Two Unconstrained Biometric Databases

Hélder P. Oliveira* and Filipe Magalhães

INESC TEC (formerly INESC Porto) and Faculdade de Engenharia, Universidade do Porto, Campus FEUP, Rua Dr. Roberto Frias, 378, 4200-465 Porto, Portugal
helder.oliveira@fe.up.pt, filipe.magalhaes@inescporto.pt
http://www.inescporto.pt/~{hfpo,ftam}

Abstract. In the last few years the research community has witnessed significant progress in biometric technology, due to the availability of a wide variety of databases. However, the available databases that are currently available present significant setbacks in terms of restricted access to data, low-resolution and restrictions imposed on individuals during the acquisition phase.

In this paper, two new public databases are described that have been created, with fingerprint and palm print images and their characteristics are compared with other databases available in the research community. The advantages of these databases are the great variety of individual characteristics, they have no restrictions during acquisition and they have manual ground truth annotation. They were presented in two different international competitions and have been used in research by different authors.

Keywords: Database, Biometrics, Fingerprint, Palmprint.

1 Introduction

Studies of physical and behavioral traits for recognition purposes are known as biometrics. As these traits are specific to an individual, they guarantee one's identity in security control situations. A well known example of a biometric characteristic are fingerprints and they are the most widely used [1]. It is theoretically impossible to find any two individuals with the same fingerprint [2]. This is a crucial feature of a biometric characteristic: it is unique to each person. Other equally important aspects regarding biometric characteristics are universality, as they have to be present in all individuals; and permanence, so they are constant during one's life. Moreover, they should be easy to extract. At the present time, there is research being conducted on a broad range of biometric characteristics which can be divided into physical and behavioral characteristics. Physical are, for instance, fingerprints, the iris, the retinal capillary structure, the face and hand recognition. Examples of behavioral traits are voice and handwriting. Fig. 1 illustrates several biometric characteristics.

* The authors would like to thank Fundação para a Ciência e Tecnologia (FCT) for the Financial support of the PhD grants SFRH/BD/45380/2008 and SFRH/BD/43772/2008.

A. Campilho and M. Kamel (Eds.): ICIAR 2012, Part II, LNCS 7325, pp. 11–19, 2012.

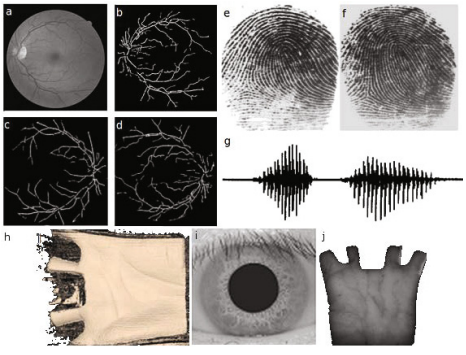

Fig. 1. Biometric characteristics. a),b),c) and d) Retinal data; e) and f) Fingerprint data; g) speech data; h) 3D palm data; i) iris data; j) palm veins data.

Biometric systems can be used for identification and verification purposes. In all cases there should be a database where biometric features from a set of individuals are stored. In an identification task, the role of the system is to compare an input with all the entries in the database and verify if there is a match, thus detecting the presence of the individual in the database. In a verification task, the algorithm checks if an individual is who he claims to be. To compare any kind of biometric characteristics it is necessary to represent them in a stable fashion. For instance, it is not feasible to directly compare images from two palm prints, as it is practically impossible to place the hand in the exact same position in different occasions, thus producing slightly different images that have to be compared in some way. This is the most crucial aspect and it can be divided into two tasks: firstly, to represent a characteristic trait in reproducible and stable features that resist input variability; secondly, to compare such features so users can be accurately recognized. These two questions are in the core of a biometric system and are addressed by most of the research in the field. Its importance is highlighted in Fig. 2 where the layout of a biometric system is depicted.

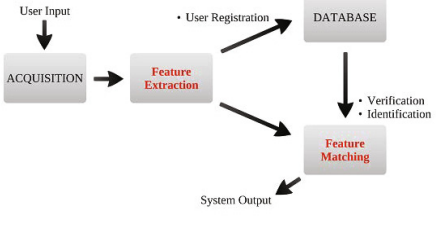

Fig. 2. A typical scheme for a biometric system

1.1 Biometric Databases

One of the main problems found in the development of biometric recognition systems, in both identification and verification tasks, is the shortage of unconstrained large public databases acquired under real working conditions. The availability of unconstrained biometric features corresponding to a large population of individuals makes database collection a complicated process, in which a high degree of co-operation from the participants is needed. For that reason, nowadays, the number of existing public databases that can be used to evaluate the performance of fingerprint and hand geometry-based biometric recognition systems is quite limited.

