International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.8 (2015), pp.263-272
http://dx.doi.org/10.14257/ijsip.2015.8.8.28

SIFT-Based Low-Quality Fingerprint LSH Retrieval and Recognition Method

Yunfei Zhong[1,2*] and Xiaoqi Peng[1]

[1] *School of Information Science and Engineering, Central South University, Changsha 410083, China*
[2] *School of Packaging Materials and Engineering, Hunan University of Technology, Zhuzhou 412007, China*
maczone@163.com, pengxq126@126.com

Abstract

Most of the existing fingerprint retrieval systems are based on the overall characteristics and detailed features of fingerprints, and their performance is poor in the cases of low-quality fingerprint images, such as incomplete fingerprint images. In order to improve the recognition speed, accuracy, and robustness of automated fingerprint recognition systems based on large-scale fingerprint databases, in this paper, we propose a fast fingerprint classification retrieval and identification method based on Scale Invariant Feature Transform (SIFT) and Local Sensitive Hash (LSH) algorithms. A method based on scale space theory extracts SIFT feature point descriptors of relatively high-quality fingerprint images in accordance with the principle of a greater matching contribution rate, uses multi-template image feature fusion technology to build a stable fingerprint feature template database, achieves the storage and retrieval of high-dimensional SIFT features using the LSH algorithm, and carries out matching progressively by level on the basis of SIFT's matching principle of close neighboring priority scale. Experimental results show that the proposed method has strong penetration, high retrieval efficiency, good recognition accuracy, and strong robustness, thereby providing a fast and efficient retrieval and matching mechanism for the automated recognition of the large-scale fingerprint database, with strong practicality.

Keywords: *SIFT; LSH; fingerprint recognition; scale space; matching contribution rate*

1. Introduction

Biometric recognition refers to the use of computer technology on the biometrics of an organism (generally refers particularly to people) to distinguish individual organisms. This has gradually become a publicly acceptable body authentication technology, in which the automated fingerprint recognition technology still occupies a very important position in the market [1]. In large-scale automatic recognition applications, a large database of fingerprint templates will severely affect the efficiency of the system [2]. The templates used in fingerprint recognition are not the fingerprint images collected but the key features extracted from the fingerprint images. Feature selection and feature representation show the characteristics of the fingerprint itself and are closely related to the specific classification, retrieval, and matching algorithms. Fingerprint recognition methods based on fingerprint features such as global and local features have two main shortcomings [3-5]: The first is due to the uneven distribution probability of the fingerprint classification in natural environments, resulting in reduced classification efficiency. If the number of categories is increased, then the distinction between the

ISSN: 2005-4254 IJSIP

classes is reduced, and the difficulty of accurate classification increases and may lead to non-unique classification results. Second, for the global features method, the use of the unified model makes it difficult to achieve matching on incomplete or deformed low-quality fingerprint images, while the local features method based on details is heavily dependent on the number of detailed features in the retrieved information.

Scale Invariant Feature Transform (SIFT), which is widely used in image retrieval, can extract the repeated image feature points and carry out effective description on the texture structure around each feature point. Literature [6] proposed a fingerprint matching method based on SIFT; literature [7] proposed a streamlined SIFT feature collection fingerprint indexing method; and literature [8-9] presented an incomplete fingerprint matching method based on SIFT.

In this study, SIFT is used on the feature quality of fingerprint images to carry out analysis and feature selection. The E^2LSH method is used for establishing fingerprint templates for database indexing, while querying, as long as the calculation of the position of the hash table and the collection of the fingerprint template to be matched can be selected. Then, the method uses the scale grouping principle to perform reduction recognition, thus providing a fast and efficient retrieval and matching mechanism for the automatic recognition of large-scale fingerprint databases.

2. SIFT Features Representation of Fingerprint Image

2.1. Brief Description of SIFT

SIFT is a local features descriptor of images based on the scale space that remains invariant with image scaling, rotation, and even affine transformation, proposed by David Lowe in 1999; more in-depth development and improvements were carried out in 2004 [10-11].

