Fast palmprint retrieval using speed up robust features

G Kiran Kumar, D Malathi Rani, E Guru Mohan Rao

1 Professor, Department of Information Technology, MLR Institute of Technology, Hyderabad, 500043, India
2 Assistant Professor, ECE Department, Marri Laxman Reddy Institute of Technology and Management, Hyderabad, 500043, India
3 Assistant Professor, Department of Information Technology, MLR Institute of Technology, Hyderabad, 500043, India

Abstract

Background/Objective: Biometric usage is increasing in exponential series in all organisations for multiple purposes like employee attendance, Aadhaar based authentication and secure login using finger print etc. This biometric process should be as quick as possible without making much delay to retrieve the respective finger print. So an efficient quick retrieval procedure is required, in this regards a fast retrieval method for palm prints is proposed in this article.

Method: This method uses Speed up Robust Features (SURF) and an efficient look up table for fast retrieval of palm prints. A key is computed for each palmprint by matching with a pre-selected palmprint called representative. This key is used, to place the palmprint into the look up table like traditional database record. To identify a query palmprint, key is computed and selects a set of palm prints from the look up table which are having similar key as possible matches.

Findings: This proposed solution is experimented with multiple representative images to check the improved performance. As an outcome we could achieve better hit rate by comparing with existing system.

Novelty: This proposed method makes the new palm prints dynamically without disturbing the current records in the system. The entire solution is experimented on benchmark PolyU palmprint database of 7,753 images and significant performance is shown in results. This proposed solution shows better results with respect to hit rate and miss rate.

Keywords: Palmprint; index key; SURF; similarity score; representative

1 Introduction

Nowadays security is an important issue in every sector including government, business, etc. Therefore, the use of biometric systems has been increased enormously in every field. However, many of these biometric frameworks has to manage huge databases and its size is expanding at a quick pace. For example, India’s national ID program known as Unique Identification Authority of India (UIDAI) has a database of more than 700 million individuals. It may reach 1.25 billion individuals in last couple
of years.

Generally, recognition of a person in these biometric applications is done by comparing the person's biometric sample with all registered samples in the dataset. This linear matching process increases the search time of the system. Hence, an efficient indexing procedure or method is required and those methods can enable searches over a large database in less time and without compromising recognition accuracy. The indexing method should be designed such that, to identify a query, it has to immediately fetch a small group of candidates which are most alike from the database. Then, as a replacement of comparing the query with the whole database, only the samples in the candidate set are compared which increases the query speed and reduces the response time of the system

An indexing technique generates an index key to all registered image in the database. How-ever, the below issues to be viewed before to construct index keys: (1) As biometric image accession and refining is liable to noise, the biometric sample acquired from a user in the process of spotting may be non-identical from the sample which was utilized to create the framework during enrolment. Which in turn results in dissimilar index keys for the same user and the recognition process becomes difficult. (2) Traditionally, database records are arranged in sorting order for efficient retrieval based on either alphabetically or numerically. But, we cannot set out the biometric pictures similar to traditional records as they won't have any regular sorting sequence to index

Remaining manuscript is arranged as follows. The current biometric indexing techniques will be explained in Section 2. Section 3 explains the proposed indexing technique. Experimental outputs are provided in Section 4. Section 5 concludes the paper.

2 Related work

The current biometric indexing techniques can be vastly divided into: i) point, ii) triplet or quadruplet of points, or iii) match score based approaches. In point based approaches, authors taken the important factors of the biometric samples and utilized geometric hashing or its variants to do indexing. In triplet based indexing approaches, author first compute a set of triangles from the extracted feature points of the biometric samples. The computed triplets are assigned to a hash table utilizing their geometric features like angles, lengths of the sides, orientation, type and many more. Iloanusi, computed the quadruplets from the key feature points and used the geometric properties calculated from them for indexing purpose. However, the drawback of all the methods described above is that, the features extracted for different samples are not equal in number which makes the system unreliable.

Further, there are some logics apply on similarity scores for indexing which deal with fixed length keys. However, most of these methods fail to register the new users dynamically and used sequential search over the index space to identify a query. This paper addresses the above problems and investigates an accurate technique to index palmprints using SURF features. A look up table is created to index the biometric samples like traditional database records. This allows the identification approach to avoid the linear search during identification of a query and performs a quick search. Further, this approach creates entry to the new users dynamically without disturbing the existing system.

