International Journal of Innovative
Computing, Information and Control
Volume 8, Number 1(B), January 2012
 ISSN 1349-4198
pp. 989–1000

FEATURE SELECTION AND CLASSIFICATION OF SELDI-TOF MASS SPECTRA OF HEPATOMA USING GENE-WEIGHTED GENETIC ALGORITHM

Leehter Yao and Tzu-Yi Pan

Department of Electrical Engineering
National Taipei University of Technology
No. 1, Sec. 3, Chung-hsiao E. Rd., Taipei, Taiwan
ltyao@ntut.edu.tw

Received October 2010; revised February 2011

Abstract. *A classifier to classify the normal sample or sample with hepatoma based on the sample's surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) mass spectra is designed in this paper. A modified genetic algorithm (GA) called gene-weighted GA (GWGA) is proposed to design the classifier based on the SELDI-TOF mass spectra of hepatoma. To reduce the computation efforts, an approach dividing the measurement intensities within different range of m/z values into several data sectors and finding the peak intensity within each data sector is proposed. The peak intensity at each data sector is taken as features for classification. The proposed GWGA aims to select the features and minimize the number of selected features while maximizing the classification accuracy.*
Keywords: SELDI-TOF mass spectra, Genetic algorithm, Hepatoma, Support vector machine

1. **Introduction.** Cancer ranks the first among the 10 causes of death in Taiwan. Among all cancer-type causes of deaths in Taiwan, hepatoma (liver cancer) ranks the second following the lung cancer. Taiwan government has spent a lot of efforts on prognoses, diagnoses and curing of hepatoma. Recently, surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) mass spectra have been successfully used to detect protein patterns of several cancers including hepatoma. With the SELDI-TOF mass spectra, the serum proteomic patterns in the tissue affected by cancers or by the development of cancer can be analyzed. The analysis results of SELDI-TOF mass spectra can be applied to the cancer diagnosis, monitoring of cancer progression and verification of therapeutic effects of drugs. The procedure to obtain SELDI-TOF mass spectra from an examinee is simple and minimally invasive. By collecting a good number of SELDI-TOF mass spectra samples from patients with cancer and from normal persons, a classifier can be built so that an unknown sample is classified as either cancer or normal by this classifier [1,2]. However, each spectrum is composed of several thousands to several hundred thousands of measurement signals. It is both inefficient and inaccurate if all the measurement signals are utilized as features for classification. It has been a challenge for the researchers in this area finding an appropriate approach solving the problem of feature selection from the measurement signals and the pattern classification based on the selected features. The SELDI-TOF mass spectra contain different intensities at different mass-to-charge ratios (m/z values). In this paper, a grouping and peak finding approach is proposed that divides the measurement intensities within different range of m/z values into several data sectors and finds the peak intensity within each sector. Instead of taking SELDI-TOF mass spectra measurement signals as features of the prototype data points for the classifier,

peak intensity at the different m/z values is taken as features in this paper. It greatly reduces the number of features for the classification. The cancer samples tend to have a pattern of intensity distribution at certain m/z values compared with normal samples [3]. Among the peak intensities measured at different range of m/z values, it is found that it is not necessary to use all features to have good classification accuracy [4]. A feature selection approach called gene-weighted genetic algorithm (GWGA) is to be proposed in this paper so that the classification accuracy is maximized while the number of features utilized in the classification is minimized. In [5], a different approach yet based on similar concept has been used in the microarray data selection. A three-stage gene selection method was proposed in [5] to select a smaller subset of informative genes that is most relevant for the cancer classification. More accurate classification results are achieved if a smaller subset of informative genes is used for classification.

In [6,7], fuzzy logic inference has been applied to identify the normal sample and the samples with cancer. The GA has been widely applied to the classifier and/or clustering approaches design [8-12]. The classification rules or structure of the classifier are parameterized and the GA is applied to automatically learn the parameters in [8-11]. The GA with binary coded chromosomes has also been applied to feature selection [12-14] where "1" and "0" in the binary chromosome denote the feature being selected or not. In other words, every gene of a chromosome represents the selection status of the corresponding feature. With selection status of every candidate feature being encoded as a binary number of GA's chromosome, the GA can evolve to generate the best chromosome corresponding to the features finally selected. In [12-14], although the features are learned by the GA, the classification approach based on the selected features is usually not automatically designed by the GA in order to save computation time. One of the most commonly used classification schemes is the support vector machine [15-19] due to its capability of learning highly nonlinear decision regions for classification. In this paper, the support vector machine (SVM) will also be used as the classification scheme as the proposed GWGA evolves to learn the features from the candidate features. Every binary chromosome in the GWGA consists of the selected features. The SVM is utilized as the classification scheme calculating the classification rate based on the features selected within the corresponding chromosome. The classification rate associated with every chromosome is further used to calculate the cross-over probability. Recently, the GA has also been used to find the biomarkers of the SELDI-TOF mass spectra [20,21]. It is difficult for the regular GA to assign a specific gene the binary number 1 in the chromosome since every gene is generated with equal probability in the crossover operation. However, if some features lead to high classification rate, it greatly helps the feature selection if these features are assigned larger probability of appearing at the chromosomes in the following generations. To overcome the difficulty that every gene in the chromosome is generated with equal probability, the proposed GWGA assigns every gene a weighting based on classification rate associated with the chromosome. By analyzing every chromosome in the gene pool, the gene appearing more often in the chromosomes of gene pool tends to have larger weighing. Moreover, the gene appearing in the chromosome with higher classification rate is also assigned a larger weighting. The crossover approach is modified in the GWGA so that every gene generated in the crossover process is based on the weighting associated with every gene.