Mikolajczyk [12] conducted an invariance comparative experiment on ten types of local descriptors, which also includes the SIFT algorithm, and proved that SIFT and its expansion algorithm have the strongest robustness among the similar types of descriptors. To a certain extent, Mikolajczyk was able to solve the image mismatch problem caused by factors such as the target's rotation, zooming, panning, affine, projective transformation, illumination, target blocking, sundries scene, and noise.

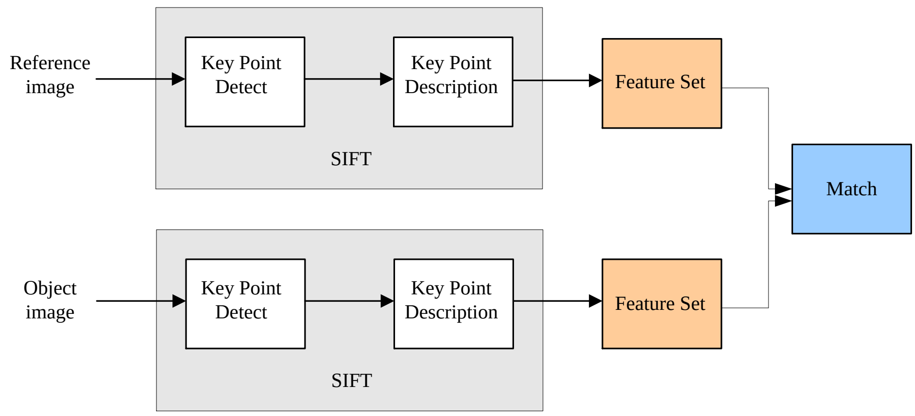


Figure 1. Feature Extract and Match of SIFT

As a local feature detection algorithm, SIFT obtains the image feature description by finding the feature points of an image (such as a point of interest or a corner) and the relevant scale and orientation of the descriptors, and used it for the image feature point matching. As shown in Figure 1, the use of the SIFT algorithm for achieving feature recognition has three major steps: extraction of feature points; additional descriptive information on the feature point (local feature), which means to identify the feature descriptor; and pairwise comparison on feature descriptors of two images to identify a number of feature points that match each other, and thereby establishes the corresponding relationship between the images.

2.2. Brief Description of SIFT

When the SIFT method is used for extracting feature points in a given scale space, some details in the fingerprint image will also be extracted as the SIFT features, as shown in Figure 2. Figure 2a is a fingerprint minutiae figure with less than 100 detail points; Figure 2b shows the SIFT feature points chart for a fingerprint with approximately 1000 feature points; Figure 2c is a preprocessed SIFT feature point figure with approximately 300 feature points.

(a) (b) (c)

Figure 2. Minutiae and SIFT features of fingerprint (Fingerprint image with a) minutiae, b) SIFT features and c) SIFT features being preprocessed.)

2.3. SIFT Adjacent Scale Priority Principle for Fingerprint Matching

In the scale space figure, the visual saliencies for different local invariant features are different, while the extraction and the description features should follow mechanisms similar to the selection principle using human visual attention [13].

During the transformation process of SIFT from fine scale to a relatively coarse scale, information is simplified and weakens gradually. It is impossible for the relatively coarse scale to produce features that are not in the fine scale. Experiments show that the characteristic scale where a SIFT feature of the fingerprint images is more than 90% is under $\delta = 2$. Simultaneously, during the SIFT matching of homologous fingerprint images, the average scale difference among the matching feature points are less than 0.2, as shown in Figures 3 and 4. Therefore, fingerprint matching based on SIFT should follow the adjacent scale priority principle; that is, the number of matching points within the same range of scales for two images to be matched is greater, and the scale difference between the matching points is minimum.

