Distorted Fingerprint Verification System

Divya KARTHIKAESHWARAN[1], Jeyalatha SIVARAMAKRISHNAN[2]
[1]Department of Computer Science, Amrita University, Bangalore, India
[2]Department of Computer Science, BITS, Pilani, Dubai, UAE
kdivya_s@yahoo.co.in, jeylatha@yahoo.com

Fingerprint verification is one of the most reliable personal identification methods. Fingerprint matching is affected by non-linear distortion introduced in fingerprint impression during the image acquisition process. This non-linear deformation changes both the position and orientation of minutiae. The proposed system operates in three stages: alignment based fingerprint matching, fuzzy clustering and classifier framework. First, an enhanced input fingerprint image has been aligned with the template fingerprint image and matching score is computed. To improve the performance of the system, a fuzzy clustering based on distance and density has been used to cluster the feature set obtained from the fingerprint matcher. Finally a classifier framework has been developed and found that cost sensitive classifier produces better results. The system has been evaluated on fingerprint database and the experimental result shows that system produces a verification rate of 96%. This system plays an important role in forensic and civilian applications.

Keywords: *Biometric, Fingerprints, Distortion, Fuzzy Clustering, Cost Sensitive Classifier*

1 Introduction

Fingerprints have been used for over a century and are the most widely used form of biometric identification. The fingerprint of an individual is unique and remains unchanged over a lifetime. A fingerprint is formed from an impression of the pattern of ridges on a finger. Usually, fingerprint verification is performed manually by professional fingerprint experts. However, manual fingerprint verification is so tedious and does not meet the performance requirements of the new applications.

As a result, an automatic fingerprint identification system is in great demand. Fingerprint verification is usually associated with criminal identification, police work and it is now become more popular in civilian applications such as access control, financial security, and verification of firearm purchasers and driver license applicants.

The main difficulty in matching two fingerprint impressions of the same finger is to deal with the nonlinear distortions, which is caused during acquisition process. There are two main reasons contributed to the fingerprint distortion. First, the acquisition of a fingerprint is a three-dimensional/two-dimensional warping process. The fingerprint captured with different contact centers usually results in different warping mode. Second, distortion will be introduced in the fingerprint image by non-orthogonal pressure exerted by the people on the sensor.

There are different attempts to deal with the nonlinear distortions in fingerprint images. Some methods measured the forces and torques on the scanner directly with the aid of specialized hardware [4]. If excessive force is applied or the estimated distortion is too large, the captured fingerprint image will be deleted. However, this method does not work with the collected images.

An alignment-based matching algorithm was used, which aligns the associated ridges of template and input fingerprint image and an adaptive elastic matching algorithm was proposed to match the aligned minutiae [5]. However, this approach results in a large template size because the associated ridges for each minutia must be saved. If only short ridges are saved, the algorithm may results in an inaccurate alignment, or much worse, a false alignment. In order to overcome this problem another technique was introduced [3], where adaptive bounding box was used in the matching process. This method is more robust to nonlinear distortion.