The organization of this paper is arranged as follows. The SELDI-TOF mass spectra of normal samples and samples with hepatoma are introduced in Section 2. The background information about how the feature selection scheme is introduced to the classifier design based on SELDI-TOF mass spectra is also introduced in Section 2. In Section 3, the technical details of GWGA for feature selection are described. The effect and efficiency of

the proposed GWGA are verified in Section 4. Finally, the conclusion is drawn in Section 5.

2. **SELDI-TOF Mass Spectra and Feature Selection.** The processed SELDI-TOF mass spectra contain different intensities at different mass-to-charge ratios (m/z values). The data in SELDI-TOF mass spectra for hepatoma ranges from 0 to 150000. Obviously, there are too many data to be analyzed for the classification. The biomarkers of certain type of cancer usually do not appear at specific m/z value, and yet appear at certain range of m/z values. Assume that the SELDI-TOF mass spectra from N samples (i.e., persons) are to be utilized as the prototype data for classifier training. Let c_i be the class that the ith sample is classified, $c_i \in \{-1, 1\}$, where $c_i = -1$, if the ith sample is classified as a patient with hepatocellular carcinoma (hcc) being detected while $c_i = 1$ if the ith sample is classified as a normal person. The SELDI-TOF mass spectra for the ith sample can be represented as $(c_i; \boldsymbol{D}_i)$, where $\boldsymbol{D}_i$ is the data set containing the measurement data of SELDI-TOF mass spectra, i.e.,

$$\boldsymbol{D}_i = \{(x_{ik}, y_{ik}) | \, k = 0, \ldots, L\}, \tag{1}$$

x_{ik} denotes the m/z value, y_{ik} denotes the intensity corresponding to x_{ik}, and L denotes the total number of measurements. To reduce the data processing effort, the data analysis range is divided into M sectors. For the jth data sector, assume that $x_{ik} \in [r_s^j, r_e^j]$. Although several data might be in a data sector, the maximum value is calculated and taken as the prototype data for further analysis. Let z_{ij} be the maximum value among all intensities in the jth data sector for the ith smaple, i.e.,

$$z_{ij} = \max_{x_{ik} \in [r_s^j, r_e^j]} (y_{ik}), \tag{2}$$

the prototype data can be simplified as $(c_i; \boldsymbol{d}_i)$, where

$$\boldsymbol{d}_i = \{z_{ij} | \, j = 1, \ldots, M\}, \quad i = 1, \ldots, N. \tag{3}$$

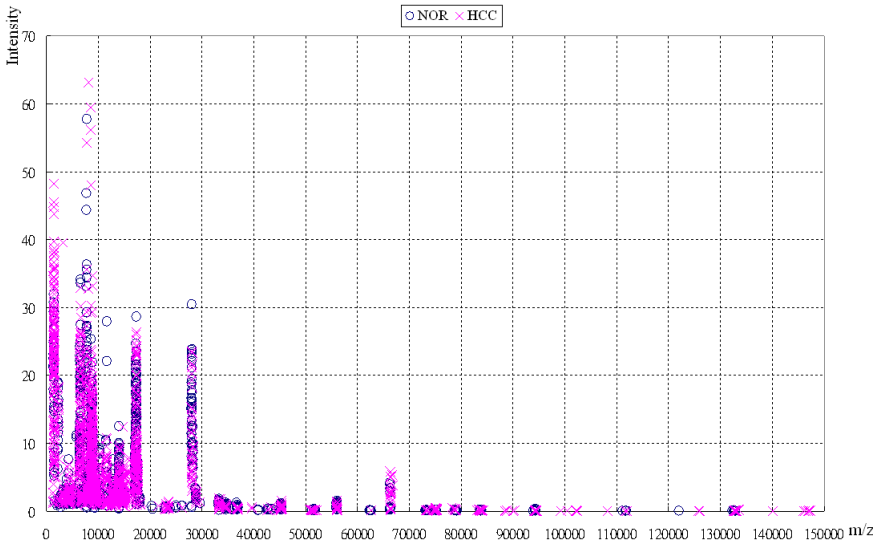
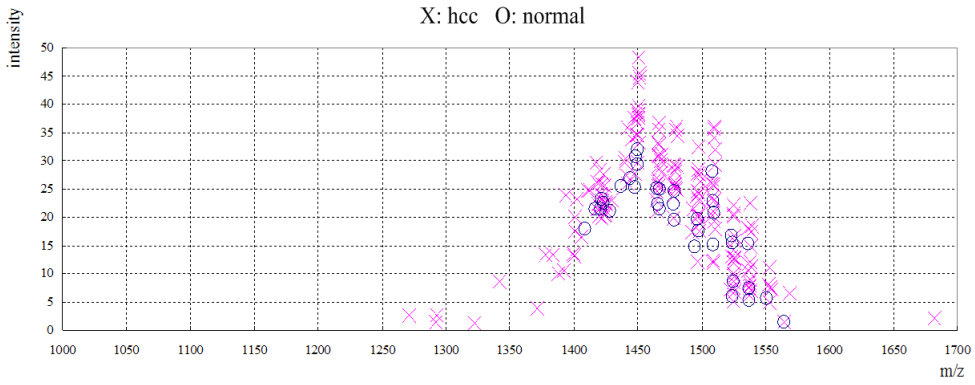


