Evolutionary Optimization Approach for Fingerprint Classification

Jing-Wein Wang, *Member*, *IEEE*

[1] *Abstract*—To test the effectiveness of GHM multiwavelets in fingerprint classification with respect to scalar Daubechies wavelets, we study the evolutionary-based algorithm to evaluate the performance of each subset of selected feature. Comparatively studies suggest that the former transform features apparently contain more fingerprint information for discrimination than the latter.

Index Terms—GHM multiwavelt, Daubechies wavelet, fingerprint classification, evolutionary-based algorithm

I. INTRODUCTION

MULTIWAVELETS have recently attracted a lot of theoretical attention and provided a good indication of a potential impact on signal processing [1]. In this paper, a novel fingerprint classification scheme is proposed both to extend the experimentation made in [1] and to test the effectiveness of the Geronimo-Hardin-Massopust (GHM) discrete multiwavelet transform (DMWT) [2] with respect to the scalar Daubechies wavelet [3]. Moreover, a point in genetic wavelet fingerprint analysis is that the chromosomes interact only with the fitness function, but not with each other. This method precludes the evolution of collective solutions to problems, which can be very powerful [4]. We further present an evolutionary framework for feature selection in which successive generations adaptively develop behavior in accordance with their natural needs. In the following sections, we give details of the propose fingerprint classification approach. The performance of the proposed method has been validated through experiments on the NIST special fingerprint database 4 (NIST-4) [5].

In the following sections, we give details of the propose fingerprint classification approach. Section II presents the example transformations for a fingerprint image. Section III describes the proposed coevolutionary feature selection scheme for classification. In section IV, we present our experimental results tested on the NIST-4. Section V concludes the paper.

[1] Manuscript received Mar 06, 2011; revised Apr 07, 2011. This work was supported in part by Taiwan, ROC, National Science Council under Grant NSC 99-2221-E-151 -057.
J. W. Wang is with the Institute of Photonics and Communications, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan, ROC (phone: +886-930-943-143; fax: +886-7-383-2771; e-mail: jwwang@ cc.kuas.edu.tw).

II. DISCRETE MULTIWAVELET TRANSFORMS

For a multiresolution analysis of multiplicity $r > 1$, (MRA), an orthonormal compact support multiwavelet system consists one multiscaling function vector $\Phi(x) = (\phi_1(x), ..., \phi_r(x))^T$ and one multiwavelet function vector $\Psi(x) = (\psi_1(x), ..., \psi_r(x))^T$. Both Φ and Ψ satisfy the following two-scale relations:

$$\Phi(x) = \sqrt{2}\sum_{k\in\mathbf{Z}} H_k \Phi(2x-k) \quad (1)$$

$$\Psi(x) = \sqrt{2}\sum_{k\in\mathbf{Z}} G_k \Phi(2x-k)\,. \quad (2)$$

Note that multifilters $\{H_k\}$ and $\{G_k\}$ are finite sequences of $r \times r$ matrices for each integer k. Let V_j, $j \in \mathbf{Z}$, be the closure of the linear span of $\phi_{l,j,k} = 2^{j/2}\phi_l(2^j x - k)$, $l = 1, 2, \ldots, r$. By exploiting the properties of the MRA, as in the scalar case, any continuous-time signal $f(x) \in V_0$ can be expanded as

$$\begin{aligned} f(x) &= \sum_{l=1}^{r}\sum_{k\in\mathbf{Z}} c_{l,0,k}\,\phi_l(x-k) \\ &= \sum_{l=1}^{r}\sum_{k\in\mathbf{Z}} c_{l,\mathrm{J},k}\,2^{J/2}\phi_l(2^J x-k) \\ &\quad + \sum_{l=1}^{r}\sum_{J\le j<0}\sum_{k\in\mathbf{Z}} d_{l,j,k}\,2^{j/2}\psi_l(2^j x-k)\,, \end{aligned} \quad (3)$$

where

$$\boldsymbol{c}_{j,k} = \left(c_{1,j,k,\ldots,}c_{r,j,k}\right)^T \quad (4)$$

$$\boldsymbol{d}_{j,k} = \left(d_{1,j,k,\ldots,}d_{r,j,k}\right)^T \quad (5)$$

and

$$c_{l,j,k} = \int f(x)\,2^{j/2}\phi_l(2^j x-k)dx \quad (6)$$

$$d_{l,j,k} = \int f(x)\,2^{j/2}\psi_l(2^j x-k)dx\,. \quad (7)$$

For the two-dimensional discrete multiwavelet transform, a 2-D MRA of multiplicity N for $L^2(R^2)$

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

can be obtained by using the tensor product of two MRA's of multiplicity N of $L^2(R)$. Fig. 1 shows a fingerprint image of size 512×512 and its one-level decomposition with the D4 wavelet transform and the GHM multiwavelet transform, respectively.

