LHS indexing

There exist many algos and approaches to deal with high dimensional data and to index the biometric data but each have their limitations.This paper proposes an lsh based space partitioning approach for indexing fingerprints independent of dimensions .In this we divide the data space into zones,blocks and bins to store data into.We first divide dataspace into zones using m hyperplanes all passing through a centre point c.And then these zones are divided into blocks using Q equicentric hyperspheres equally spaced from each other .And again these blocks are divided into bins which can have upto maximum of nine sublocations based on a search threshold t(given by user).This t also derives m and Q (no.of hyperplanes and hyperspheres we use to divide our dataspace). And to identify fingerprints we use an indexing that has three variables $<\alpha, \beta, \gamma>$ each corresponding to the zone ,block and sublocation in bin it belongs to respectively.In every fingerprint minutiae triplets are identified and for each triplet a feature vector is extracted using some geometry stated and are stored with a unique identifier.And these are mapped to the index vectors and these are used to identify the query fingerprint.In results it is shown that this approach is as good as the best one available and produces maximum hit rate at an optimised penetration rate.

Multimodal Indexing

This paper proposes a novel approach for indexing multimodal biometric data .It makes use of several biometric traits such as iris,fingerprint and face so that subject identification is more accurate .Data contains several samples of each subject for each trait . According to the said approach ,firstly feature vectors are extracted from all the samples of all subjects using some popular established techniques .For each trait we select a total of m reference subjects from all different subjects with a sample each. We select these so as to get maximum distinctiveness to generate index keys .And then with some predefined definitions for score, relative scores are calculated for all samples against all reference subjects .We'll have relative scores for each trait and now we'll use an SVM classifier to combine these scores and then also normalise them to scale the values .And again equalised scores are generated so as to make score distribution uniform and then these scores are used to generate index keys .And while storing ,an index space is created for each reference subject and in an index space a table for each trait .each table will have certain length and in each cell we have an ID list that stores identities of subjects .And when we get a query we generate its index keys and retrieve the most probable matched subjects from the database using ranking where the subject retrieved most is highest ranked and so.In results it can be seen that using multimodal traits produce lot better hit rate than unimodal and using the above indexing approach it can be made much better,accurate and efficient.