Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 1793–1803
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Efficient fingerprint search based on database clustering

Manhua Liu, Xudong Jiang*, Alex Chichung Kot

EEE, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

Received 30 June 2006; received in revised form 10 November 2006; accepted 16 November 2006

Abstract

Fingerprint identification has been a great challenge due to its complex search of database. This paper proposes an efficient fingerprint search algorithm based on database clustering, which narrows down the search space of fine matching. Fingerprint is non-uniformly partitioned by a circular tessellation to compute a multi-scale orientation field as the main search feature. The average ridge distance is employed as an auxiliary feature. A modified K-means clustering technique is proposed to partition the orientation feature space into clusters. Based on the database clustering, a hierarchical query processing is proposed to facilitate an efficient fingerprint search, which not only greatly speeds up the search process but also improves the retrieval accuracy. The experimental results show the effectiveness and superiority of the proposed fingerprint search algorithm.

Keywords: Fingerprint identification; Fingerprint search; Biometrics; Fingerprint classification; Orientation field

1. Introduction

Fingerprint as a kind of human biometric feature has been widely used for personal recognition in the commercial and forensic areas because of its uniqueness, immutability and low cost. In general, fingerprint based recognition systems work in two modes: authentication and identification [1]. In the authentication mode, the user inputs his fingerprint and claims an identity information, then the system verifies whether the input fingerprint is consistent with the claimed identity. In the identification mode, the user input his fingerprint and the system identifies the potential corresponding ones in the database without a claimed identity. Therefore, fingerprint identification requires searching the database for a match, which is more complex than the authentication. Although satisfactory performances have been reported for fingerprint authentication, both the efficiency and accuracy of identification deteriorate seriously by simple extension of a 1:1 authentication procedure to a 1:N identification system [1]. How to efficiently search the fingerprint database is a great challenge. Multi-level matching approaches are proposed to facilitate the database search by incorporating the global and local information of fingerprint [2,3]. The coarse level matching (search) is often used to reduce the search space of the time-consuming fine matching and alleviate the accuracy deterioration of identification [1]. Exclusive classification, fingerprint indexing and continuous classification have been proposed for the coarse level search of database.

Exclusive fingerprint classification is a traditional approach that has been widely investigated in the literature [4–11]. It classifies each fingerprint exclusively into one of the predefined classes such as Henry classes. Although it has some advantages such as human-interpretability, fast retrieval and rigid database partitioning, most automated classification algorithms are able to classify fingerprints into only four or five classes. Moreover, fingerprints are not evenly distributed in these classes. The natural fingerprint distribution of the Henry five classes is 3.7% plain arch, 2.9% tented arch, 33.8% left loop, 31.7% right loop and 27.9% whorl. On average, a query fingerprint still needs to be compared with about 29.48% of database templates in the fine matching of identification. Thus, the exclusive classification cannot sufficiently narrow down the search of database.

* Corresponding author. Tel.: +65 6790 5018; fax: +65 6793 3318.
E-mail address: exdjiang@ntu.edu.sg (X. Jiang).

0031-3203/$30.00
doi:10.1016/j.patcog.2006.11.007

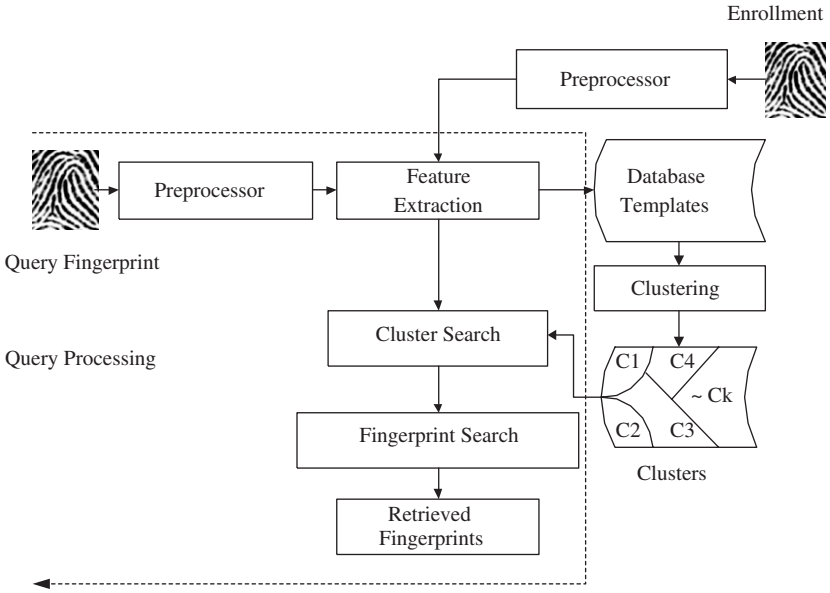

Fig. 1. The overview of the clustering based fingerprint search algorithm.

In fact, it is not necessary to classify fingerprints into human-interpretable classes for an automated identification system. Fingerprint indexing, which divides fingerprint database into a number of bins based on the minutia triplets, was proposed in Refs. [3,12]. This approach classifies fingerprints into more classes (or bins) than the exclusive classification as it exploits the more discriminating features, minutiae. However, the minutia points are the most important local features and widely used in fingerprint fine matching algorithms [2,13–15]. Although this approach can speed up the database search, it should take care to avoid a redundant representation of fingerprint in an identification system. This is because the accuracy deterioration of the identification system is hardly alleviated if the features used in the coarse search and fine matching are strongly correlated.

Continuous classification is proposed to overcome the problems of exclusive classification by representing fingerprint with numerical feature vectors [16–18]. The fingerprint search is performed by comparing the query fingerprint with all database templates and retrieving the closest ones. The tradeoff between retrieval efficiency and accuracy can be easily adapted by adjusting the size of retrieval neighborhood. Although the comparison between the query fingerprint and template is much faster than the fine matching, this full fingerprint search is still prohibitive for large database. Moreover, the continuous classification only ranks the database templates according to their similarities to the query fingerprint while neglecting the similarities among the database templates. This limits the search performance. Although some combined techniques were proposed to improve the performance of fingerprint classification [19,20], further work to facilitate an efficient search of database is still of great interest to the researchers in the area of fingerprint identification.

Data clustering is a crucial technique used in discovering the underlying structure in a data set by unsupervised grouping of the similar patterns. It accelerates the content based image retrieval by comparing the query image with a few cluster representatives instead of all database templates [21,22]. This work proposes an efficient fingerprint search algorithm based on database clustering. The data-flow chart of this algorithm is shown in Fig. 1. It differs from the continuous classification (full search) in that clustering is employed to exploit the similarities among the database templates. Fingerprint is non-uniformly partitioned by a circular tessellation to compute a multi-scale orientation field as the main feature for the search. The average ridge distance (ARD) is extracted as an auxiliary search feature. Our proposed fingerprint search algorithm consists of two phases: offline database clustering and online query processing. During the offline database clustering, a modified form of K-means clustering is proposed to partition the orientation feature space into clusters and fingerprints of each cluster are further divided into bins according to their ARDs. Based on the offline database clustering, a hierarchical online query processing is proposed to facilitate an efficient fingerprint search. In cluster search, each query fingerprint is compared with the cluster prototypes to retrieve the close clusters followed by searching the bins of the retrieved clusters. Fingerprint search is finally performed on the retrieved bins to find the templates close to the query fingerprint.

The next section presents the feature extraction, including the computation of multi-scale orientation field and ARD. In Section 3, we present the proposed fingerprint search algorithm based on database clustering. The experimental