Application of Wave Atoms Decomposition and Extreme Learning Machine for Fingerprint Classification

Abdul A. Mohammed, Q.M. Jonathan Wu, and Maher A. Sid-Ahmed

Department of Electrical Engineering,
University of Windsor, Ontario, Canada
{mohammea,jwu,ahmed}@uwindsor.ca

Abstract. Law enforcement, border security and forensic applications are some of the areas where fingerprint classification plays an important role. A new technique based on wave atoms decomposition and bidirectional two-dimensional principal component analysis (B2DPCA) using extreme learning machine (ELM) for fast and accurate fingerprint image classification is proposed. The foremost contribution of this paper is application of two dimensional wave atoms decomposition on original fingerprint images to obtain sparse and efficient coefficients. Secondly, distinctive feature sets are extracted through dimensionality reduction using B2DPCA. ELM eliminates limitations of classical training paradigm; trains data at a considerably faster speed due to its simplified structure and efficient processing. Our algorithm combines optimization of B2DPCA and the speed of ELM to obtain a superior and efficient algorithm for fingerprint classification. Experimental results on twelve distinct fingerprint datasets validate the superiority of our proposed method.

1 Introduction

Biometric verification has received considerable attention during the last decade due to increased demand for automatic person categorization. Automated classification of an individual based on fingerprints is preferred since they are less vulnerable to be copied, stolen and lost [1] and due to their uniqueness and stability [2,3]. Fingerprint detection is a technology that has been widely adopted for personal identification in many areas such as criminal investigation, access control and internet authentication.

Fingerprint classification algorithms are classified into two categories; local and global. Local feature based methods include correlation, minutiae and ridge feature based matching algorithms. Global features are obtained using mathematical transforms; a classifier compares energies of different fingerprints and classifies them based on the global trends. Correlation-based techniques utilize gray level information of an image and take into account all dimensional attributes of a fingerprint, thereby providing enough image resolution. These techniques have been successfully applied for fingerprint classification [5] but they suffer from higher computational cost. Minutiae based techniques [7] extract minutiae from two fingerprints and store them as sets of points in a two dimensional plane and execute matching by generating an alignment between the template and the input minutiae set that result in maximum

A. Campilho and M. Kamel (Eds.): ICIAR 2010, Part II, LNCS 6112, pp. 246–255, 2010.

pairings. In low quality fingerprint images minutiae extraction becomes extremely difficult and thus ridge patterns [4] are reliably extracted for classification.

Researchers have also used fast Fourier transforms (FFT) and multi-resolution analysis tools that extract global features from fingerprint images for classification. Fitz and Green [8] used a hexagonal fast Fourier transform (FFT) to transform fingerprint images into frequency domain and employed a ''wedge-ring detector'' to extract features. A fingerprint classifier based on wavelet transform and probabilistic neural network is proposed in [9]. Wilson et al. [10] developed a Federal Bureau of Investigation (FBI) fingerprint classification standard that incorporates a massively parallel neural network structure. Other neural network classification schemes, using self organizing feature map, fuzzy neural networks, radial basis function neural network (RBFNN) and ellipsoidal basis function neural networks (EBFNN) have also been proposed [11].

In this work, we present a fast and accurate fingerprint classification algorithm that extracts sparse fingerprint representation using wave atoms decomposition; these coefficients are dimensionally reduced using bidirectional two-dimensional principal component analysis (B2DPCA). An extreme learning machine (ELM) classifier, based on a fast single hidden layer feedforward neural network (SLFNN), is trained and tested using dimensionally reduced extracted features. The proposed classification algorithm requires less human interventions and can run at thousand folds faster learning speed than conventional neural networks. ELM determines network parameters analytically, avoids trivial human intervention and makes it efficient for online applications.

The remainder of the paper is divided into 5 sections. Section 2 discusses wave atoms decomposition, followed by a discussion of B2DPCA in section 3. ELM algorithm for classification is discussed in section 4 and the proposed method is described in section 5. Experimental results are discussed in section 6.