A minutiae-based fingerprint matching algorithm uses distance normalization and local alignment to deal with the nonlinear distortion [6]. Here the ridge structures are thinned which helps in extracting minutiae. However, in order to tolerate the position change of the corresponding minutiae, the size of the bounding boxes has to be increased. As a side effect, this may lead to a higher false acceptance rate by wrongly pairing the non-matching minutiae.
A fuzzy feature match based on a local triangle feature set is used in [7] to match the deformed fingerprint. In this approach, the fingerprint is represented by the fuzzy feature set of local triangles. The similarity between the feature sets is characterized by the degree to which the template and input fingerprints are similar.
A normalized fuzzy similarity measure [1] is used to compute the similarity between the template and an input fingerprint image. In this method, a local topological structure is defined from which fuzzy feature vector is computed. A Cauchy function is used has a membership function to find the similarity between two fingerprint images.
In automated fingerprint identification, a clustering algorithm [8] is used to detect similar minutiae group from multiple template images generated from the same finger and create the cluster core set. The minutiae of test fingerprint are compared with cluster core set and decision is made whether it is from same finger or not. Another algorithm developed first estimated the local ridge frequency in the entire fingerprint and then converted a distorted fingerprint image into an equally ridge spaced fingerprint [9]. The stricter matching condition slightly decreases the algorithm performance, this method only solves a part of the deformations.
In the filter-based algorithm, a bank of Gabor filters is used to capture both the local and global details of the fingerprints in a compact fixed length finger code [12]. The fingerprint matching is based on the Euclidean distance between the two corresponding Finger Codes. Although verification rate is considerable, the process of matching is slower since a set of Gabor filter is involved.
A minutiae matching method that describes elastic distortion by means of a thin plate spline model [13] is developed and this model normalizes the shape of the test fingerprint with respect to the template. One of the fingerprints deformed and registered according to the estimated model and then matching minutiae are found. The drawback of this algorithm is that it detects only small elastic deformations.
Global features [14] shape a special pattern of ridges and valleys called Singular points. This algorithm makes uses of short Fourier transform analysis followed by calculating the orientation field reliability and locates the singular points.
In this paper an alignment based fingerprint matcher is presented. The fingerprint image is enhanced to improve the quality of the image which facilitates better minutiae extraction. The fingerprint images are aligned and matching score is computed. From the fingerprint matcher two features are taken. A fuzzy clustering based on distance and density is used to cluster these features into genuine and imposter cluster. Finally, a classifier framework is developed using bagging, boosting and cost sensitive classifier and their performance is evaluated.
In the following sections, details of distorted fingerprint verification system are provided. Section 2 gives a brief overview of the system design. Section 3 discusses about alignment based fingerprint matcher. Section 4 discusses about fuzzy clustering. Section 5 presents the classifier framework. Section 6 provides the experimental results and Section 7 contains the conclusion and future work.

2 Overview of the System

The distorted fingerprint verification system as shown in fig 1 operates in five stages namely, fingerprint enhancement, minutiae extraction, fingerprint matcher, fuzzy clustering and classifier framework.
In the fingerprint enhancement block, quality of an input grayscale fingerprint image is improved thereby facilitating better

extraction of reliable minutiae. A fast fingerprint enhancement algorithm [10] is used which can adaptively improve the clarity of ridge and valley structures of the input fingerprint image based on the estimated local ridge orientation and frequency.
The enhanced binary image is thinned and minutiae are extracted. The ridge ending and ridge bifurcation are the minutiae considered.
All minutiae may not be genuine and significant amount of spurious minutiae are detected during extraction process. Therefore a methodology is used to remove all the spurious minutiae detected in the minutiae extraction process. A minutia is considered to be spurious if it lies in the fingerprint border or forms small clusters in the middle area of the fingerprint image.

