Application of Data Mining Techniques to Palmprint Recognition

Prachi Janrao, Vikas Singh
Assistant Professor
Department of Computer Engineering, Thakur College of Engineering and Technology, Mumbai, Maharashtra.

ABSTRACT
Data mining is a powerful technology that extracts and analyzes data for finding correlations or patterns from the large sets of databases. It performs pattern analysis on large sets of data, using tools like association, clustering, segmentation, and classification for helping better manipulation of the data. Palmprint is considered to be one of the most stable biometric characteristics. Over the last decade implementation of Palmprint recognition for security purpose has been increased all over the world. The paper proposes the application of Data mining techniques to the Palmprint recognition for improving the performance of retrieval.

General terms
Data mining, Biometrics.

Keywords
K-means algorithm, Palmprint recognition.

1. INTRODUCTION
There is increasing need in the development of reliable, rapid, efficient and non-intrusive security control systems. A wide variety of systems requires reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services [1]. Amongst these systems is the biometric system, which recognizes the individual based on their physical characteristics or the behavioral characteristics.

In the physical characteristics of an individual, palmprint has the most stable and unique characteristics and is rich in the feature information. [2] Even two identical twins have different palmprints. The features of a palm are palm geometry, flexion creases (principle lines), secondary creases (wrinkles), ridge and valley features, delta features, minutiae features, singularity point and texture. By combining all these features, it is possible to build a highly accurate biometrics system [3].

There are 5 modules of the biometric systems for palmprint recognition as Sensor module, Feature extraction module, Matching module, Decision module and System Database module.

This paper proposes the application of data mining techniques to the recognition of palmprint from the large database. The rest of the paper is organized in following sections. Section 2 gives the different types of the biometric system for palmprint identification. Section 3 gives the detailed palmprint recognition procedure. Section 4 discusses the different data mining techniques that can be applied to the palmprint recognition system. Section 5 summarizes the results of the paper.

2. BIOMETRIC SYSTEM
A biometric system is essentially a pattern recognition system which recognizes the individual by determining the authenticity of a specific anatomical or behavioral characteristic possessed by an individual. The system acquires the biomedical data from the individual, extracts the features from the acquired data and compares it against the database. Any human physiological and/or behavioral characteristic can be used as a biometric characteristic as long as it satisfies the following requirements [1]:

- Universality: each person should have the characteristic.
- Distinctiveness: any two persons should be sufficiently different in terms of the characteristic.
- Permanence: the characteristic should be sufficiently invariant (with respect to the matching criterion) over a period of time.
- Collectability: the characteristic can be measured quantitatively.

There are different types of biometric systems.
Unibiometric system is a system that uses only a single biometric identifier

- **Unimodal biometric system** is a subset of a unibiometric system that uses a single instance (snapshot), a single representation, and a single matcher for a recognition decision.

- **Multi-biometric system** is a biometric system that uses more than one independent or weakly correlated biometric identifier taken from an individual (e.g., palmprint and face of the same person, or fingerprints from two different fingers of a person, respectively).

- **Multimodal biometric system** is a superset of a multi-biometric system that may use more than one correlated biometric measurement, e.g., multiple impressions of a finger, multiple images of a face in a video, multiple representations of a single input, multiple matchers of a single representation, or any combination thereof.

A typical design of biometric system has features shown in fig.2 [4].

3. **Palmprint Recognition**

There are four different stages in which recognition procedure is carried out: palmprint acquisition, preprocessing, feature extraction and matching with the database.

Usually acquisition of palmprint is done with the help of different types of cameras and scanners. Contactless image acquisition can be done with the help of digital cameras and video cameras [5]. But quality of such images may not be good as it is contactless.

For preprocessing, usually number of steps are involved such as binarizing the image, obtaining boundary points, curvature points and thus finding the region of interest (ROI) [6]. For preprocessing many algorithms can be used but all the algorithms work one or the other ways same.