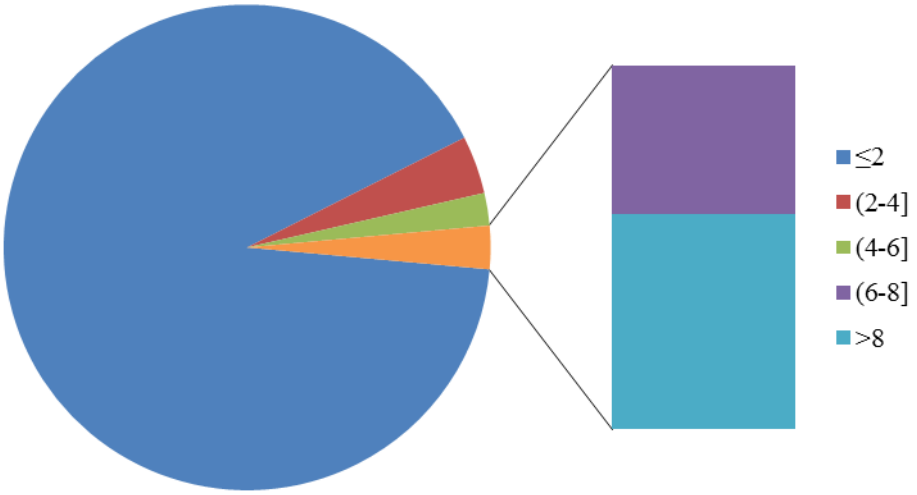


Figure 3. Scale Distribution Statistics for SIFT Features of Fingerprint

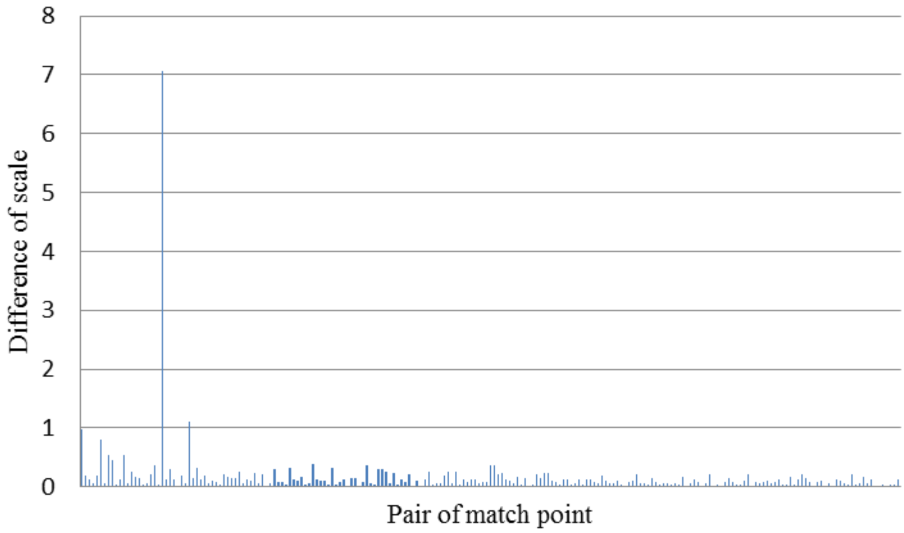


Figure 4. Scale Difference between SIFT Matching from Same Finger

2.4. SIFT Adjacent Scale Priority Principle for Fingerprint Matching

When there are more fingerprint feature points, the matching is very time consuming; therefore, in this study, we adopt the multi-fingerprint image features fusion method that is based on the matching contribution rate to construct a fingerprint template to streamline the SIFT features.

2.4.1. Preprocessing: Randomly select three homologous fingerprint images from the multiple images and compare them with the fingerprint images that did not preprocess and adopt a Gabor filtering and adaptive binarization process. The number of SIFT feature point descriptors obtained by the latter is approximately 40% of that obtained by the former. This significantly reduced the number of feature points to be matched and the complexity of the algorithm.

2.4.2. Mutual Cross Matching and Matching Contribution Rate: Extract the SIFT features of the selected three homologous fingerprints and conduct mutual cross checking to determine the number of feature points obtained from the three matching instances (pairwise exchange matching, take the one with higher matching feature points). Then, calculate the matching contribution rate η for these three fingerprints for every matching according to Formula (1):

$$\eta = \frac{M}{N} \times 100\% , \qquad (1)$$

where M is number of matched feature point pairs, and N is total number of feature points.