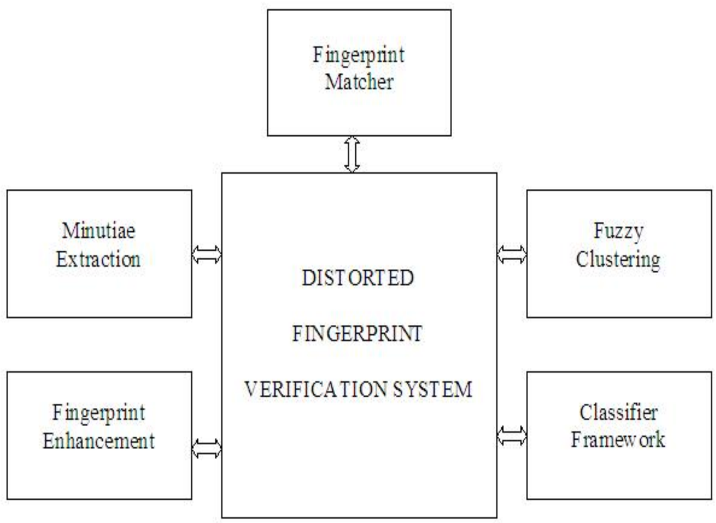


Fig. 1. System Architecture of Distorted fingerprint Verification System

In the fingerprint matcher block, an input fingerprint image is aligned with respect to the template fingerprint image and both the minutiae set are converted to the polar coordinate system. The two fingerprint images are matched using adaptive bounding box technique and match score is found.
An adaptive bounding box is more robust in matching distorted fingerprint images. Due to the presence of distortion, there is lot of vagueness in such fingerprint images.
Fuzzy clustering based on distance and density is used to group the fingerprint data set into two clusters namely, genuine cluster and imposter cluster. For the process of clustering two features are considered.
In the classifier framework block, a comparative study of different machine learning models is made and the most appropriate algorithm for the fingerprint dataset is selected.

3 Alignment Based Fingerprint Matcher

Alignment based fingerprint matcher [3] operates in two stages namely, alignment stage and matching stage.

3.1 Alignment Stage

In the alignment stage [3] transformation such as translation, rotation and scaling between an input and a template fingerprint image are estimated and input minutiae are aligned with the template minutiae accordingly to the estimated parameters. Each minutia in a fingerprint is associated with a ridge. It is clear that alignment can be achieved by aligning corresponding ridges. For each detected minutiae, the following parameters are recorded:

- x and y coordinates of the minutiae point.
- Orientation which is defined as the local ridge orientation of the associated ridge.
- The type of the minutiae point whether it

is ridge ending or ridge bifurcation.

- Associated ridge.

The associated ridge is sampled at regular intervals and entire ridge is not taken for the process of alignment. This speeds the process of alignment and reduces the unnecessary computation of distance and angle of minutiae.

Based on the distance and angle, the similarity of ridges is computed and the respective minutia on corresponding ridge is considered to be the reference points of template and input fingerprint image. The angle to which the input image has to be rotated with the template image is estimated.

To align the input minutiae set with the template minutiae set in the polar coordinate, all that needs to be done is to translate the input minutiae and the template minutiae to polar coordinate with respect to the reference minutiae and then add the angle rotate to the radial angle of the polar coordinate of every input minutia. The minutiae set are converted into polar coordinates using the formula:

$$r = ((x_i - x_r)^2+(y_i - y_r)^2)^{1/2}$$
$$e = \tan^{-1}((y_i - y_r)/(x_i-x_r))+\text{rotate} \quad (1)$$
$$\theta = \theta_i-\theta_r$$

where (x_r, y_r, θ_r) represents the coordinates of reference minutiae, (r, e, θ) represents the minutiae in polar coordinate system and rotate represents the angle to which input image as to rotated. Rotate can be computed using this formula:

$$\text{Rotate} = \text{dir\_temp} - \text{dir\_in} \quad (2)$$

where, dir_temp is the orientation of template minutiae and dir_in is the orientation of input minutiae.

Minutia matching in the polar coordinate has several advantages. It has been observed that the nonlinear deformation of fingerprints has a radial property. In other words, the nonlinear deformation in a fingerprint impression usually starts from a certain point (region) and nonlinearly radiates outward. Therefore, it is beneficial to model it in the polar space. At the same time, it is much easier to formulate rotation, which constitutes the main part of the alignment error between an input image and a template, in the polar space rather than in the Cartesian space.

3.2 Matching Stage

If two identical ridge patterns are exactly aligned with each other, each pair of corresponding minutiae is completely coincident. In such a case, a point pattern matching can be simply achieved by counting the number of overlapping pairs.

However, in practice, such a situation is not encountered. On the other hand, the error in determining and localizing minutia hinders the alignment algorithm to recover the relative pose transformation exactly.