There are some fingerprint-based databases available but some of them are not public as is the case for the East Shore Fingerprint Image Database [3] or the images were not acquired under favorable condition, such as in the FVC200x Fingerprint verification competition databases [4] and the CASIA Fingerprint Database [5]. The East Shore database [3] has over 100,000 fingerprint images with totally flat impressions. All fingerprint images were recorded at 500 dpi and in 256 gray tone colors. The finger images have been captured from ten print inked fingerprint cards that were randomly selected to insure a representative mix of print with varying quality, ranging from those of extremely poor quality to those of excellent quality. The fingerprint verification competition databases [4] contain four disjoint databases, each collected with a different sensor/technology. Each database has 150 fingers and 12 in depth samples per finger (i.e., it consists of 1800 fingerprint images). The image format is BMP, 256 gray-levels, uncompressed. Data collection was performed without introducing difficulties, such as exaggerated distortion, large amounts of rotation and displacement and wet/dry impressions. The CASIA Fingerprint Image Database [5] contains 20,000 fingerprint images of 500 subjects. Each volunteer contributed with 40 fingerprint images, and was only asked to exert different pressure on the sensor, without introducing any other distortions. All fingerprint images are 8 bit gray-level BMP files and the image resolution is 328×356 pixels.

There are also some examples of palm print-based databases, such the PolyU Palmprint Database [6], the CASIA Palmprint Database [5] and the IIT Delhi Touchless Palm print Database [7]. The PolyU [6] is the most widely used low-resolution palm print database for algorithmic research for recognition purposes. It is comprised of 7752 images from 386 different users. Users provide either the left or the right hand but not both. On average, there are 20 samples per user, taken over two sessions. Visually, it is possible to identify more variety between images of different sessions. This happens because there is a time lapse between the sessions, which makes the results closer to what happens in operating systems. Due to the use of pegs, hand positions are restricted, thus, simplifying preprocessing stages. The CASIA palm print database [5] is similar to the PolyU although it is smaller in size. There are no pegs to restrict hand positions. Users lay the back of the hand on a dark surface and a fixed CMOS camera is used for image acquisition. The IIT Delhi Touchless Palm print database [7] uses a normal camera to take pictures of upright hands without any support. This

database has two identified drawbacks due to its low-resolution and the fact that is difficult to see hand lines.

Existing databases have important drawbacks, namely: low-resolution, controlled acquisition conditions (for example, auxiliary objects to maintain the finger or hand in a specific position) and access restrictions. Therefore, the BioStar Group[1], has promoted a plan of action to, design and acquire image for two different large-scale biometric databases with fingerprint [8] and palm print [9] images, in an unconstrained environment.

2 Databases Overview

The importance of having public datasets is recognized in the biometrics community, both to evaluate the performance of existing biometric systems and to encourage the development of new algorithms. One of the trends is related to the database size. For example, a database with an enrollment of over 100 subjects is common. Another trend is that databases have become more diversified, with respect to restrictions on acquisition methods, without the presence of auxiliary objects to help on hand or finger position.

The BioStar's Databases had four fundamental requirements: a great variety of individual characteristics (i. e. age, sex, ethnicity, etc.); no restrictions on finger or hand positioning; the possibility of using images in different biometric methods and manual ground truth annotation. Fingerprint and palm print research has drawn lots of attention from the researchers and it enabled the production of several papers and technical reports. The BioStar's databases are widely employed, since they are publicly available, and include ground truth annotation.

2.1 Fingerprint Singular Point Detection Competition Database

The BioStar fingerprint image database [8] consists of 500 samples collected from several individuals, from both sexes, aged between 20 and 62 years and various nationalities (Portuguese, Chinese, Brazilian, Nigerian, Iranian, Tanzanian and Romanian). All of the images were collected indoors and captured with an optical scanner (Microsoft Fingerprint Reader - model 1033). The resolution of these images is 355×390 pixels and all of these images are available in .bmp format. Data collection was performed without any restrictions, individuals were simply asked to place their fingers naturally on the acquisition device but no constraints were enforced to guarantee a minimum quality in the acquired images. Due to this fact, there is a great variety of images and fingerprint types, such as: image distortion, large amounts of rotation and displacement, wet/dry impressions, presence of scars and low-quality images. (see Fig. 3).