2 Wave Atoms Decomposition

Wave atoms [12] are the most recent mathematical transforms for harmonic computational analysis. They are a variant of 2D wavelet packets that retain an isotropic aspect ratio. Wave atoms encompass a sharp frequency localization that cannot be achieved using a filter bank based on wavelet packets and offer a significantly sparser expansion for oscillatory functions than wavelets, curvelets and Gabor atoms. Wave atoms capture coherence of a pattern across and along oscillations whereas curvelets capture coherence only along the oscillations. Wave atoms precisely interpolate between Gabor atoms and directional wavelets in the sense that the period of oscillations of each wave packet (wavelength) is related to the size of essential support via parabolic scaling i.e. wavelength is directly proportional to square of the diameter.

Two distinct parameters α; indexing multiscale nature, and β representing directional selectivity are adequate for indexing all known forms of wave packet architectures namely wavelets, Gabor, ridgelets, curvelets and wave atoms. The triangle formed by wavelets, curvelets and wave atoms, as shown in the Fig. 1, indicates the wave packet families for which sparsity is preserved under transformation. Wave atoms are defined for $\alpha=\beta=1/2$, where α indexes the multiscale nature of the transform, from α = 0 (uniform) to α = 1 (dyadic). β measures the wave packet's directional selectivity (0 and 1 indicate best and poor selectivity respectively). Wave atoms represent a class of wavelet packets where directionality is sacrificed at the expense of preserving sparsity of oscillatory patterns under smooth diffeomorphisms. Essential support of wave packet in space (left) and in frequency (right) is shown in Fig. 2.

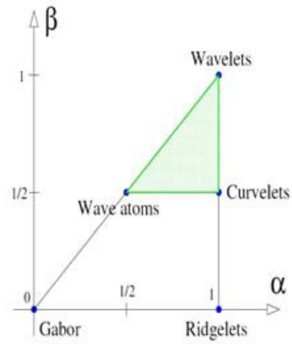


Fig. 1. Comparison of different wave packets architectures with respect to multiscale nature and directional selectivity [12]

2.1 1D Discrete Wave Atoms

Wave atoms are constructed from tensor products of adequately chosen 1D wave packets. Let $\psi^{j}_{m,n}(x)$ represent a one-dimensional family of wave packets, where $j, m \geq 0$, and $n \in Z$, centered in frequency around $\pm\omega_{j,m} = \pm\pi 2^{j} m$, with $C_1 2^j \leq m \leq C_2 2^j$ and centered in space around $x_{j,n} = 2^{-j} n$. Dyadic scaled and translated versions of $\hat{\psi}^{0}_{m}$ are combined in the frequency domain and the basis function is written as:

$$\psi^{j}_{m,n}(x) = \psi^{j}_{m}(x - 2^{-j} n) = 2^{j/2} \psi^{0}_{m}(2^{j} x - n). \tag{1}$$

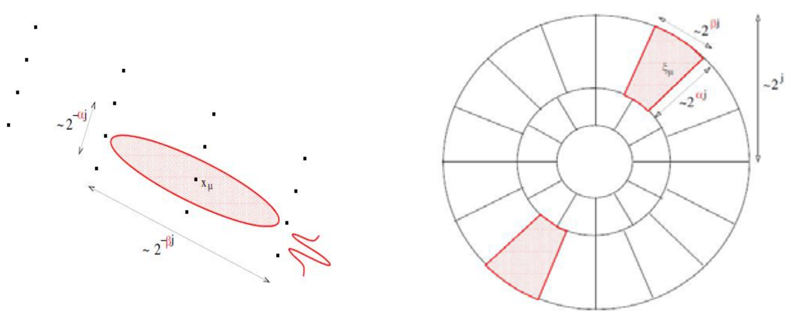


Fig. 2. Wave atoms tiling in space and frequency [12]