Therefore, the aligned point pattern matching algorithm needs to be elastic which means that it should be capable of tolerating, to some extent, the deformations due to inexact extraction of minutia positions and nonlinear deformations. Usually, such an elastic matching can be achieved by placing a bounding box around each template minutia, which specifies all the possible positions of the corresponding input minutia with respect to the template minutia, and restricting the corresponding minutia in the input image to be within this box. The steps of minutiae matching algorithm is as follows:

1. Select a reference minutia point both in the template and input fingerprint image.
2. Convert each minutia point in the template minutia set and the input minutia set to the polar coordinate system as given in equation 1.
3. Represent the template and input minutiae in the polar coordinate system as symbolic strings by concatenating each minutia in the increasing order of radial angle.
4. Match the resulting strings using adaptive bounding box technique and compute the matching score.
5. If the matching score is higher than a threshold value, then the input image is considered to come from the same finger as the template image.

The purpose for using an adaptive bounding box [3] is to deal with nonlinear deformation more robustly. When the radius of the minutia is small, a small deformation will mean a large change of the radial angle while the change of radius remains small.
On the other hand, when the radius of the minutia is large, a small change in radial angle will cause a large change in the position of the minutia. While the radius can have larger deformation as it is the accumulation of deformation from all the regions between this minutia and the reference minutia.

4 Fuzzy Clustering Based on Distance and Density

Cluster analysis has been a fundamental research area in pattern recognition. Clustering helps to find natural boundaries in the data. Since the fingerprint is distorted there is lot of vagueness involved during matching.
To overcome this problem fuzzy clustering [2] technique is used to find vague boundaries of genuine and imposter cluster. In fuzzy clustering, the requirement of crisp partition of the data is replaced by a weaker requirement of fuzzy partition. For the process of fuzzy clustering, two features are selected.

- Number of matched minutiae points of the template and input fingerprint image
- The mean distance difference of the matched minutiae pairs.

Table 1 represents the sample fingerprint dataset where first three data are from image of same finger and the last three data are obtained from different finger. If the fingerprints are from different finger then the value of n will be small and d will be large.

Table 1. Sample Fingerprint dataset

No of matching minutiae (n)	Mean distance difference (d)
83	8.6
56	14
63	10
12	78
7	89
10	74

The fingerprint feature set is normalized in the range of [0, 1]. The dataset is clustered into two classes namely, genuine cluster and imposter cluster.
Genuine cluster represents the data of matched fingerprint images from same finger. Imposter cluster represents the fingerprint images from different fingers. In this technique two parameters have to be specified namely, density radius and select density. The rationale here is that the two data points whose distance is less than density radius and density difference is less than a threshold, is in the same cluster.
To compute the distance between the data point's x_i and x_j which has two attributes each is given in formula below,

$$D_{ij}= ((x_{i1}-x_{j1})^2+ (x_{i2}-x_{j2})^2)^{1/2} \qquad (3)$$

The maximal distance D_{max} and minimal distance D_{min} for the entire data set has to be computed. The value of the density radius chosen should be within the range of D_{min} and D_{max}. In order to satisfy this condition, the density radius r is calculated using the formula where value of α lies within the range of 0 to 1.

$$r = D_{min} + \alpha\,(D_{max} - D_{min}) \qquad (4)$$

Using the density radius r, density and density set of each data point in the feature set is calculated using the formula

$$\rho_i = \mathrm{Num}\,(D_{ij} <= r) \qquad (5)$$

The maximal density ρ_{max} and minimal density ρ_{min} for the data set is computed. The value of the select density chosen should be within the range of ρ_{min} and ρ_{max}. In order to satisfy this condition, the select density ρ is calculated using the formula where value of β lies within the range of 0 to 1.

$$\text{Select density} = \rho_{min} + \beta\,(\rho_{max} - \rho_{min}) \qquad (6)$$

