

# Summary

## LSH Indexing for fingerprint data

---

This paper has given an insight into how to use indexing to reduce time complexity and improve the accuracy of search through a database consisting of data with numerous features/dimensions. In particular, the database here consists of scanned fingerprint images. This paper has also provided us with an idea for generating a 4-dimensional feature vector from a raw fingerprint image by first extracting minutiae from it.

### **Indexing part:**

We basically use two hash functions for indexing data with any number of dimensions in its own domain. The first hash function splits the space through hyperplanes, thus minimizing the search space substantially. The second hash further divides these spaces through hyperspheres, further minimizing the space. This is almost similar to splitting a globe through longitudes and latitudes.

### **Feature Vector Extraction from fingerprint**

First, we process the scanned fingerprint through several steps such as sharpening and enhancing, increasing contrast, and converting it into binary, thinning, etc. to minimize the effect of noise and distortion. Thus a Minutiae is extracted. Through this minutiae, we deduce stress points where we find ridge ending, ridge island, ridge dot, etc. For every such point, we find its two nearest neighbor points, thus creating a triangle. We generate a feature vector consisting of 4 dimensions  $\langle r, \theta_a, \theta_b, \theta_c \rangle$ , where  $r$  is the ratio between circumradius and inradius of the formed triangle.  $\theta_a, \theta_b, \theta_c$  are relative phases of the three points with respect to one another.