Feature extraction can be done with different approaches. Some of the approaches are SIFT (Scale Invariant Feature Transformation), a time series technology, SAX [7] and its extension 2D SAX which represents a real valued 2D data as 2D matrix of symbols, 2D gabor filters, phase based palmprint matching, Principle component analysis based approaches which includes PCA on PCA and 2D PCA analysis for Gabor wavelets, Kekre’s wavelet’s energy entropy based feature vector [8], different types of filiformity technique which is a pixel level operation [9], fisher palms which suggests that images with resolution 32x32 are optimal for medium security biometric systems while those with resolution 64x64 are optimal for high security biometric systems [10], a multimodal approach called “Laplacianpalm” in which the palmprint and palm vein images are fused by a novel integrated line preserving and contrast enhancing fusion method [11], for speed, accuracy and convenience embedded online palmprint verification system based on Ethernet can be used [12] etc.

Palmprint matching can be done with the help of different algorithms. These algorithms determine the similarity between the two data sets. The palmprint is said to be authentic if the result obtained after matching is more than the preset threshold value [9]. The palmprint recognition algorithms can be roughly divided into four categories: structure based, statistics based, subspace based and coding based methods [13].

A typical palmprint recognition steps are as shown in figure 5.
4. DATA MINING TECHNIQUES FOR PALMPRINT RECOGNITION

Since decade different techniques for palmprint feature extraction and matching has been used to improve the accuracy of palmprint matching from the databases. All the algorithms uses different filters, classifiers, feature vectors, which gives excellent results for usual, medium size databases. But these same techniques slow down the system for the large data bases. For real time system where in thousands and lacs of verification and matching is required for the authentication, a very huge data bases are used.

For matching of one palmprint it is required to search the whole database. This exhaustive matching significantly reduces the system’s throughput as well as the accuracy because misidentification is more likely in large database. In real time systems speed is equally important as accuracy; one cannot compromise with the speed. ZHANG Li-cong, DING Xiao-ming, ZHANG Yan-qiang, LI Qiang, WANG Cheng-qi has suggested the use of Ethernet for speed, where server and client system is used. But it can be used with low-resolution palmprint images only [12].

Data mining techniques can be applied to large databases for speedy retrieval. Classification or Clustering approaches can be used for the retrieval from the database. Classification approach will classify the palmprint images from database to different classes within the database as per their features. For the incoming image, features will be extracted and these features will be matched with the existing classes within the database, wherever the similar feature class is available, the incoming image will be inserted. This approach will reduce the time required to retrieve the image from the database while matching.

In 2004, Wu et al’s [14] proposed classification method which divided palmprints into six different palmprint categories further Li Fang [15] proposed a novel classification method to further categorize Category 5 of the previous classification into five sub-categories which is based on two-stage classifier. Clustering approach will make clusters (group of similar feature images) around the defined feature, within the database. Following is the proposed model for the palmprint recognition with the application of clustering.

In the proposed model, instead of unimodal biometric use of multimodal biometric can be there. Two different feature extraction algorithms can be used simultaneously and their result can be fused. These images after feature extraction can then be formed a cluster by applying the clustering algorithm. Application of K-means clustering algorithm would be beneficial as it is calculates around the mean value, which would be easy to calculate. The procedure for application of K-means is as follows.

a. First, the feature vector of the palmprint image can be extracted in the form of dimensional vector.

b. Form the decided number of clusters of the images. For formation of clusters let first K clusters are the individual single image cluster. Then for rest of the images calculate their distances from each of the cluster which are already formed. Place the image in cluster with the nearest centroid.

c. Take each feature vector serially and compute its euclidean distance from centroid of each cluster. If it is not present in the cluster with closest centroid, place the feature vector in that cluster and calculate the centroid with the addition of the new feature vector as well the losing feature vector.

Apply the algorithm until no new image is remaining.

In this process the total search space gets reduced thereby reducing the time to retrieve from the database.
5. CONCLUSION

Application of K-mean clustering algorithms can surely reduce the time to retrieval there by increasing the efficiency. K-mean Clustering algorithm can be used on multimodal biometrics with more than one physiological feature also.

6. REFERENCES


[3] Xiangqian Wu, David Zhang, Kuanquan Wang, Bo Huang, “Palmprint classification using principal lines” 2004 Pattern Recognition Society. Published by Elsevier Ltd.


[15] Li Fang, Maylor K.H. Leung, Tejas Shikhare, Victor Chan, Kean Fatt Choon, “Palmprint Classification”