The coefficients $c_{j,m,n}$, for each wave $w_{j,m,n}$, are obtained as decimated convolution at scale 2^{-j}. Input sample u is discretized at $xk=kh$, $h=1/N$, $k=1,\ldots N$, and discrete coefficients $c^D_{j,m,n}$ are computed using a reduced inverse FFT inside an interval of size $2^{j+1}\pi$ centered about origin:

$$c^D_{j,m,n} = \sum_{k=2\pi(-2^j/2+1:2^j/2)} e^{i2^{-j}nk} \sum_{p\in 2\pi Z} \overline{\hat{\psi}^j_m(k+2^j p)}\,\hat{u}(k+2^j p). \tag{2}$$

A simple wrapping technique is used for the implementation of discrete wavelet packets and the steps involved are:

1. Perform an FFT of size N on the samples of $u(k)$.
2. For each pair (j,m), wrap the product $\overline{\hat{\psi}^j_m}\hat{u}$ by periodicity inside the interval $[-2^j\pi, 2^j\pi]$ and perform an inverse FFT of size 2^j to obtain $c^D_{j,m,n}$.
3. Repeat step 2 for all pairs (j,m).

2.2 2D Discrete Wave Atoms

A two-dimensional orthonormal basis function with 4 bumps in frequency plane is formed by individually taking products of 1D wave packets. 2D wave atoms are indexed by $\mu=(j,m,n)$, where $m=(m_1,m_2)$ and $n=(n_1,n_2)$. Construction is not a simple tensor product since there is only one scale subscript j. This is similar to the non-standard or multi-resolution analysis wavelet bases where the point is to enforce same scale in both directions in order to retain an isotropic aspect ratio. In 2D eq. (1) is modified accordingly.

$$\varphi^+_\mu(x_1,x_2) = \psi^j_{m1}(x_1-2^{-j}n_1)\,\psi^j_{m2}(x_2-2^{-j}n_2). \tag{3}$$

Combination of (3) and its Hilbert transform provides basis functions with two bumps in the frequency plane, symmetric with respect to the origin and thus directional wave packets oscillate in a single direction.

$$\varphi^{(1)}_\mu = \frac{\varphi^+_\mu + \varphi^-_\mu}{2},\quad \varphi^{(2)}_\mu = \frac{\varphi^+_\mu - \varphi^-_\mu}{2} \tag{4}$$

$\varphi^{(1)}_\mu$ and $\varphi^{(2)}_\mu$ together form the wave atoms frame and are jointly denoted by φ_μ. Wave atoms algorithm is based on the apparent generalization of the 1D wrapping strategy to two dimensions.

3 Bidirectional Two Dimensional Principal Component Analysis

Principal Component Analysis (PCA) is a data representation technique widely used in pattern recognition and compression schemes. In the past researchers used PCA and bunch graph matching techniques for enhanced representation of face images. PCA cannot capture even a simple variance unless it is explicitly accounted in the training data. In [13] Yang et al. proposed two dimensional PCA for image representation. As opposed to PCA, 2DPCA is based on 2D image matrices rather than 1D vector so the image matrix does not need to be vectorized prior to feature extraction. Instead an image covariance matrix is computed directly using the original image matrices.

Let X denote a q dimensional unitary column vector. To project a p x q image matrix A to X; linear transformation $Y=AX$ is used which results in a p dimensional projected vector Y. The total scatter of the projected samples is characterized by the trace of the covariance matrix i.e. matrix of the projected feature vectors, $j(X)=tr(S_x)$, where $tr(\ )$ represents the trace of S_x, and S_x denotes covariance matrix of the projected features.

$$S_x = E(Y-E(Y))(Y-E(Y))^T = E[(A-EA)X][(A-EA)X]^T. \quad (5)$$

$$tr(S_x) = X^T[E(A-EA)^T(A-EA)]X. \quad (6)$$

$G_t = E[(A-EA)^T(A-EA)]$ is q x q nonnegative image covariance matrix. If there are M training samples, the j^{th} image sample is denoted by p x q matrix A_j.

$$G_t = \frac{1}{M}\sum_{j=1}^{M}\left(A_j-\overline{A}\right)^T\left(A_j-\overline{A}\right). \quad (7)$$

$$J(X) = X^T G_t X, \quad (8)$$

where $\overline{A}$ represents an average image of all the training samples. Above criterion is called the generalized total scatter criterion. The unitary vector X that maximizes the criterion $j(X)$ is called the optimal projection axes. An optimal value represents a collection of d orthonormal eigen vectors $X_1,X_2,\ldots.X_d$ of G_t corresponding to d largest eigen values. A limitation of 2DPCA based dimension reduction is the processing of higher number of coefficients since it works along row directions only. Zhang and Zhou [14] proposed $(2D)^2$ PCA based on the assumption that training sample images are zero mean, and image covariance matrix can be computed from the outer product of row/column vectors of images.