2.4.3. Creation of a Fingerprint Feature Template: According to the matching contribution rate of a fingerprint image, the quality of each image selected for this matching can be determined. While constructing the template, the feature scale of this selection should be used and the best-quality image should provide the basis. Meanwhile, taking into account that the other two images also contribute to this template selection, determine the SIFT feature points that constitute the fingerprint feature template as follows: ①feature point that still exists during the mutual matching of three homologous fingerprints;②feature point of the best-quality fingerprint among the three homologous fingerprints; and ③feature points of the other two fingerprint images.

3. LSH Retrieval and Matching

The data for the fingerprint feature template that was built using the SIFT method have high dimensionality; however, carrying out adjacent searching on it using a traditional indexing method will pose the problem of dimensionality. Therefore, similarity searches are generally carried out using K-adjacent or similar adjacent queries. Taking into consideration that the Location Sensitive Hash (LSH) algorithm proposed in literature [14] can achieve similar searching, classification, or clustering against the high-dimensional data, in this study, we use the LSH of the Euclidean space proposed in literature [15] to conduct Hash storage on a fingerprint template.

3.1. LSH

3.1.1. Algorithm Idea: Treat the element of the high-dimensional space as a point and assign a coordinate value to it; the coordinate value is a positive integer. Use the family of the hash function F to map all the points in the space to the n hash table, $n = |F|$. That is, every hash function $f \in F$ corresponds to a hash table and each hash table is stored along with all the points of the space. Given a subquery q, calculate $f_1(q)$, $f_2(q)$, …, $f_i(q)$, …, $f_n(q)$, $i = 1, 2, \cdots, n$. Considering all the points in the bucket of hash table T_i where all $f_i(q)$ fall into a candidate set, compare its distance with q and select point K with the shortest distance.

3.1.2. Algorithm Steps: Preprocessing. For any point $p \in P$, P as d-dimensional space, note that $p = \{x_1, x_2, \cdots, x_d\}$ and let the space P map to the d' Hamming dimensional space $H^{d'}$ as follows:

$p \to p' \in P$,

$$p' = Unary_c(x_1)Unary_c(x_2) \cdots Unary_c(x_d) . \qquad (2)$$

where $Unary_c(x_i)$ represents x_i 1s followed by $c - x_i$ 0s and c represents the maximum value of any coordinate p in the space. This mapping is distance preserving; that is, $\forall p, q \in P$, $d_1(p,q) = d_H(p,q)$, where d_1 represents the defined Euclidean distance under the guideline l_1 in space P and d_H represents the defined Hamming distance of space $H^{d'}$. Therefore, the problem to find point K that has the shortest distance with subquery q in space P is converted to the ε-NNS problem in space $H^{d'}$.

Building hash table: define $I=\{1,2,\cdots,d'\}$, take l subsets of I (l is a positive integer), denoted as $I_1,I_2,\cdots,I_l$. Define $p|I$ as the projection of vector p on the coordinate set I; that is, use every coordinate of the coordinate set as the location index, get the bit value of vector p corresponding to the location, and connect the result in series. Note that $g_j(p)=p|I$ stores all the points in space P into the hash bucket of the hash table by using hash function $g_j(p)$. The respective hash family has the total number of l hash tables, and the i-th hash table has k_i hash buckets, that is the hash table number i-th has mapped the points in space P to k_i different locations and has a total of $\sum_{i=1}^{l} k_i$ hash buckets.

Query: Given subquery q, calculate $g_1(q),\cdots,g_l(q)$, take all the points in the hash bucket that correspond to $g_i(q)$ (i.e., collection of points p that satisfy $g_i(q)=g_i(p)$) as the candidate set, and select k points that are nearest to subquery q from the candidate set.

3.2. Retrieval Process

The SIFT-based fingerprint retrieval and recognition process is as follows:

3.2.1. Fingerprint Template Storage: Construct a hash index table according to the method mentioned earlier. Use the family of hash function F to map the SIFT feature descriptor of each fingerprint template to n hash tables T_i, $n=|F|$; that is, each hash function $f\in F$ should correspond to a hash table, and every hash table should store all the SIFT descriptors of the fingerprint template space.

