Expert Systems with Applications 40 (2013) 2858–2871

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Indexing and retrieving in fingerprint databases under structural distortions

Andrés Gago-Alonso *, José Hernández-Palancar, Ernesto Rodríguez-Reina, Alfredo Muñoz-Briseño

Advanced Technologies Application Center (CENATAV), Havana, Cuba

ARTICLE INFO

Keywords:
Fingerprint identification
Fingerprint indexing
Triangular hull
Redundant feature vector
Delaunay triangulation

ABSTRACT

This paper presents a new algorithm for fingerprint indexing, which is based on minutia triplets, and it is very tolerant to missing and spurious minutiae. In this sense, a novel representation for fingerprints is proposed by defining a triangle set based on extensions of Delaunay triangulations. Moreover, a set of robust features is used to build indices. Finally, a recovery method based on calculating the recommendation score is introduced, using a new similarity function between geometric transformations. Our proposal was tested on well known databases, showing that it outperforms most of the already reported methods, especially under conditions of distortions.

1. Introduction

Biometrics can be defined as the automated use of physiological or behavioral characteristics to identify or verify the identity of a person. One of the most widely used techniques in biometric systems is the comparison of fingerprints. The ridge patterns found on fingers and other body parts are unique, and they provide enough information to distinguish a specific person from the rest. Also, these patterns can be extracted very easily and are very reliable compared with other biometric features.

There are two kinds of general problems related to fingerprint recognition systems: verification and identification. The purpose of verification systems is to confirm the identity of a particular individual, so comparisons are only necessary with fingerprints that belong to that person (Girgisa, Sewisyb, & Mansourb, 2009; Nanni & Lumini, 2009). On the other hand, the purpose of identification is to establish the identity of a specific person, given a query impression and a dataset of fingerprints of different individuals. As we can see, identification requires a search on all possible fingerprint candidates. However, a comparison between the query and every candidate stored in the dataset is impracticable, since modern fingerprint collections usually have millions of entries.

There are some proposed approaches in literature that try to reduce the search space in which the comparisons are made (Cappelli, Ferrara, & Maltoni, 2011). One of these solutions is the classification of the fingerprints stored in the dataset, in the five classes of Henry (left loop, right loop, arch, tended arch and whorl) (Hung, Liung, & Yi, 2011; Liu, 2010; Rajanna, Erol, & Bebis, 2010). These classes divide the impressions in groups according to ridge patterns. In this way, comparisons are only made with fingerprints in the dataset that have the same classification as the query. This method has serious disadvantages mainly because the number of classes in which the search space is divided is small. In addition 90% of impressions belong to three classes (Liang, Bishnu, & Asano, 2007), so, in most cases, the reduction of potential candidates is insignificant.

Another group of algorithms uses indexing in order to return a subset of the dataset, ordered by a recommendation score. This approach, also known as continuous classification, allows choosing the number of impressions of the dataset that will be compared to the query. However, most of these solutions do not have robust strategies to deal with missing or spurious minutiae, and the commonly used mechanisms to reduce the negative effect of noise are insufficient.

This paper proposes an indexing algorithm which is prepared for dealing with the problem of missing and spurious minutiae. This algorithm is based on minutia triplets, and it introduces a novel fingerprint representation based on an expanded triangle set obtained from Delaunay triangulations. From each of these triangles, indices are formed using fingerprint features such as ridge counters, minutia directions and triangle sign. With these indices, an index table is built in preprocessing time. In the retrieving stage, a novel method for calculating the recommendation score and a mechanism to deal with noise are also defined. Thus, we can get a list of candidates with the highest degrees of affinity with the query, considering the best geometric transformation.

The rest of this paper is organized as follows. In Section 2, some basic concepts, which are useful for understanding the rest of the document, are presented. Also, a general scheme of fingerprint indexing algorithms is given. Next, in Section 3, a description of the main state of the art algorithms is exposed. In Section 4, a

* Corresponding author. Tel.: +537 272 1670; fax: +537 273 0045.
E-mail addresses: agago@cenatav.co.cu (A. Gago-Alonso), jpalancar@cenatav.co.cu (J. Hernández-Palancar), erreina@cenatav.co.cu (E. Rodríguez-Reina), amunoz@cenatav.co.cu (A. Muñoz-Briseño).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.12.004

new feature extraction strategy is defined using a new criterion for selecting the set of triangles. Section 5 introduces the index function, the indexing process, and index table construction. In Section 6, a novel method for recovering a list of candidates is proposed. Finally, experimental results are shown in Section 7, and our conclusions are given in Section 8.

2. Background

In this section, we present some basic concepts and a general scheme of fingerprint indexing algorithms. Thus, we declare the necessary background for understanding our proposal and the rest of the paper. Finally, we describe the Delaunay triangulation and its properties, considering that this kind triangulation is used for many indexing algorithms, including our approach.

2.1. Fingerprint related concepts

Fingerprints are marks produced by the contact of a finger with a surface in a controlled environment. These marks reflect the different patterns formed by the ridges that are visible in the epidermis. For many years, fingerprint image acquisition has been accomplished from different sources like: inked finger on paper or ink-less fingerprint scanners. In Section 7.1, we can see examples of some datasets of fingerprints obtained in different sources.