4 Extreme Learning Machine

Feedforward neural networks (FNNs) are widely used in classification techniques due to their approximation capabilities for non-linear mappings. Slow learning speed of FNNs is a major bottleneck encountered, since input weights and hidden layer biases are updated using a parameter tuning approach such as gradient descent algorithm. Huang et al. [16] proposed an extremely fast learning algorithm called ELM for training a SLFNN. ELM randomly assigns input weights and hidden layer biases if the hidden layer activation function is infinitely differentiable. In ELM a learning paradigm is converted to a simple linear system whose output weights are analytically determined through a generalized inverse operation of the hidden layer weight matrices.

An N dimension random distinct sample (x_i,t_i) where $x_i=[x_{i1},x_{i2},\ldots.x_{in}]^T\in \Re^n$ and $t_i=[t_{i1},t_{i2},\ldots.t_{im}]^T\in \Re^m$, ELM with L hidden nodes and an activation function $g(x)$ is modeled as:

$$\sum_{i=1}^{L} \beta_i g_i(x_j) = \sum_{i=1}^{L} \beta_i g_i(w_i.x_j + b_i) = o_j,\ j = \{1,2,...N\}\ , \tag{9}$$

Where $w_i=[w_{i1},w_{i2},....w_{in}]^T$, $\beta_i=[\ \beta_{i1},\ \beta_{i2},....\beta_{iL}]^T$ represent weight vectors connecting input nodes to an i^{th} hidden node and from the i^{th} hidden node to all output nodes. b_i indicates threshold for i^{th} hidden node whereas $w_i.x_j$ represents an inner product of w_i and x_j. An ELM can reliably approximate N samples with zero error.

$$\sum_{i=1}^{L} \beta_i g_i(w_i.x_j + b_i) = t_j,\ j = \{1,2,...N\}\ . \tag{10}$$

Eq. (10) is modified as $H\beta=T$, $H=(w_1,...,w_L,b_1,...,b_L,x_1,...,x_N)$, such that i^{th} column of H is the output of i^{th} hidden node with respect to inputs $x_1,x_2,....x_N$. If the activation function $g(x)$ is infinitely differentiable, it is proved that the number of hidden nodes are such that $L<<N$. Training of SLFNN requires minimization of an error function E.

$$E = \sum_{j=1}^{N}\left(\sum_{i=1}^{L} \beta_i g(w_i x_j + b_i) - t_j\right)^2 = E = \|H\beta - T\|\ . \tag{11}$$

H is determined using gradient descent and the weights w_i, β_i and bias parameters b_i are tuned iteratively with a learning rate ρ. A small value of ρ causes the learning algorithm to converge slowly whereas a higher rate leads to instability and divergence to local minima. To avoid these limitations, ELM incorporates a minimum norm least-square solution, and instead of tuning the entire network parameters a random allocation of input weights and hidden layer biases help to analytically determine the hidden layer output matrix H and curtail the problem to a least-square solution of $H\beta=T$. H is a non-square matrix, the norm least-square solution of the above linear system becomes $\beta=H*T$, where H^* is the moore-penrose generalized inverse of H. The above relationship holds for a non-square matrix H whereas the solution is straightforward for $N=L$. An infinitely small training error is achieved using the above model since it represents a least-square solution of the linear system.

$$\left\|H\hat{\beta} - T\right\| = \left\|HH^*T - T\right\| == \min_{\beta}\|H\beta - T\|. \tag{12}$$