3.2.2. Fingerprint Queries: ①Carry out preprocessing of the fingerprint images to be queried.②Extract the SIFT features of the fingerprint images to be queried.③Carry out classification of the SIFT feature descriptors of the fingerprint images to be matched according to scale, and form a SIFT feature template for the fingerprint to be queried.④ For a given subquery q, calculate $f_1(q)$, $f_2(q)$, …, $f_n(q)$, $f_i\in F$, $i=1,...,n$, respectively. Considering all the points in the bucket of hash table T_i where all $f_i(q)$ fall into the candidate set, compare its distance with q and select the set of features of point K that has the shortest distance.

3.3. Matching Strategy

After retrieving a template image that is similar to the fingerprint image to be queried, execute matching between classes as follows:

①Carry out matching between classes in accordance with the scale classification marks. If the cumulative matching points reach the threshold value, which means that the matching is successful. End this query and return to the query result. If it has yet to reach the predetermined threshold value, the matching is continued until cumulative matching points reach the threshold value or all matching is unsuccessful.

②If all the matching points of ① are unsuccessful, take the next fingerprint template of the current candidate set to conduct matching, until the confirmation of success or failure, and then, end the fingerprint recognition process.

4. Experiment Results and Performances Analysis

Fingerprint Verification Competition (FVC) is the fingerprint algorithm competition organized by International Association of Pattern Recognition (IAPR), where FVC2002 is one of the fingerprint authentication databases used in the competition [16]. In this study, we selected FVC2002 DB1 as the experimental subject, on which many published results are based; it has a total of 100 sets of fingerprint images, each set containing 8 homologous fingerprints. The sizes of images are 364 × 256.

4.1. Retrieval Performance

Two main indications of the retrieval performance are accuracy and efficiency. The retrieval accuracy is calculated by the percentage of the input fingerprints whose corresponding ones in the database to be correctly retrieved. The retrieval efficiency is indicated by a so-called penetration rate, which is the average percentage of fingerprints in the database retrieved over all input fingerprints.