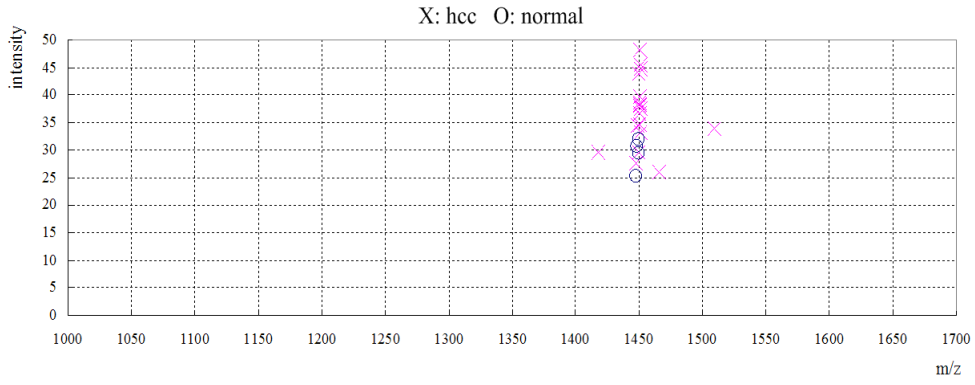
FIGURE 1. The processed SELDI-TOF mass spectra of all 97 samples

Figure 1 shows the SELDI-TOF mass spectra of a group of 97 samples, among which 55 samples are with hepatocellular carcinoma (hcc) and 42 samples are normal. In other words, the prototype data $(c_i; \boldsymbol{D}_i)$, $i = 1, \ldots, 97$, are shown in Figure 1. Zooming into

the range 1000-1700, the data in Figure 1 are shown as in Figure 2(a). With the simplification stated in (3), the simplified data $(c_i; \boldsymbol{d}_i)$, $i = 1, \ldots, N$, are shown in Figure 2(b). Comparing Figures 2(a) and 2(b), it is obvious that the data $(c_i; \boldsymbol{d}_i)$ in the data sector greatly reduce the amount of data to be processed compared to the data $(c_i; \boldsymbol{D}_i)$, $i = 1, \ldots, N$. Reviewing distribution of the processed SELDI-TOF mass spectra of all 97 samples in Figure 1, the data sectors chosen for further simplification are list in Table 1.



(a)



(b)

FIGURE 2. Comparison of the original with the simplified version of the processed SELDI-TOF mass spectra, (a) the original SELDI-TOF mass spectra and (b) the simplification of the SELDI-TOF mass spectra

3. **Gene-weighted Genetic Algorithm.** The GWGA is proposed to select appropriate features among all candidate features. Based on the features evolved in the GWGA, the SVM is included in the GWGA for the classifier training using the prototype data $(c_i; \boldsymbol{d}_i)$, $i = 1, \ldots, N$. Assume that Q chromosomes are used in the gene pool in every generation. Let the jth chromosome in the kth generation of the GWGA be denoted as $h_j(k)$, then

$$h_j(k) = \{g_{j1}(k), \ldots, g_{jM}(k)\}, \quad j = 1 \ldots, Q, \tag{4}$$

where $g_{jr}(\cdot) \in \{0, 1\}$ is the rth gene in the chromosome $h_j(\cdot)$, $r = 1, \ldots, M$. Since GA has the characteristics of keeping strong and eliminating weak, the feature that dominates the classification results tends to appear more often as the GWGA evolves. The dominating importance of the rth feature, i.e., the gene $g_{jr}(\cdot)$, $r = 1, \ldots, M$, to the classification result

TABLE 1. Data sectors selected for classification

	data sector		data sector		data sector
1	$1250 \sim 1700$	16	$8675 \sim 8790$	31	$27950 \sim 28300$
2	$2050 \sim 2450$	17	$8800 \sim 8860$	32	$28600 \sim 29100$
3	$2750 \sim 3500$	18	$8900 \sim 9020$	33	$33200 \sim 33700$
4	$3850 \sim 3950$	19	$9100 \sim 9175$	34	$34450 \sim 34700$
5	$4300 \sim 4375$	20	$9275 \sim 9335$	35	$36450 \sim 37200$
6	$4450 \sim 4500$	21	$9405 \sim 9470$	36	$42200 \sim 44200$
7	$5750 \sim 6000$	22	$9700 \sim 9730$	37	$44950 \sim 45450$
8	$6400 \sim 6500$	23	$10245 \sim 10310$	38	$50700 \sim 52200$
9	$6600 \sim 6700$	24	$10950 \sim 11900$	39	$55700 \sim 56200$
10	$6800 \sim 6900$	25	$12500 \sim 12700$	40	$65750 \sim 66750$
11	$7550 \sim 7665$	26	$12750 \sim 12950$	41	$73000 \sim 84300$
12	$7750 \sim 7875$	27	$13250 \sim 14200$	42	$88200 \sim 102700$
13	$7910 \sim 8005$	28	$14580 \sim 16200$	43	$107700 \sim 150000$
14	$8120 \sim 8230$	29	$16700 \sim 18200$		
15	$8550 \sim 8650$	30	$22700 \sim 23900$		