In this paper, they mainly worked on 2-D palm print images. To recognize the palm print they proposed a method where touch of palm doesn't require. Convolutional Neural Network (CNN) approach is used to extract features of the palm that are highly discriminated. Earlier methods are based on supervised learning for which class labels are required. But in this approach unsupervised procedures are adopted through Gabor responses and Principle Component Analysis (PCA). Though there are many approaches such as coding – based approaches, Local Texture Descriptor based approaches and Deep Learning based approaches, in this paper touch less palm prints are identified by PalmNet and it is a novel CNN based method.

In this paper, the authors state that, basically there are two types of palmprint recognition: First is online palm print identification and second is latent palmprint identification. The first one will happen with a digital camera where the user of palm print will cooperate to take palm print. Second is palm print collected from crimes scenes, where user obviously does not cooperate. In second case usually criminals may hide their faces, but we can use their palm print from the whole image to identify the person. In this paper, a new palm print database is established and applied a new deep learning method.

In this paper, the authors proposed a method to retrieve the palmprint using Deep Convolutional Neural Network (DCNN). In this method, large database set has been taken for experiment their proposed idea. Few research articles summarised all the palmprint retrieval methods.

3 Proposed indexing technique

This technique relies mainly on similarity scores between the palmprints. The similarity score between 2 palmprints is calculated by comparing their SURF features in Euclidean space.
The concept behind using similarity scores is that, generally palmprints that belong to same user will be having similar features and their similarity scores with a third palmprint (say representative) are also equal (Equation 3.1).

\[
\text{if } (A = B) \Rightarrow [S(A, r) = S(B, r)]
\]  

(3.1)

where A and B be are two palmprints of a user, r is a representative palmprint, S(A, r) and S(B, r) are the similarity scores of palmprint A and palmprint B with r respectively. This motivates us to use the similarity score of each palmprint as its index key and we use this key to make entry into a look up table (Figure 1). Let A be a palmprint and s is its key, then A is registered to s\(^{th}\) location of A. It can be observed from Figure 1 that, the palmprint IDs are organized in sorted order based on their keys like traditional database records. During identification, first we compute the key of the query by matching it with r. Next, using this key we retrieve the palmprints that are having similar key in the look up table (Figure 1). This is given in Equation 3.2:

\[
C + \{I_{id} \text{ List } | [S(q, r) = S(I_{id} \text{ List, r})]\}
\]  

(3.2)

where C is the set of candidates retrieved as similar matches, q is the query palmprint and I\(_{id}\)List is the palmprint IDs in the look up table, those are similar to q.

The list of similar palmprints for a query can be retrieved in O(1) time, as the palmprints in the look up table are already arranged in sorted manner based on their keys. However, this retrieved list contains some palmprints that are dissimilar (false matches) to q but have similar key. For example, alphabets X and Z have same distance (i.e., score) to Y in the sequence even though X and Z are different. These false matches can be filtered out by choosing more number of representative palmprints. For each representative, a look up table is created. The user palmprints are made and entry into each table according to their score (key) with the corresponding representative. During identification of a query palmprint, each look up table produces a various set of candidate palmprints as similar matches. We select the palmprints which are repeated in most of the sets as final candidates for the query.

Note that, this process filter out the false matches and increases the identification accuracy of the system.

### Algorithm 1 User enrolment or make an entry process

1: **INPUT:** palmprint x, representative image set \(R = \{r_1, r_2, ... r_k\}\)  
   Index space ISP ACE = \{Table\(_1\), Table\(_2\), ..., Table\(_k\)\}  
2: **OUTPUT:** Updated ISP ACE  
3: // Enroll or make entry palmprint x into the ISP ACE  
4: for each \(r_i \in R\) do  
5: \(s \leftarrow s(x, r_i) // S(x, r_i)\) is the similarity score (i.e., key) of x against \(r_i\) 
6: \(\text{Table}_i[s], I_{id}\)List \(\leftarrow x //\)make an entry x into \(I_{id}\)List at location s of Table\(_i\).  
7: end for  
8: **RETURN** Updated ISP ACE.