This database was initially used in the Fingerprint Singular Points Detection Competition [8]. It is now publicly available and can be downloaded from the competition website. The goal of the competition was to compare different

[1] `http://biostar.fe.up.pt/`

Fig. 3. Fingerprint images database examples

methodologies for software-based fingerprint singular point detection within a dataset composed of real fingerprint images. A strong contribution to the state of the art of this particular subject was expected, as well as an improvement in the reliability of this crucial step in fingerprint-based biometric security systems. The challenge presented in the competition was the development of an algorithm that should be optimized to maximize the accurate detection of singular points – cores and deltas – that are the most important topological features of a fingerprint [10] (see Fig. 4).

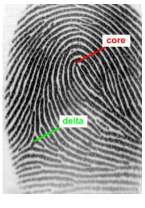

Fig. 4. Core and Delta example

Automatic biometric identification based on fingerprints is still one of the most reliable identification methods in criminal and forensic applications [11]. Fingerprint images are directionally oriented patterns formed by ridges and valleys that can be captured from a finger with a multitude of sensors, particularly, capacitive, optical, thermal sweeping and ultrasonic. Fingerprints are unique to each individual and have been widely used for biometric identification. A critical step in fingerprint analysis without human intervention is to automatically and reliably extract singular points from the input fingerprint images. The singular point area is defined as a region where the ridge curvature is higher than normal and where the direction of the ridge changes rapidly. These singular points not only represent the characteristics of local ridge patterns but they also determine the topological structure (i.e., fingerprint type). A core is defined by a turning point of an inner-most ridge and a delta is a place where two ridges running

side-by-side diverge. These singular points can constitute a powerful basis for the classification of fingerprint types [12], as well as for fingerprint alignment and orientation field modeling [13].

Our database was manually annotated with the position of the cores and deltas presented on each fingerprint image. There are many different ways to identify these singular points; our manual labeling process was based on E. R. Henry's definition [14]. The ground truth file is also available on the conference website. Although, this database was created to be applied in singular point detection (cores and deltas), it can also be used in problems relating to other fingerprint features.

2.2 Hand Geometric Point Detection Competition Database

The BioStar palmprint image database [9] contains 460 images captured from 46 individuals, 10 images from each individual. These images correspond to the right hand of individuals of both sexes, with aged between 10 and 75 years and of various nationalities (Portuguese, Chinese, Indian, Cape Verdean, Brazilian, Polish, Nigerian, Iranian and Egyptian). Usually, there are many restrictions on the acquisition and these, affect feature extraction and acquisition convenience. In many hand recognition systems it is also common to use pegs during image acquisition. This method causes the shape of the hand to alter, thus the features extracted are not genuine. Recent publications use a peg free acquisition system but these systems still limit hand positioning and orientation as they force the user to keep their hand open and their fingers spread.

This database was created to develop a revolutionary biometric recognition system therefore it has no constraints with regards to hand positioning, orientation or finger arrangement. All of the images were acquired with a normal flatbed scanner, the Epson Perfection V300 Photo, with the following characteristics: resolution of 200 ppp; .tif format; 24-bits color map. During the acquisition, there are no pegs to restrict hand position and the user is free to put his/her hand in any possible position (see Fig. 5).

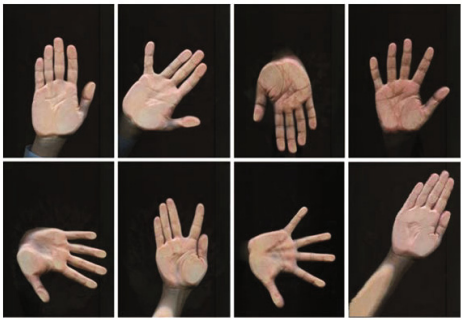

Fig. 5. Palm print image database examples

This database was initially used in the Hand Geometric Points Detection Competition [9]. It is now publicly available and can be downloaded from the competition website. The resolution of these images is 510×702 pixels and all of them are available in .tif format. The goal of the competition was to compare different methodologies for software-based hand characteristic point detection within a data set composed of real hand images. The aim was to become an internationally recognised event for academic and industrial researchers. The challenge presented in the competition was the development of an algorithm that should be optimized to maximize the accurate detection of the 9 characteristic points on the hand: 5 finger tips and 4 valleys (see Fig. 6).

Fig. 6. Hand geometry characteristics points

Hand geometry has been used in biometric verification and identification because of its acquisition convenience and identification performance [15]. From an anatomical point of view, the human hand can be characterized by its length, width, thickness, geometrical composition, shape of the palm and shape and geometry of the fingers [16]. Several researchers have been using combinations of these features for recognition with varying degrees of success [17]. Normally, the way to define these previously stated features is through the detection of the 9 characteristic points shown in Fig. 6, that is why correctly detecting their location is so important. Our database was manually annotated with these characteristic points. The position of the 5 finger tips is located at the point with the bigger curvature and the 4 valleys are located between neighbor fingers and defined at the point where it seems that there is an inversion of the direction of the hand contour. The ground truth file is also available on the conference website. Although, this database was created to be applied in the analysis of geometric characteristics, due to its high resolution, it can be also used in problems related to hand line detection or palm print singular points detection.