can thus be weighted by calculating the number of 1's appeared at the corresponding jth gene in the rth chromosome of the gene pool, $r = 1, \ldots, Q$, as GWGA evolves. The gene's weighting can be further refined by multiplying the classification rate $\alpha_j(\cdot)$ corresponding to the chromosome $h_j(\cdot)$ with the value of gene $g_{jr}(\cdot)$, and then accumulating all the multiplied weighting among all the chromosomes in the gene pool.

Let $\Gamma(s)$ be a threshold function for a logic statement s as following:

$$\Gamma(s) = \begin{cases} 1, & \text{if } s \text{ is true;} \\ 0, & \text{if } s \text{ is false.} \end{cases} \tag{5}$$

The classification approach SVM is utilized based on the features determined by the chromosome. Let $\hat{c}_i^j(k)$ be the class of the ith sample determined by SVM based on the features determined by the chromosome $h_j(k)$ in the kth generation. Denote $\alpha_j(k)$ as the classification rate associated with the chromosome $h_j(k)$, then

$$\alpha_j(k) = \frac{\sum_{i=1}^{N} \Gamma(\hat{c}_i^j(k) = c_i)}{N}, \quad j = 1, \ldots, Q. \tag{6}$$

Denote $b_r(k)$ as the weighting of the rth gene in the kth generation as following:

$$b_r(k) = \sum_{j=1}^{Q} g_{jr}(k)\alpha_j(k). \tag{7}$$

Since the value of $b_r(k)$ varies from gene to gene, it is normalized between $[-\delta, \delta]$ for every gene. Denote $\bar{b}(k)$ and $\underline{b}(k)$ as the maximum and minimum of $b_r(k)$ for $r = 1, \ldots, M$, i.e.,

$$\bar{b}(k) = \max_{r=1,\ldots,M}(b_r(k)); \tag{8}$$

$$\underline{b}(k) = \min_{r=1,\ldots,M}(b_r(k)). \tag{9}$$

The value $b_r(k)$ can be normalized between $[-\delta, \delta]$ as $h_r(k)$ based on $\bar{b}(k)$ and $\underline{b}(k)$ according to the following:

$$h_r(k) = \frac{2\delta}{\bar{b}(k) - \underline{b}(k)} b_r(k) - \frac{\bar{b}(k) + \underline{b}(k)}{\bar{b}(k) - \underline{b}(k)}\delta. \tag{10}$$

The weighting of every gene in the previous generation is taken into consideration because the domination of features rarely shows right away in the current generation. It usually takes several generations for the feature to become dominated. However, the weighting of every gene calculated at the generation farther away from the current generation should have less impact to the gene weighting calculated in the current generation. Therefore, a forgetting factor λ is assigned and multiplied with the gene weighting calculated in the previous generation. Let $w_r(k)$ be the accumulated weighting of the rth gene, then

$$w_r(k) = h_r(k) + \lambda w_r(k-1) \tag{11}$$

where $h_r(k)$ is calculated as in (10) and the forgetting factor $\lambda < 1$.

After calculating the accumulated weighting of every gene as in (11), the weighting is utilized to refine the probability of generating 1's in the crossover operation. The crossover in GWGA is operated bit by bit (or gene by gene). Assume that β parent chromosomes are selected for the crossover operation and assume that the probability of generating 1 at every gene is in uniform distribution. Given that the chromosomes numbered $j_1^*, j_2^*, \ldots, j_\beta^*$ are selected as the parent chromosomes for crossover at the $(k-1)$th generation, the probability of generating 1 at the rth gene in the jth chromosome for the kth generation, denoted by $f_{jr}(k)$, is defined as:

$$f_{jr}(k) = (g_{j_1^* r}(k-1) + g_{j_{21}^* r}(k-1) + ... + g_{j_\beta^* r}(k-1))/\beta. \tag{12}$$

The probability of generating 1 at every gene given that 3 parent chromosomes are selected for crossover operation is illustrated in Figure 3.