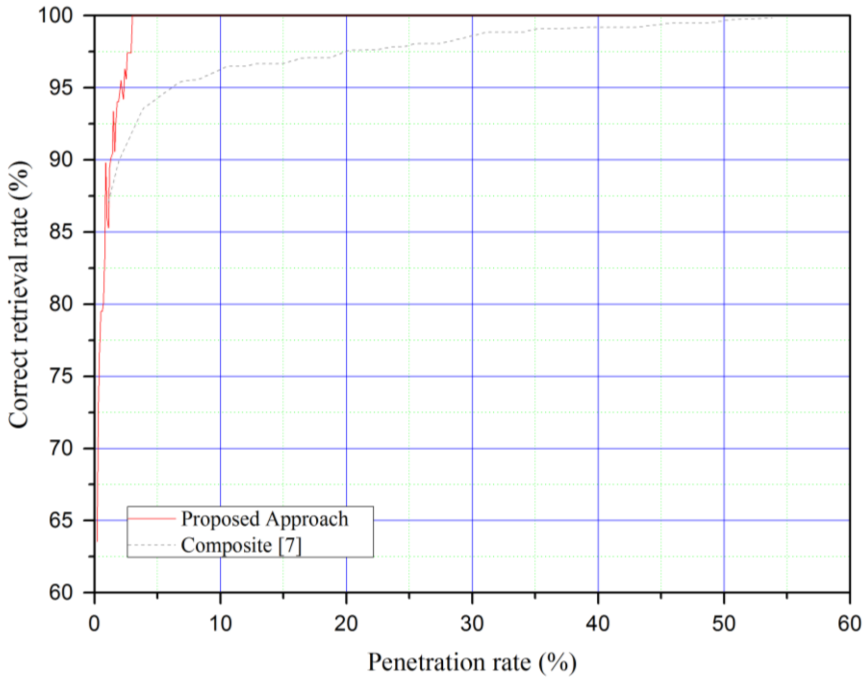


Figure 5. Comparison of Retrieval Performances

In the process of building the template, three images were selected randomly from each set of fingerprint images as described in section 2.4.3, and the remaining 500 pieces were used for testing the performance of the proposed algorithm. Figure 5 shows the experimental results via these two approaches on FVC2002 DB1. One is “Composite” approach (using composite set of key points from three impressions for index construction) in literature [7], the other is proposed in this paper. The proposed approach is much better than the “Composite”. Our approach can significantly reduce the number of queries and improve the recognition accuracy rate, which means the method of building the template according to the matching contribution rate is more efficiently.

4.2. Matching Performance

Matching between classes is performed according to the retrieval result obtained using the error rate-rejection rate to represent the matching performance. It can be seen from Figure 6 that the performance of the proposed method is between that of the single SIFT method and that of the minutiae fusion method, is better than that of the method described in literature [6].

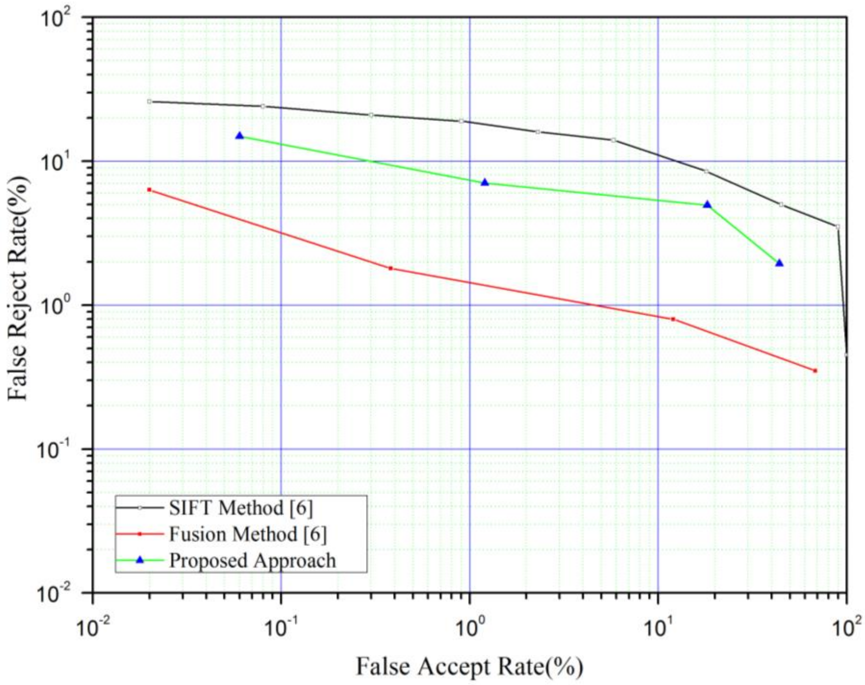


Figure 6. Comparison of Retrieval Performances

5. Conclusion

The development of a rapid fingerprint recognition technology for a large-scale fingerprint database has attracted considerable attention in the field of fingerprint recognition research. An efficient classification retrieval and recognition method is the key. The SIFT-based fingerprint retrieval and recognition method proposed in this paper increased the recognition efficiency of an automatic fingerprint recognition system as the target by assessing the quality of fingerprint features and adopted a multi-template image feature fusion technology to build a stable fingerprint feature template database to lay the foundation for efficient fingerprint retrieval and matching. It used the LSH algorithm to achieve SIFT storage retrieval with high-dimensional features and followed gradual matching by the level method of SIFT’s matching principle of close neighboring priority scale to increase the matching speed. Experimental results revealed that the proposed method had good penetrating ability, high retrieval efficiency, and good robustness; provides a fast and efficient retrieval mechanism for the automatic recognition of a large-scale fingerprint database; and has a relatively strong practicality.