III. Feature Selection Algorithm

In the proposed method that is derived from the principles of the natural species evolution theory [6], individuals grouped in populations and thereafter referred to as *inter* population P_b and *intra* population P_w are randomly created. The two populations have interdependent evolutions (coevolution). The term *inter* reflects the reluctance of this individual for the opposite class. This reluctance is quantified by the mean square distance between pattern points that belong to different classes. An individual of the population P_b, I_x, will compete with each individual of the population kernel K_b which is the collection of individuals with best inter distances. The term *Inter* is formulated as follows:

$$\left[\,Inter\right]_{I_x \in P_b} = \sum_m (I_x \Leftrightarrow I_m)$$
$$\text{with } I_m \in K_b,\ m = 1,\ldots, M,$$

$$(I_x \Leftrightarrow I_m) = \begin{cases} D_b^x - D_b^m & \text{if } D_b^x > D_b^m, \\ p & \text{if } D_b^x \le D_b^m, \end{cases} \qquad (8)$$

where D_b is the Euclidean distance between classes and p is a penalty. Conversely, the term *Intra* reflects the attraction of this individual for its own class. An individual of population P_w, I_x, will compete with each individual of the population kernel K_w which is the collection of individuals with best intra distances. A best individual of the population kernel K_b will compete with each of the best individuals of the opposite population kernel K_w. The combined results of these competitions directly provide the fitness function, and therefore the fitness function Θ is defined as a number composed of two terms:

$$\Theta = (\,1 - \xi \cdot \delta/\chi\,) \cdot (\,[Inter] - [Intra]\,), \qquad (9)$$

where ξ is the weighting constant greater or equal to one, δ is the number of features selected, χ is the number of training samples. The evaluation process of Θ is randomly combined with the *Inter* individual of the population kernel K_b and the *Intra* individual of the population kernel K_w.

After computation of the fitness function for all the combination of the two kernel individuals, a feature selection step is activated for choosing the individuals allowed reproducing at the next generation. The strategy of feature selection involves selecting the best subset A_q,

$$\mathrm{A}_q = \{\alpha_u \,|\, u = 1, \ldots, q;\ \alpha_u \in \mathrm{B}\} \qquad (10)$$

from an original feature set B,

$$\mathrm{B} = \{\beta_v \,|\, v = 1, \ldots, Q\},\ Q > q. \qquad (11)$$

In other words, the combination of q features from A_q will maximize equation (9) with respect to any other combination of q features taken from Q, respectively. The new feature β_v is chosen as the (λ+1)st feature if it yields

$$\max_{\forall \beta_v} \max_{\forall \alpha_u} \Delta[Inter]\,(\alpha_u, \beta_v), \qquad (12)$$

where $\alpha_u \in \mathrm{A}_\lambda$, $\beta_v \in \mathrm{B} - \mathrm{A}_\lambda$, and $\Delta\,[Inter](\alpha_u, \beta_v) = [Inter](\alpha_u, \beta_v) - [Inter](\alpha_u)$. $[Inter](\alpha_u)$ is the evaluation value of equation (8) while the feature α_u is selected and $[Inter](\alpha_u, \beta_v)$ is the evaluation value of equation (8) while the candidate β_v is added to the already selected feature α_u. In a similar way, the feature selection mechanism minimizes *intra* measure and helps to facilitate classification by removing redundant features that may impede recognition. The proposed schemes consider both the accuracy of classification and the cost of performing classification.

To speed up such a selection process, we present a *packet-tree selection scheme* that is based on fitness value of equation (9) to locate dominant wavelet subbands. Following this innovative idea, the decomposed subbands at the current level, which can be viewed as the parent and children nodes in a tree, will be selected only if the predecessor at the previous level was selected. Otherwise, the scheme skips the successors and considers the next subbands. For each textured fingerprint, a representative tree by averaging the selected feature vectors over all the training samples is generated.

IV. Genetic Operations

With a direct encoding scheme, the genetic representation is used to evolve potential solutions under a set of five-class 512 × 512 images with 256 gray levels (see Fig. 2) found in the NIST-4 database. According to the roulette wheel selection strategy [7], the combination of populations P_b and P_w individuals with higher fitness value in equation (9) will survive more at the next generation. The combinative individuals selected in the previous step are used to as the parent individuals and then their

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

chromosomes are combined by the following proposed *combinative crossover criterion* so as to toward the chromosomes of two offspring individuals. If the *i*-th genes of the inter and intra individuals are the same, then the *i*-th gene of the offspring individual is set as either individual. If not, the *i*-th gene of the offspring individual will be set as either individual at random. The size of each of the population remains constant during evolution. The mutation operation randomly changes a bit of the chromosome.