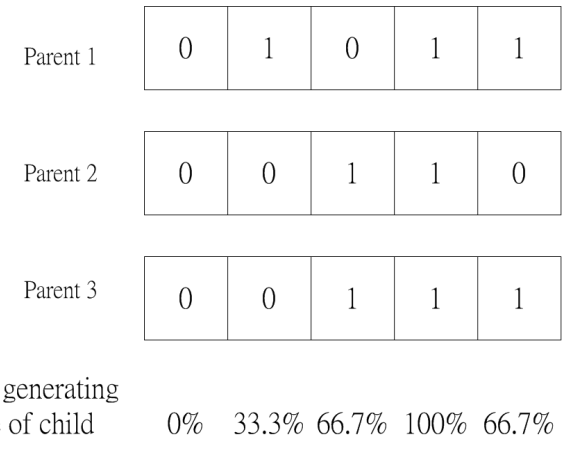


FIGURE 3. Probability of generating 1 at each gene of child chromosome for 3 parent chromosomes

The probability of generating 1's in (12) for crossover is refined by the accumulated gene weighting in (11). Let the updated probability for the rth gene in the jth chromosome of the kth generation be $\sigma_{jr}(k)$, then

$$\sigma_{jr}(k) = f_{jr}(k) + w_r(k). \tag{13}$$

The value of $\sigma_{jr}(k)$ cannot be directly utilized as the probability of generating 1's for crossover operation because it might be negative or greater than 1. It is normalized by a sigmoid function

$$p_{jr}(k) = \frac{1}{1 + e^{-\kappa(\sigma_{jr}(k)-0.5)}}, \tag{14}$$

where κ is a tuning factor for sigmoid function. The probability of generating 1 at every gene in the crossover operation is determined by $p_{jr}(\cdot)$ in (14).

The elitist scheme is adopted in the crossover operation, i.e., only η chromosomes in the gene pool with comparatively larger classification rates are allowed to be selected randomly as the parent chromosomes. To assure that the classification rate is monotonically increasing, the best chromosome with largest classification rate is directly passed into next generation along with child chromosomes generated by crossover. The GWGA increases weighting as the gene is assigned 1's more often in the latest generations. According to (13) and (14), the gene is assigned a larger probability of generating 1 with larger gene weighting. Although this mechanism is effective, it might lead to a pre-mature gene pool, i.e., the gene pool converges at an early stage. An operation called extinction and immigration is also introduced in the GWGA preventing the η best chromosomes chosen as parent chromosomes from becoming uniform. If the η best chromosomes for parent chromosomes evolve to be uniform, the generated child chromosomes tend to also become uniform leading to the feature selection process being trapped at local minimum. The extinction and immigration operator keeps the best chromosome in the gene pool and replaces the rest of chromosomes by random generation [22]. It is triggered when the classification rate of the best chromosome and the ηth best chromosome differs less than a threshold ε_{ex}, i.e., when the following condition is satisfied:

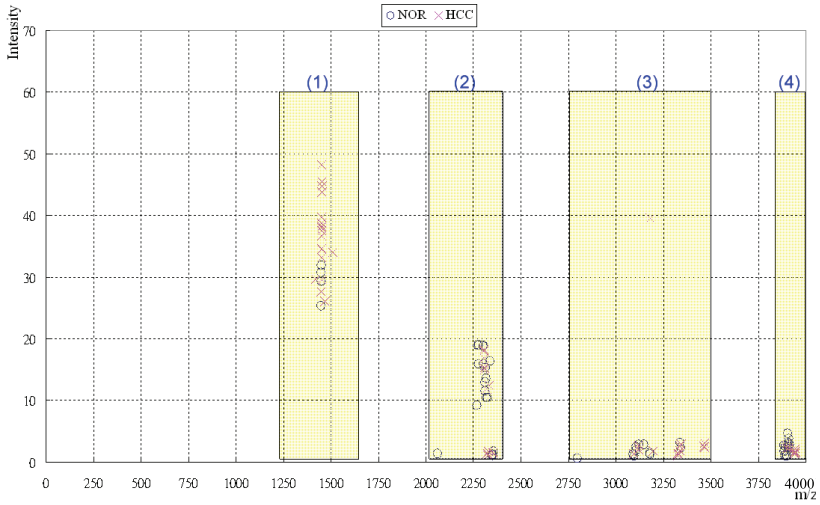
$$|\alpha_1(k) - \alpha_\eta(k)| < \varepsilon_{ex}. \tag{15}$$

The stopping criteria is set as the situation when both classification rate and the number of selected features, i.e., the number of 1's, associated with the best chromosome are less than the threshold ε_{ca} and ε_{fe}, respectively, for T generations. Let $\upsilon_r(k)$ be the number of features corresponding to the rth chromosome in the kth generation. The stopping criteria can be defined as

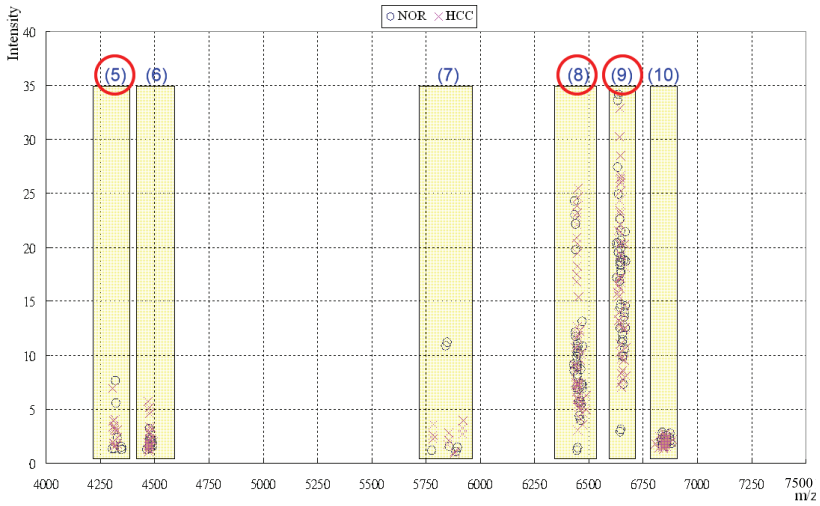
$$|\alpha_1(k) - \alpha_1(k-1)| < \varepsilon_{ca} \text{ and } |\upsilon_1(k) - \upsilon_1(k-1)| < \varepsilon_{fe} \text{ for T generations.} \tag{16}$$