3 Conclusion

In this paper, we describe two biometric databases that consist of fingerprint and palm print images. We highlighted the most discriminating points between these databases and others with similar purposes, such as the fact that they contain images acquired without restrictions during acquisition, they are publicly available and were manually annotated, allowing researchers to compare their results

objectively. This type of dataset is more challenging and has characteristics that permit the user to conduct a wide variety of studies.

The BioStar databases were first introduced in two world-wide competitions. They are indexed on specialized websites with research concerns[2] and have been used in different research [18–22].

References

1. Jain, A., Lin, H., Bolle, R.: On-line fingerprint verification. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(4), 302–314 (1997)
2. Maltoni, D., Cappelli, R.: Fingerprint recognition. In: Handbook of Biometrics, pp. 23–42. Springer US (2008)
3. East Shore Technologies, `http://www.east-shore.com/data.html`
4. Cappelli, R., Ferrara, M., Franco, A., Maltoni, D.: Fingerprint verification competition 2006. Biometric Technology Today 15(7-8), 7–9 (2007)
5. Casia fingerprint database, `http://biometrics.idealtest.org/`
6. Polyu palmprint database, `http://www.comp.polyu.edu.hk/~biometrics`
7. IIT Delhi Touchless Palmprint Database, `http://www.comp.polyu.edu.hk/~csajaykr/IITD/Database_Palm.htm/`
8. Magalhaes, F., Oliveira, H. P., Campilho, A.: SPD 2010 - Fingerprint Singular Points Detection Competition Database, `http://paginas.fe.up.pt/~spd2010/`
9. Magalhaes, F., Oliveira, H. P., Matos, H., Campilho, A.: HGC 2011 - Hand Geometric Points Detection Competition Database, `http://www.fe.up.pt/~hgc2011/`
10. Srinivasan, V.S., Murthy, N.N.: Detection of singular points in fingerprint images. Pattern Recognition 25(2), 139–153 (1992)
11. Pankanti, S., Prabhakar, S., Jain, A.K.: On the individuality of fingerprints. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(8), 1010–1025 (2002)
12. Karu, K., Jain, A.K.: Fingerprint classification. Pattern Recognition 29(3), 389–404 (1996)
13. Jinwei, G., Jie, Z.: A novel model for orientation field of fingerprints. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 493–498 (2003)
14. Henry, E.R.: Classification and Uses of Fingerprints. George Routledge and Sons, London (1900)
15. Kong, A., Zhang, D., Kamel, M.: A survey of palmprint recognition. Pattern Recognition 42(7), 1408–1418 (2009)
16. Wong, R., Shi, P.: Peg-free hand geometry recognition using hierarchical geometry and shape matching. In: IAPR Workshop on Machine Vision Applications, pp. 281–284 (2002)
17. Montalvao, J., Molina, L., Canuto, J.: Robust hand image processing for biometric application. Pattern Analysis and Applications 13, 397–407 (2010)
18. Magalhães, F., Oliveira, H.P., Campilho, A.C.: A New Method for the Detection of Singular Points in Fingerprint Images. In: Proceedings IEEE Workshop on Applications of Computer Vision, pp. 157–162 (2009)

[2] `http://homepages.inf.ed.ac.uk/rbf/CVonline/`

19. Sierra, A., Avila, C., del Pozo, G., Casanova, J.: Unconstrained and Contactless Hand Geometry Biometrics. Sensors 11(11), 10143–10164 (2011)
20. de Santos Sierra, A., Sánchez Ávila, C., Guerra Casanova, J., del Pozo, G.B.: Invariant Hand Biometrics Feature Extraction. In: Sun, Z., Lai, J., Chen, X., Tan, T. (eds.) CCBR 2011. LNCS, vol. 7098, pp. 108–115. Springer, Heidelberg (2011)
21. Mestetskiy, L., Bakina, I., Kurakin, A.: Hand Geometry Analysis by Continuous Skeletons. In: Kamel, M., Campilho, A. (eds.) ICIAR 2011, Part II. LNCS, vol. 6754, pp. 130–139. Springer, Heidelberg (2011)
22. Matos, H., Oliveira, H.P., Magalhães, F.: Hand-Geometry Based Recognition System - A Non Restricted Acquisition Approach. In: Campilho, A., Kamel, M. (eds.) ICIAR 2012, Part II. LNCS, vol. 7325, pp. 38–45. Springer, Heidelberg (2012)