V. Experiments Results and Discussions

The reported results have the following parameter settings: population size = 20, number of generation = 1000, and the probability of crossover = 0.5. A mutation probability value starts with a value of 0.9 and then varied as a step function of the number of iterations until it reaches a value of 0.01. Due to the curse of dimensionality, one hundred 256 × 256 overlapping subimages each class as training samples are used for the D4 wavelet and one thousand samples are used for the GHM multiwavelet. Textural features are given by the extrema number of wavelet coefficients [8], which can be used as a measure of coarseness of the fingerprint at multiple resolutions. Then, fingerprint classifications with feature selection were performed using the simplified Mahalanobis distance measure [9] to discriminate fingerprint textures and to optimize classification by searching for near-optimal feature subsets. The mean and variance of the decomposed subbands are calculated with the leave-one-out algorithm [9] in classification.

The performance of the classifier was evaluated with three different randomly chosen training and test sets. Algorithms based on the two types of wavelets have been shown to work well in fingerprint discrimination. The classification errors in Tables 1 and 2 mostly decrease when the used features are selectively removed from all the features at the decomposed levels 4 and 3, respectively. This decrease is due to the fact that less parameter used in place of the true value of the class conditional probability density functions need to be estimated from the same number of samples. The smaller the number of the parameters that need to be estimated, the less severe the curse of dimensionality can become. In the meanwhile, we also noticed that the multiwavelet outperforms the scalar wavelet with the packet-tree feature selection. This is because the extracted features in the former are more discriminative than the latter and, therefore, the selection of a subband for discrimination is not only dependent on the wavelet bases, wavelet decompositions, and decomposed levels but also the fitness function.

To explore the performance of the proposed system, we report classification results on NIST-4 database with seven categories: right loop (437 images), left loop (484 images), tented (149 images), arch (530 images), S-type (twin loop) (110 images), whorl (241 images), and eddy (49 images). We compare our method to a few modern techniques as shown in Table 3. Our method not only achieves more accuracy in 4 and 5 classes than referred methods [10]-[13] with lower rejection rate, 7-class offers new report in classifying whorl, S-type, and eddy types, as well. On the other hand, we notice that there are some failures occurred in the experiments and the reasons can be summarized as the following. The indistinct ridges and valleys due to bad quality of the fingerprint image may lead to the fatal errors of wavelet extrema detection.

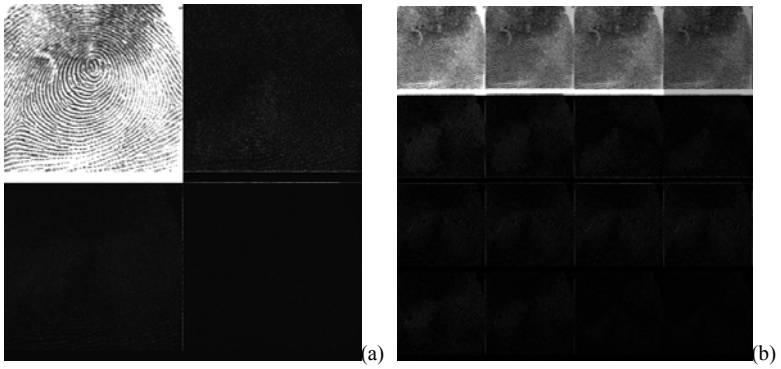
VI. Conclusions

This paper introduces a promising evolutionary algorithm approach for solving the fingerprint classification problem with the coevolving concept. While much of the researches are fighting to work out on the classification of four or five categories, even one or two seven classes, we have not joined the drive instead of reporting a new result on whorl, S-type, and eddy classes except general arch, tented arch, right and left loops.