4. **Experiment.** A group of 97 examinees' SELDI-TOF mass spectra processed by (2) and (3) for the detection of hepatocellular carcinoma (hcc) are shown as in Figure 4 for the experiment. The mass spectra in Figures 4(a)-(i) is actually the detailed version of the spectra in Figure 1. In other words, Figures 4(a)-(i) show the spectra obtained by zooming into different range of m/z values. Among 97 samples, 55 samples have been identified as the patients with hcc and 42 samples are normal persons. As stated in the previous section, SVM is utilized as the classification approach based on the features corresponding to every chromosome in the GWGA. The classification rate calculated by SVM for every chromosome is based on the 5-fold scheme. For the 5-fold training and testing scheme, 5 groups of normal samples and samples with hcc are evenly divided within 97 samples. A total of 43 candidate features have been determined as in Table 1. The GWGA is applied to search among these 43 features aiming to minimize the number of selected features while maximizing the classification rate base on the selected features. The tuning parameters of GWGA are set as following:

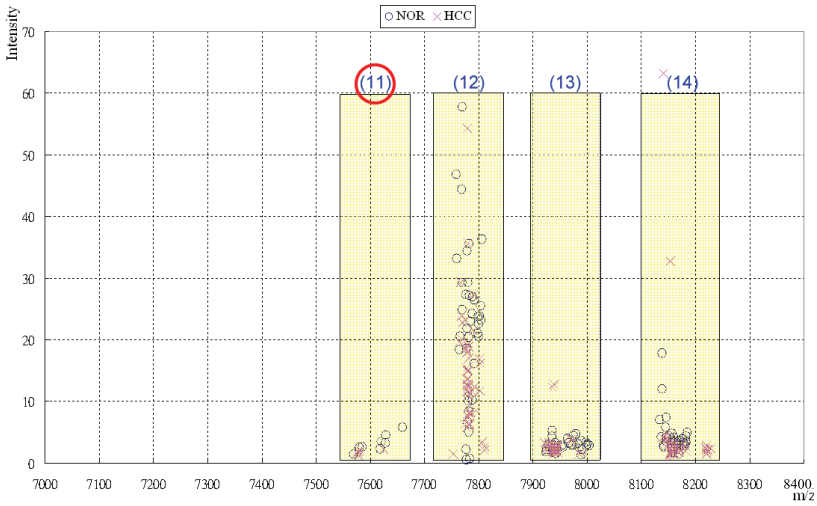
Number of genes in every chromosome M = 43,
Number of chromosomes in every generation Q = 50,
Number of parent chromosomes $\eta = 20$,
Normalization interval for gene weighting $\rho = 0.1$,
Tolerance for extinction and immigration $\varepsilon_{ex} = 0.01$,
Tolerance of classification rate $\varepsilon_{ca} = 10^{-6}$,
Tolerance of number of features $\varepsilon_{fe} = 1$.