Most of the indexing algorithms use minutiae as basis for representing fingerprint and building indices. Minutiae are singularities in the ridge patterns, which are commonly classified in two types: bifurcations and endings. A bifurcation is a point where a ridge splits into two ridges, while an ending is an endpoint of a ridge. In Fig. 1, we can see examples of bifurcations, endings, and ridges in a real fingerprint.

The direction of a minutia is another commonly used feature in fingerprint indexing algorithms. This feature is defined as the angle formed between the horizontal axis and the tangent of the ridge associated to the corresponding minutia, in counter clockwise. In Fig. 1, a bifurcation with its respective direction is shown.

There are several published minutia extraction algorithms which have shown an allowable performance (Kasaei & Boashash, 1997; Rajanna et al., 2010). However, in almost all of these methods, the possibility of finding false minutiae always exists. False minutiae are the points which are incorrectly identified as minutiae. In Fig. 1, we can see a false minutiae caused by a scar in the finger.

2.2. Indexing based systems

The general scheme of all indexing based fingerprint identification systems, is the same, see Fig. 2. Such scheme is made up of an indexing stage and a retrieving stage. Indexing stage is also known

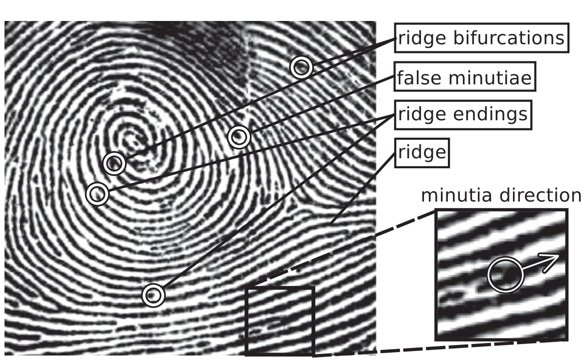

Fig. 1. Fingerprint parts.

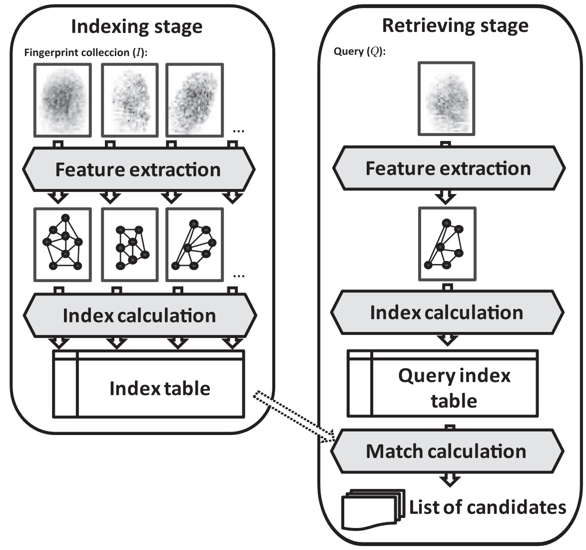

Fig. 2. General scheme of an indexing based fingerprint identification system.

as offline stage since it is executed while the fingerprint collection is preprocessed. The queries are attended in a retrieving stage, which detects the query occurrence in the fingerprint collection.

The indexing stage can be described as follows in almost all reported methods. Let $I = \{F_1, F_2, \ldots, F_N\}$ be a collection of fingerprints, where F_i represents the i-th stored impression, and N is the number of impressions in I. Each $F_i \in I$ is preprocessed for extracting a set of features, which is used as model for representing it. These models are the basis for calculating indices which are stored in the index table. This table is used for reducing the search space during the retrieving stage.

The retrieving stage also has the same structure in almost all reported methods. Given a query fingerprint Q, it is required to detect if there is any $F_h \in I$ such that F_h and Q represent the same finger of the same person. This stage finds a list of candidates $C_Q \subset I$, such that the probability of finding F_h in C_Q is very high, while the probability of finding F_h in $I \backslash C_Q$ is very low.

Each query Q is processed in a similar way as it was done for each fingerprint in I during the indexing stage. Thus, a set of features is calculated, and it is used as model for representing Q. This model is used for calculating the query indices, which are used for finding the matches with the already calculated index table. Finally, the list of candidates is obtained by combining these match results.

2.3. Delaunay triangulation

In general, a triangulation of a set of points, $P = \{p_1, p_2, \ldots, p_N\}$, in the plane is the set of triangles that conforms a maximal planar subdivision whose vertex set is P. A maximal planar subdivision is a subdivision S such that no edge connecting two vertices can be added to S without destroying its planarity (Berg, Krevelt, Overmars, & Scharzkopf, 1997).

Delaunay triangulation is a specific kind of triangulation, which has been used for representing fingerprints, in some of the reported indexing algorithms (Bebis et al., 1999; Liang, Asano, & Bishnu, 2006; Liang et al., 2007). This concept is defined as follows.

Definition 1 (*Delaunay triangulation*). Let $P = \{p_1, p_2, \ldots, p_N\}$ a set of points in the plane, and let T be a triangulation of P. Then, T is a Delaunay triangulation if and only if every triangle $\Delta P_i P_j P_k$ that belongs to T satisfies that its circumcircle contains no other point of P.