The two data point's x_i and x_j are considered to be in the same cluster if the following condition is satisfied.

$$\rho_i - \rho_j <= \text{select density} \qquad (7)$$

where ρ_i is the density of x_i and ρ_j is the density of x_j. The un-chosen set represents the data points which have never been selected for processing. The various steps of the clustering algorithm is as follows

1. The number of clusters and un-chosen set is given has input to the process.
2. Calculate the distance between every two data points and find D_{min} and D_{max}.
3. Specify the density radius according to D_{min} and D_{max}..
4. Calculate the density of each data point and find ρ_{min} and ρ_{max}.
5. Specify the select density according to ρ_{min} and ρ_{max}.
6. Select a data point x_i which has never been chosen.
7. Select any data point x_j from x_j's density set and check for the same cluster condition
8. Delete x_i from the un-chosen set.
9. Select the next data point for clustering from the un-chosen set using the following criteria
   a. If intersection of un-chosen and density set is not empty, select the next data from the intersection set.
   b. Otherwise, select any data from un-chosen set.

5 Classifier Framework

The experience shows that no single machine learning scheme is appropriate for all data mining problems. In this section, different learning models are compared and the most appropriate algorithm is selected. The WEKA [11] is an open source data mining package which provides a collection of machine learning algorithms. During training, all base classifiers are evaluated by cross-validation on the fingerprint dataset.

The various classifiers used for comparative study for effective classification of fingerprint dataset are random forest, J48, NB tree, Bagging, adaboost and cost sensitive classifier.

Random forest is an ensemble decision trees that offers good predictive performance. Random forests construct a series of tree-based learners. Each base learner receives different training set which are drawn independently with replacement from the original learning set. Each tree predicts a class which is considered as a vote and the forest selects the class which as the most votes.

J48 tree is a reimplementation of C4.5 tree. NB Tree is a hybrid between decision trees and Naïve Bayes. It creates trees whose leaves are Naïve Bayes classifiers for the instances that reach the leaf. When constructing the tree, cross-validation is used to decide whether a node should be split further or a Naive Bayes model should be used instead.

Ensembles of classifiers are groups of classifiers in which the individual classifiers predictions are combined to classify new samples. Bagging and boosting is ensemble of classifiers used to improve the performance of classification.

In bagging, many bootstrap samples are drawn from the available data set, and some prediction method like decision tree is applied to each bootstrap sample, and then the prediction of individual classifiers are combined by simple voting to obtain the overall prediction. Here REP trees are used as base classifier. REP Tree builds a decision tree using information gain/variance reduction and prunes it using reduced-error pruning.

Boosting, like bagging, is a committee-based approach that can be used to improve the accuracy of classification. In adaboost, decision stump is used as base classifier. The Decision Stump is a one-level binary decision trees that can generate tree using categorical or numeric class. It deals effectively with missing values.

The cost-sensitive learning addresses the issue of classification in the presence of

varying costs associated with different types of misclassification. In particular, misclassification errors typically have non-uniform costs. These misclassification costs are often determined by the class associated with an example, such that for two-class problems, the cost of a false positive prediction is not equal to the cost of a false negative prediction.
Table 2 specifies the cost matrix used for cost-sensitive classifiers. For cost-sensitive learning, one typically specifies only the costs for the false positives (C10) and false negatives (C01) and assigns a cost of zero to the true positives (C11) and true negatives (C00).
Conceptually, the cost of labeling an example incorrectly should always be greater than the cost of labeling it correctly.

Table 2. Cost Matrix

	Actual Negative	**Actual Positive**
Predicted Negative	C00	C01
Predicted Positive	C10	C11

Table 3 illustrates the comparative results of various classifiers like bagging, Adaboost, J48, Random Forest, Cost sensitive classifiers and NB trees. The cost sensitive classifier produces a better result with errors minimized.