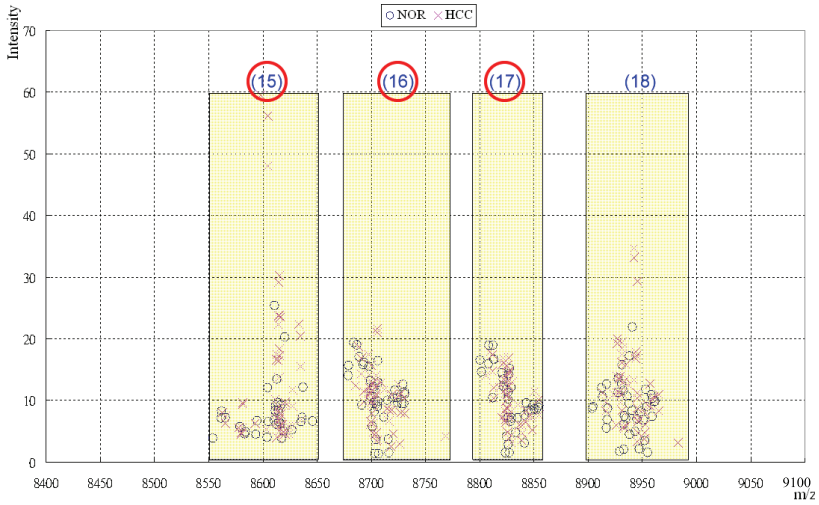
(a) Mass spectra from 0 to 4000 m/z



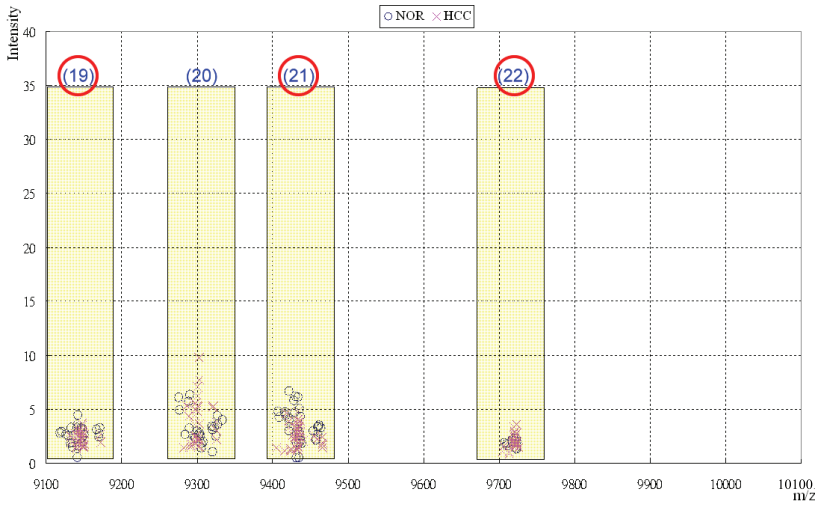
(b) Mass spectra from 4000 to 7500 m/z



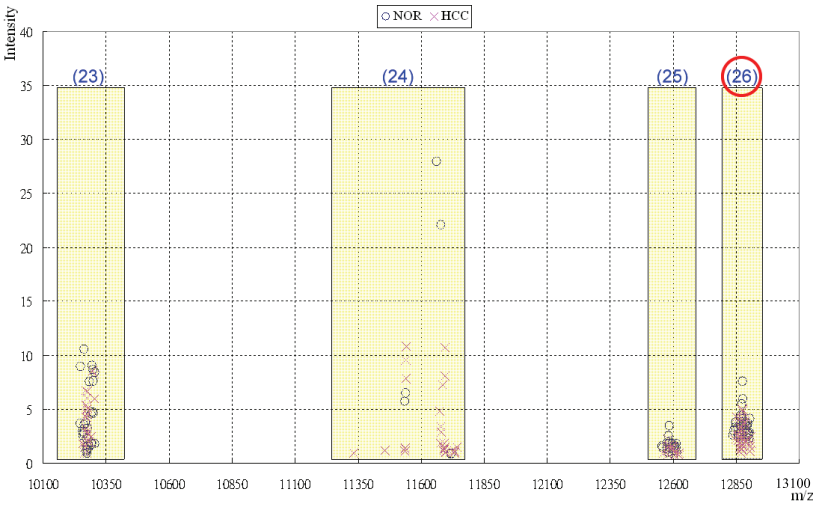
(c) Mass spectra from 7000 to 8400 m/z



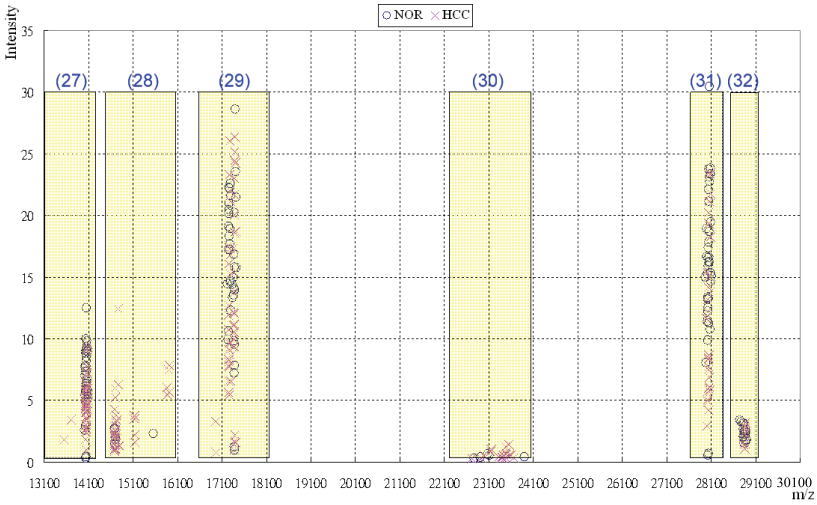
(d) Mass spectra from 8400 to 9100 m/z



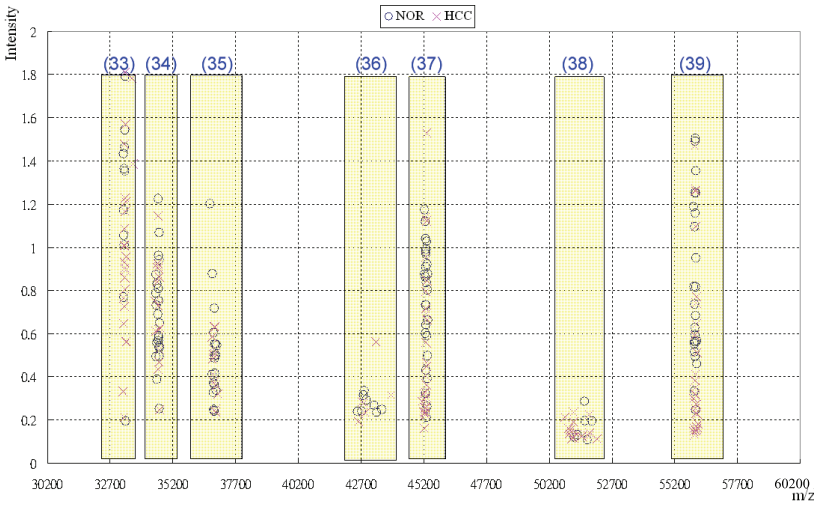
(e) Mass spectra from 9100 to 10100 m/z



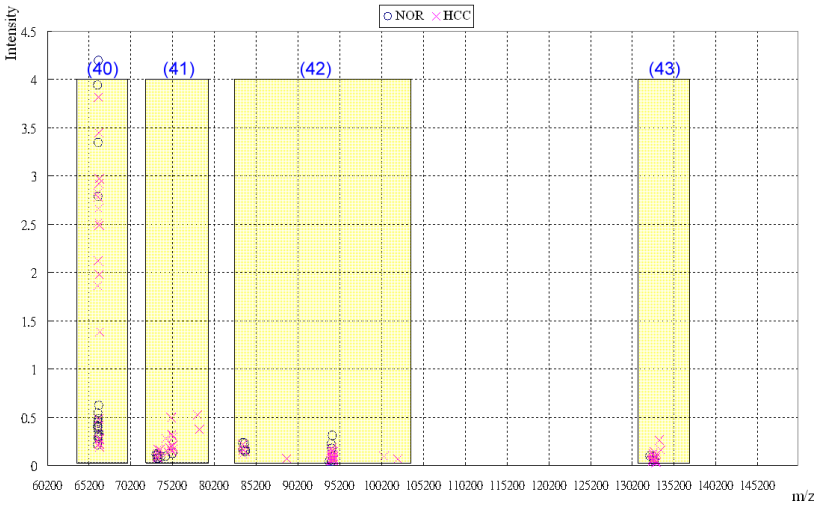
(f) Mass spectra from 10100 to 13100 m/z



(g) Mass spectra from 13100 to 30100 m/z



(h) Mass spectra from 30200 to 60200 m/z



(i) Mass spectra from 60200 to 145200 m/z

FIGURE 4. Mass spectra of 97 examinees SELDI-TOF mass spectra and selected features

The classification rate is 82% if all 43 features are used for the classification using SVM. The experiment is run 10 times as shown in Table 2 to justify the effect and efficiency of GWGA. Table 2 shows that the average number of selected features is 13.6 and the associated classification rate is 99.8%. The average times of extinction and immigration operation triggered is 7.8. It takes 192.9 generations in average for the GWGA to stop given that the stopping criteria are set as stated above. Table 2 also shows that the minimum number of features the GWGA can achieve is 11 features. Taking one of the best experiment result in Table 2 as an illustration example, the 11 features selected by the GWGA number 5, 8, 9, 11, 15, 16, 27, 19, 21, 22, 26 according to the features list in Table 1. The selected features are shown in Figure 4 with the feature number being circled.

TABLE 2. Experiment results of GWGA

Experiment	Classification rate	No. of features	No. of generations	Times of E&I
1	100	11	150	7
2	99	16	285	13
3	100	13	151	7
4	99	12	145	3
5	100	16	126	5
6	100	14	165	3
7	100	16	293	14
8	100	11	133	6
9	100	16	200	10
10	100	11	281	10
average	99.8	13.6	192.9	7.8