References

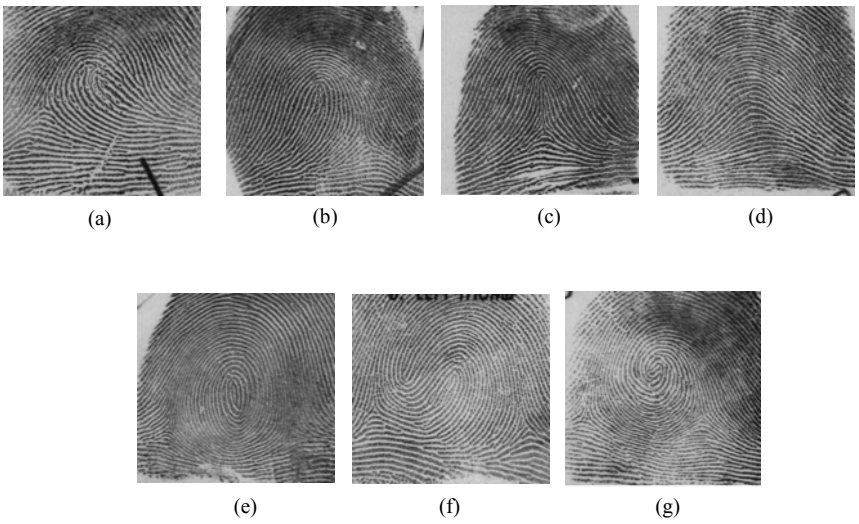
[1] V. Strela, P. N. Heller, G. Strang, P. Topiwala, and C. Heil, "The application of multiwavelet filter banks to image processing," *IEEE Trans. Image Process*., vol. 8, pp. 548-563, 1999.
[2] F. Keinert, *Wavelets and Multiwavelets*, Chapman & Hall, CRC, 2003.
[3] I. Daubechies, *Ten lectures on wavelets*. SIAM, Philadelphia, Penn., 1992.
[4] W. Siedlecki and J. Sklansky, "A note on genetic algorithm for large-scale feature selection," *Pattern Recognition Letters*, vol. 10, pp. 335-347, Nov. 1989.
[5] C. I. Watson and C. L. Wilson, *NIST special database 4, fingerprint database*. National Institute of Standards and Technology, March 1992.
[6] T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press, New York, 1996.
[7] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. MA: Addison-Wesley, 1989.
[8] S. Mukhopadhyay and A. P. Tiwari "Characterization of NDT signals: Reconstruction from wavelet transform maximum curvature representation," *Signal Processing*, vol. 90, no. 1, pp. 261-268, 2010.
[9] R. O. Duda, P. E. Hart, and David G. Stork, *Pattern Classification*. Wiley-Interscience, 2000.
[10] A. K. Jain, S. Prabhaker, and L. Hong, "A multichannel approach to fingerprint classification," *IEEE Trans. on Pattern Recognition and Machine Intell.*, vol. 21, no. 4, pp. 348-359, 1999.
[11] Y. Yao, G. L. Marcialis, M. Pontil, P. Frasconi, and F. Roli, "Combining flat and structured representation for fingerprint classification with recursive neural networks and support vector machines," *Pattern Recognition*, vol. 36, pp. 397-406, 2003.
[12] Q. Zhang and H. Yan, "Fingerprint classification based on extraction and analysis of singularities and pseudo ridges," *Pattern Recognition*, vol. 37, pp. 2233-2243, 2004.
[13] J. Li, W. Y. Yau, and H. Wang, "Combining singular points and orientation image information for fingerprint classification," *Pattern Recognition*, vol. 41, pp. 353-366, 2008.

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



(a) (b)

Fig. 1. One-level decomposition for the NIST-4 fingerprint: (a) D4, (b) GHM.



(a) (b) (c) (d)

(e) (f) (g)

Fig. 2. Fingerprint examples defined in Henry system: (a) right loop, (b) left loop, (c) tented arch, (d) arch, (e) whorl, (f) S-type (twin loop), (g) Eddy.

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE 1
CLASSIFICATION RESULTS (CORRECT RATE IN %) USING THE D4 WAVELET PACKET DECOMPOSITION WITH COEVOLUTIONARY FEATURE SELECTION

Sample Set	$\xi = 1$	$\xi = 2$	$\xi = 3$	$\xi = 4$	$\xi = 5$
1	90.47	90.48	90.29	90.49	90.43
2	90.41	90.38	90.62	90.29	90.49
3	90.41	90.43	90.38	90.51	90.49
Average	90.43	90.43	90.43	90.43	90.47

TABLE 2
CLASSIFICATION RESULTS (CORRECT RATE IN %) USING THE GHM MULTIWAVELET PACKET DECOMPOSITION WITH COEVOLUTIONARY FEATURE SELECTION

Sample Set	$\xi = 1$	$\xi = 2$	$\xi = 3$	$\xi = 4$	$\xi = 5$
1	90.83	90.80	90.54	90.82	90.58
2	90.85	90.72	90.80	90.71	90.66
3	90.96	90.73	90.90	90.79	90.73
Average	90.88	90.75	90.75	90.77	90.79

TABLE 3
CLASSIFICATION RESULTS (CORRECT RATE IN %) COMPARED TO THE RELATED WORKS

Fingerprint class	Wang	Jain et al.[10]	Yao et al.[11]	Zhang and Yan [12]	Li et al. [13]
7-class	90.88% (2.3%)	*	*	*	*
5-class	94.71%	90% (1.8%)	90% (1.8%)	84.3%	93.5%
4-class	95.36%	95.8%	94.7%	92.7%	95%

* Unavailable

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)