Table 3. Comparison of classifier performance

Comparative Measure	**Bagging**	**Adaboost**	**J48**	**Random Forest**	**Cost Sensitive**	**NB tree**
Percentage correct	95.06	95.06	93.87	94.03	96	91.23
Mean absolute error	0.08	0.08	0.09	0.09	0.07	0.13
RMSE	0.17	0.19	0.19	0.19	0.17	0.21
Precision	0.93	0.93	0.93	0.93	0.94	0.90
Recall	0.89	0.88	0.86	0.87	0.90	0.85
F_measure	0.82	0.84	0.85	0.85	0.82	0.80
Training time	0.02s	0.02s	0.01s	0.02s	0.02s	0.01s

6 Results and Discussion

The proposed system has been evaluated on the fingerprint database of fingerprint verification competition. The emphasis is on distorted and dry and wet fingerprints. These fingerprints were acquired through optical sensor.
For matching, 100 images captured from 20 different fingers, five images form each finger is used. The size of images is 640*480 pixels with a resolution of 512 dpi. The fingerprint matcher is evaluated using two parameters namely, verification rate and rejection rate.

Verification rate = correct-num / (correct-num + false-num) * 100 % (8)
Reject rate = reject-num /total number of matches * 100 % (9)
where correct-num denotes no of correct

matches obtained, reject-num denotes no of genuine images rejected and false-num denotes no of false matches produces. Table 4 represents the verification rate and rejection rate for the various input fingerprint images.

Table 4. Verification Rate and Reject Rate

Verification Rate	Rejection Rate
100%	11.1%
98.3%	9.1%
91.7%	7.3%
87.52%	5%
81%	3%
80%	2%

In order to evaluate the performance of the fuzzy clustering algorithm, the clustering results are compared with class labels of the dataset. The fuzzy clustering increases the performance of the alignment based algorithm to 93% of accuracy.

The performance of a biometric system can be represented as receiver operating characteristic (ROC) curve showing the genuine acceptance rate against false acceptance rate. Figure 2 illustrates the comparison of the performance of fingerprint matcher using alignment based algorithm and fuzzy clustering of fingerprint dataset.

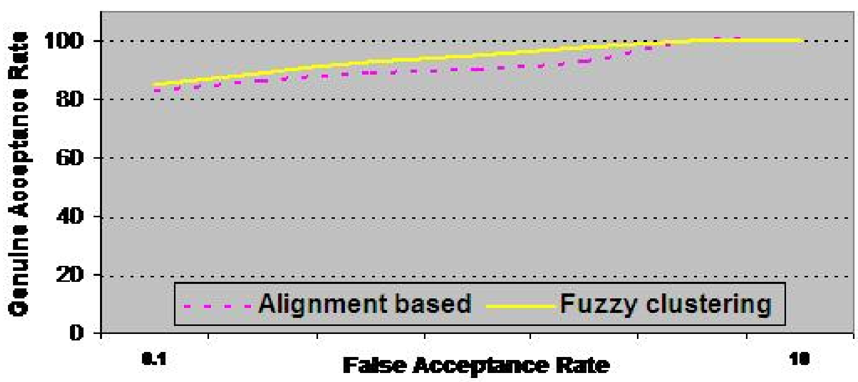


Fig. 2. Performance comparison of fingerprint matcher

7 Conclusion

The design and implementation of the distorted fingerprint verification system is presented in this paper. The performance of fingerprint identification system relies critically on the image quality. Hence, good quality images make the system performance more robust. However, it is always very difficult to obtain good quality images in practical use.

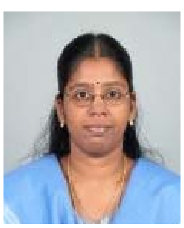
To overcome this problem, image enhancement step is performed to improve the clarity of ridges and valleys. An alignment based matching algorithm is used to match the fingerprint images. The vagueness present in the distorted fingerprint images is effectively handled by the fuzzy clustering based on distance and density technique. In order to improve the performance of the system, a classifier framework is developed and the cost sensitive classifier is found to have higher accuracy.

