardware requirements and; (iii) its acceptability to the public. While hand-shape biometrics is a well-investigated subject in the literature, to the best of our knowledge, there are no proposals for automatic generation of hand-shape images. In this paper we introduce a novel method for the generation of synthetic hand-shape images for biometric applications.

The main contributions in this work can be summarized as follow: (i) a comprehensive method to synthetically generate hand-shape images which comprise the natural intra-person and inter-person

☆ This paper has been recommended for acceptance by Jose Ruiz-Shulcloper

* Corresponding author. Tel.: +34659860245.
E-mail addresses: aythami.morales@uam.es (A. Morales), miguelangel.ferrer@ulpgc.es (M.A. Ferrer), raffaele.cappelli@unibo.it (R. Cappelli), davide.maltoni@unibo.it (D. Maltoni), julian.fierrez@uam.es (J. Fierrez), javier.ortega@uam.es (J. Ortega-Garcia).

http://dx.doi.org/10.1016/j.patrec.2015.09.011
0167-8655/

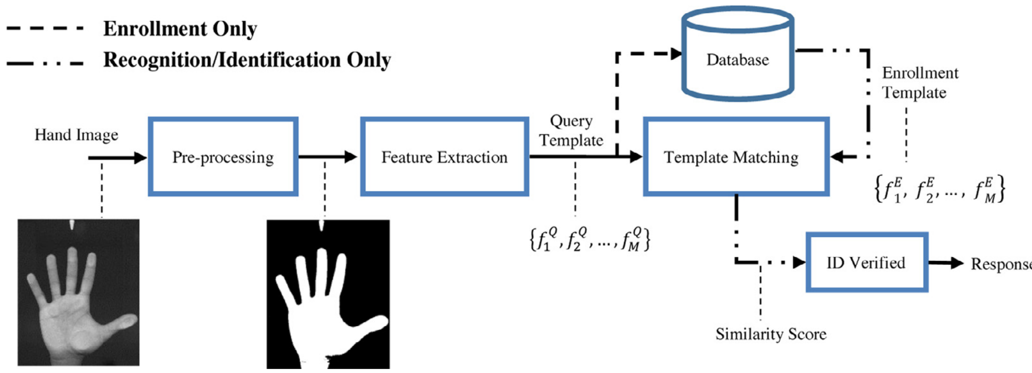

Fig. 1. Block diagram of a traditional hand-shape biometric recognition system.

variability; (ii) an extensive analysis of the performance of two biometric recognition algorithms when using synthetic and real samples; (iii) two novel applications of synthetic databases which reveal new insights on hand-shape recognition approaches; (iv) a synthetic database including 100,000 identities and 500,000 samples is made freely available for the research community[1].

The rest of this paper is organized as follows: Section 2 describes related works; Section 3 introduces the hand-shape generation method and Section 4 presents the evaluation of the synthetic data; and Section 5 introduces two novel applications of the synthetic samples. Finally, Section 6 draws some conclusions.

2. Related works

The generation of hand-shapes is not new and it has been proposed previously. In [6] authors used the concavities of the hand to show the potential of the Active Shape Models (ASM) for border detection in images. The generation of hand-shape was proposed in [17] for vulnerability assessment of hand-shape recognition systems. This work proposed a hill-climbing algorithm, in combination with an ASM hand-shape synthesizer, to reconstruct the hand-shape from the templates stored in the databases of recognition systems. The main aim of the generation method proposed there was to reconstruct hand images from the templates, so it is a special case of synthesis (new identities are not generated). Finally, hand-shape synthesis [9,22] has been integrated as part of hand-print generation focused on ridge-line patterns. The aim of the hand-shape synthesizer in the previous work [22] was the generation of realistic hand-shape contours (in terms of appearance exclusively) to be integrated in the hand-print model. The hand-shape synthesizer was only briefly described and not studied as a biometric modality.

Hand-shape generation has been explored and proposed in previous work but for the best of our knowledge, its usefulness for database generation (composed by synthetic identities) and hand-shape applications remains unknown.

3. Hand-shape synthesis

Traditional hand-shape recognition approaches can be divided into three main modules: preprocessing, feature extraction and matching. The main aim of the preprocessing phase is to prepare the images for the subsequent feature extraction. The first steps of the preprocessing involve the conversion from color/grayscale hand images into black and white images, see Fig. 1. Most of the feature extraction and matching techniques in the hand-shape recognition literature [7,12,19,26] are based on the silhouette of the hand or geometric features derived from it (e. g. widths, lengths, angles, aspect ratios, areas). All these approaches extract features from black and white hand images.

Instead of grayscale/color hand images, the generation method proposed in this work focus on the synthesis of binary hand-shapes which can be used in biometric applications. The generated images lack texture (more appropriate for palmprint/fingerprint approaches). The hand-shape generation process proposed in this work can be divided into two main tasks:

- **Generation of master hand-shapes** (modeling the inter-person variability): the aim of this task is to generate a realistic master hand-shape which includes the biometric information of a specific synthetic identity. The biometric information comprises the patterns related with each individual as aspect of the hand, sizes and shapes, among others.
- **Generation of multiple samples from a master hand-shape** (modeling the intra-person variability): the aim of this task is to model the natural variability between acquisitions of hand images from the same person. These variations include rotation, finger movement and minor changes in the shape due to elasticity of the skin.

The next subsection will detail the methodology proposed to automatically generate large scale synthetic databases.

3.1. Generation of a master hand-shape

Our hand-shape synthesis is based on the Active Shape Model algorithm [6]. The parameters of the synthesizer are calibrated using the publicly available GPDSHand database [10]. This database includes 10 images from 144 people acquired using a commercial document scanner. Segmentation can be performed by applying Otsu thresholding. Once we obtain the binary images, the contour is represented as a $2n$ element vector composed by the $\{x, y\}$ Cartesian coordinates of $n = 1300$ equally-spaced points. The contours are aligned by placing the hand geometric center as the coordinate origin and rotating the hand contour by an angle defined by the first and third finger-web (see Figs. 2 and 3). This normalization allows to reduce the effects of translation, rotation and scale in the generation model.

The training set is therefore composed by 1440 aligned contours of dimension equal to 2600. The mean contour $\bar{\mathbf{c}}$ (see Fig. 2) is obtained as:

$$\bar{\mathbf{c}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{c}_i \quad (1)$$

being $\mathbf{c}_i$ the ith aligned contour (as a column vector) of the training set and N the number of training contours (N is equal to 1440 in our model). The principal axes of a $2n$ dimensional ellipsoid fitted to the data can be calculated by applying a Principal Component Analysis (PCA) to the data [18]. Each axis gives a mode of variation, a way in

[1] http://www.gpds.ulpgc.es/