### 3.1 User enrolment or make an entry

Let \(R = \{r_1, r_2, ... r_k\}\) be the representative image set. We create a separate look up table Table\(_i\) for each \(r_i\), where \(1 \leq i \leq k\). The palmprints are enrolled into each Table\(_i\) based on their key (i.e., score) with the corresponding \(r_i\). The process of enrolling a
palmprint is described in Algorithm 1. Let \( a \) be a user palmprint to make an entry and let \( s = S(a, r_1) \) be the key of it with \( r_1 \). We access Table 1 and registered palmprint \( a \) into location \( s \). We repeat this process with other representative palmprints and register \( a \) into corresponding look up tables.

### 3.2 Retrieval of similar palmprints for a query

In this section, the proposed is discussed i.e retrieval technique that identifies a set of potential (i.e.; similar) candidates from its database for a query palmprint. To identify the similar candidates for a query palmprint \( q \), it is matched against each representative image and computes the corresponding index key. Let \( x (= \text{Key}_i) \) be the index key of the query against \( r_i \), we access \( x^{th} \) location in Table \( i \), and retrieve the \( I_{id} \) List from there. We repeat this process with all keys \([\text{Key}_i]_{i=1}^k\) of the query palmprint and retrieve the \( I_{id} \) List from respective Table \( i \). All the retrieved \( I_{id} \) Lists are stored in a temporary set. Finally, we count the number of occurrences (i.e., votes) of each palmprint ID and select the palmprints that are received more votes than a predefined threshold \( T \) as possible matches (i.e., candidate set).

Let \( N \) is the database size, a sequential search process takes \( O(N) \) time to identify a query. On the other hand, this technique compares the query with only \( k \) representatives and hence require \( O(k) \) time where \( k << N \).

### 3.3 Representatives

In this approach, selection of representatives is an important concern. The representative palmprints must be different from one another and should reflect the qualities of all classes of users in the database. In this work, we use a clustering technique to choose the representatives. Clustering partitions all user palmprints into set of groups such that palmprints in same group are equal or similar to each other whereas palmprints in different groups are dissimilar. As each group has similar palmprints, we select one palmprint from the group as its representative. Finally, the representative set contains palmprints that are from different classes of users.

![Fig 1. Sample palmprint image and its extracted SURF features](image-url)
also fixed. On the other hand, with Leader algorithm the new enrollments can be done dynamically and without affecting the existing system.

The Leader algorithm randomly chooses one of the palmprint as starting Leader and scans the database images one by one, and selects set of Leaders as cluster representatives. Every time, if a new palmprint to to make entry, it is assigned to the most similar cluster if exist. Otherwise, a new cluster is created and this palmprint becomes the Leader of that cluster.

4 Experiments

Experiments are conducted based on benchmark PolyU palmprint database. This database contains 7752 palmprints of 386 different users approximately 20 palmprints per user. We have chosen randomly 10 palmprints per user for training and leftout things for testing. We segmented the palmprints into 151 x 151 pixels. Figure 2 show a sample palmprint image and its extracted SURF features.

4.1 Performance measures

We evaluated the proposed system performance using three measures: Hit Rate (HR), Miss Rate (MR) and Penetration Rate (PR). HR shows the percentage of test palmprints to which the genuine palmprint is found in the candidate set.

\[ HR = \left( \frac{X}{M} \right) \times 100\% \] (4.1)

where X represents the query palmprints count to which the genuine match is found in the retrieved candidate set and M is the count of palmprints tested.

MR represents the percentage of test palmprints to which the genuine match is not found in the candidate set.

\[ MR = 100 - HR \] (4.2)

PR describes the average size of candidate set generated for a query palmprint by the indexing technique.

\[ PR = \left( \frac{1}{M} \sum_{i=1}^{M} \frac{|L_i|}{N} \right) \times 100\% \] (4.3)

where \( L_i \) represents candidate set size for \( i^{th} \) test palmprint, N is the total training palmprints taken and M shows the testing palmprints count.

4.2 Selection of parameters

The experimental results in this solution are involved with various parameters. In this proposed work, mainly the parameters with respect to Miss Rate and Penetration Rate have been considered. By considering threshold value the MR and PR parameters has been analysed. This section validates the optimal value for these parameters for which the system achieves its better performance. First, we validate the optimal value for the number of representatives. Next, we validate the selection rules for the representative images.

4.2.1 Number of representatives

The number of representatives i.e., k plays a primary role on the system performance. The low number of images may not represent the entire database. In contrast, very large number of images may leads to redundancy in the representative set. An optimal value should be selected so that the system achieves its better performance. An experiment is conducted by varying the number of representatives. As shown in Figure 3, the MR and PR of the system values decrease with k. Further, it can be seen that for the number of representatives greater than 130, the decrement in MR and PR is not that much significant. Hence, the optimal value for k is chosen as 130.