Based on the experimental results, it is observed that the matching errors in the proposed system mainly results from incorrect minutiae extraction and inaccurate alignment. Further improvements in the distorted fingerprint verification system can be made by using more ridge information like pores and ridge width in the matching process.

References

[1] X. J. Chen, J. Tian, and X. Yang, “A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure,” IEEE Transaction on

Image Processing Vol. 15, no. 3, pp. 767–776, Mar. 2006.
[2] X. Qiu, Y. Tang and Dan Meng, " A new fuzzy clustering method based on distance and density", IEEE Conference on System, Man & Cybernetics, Oct. 2002,Vol 7, pp 6- 9.
[3] X. P. Luo, J. Tian, and Y.Wu, "A minutia matching algorithm in fingerprint verification," in Proc. 15th ICPR, Sep. 2000, Vol. 6, pp. 1039–1045
[4] N. K. Ratha and R. M. Bolle, "Effect of controlled acquisition on fingerprint matching," in Proc. 14th ICPR, 1998, Vol. 2, pp. 1659–1661.
[5] A. K. Jain, L. Hong, and R. Bolle, "On-line fingerprint verification," IEEE Transaction on Pattern Analysis and Machine Intelligence*,* Vol. 19, no. 4, pp. 302–313, Apr. 1997.
[6] D. Lee, K. Choi, and J. Kim, "A robust fingerprint matching algorithm using local alignment," in Proc. 16th ICPR, Aug. 2002, Vol. 3, pp. 803–806
[7] X. Chen, J. Tian*,* X. Yang, and Y. Zhang, "An Algorithm for Distorted Fingerprint Matching Based on Local Triangle Feature Set," IEEE Transaction on Information Forensics & security, Vol. 1, pp. 169-177, June 2006
[8] R. Qun, T. Jie, H. Yuliang and C. Jiangang, "Automatic fingerprint identification using cluster algorithm", in Proc. 16th ICPR, Vol. 2, pp. 398-401, Aug. 2002
[9] A. Senior and R. Bolle, "Improved fingerprint matching by distortion removal," IEICE Trans. Information System, Special issue on Biometrics, Vol. E84-D, no. 7, pp. 825–831, Jul. 2001.
[10] L. Hong, Y. Wan, and A. K. Jain, "Fingerprint image enhancement: algorithms and performance evaluation," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, no. 8, pp. 777–789, Aug. 1998.
[11] The WEKA package, http://www.cs.waikato.ac.nz/ml/weka/
[12] A. K. Jain, S. Prabhakar, L**.** Hong, and S. Pankanti, "Filterbank-based fingerprint matching," IEEE Trans. Image Process., vol. 9, no. 5, pp. 846–859 , May 2000.
[13] A. M. Bazen and S. H. Gerez, "Fingerprint matching by thin-plate spline modeling of elastic deformations," Pattern Recognition., vol. 36, no. 8, pp. 1859–1867, 2003.
[14] M.S.Khalil1, D. Muhammad, M.K. Khan and K. Alghathbar, "Singular points detection using fingerprint orientation field reliability", International Journal of Physical Sciences Vol.5, pp 352-357, April 2010



Divya KARTHIKAESHWARAN holds a M.E., in Multimedia Technology from College of Engineering, Anna University, Chennai, India. She has 2 years of experience in software industry and 1.6 years of teaching experience in University. Her areas of interest include image processing, pattern recognition and data mining.

Jeyalatha SIVARAMAKRISHNAN holds a M.E., in Computer Science from Anna University, Chennai, India and a PhD student from BITS, Pilani-India. She has 10 years of teaching experience. Presently, she is working as Senior Lecturer, CS, BITS, Pilani-Dubai. Her areas of interest include Web Mining, Data Mining and Database Systems. She is a member of Computer Society of India.