5 Proposed Fingerprint Classification Algorithm

The proposed classification scheme is independent of fingerprint patterns and is based on individual features and the number of trained fingerprint classes. Table 1 consists of detailed steps that demonstrate our proposed technique. Our system classifies fingerprint images into one of the trained classes; therefore, only one verification process is required per image. Our proposed scheme deals with classification of fingerprint images using ELM design and utilizes dimensionally reduced feature vectors obtained from wave atoms decomposition. Wave atoms decomposition is used for sparse representation of fingerprint images since they belong to a category of images that oscillate smoothly in varying directions. Discrete 2D wave atoms decomposition is applied on the original fingerprint image to efficiently capture coherence patterns along and across the oscillations. Fingerprint images are digitized using 256 gray levels therefore a transformation in color space is not required. Dimension of fingerprint images within each database were reduced to *64×64* prior to wave atoms decomposition. Image resizing was the only pre-processing performed on all datasets to minimize computations and to guarantee uniformity with other methods used for comparison. An orthonormal basis is used instead of a tight frame since each basis function oscillates in two distinct directions instead of one. This orthobasis variant property is important in applications where redundancy is undesired.

In addition to the aforementioned alterations there were no further changes made to the images as it may lead to image degradation. We randomly divide image database into two sets namely training set and testing set. All images within each database have the same dimension, i.e. $R \times C$. Similar image sizes support the assembly of equal sized wave atoms coefficients and feature vector extraction with identical level of global content. 2D wave atoms decomposition of every image is computed and coefficients are saved as initial feature matrix. Wave atoms decomposition is a relatively new technique for multiresolution analysis that offers significantly sparser expansion, for oscillatory functions, than other fixed standard representations like wavelets, curvelets and Gabor atoms.

Table 1. Outline of our Proposed Classification Scheme

INPUT: Randomly divide image database into two subsets TR_i and TE_j where i={1,2,…,n} and j={1,2,…,m} representing training and test image sets respectively.
OUTPUT: Classifier - f(x)

1. Resize fingerprint images from all databases to RxC.
2. Compute the wave atoms decomposition of each training and test images and extract feature sets. Each feature set is of dimension RxC. (Refer to section 2 for details of wave atoms decomposition)
3. Calculate image covariance matrix of test and train images to obtain intermediate featue matrix.
   $$G_{tR} = \frac{1}{n}\sum_{i=1}^{n}(A_i - \overline{A_R})^T (A_i - \overline{A_R})$$
   $$G_{tE} = \frac{1}{m}\sum_{j=1}^{m}(A_j - \overline{A_E})^T (A_j - \overline{A_E})$$
4. Evaluate the maximizing criteria *J(X)* for train and test images.
   $$J(X_R) = X_R^T G_{tR} X_R$$
   $$J(X_E) = X_E^T G_{tE} X_E$$
5. Repeat steps 3-4 on the transposed intermediate feature matrix to obtain B2DPCA based feature vectors, f_p of size UxV.
6. Train Extreme Learning Machine (ELM) classifier: Generate set of 2DPCA based feature vectors (vectorized feature vectors obtained in previous step) for training.
7. Classify images with test feature vectors using ELM trained in step 6.

Application of ELM based classification on original wave atoms coefficients is computationally expensive due to higher dimensionality of data originating from large image datasets. Outliers and irrelevant image points being included into classification task can also affect the performance of our algorithm; hence B2DPCA is employed to reduce dimensionality of initial feature vectors. Features are extracted by computing 2DPCA of initial feature matrix along image rows, called as intermediate feature matrix. 2DPCA is again applied on the transposed intermediate feature matrix along its rows so as to generate a final feature matrix. Application of 2DPCA using the modified approach retains better structure and correlation information amongst neighboring pixel coefficients. Dimensionally reduced wave atoms coefficients are vectorized into a $U \times V$ dimension vector, final feature vector, where $U \times V << R \times C$.

B2DPCA based feature vectors better retain the global structure of input space and facilitate accurate classification with lower computational complexity, diminished outliers and irrelevant information. ELM is trained using labeled B2DPCA feature vectors and classified using the trained network.