4.2.2 Selection rule for representatives

The representative's selection rule is also an important concern of the system performance. We experimented with three different selection rules: randomly selected k images, first k images and Leader clustering (Figure 3). It can be observed that clustering performs well than the other approaches. In other way, MR and PR of the system are low which shows the efficacy of the Leader
algorithm for the representative's selection. Leader algorithm is an incremental clustering algorithm generally used to cluster large data sets. This algorithm is order dependent and may form different clusters based on the order the data set is provided to the algorithm.

4.3 Results

Once the optimal parameter values are chosen, we conducted an experiment by varying the threshold value from 0 to 100 and recorded the MR and PR values for the proposed approach (Figure 5). It can be observed that there exists a trade-off between the PR and MR. The PR decreases with the threshold whereas MR increases.

![Fig 2. Indexing performance with the increasing representative set size (k)](image)

![Fig 3. Indexing performance with different selection rules](image)
Hence we have to choose an optimum threshold value where the system achieves its better performance. But the optimum threshold value depends on the application where we are using this system. For a critical application like boarder security, identifying accuracy (i.e., genuine identification) should be high (i.e., HR is 100% and MR is 0%). Hence for this type of applications, the optimum threshold value is chosen as threshold where the system can get MR is 0%. On the other side, for general applications both PR and MR should be less. Based on different threshold value the MR and PR parameters have been analysed in Figure 4.

Fig 4. PR and MR of the system at different thresholds

Hence for these applications, an optimum threshold is selected as the threshold where the two curves (i.e., PR=MR) intersect. It can be observed from Table 2 and Table 3, that MR=0%, the PR is 16.46%. In other terms, proposed approach searches only 16.46% of the database to identify a query with 100% accurate. Further, at MR=PR our approach searches only 6.48% and...
identifies the query with 93.51% accuracy.

<table>
<thead>
<tr>
<th>PR %</th>
<th>MR=PR</th>
<th>MR=0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR (%)</td>
<td>6.48%</td>
<td>16.46%</td>
</tr>
</tbody>
</table>

Further, we also compare our approach with SIFT (Scale Invariant Feature Transform) features which we are used in our earlier work. It can be observed from Figure 5 that the SURF features perform better than SIFT features.

<table>
<thead>
<tr>
<th>Approach</th>
<th>HR (%)</th>
<th>PR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed (14)</td>
<td>98.28%</td>
<td>—</td>
</tr>
<tr>
<td>Proposed (4)</td>
<td>100%</td>
<td>31.89%</td>
</tr>
</tbody>
</table>

### 4.4 Comparison with other approaches

We compare the proposed approach with technique which is also based on match scores and geometric hashing based method. In approach uses a vector estimated file to store the similar and matched scores. They have choses 170 images for the representative set and achieve only a maximum of 98.18% HR, for PolyU database. On the other side, our proposed solution used one thirty reference images and obtained 100% HR. In authors retrieved the key properties from the biometric images and it has been used the geometric hashing for indexing. But, the authors could gain a PR of 31.89% for PolyU database and shown the performance of various approaches.

### 4.5 Effectiveness for new enrolments

To enrol or register a new palmprint x to the existing system, we first compute its index keys against each representative and make an entry into the respective look up tables (see Section 3.1). Next, x is assigned to a closest cluster using its similarity score, if exists. Else, a new cluster is created and x is selected as its Leader. Further, a new look up table is created corresponding to this new representative and all palm prints are enrolled into it. This approach without effecting existing entries creates entry to the new users.

### 5 Concluding remarks

In this study, an efficient indexing technique is proposed for palm prints using SURF features and similarity scores. The defined look up table reduced the search time of the system by arranging the palm prints in sorted order. The use of Leader clustering for representative selection allows enrolling or creating the new users efficiently and without affecting the existing records. The retrieval time of the proposed approach is only O(k), where k is the number of representatives. On the other hand side, a linear search method requires O(N), here N is the database size and k ≪ N. The proposed could experiment on small database which contains of 7,753 images. This solution may able to show 100% accuracy when large databases considered. As a future work huge database can be considered using advanced techniques like BigData with suitable tools.

### References


https://www.indjst.org/