

Short Synopsis of the thesis

Minutiae Local Structures for Fingerprint Matching and Indexing

Human beings use specific characteristics of people such as their facial features, voice and gait to recognize people who are familiar to us in our daily life. The fact that many of the physiological and behavioral characteristics are sufficiently distinctive and can be used for automatic identification of people has led to the emergence of *biometric recognition* as a prominent research field in recent years. Several biometric technologies have been developed and successfully deployed around the world such as fingerprints, face, iris, palmprint, hand geometry, and signature. Out of all these biometric traits, fingerprints are the most popular because of their ease of capture, distinctiveness and persistence over time, as well as the low cost and maturity of sensors and algorithms.

This thesis is focused on improving the efficiency of fingerprint recognition systems using local minutiae based features. Initially, we tackle the problem of large scale fingerprint matching called fingerprint identification. Large size of databases (sometimes containing billions of fingerprints) and significant distortions between different impressions of the same finger are some of the major challenges in identification. A naive solution involves explicit comparison of a probe fingerprint image/template against each of the images/templates stored in the database. A better approach to speed up this process is to index the database, where a light-weight comparison is used to reduce the database to a smaller set of candidates for detailed comparison.

In this thesis, we propose a novel hash-based indexing method to speed up fingerprint identification in large databases. For each minutia point, its local neighborhood information is computed with features defined based on the geometric arrangements of its neighboring minutiae points. The features proposed are provably invariant to distortions such as translation, rotation and scaling. These features are used to create an affine invariant local descriptor called an *Arrangement Vector*, which completely describes the local neighborhood of a minutiae point. To account for missing and spurious minutiae, we consider subsets of the neighboring minutiae and hashes of these structures are used in the indexing process. Experiments conducted on FVC 2002 databases show that the approach is quite effective and gives better results than the existing state-of-the-art approach using similar affine features.

We then extend our indexing framework to solve the problem of matching of two fingerprints. We extend the proposed arrangement vector by adding more features to it and making it more robust. We come up with a novel fixed-length descriptor for a minutia that captures its distinctive local geometry. This distinctive representation of each minutiae neighborhood allows us to compare two minutiae points and determine their similarity. Given a fingerprint database, we then use unsupervised K-means clustering to learn prominent neighborhoods from the database. Each fingerprint is represented as a collection of these prominent neighborhoods. This allows us to come up with a binary fixed length representation for a fingerprint that is invariant to global distortions, and handle small local non-linear distortions. The representation is also robust to missing or spurious minutiae points. Given two fingerprints, we represent each of them as fixed

length binary vectors. The matching problem then reduces to a sequence of bitwise operations, which is very fast and can be easily implemented on smaller architectures such as smart phones and embedded devices. We compared our results with the two existing state-of-the-art fixed length fingerprint representations from the literature, which demonstrates the superiority of the proposed representation.

In addition, the proposed representation can be derived using only the minutiae positions and orientation of a fingerprint. This makes it applicable to existing template databases that often contain only this information. Most of the other existing methods in the literature use some additional information such as orientation flow and core points, which need the original image for computation. The new proposed binary representation is also suitable for biometric template protection schemes and is small enough to be stored on smart cards.