6 Results and Discussion

Extensive experiments were performed using 3 standard and distinctive collections of fingerprint datasets; FVC2000, FVC2002 and FVC2004 [2] to test the practicality of our proposed method. Each dataset consists of four diverse databases generated using various fingerprint acquisition techniques. Each database contains 8 fingerprints of each of the 100 distinctive subjects.

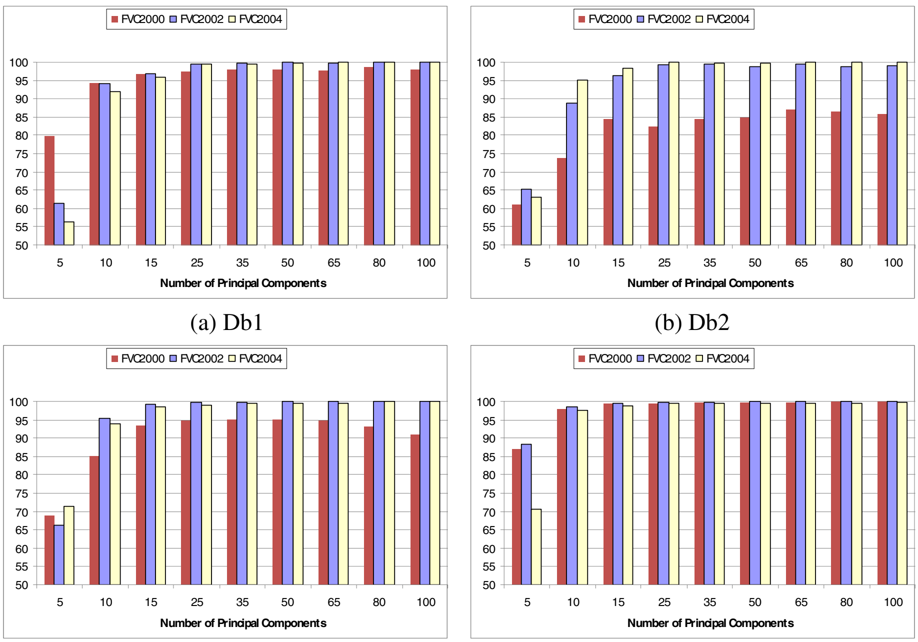
All images were resized to 64×64 in our experiments and 5 images from each database were used as prototypes and the remaining 3 for testing to ensure consistency with other methods used for comparison. Experiments were also performed on original fingerprint image without resizing and consistently better results were obtained since detailed fingerprint information is incorporated at the expense of large feature vectors. Both the testing and training sets of images are decomposed using 2D wave atoms transform using an orthonormal basis function and dimensionally reduced through application of B2DPCA. Dimensionally reduced features are vectorized and classification is performed by using ELM. The above process was repeated 10 times for all the databases and averaged results of few experiments are documented in the paper. The classification accuracy for Db1 database from FVC2000, FVC2002 and FVC2004 is compared with wavelet transform (WT) based RBFNN and EBFNN fingerprint classification algorithms. Results, obtained with the proposed method (only 6 principal components are used for consistency with other methods), are compared with the classification accuracy reported in [11] using WT-2DPCA-RBFNN and WT-2DPCA-EBFNN.

Table 2. Fingerprint classification rates (%) for different techniques

Database	WT-2DPCA-RBFNN	WT-2DPCA-EBFNN	Proposed Method
FVC 2000	91	91	93.25
FVC 2002	87	87	92.63
FVC 2004	86.5	87	89.62

We conclude from the results in Table 2 that our proposed fingerprint classification algorithm performs significantly better than the wavelet based RBFNN and EBFNN fingerprint classification algorithms. In addition to the improved classification accuracy, our proposed ELM based classifier performs training and testing thousands folds faster than conventional neural network based classification algorithms [15].

From Fig. 3 it is evident that several factors influence classification accuracy, namely, fingerprint acquisition techniques, climatic and environmental conditions and most notably the number of principal components. Dataset Db4 from each of the databases is generated using a synthetic fingerprint generator; consequently the effects of environment and other irrepressible conditions are trifling and are substantiated by improved classification accuracy at low principal components.