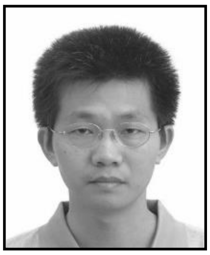
Acknowledgements

This work was supported by Natural Science Foundation of Hunan Province (Grant No. 10JJ2048 and 12JJ9043), National Natural Science Foundation of China (Grant No. 61134006) and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 61321003).

References

[1] T. Jie and Y. Xin, "Theories and Applications of Biometrics Recognition", Tsinghua University Press, Beijing, **(2009)**.
[2] D. Maltoni, D. Maio, A. K. Jain and S. Prabhakar, "Handbook of Fingerprint Recognition", Springer", London**, (2009)**.
[3] J. Tian, X. Chen and Y. Zhang, "New Progress in Fingerprint Recognition Technology", Progress in Natural Science, vol. 16, **(2006)**, pp. 400-408.
[4] N. Yager and A. Admin, "Fingerprint Classification: A Review", Pattern Analysis and Application, vol. 7, **(2004)**, pp. 77-93.
[5] H. Chen, J. Yin and E. Zhu, "A Method to Adjust Minutiae Location and Direction in Nonlinear Distorted Fingerprint Image", Journal of Computer Research and Development, vol. 47, **(2010)**, pp. 2141-2148.
[6] U. Park, S. Pankantia and A. K. Jain, "Fingerprint Verification Using SIFT Features", Proceedings of SPIE on Biometric Technology for Human Identification, (**2008)**, March 18-19; Orlando, Florida, USA.
[7] X. Shuai, C. Zhang and P. Hao, "Fingerprint Indexing Based on Composite Set of Reduced SIFT Features", Proceedings of 19th International Conference on Pattern Recognition, **(2008)**, December 8-11; Tampa, Florida, USA.
[8] S. Malathi and C. Meena, "Partial Fingerprint Matching Based on SIFT Features", Proceedings of International Journal on Computer Science and Engineering, vol. 2, **(2010)**, pp. 1411-1414.
[9] S. Malathi and C. Meena, "Improved Partial Fingerprint Matching Based on Score Level Fusion Using Pore and SIFT Features", Proceedings of International Conference on Process Automation, Control and Computing, **(2011)**, July 20-22; Coimbatore, India.
[10] D. Lowe, "Object Recognition from Local Scale-Invariant Features", Proceedings of the International Conference on Computer Vision, **(1999)**, September 20-25; Corfu, Greece.
[11] D. Lowe, "Distinctive Image Features from Scale-Invariant Key Points", International Journal of Computer Vision, vol. 60, **(2004)**, pp. 91-110.
[12] K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, **(2005)**, pp. 1615-1630.
[13] J. Li, M. D. Levine, X. An, X. Xu and H. He, "Visual Saliency Based on Scale-Space Analysis in The Frequency Domain", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, **(2013)**, pp. 996-1010.
[14] A. Gionis, P. Indyk and R. Motwani, "Similarity Search in High Dimensions Via Hashing", Proceedings of International Conference on Very Large Data Bases, **(1999)**, September 7-10; Edinburgh, Scotland, UK.
[15] A. Andoni and P. Indyk, "Near-Optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions", Communications of the ACM, vol. 51, **(2008)**, pp. 117-122.
[16] D. Maio, D. Maltoni, J. L. Wayman and A. K. Jain, "FVC2002: Second Fingerprint Verification Competition", Proceedings of 16th International Conference on Pattern Recognition, **(2002)**, August 11-15; Quebec, Canada.

Authors



Zhong Yunfei, he received B.Sc. in 1998 from Zhuzhou Institute of Technology and M.Sc. in 2005 from National University of Defense Technology. Now he is a Ph.D. candidate in Central South University. His research direction is image process and pattern recognition.

Peng Xiaoqi, he received B.Sc. in 1983 from Chongqing University, M.Sc. in 1988 from Harbin Institute of Technology and Ph.D. in 1998 from Central South University of Technology. Now he is a professor in Hunan First Normal University and Ph.D. supervisor in Central South University. His research direction is optimization, decision-making, control of complex industrial process and pattern recognition.