(a) Db1 (b) Db2

(c) Db3 (d) Db4

Fig. 3. Classification accuracy vs. number of principal components

7 Conclusion

An original supervised fingerprint classification algorithm for multiclass categorization based on wave atoms decomposition and bidirectional two-dimensional principal component is proposed. Improvements in classification accuracy validate the fact that wave atoms multiresolution analysis offers significantly sparser expansion, for oscillatory functions, than other fixed standard representations like wavelets, curvelets and Gabor atoms. The proposed classifier is capable of handling marginal fingerprint orientations, illumination variations, moderate pressure changes against the sensor surface and climatic conditions. The algorithm combines the strengths of both B2DPCA and ELM; creates distinctive and improved feature set, an efficient and fast algorithm for fingerprint classification. The proposed fingerprint classification algorithm is independent of the number of prototypes used for training and or testing and is also free of the amount of hidden neurons used for classification, unlike traditional neighborhood based classifiers whose accuracy is greatly affected by the number of prototypes and neighborhood size. Law enforcement, multimedia, and data mining related applications can benefit from our proposed classification scheme.

References

1. Jain, A.K., Ross, A., Pankanti, S.: Biometrics: A tool for information security. IEEE Transaction on Information Forensics Security 1(2), 125–143 (2006)
2. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of fingerprint recognition. Springer, New York (2003)
3. Yager, N., Amin, A.: Fingerprint classification: a review. Pattern Analysis and Applications 7(1), 77–93 (2004)
4. Jain, A., Chen, Y., Demitrius, M.: Pores and Ridges: Fingerprint matching using Level 3 features. In: Proc. 18th International Conference on Pattern Recognition, vol. 4, pp. 477–480 (2006)
5. Maio, D., Maltoni, D., Cappelli, R., Wayman, J., Jain, A.K.: FVC 2004: Third fingerprint verification competition. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 1–7. Springer, Heidelberg (2004)
6. Bazen, A., Gerez, S.: Systematic methods for the computation of the direction fields and singular points of fingerprints. IEEE Transaction on Pattern Analysis and Machine Intelligence 24(7), 905–919 (2002)
7. Bhanu, B., Tan, X.: Fingerprint indexing based on novel features of minutiae triplets. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(5), 616–622 (2003)
8. Fitz, A., Green, R.: Fingerprint classification using a hexagonal fast Fourier transform. Pattern Recognition 29(10), 1587–1597 (1996)
9. Seokwon, L., Boohee, N.: Fingerprint Recognition using Wavelet Transform and Probabilistic Neural Network. In: Proc. of International Joint Conference on Neural Networks, vol. 5, pp. 3276–3279 (1999)
10. Wilson, C., Candela, G., Grother, P., Watson, C., Wilkinson, R.: Massively parallel neural network fingerprint classification system. National Institute of Standards and Technology; NISTIR 4880 (1992)
11. Luo, J., Lin, S., Lei, M., Ni, J.: Application of dimensionality reduction analysis to fingerprint recognition. In: ISCID, vol. 2, pp. 102–105 (2008)
12. Demanet, L., Ying, L.: Wave Atoms and Sparsity of Oscillatory Patterns. Applied and Computational Harmonic Analysis 23(3), 368–387 (2007)
13. Yang, J., Zhang, D., Frangi, A., Yang, J.: Two-dimensional PCA: a new approach to appearance based face representation and recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence 26(1), 131–137 (2004)
14. Zhang, D., Zhou, Z.: $(2D)^2$ PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. Neurocomputing 69(1), 224–231 (2005)
15. Pankanti, S., Prabhakar, S., Jain, A.K.: On the Individuality of Fingerprints. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(8), 1010–1025 (2002)
16. Huang, G., Zhu, Q., Siew, C.: Extreme learning machine: Theory and applications. Neurocomputing 70(1), 489–501 (2006)
17. Ross, A., Jain, A.K., Reisman, J.: A Hybrid Fingerprint Matcher. Pattern Recognition 36(7), 1661–1673 (2003)