The same experiment is also conducted 10 times using simple GA (SGA) with SVM being utilized as the classification scheme. The average classification rate is 95.9% and the average selected features is 18.2. This justifies the fact that although SGA is widely applied to the feature selection problem, the number of features cannot be further reduced because the learning is constrained by SGA's regular uniform gene weighting in the chromosome and by SGA;s regular crossover operation. In GWGA, delicate manipulation of weighting for every individual gene leads to a much smaller number of selected features while achieve larger classification rate compared to the SGA.

5. **Conclusion.** The GWGA has been successfully applied to solve the feature selection problem while obtaining high classification rate. Along with SVM, the GWGA is applied to select appropriate data sectors and to learn the classifier to identify the sample with or without hepatoma from the SELDI-TOF mass spectra. The candidate features or the candidate data sectors for the feature selection performed by GWGA are manually determined by reviewing the data distribution of the SELDI-TOF mass spectra. Automatic learning of candidate features from the SELDI-TOF mass spectra bodes well for the future research. Actually, the proposed GWGA cannot only be applied to the feature selection and classification of SELDI-TOF mass spectra, it can also be applied to the data with large number of features in other applications.

REFERENCES

[1] M. Wagner, D. N. Naik, A. Pothen, S. Kasukurti, R. R. Devineni, B. L. Adam, O. J. Semmes and G. L. W. Jr, Computational protein biomarker prediction: A case study for prostate cancer, *BMC Bioinformatics*, vol.5, pp.26-35, 2004.

[2] R. H. Lilien, H. Farid and B. R. Donald, Probabilistic disease classification of expression-depentent proteomic data from mass spectrometry of human serum, *J. of Computational Biology*, vol.6, pp.925-946, 2003.
[3] B. L. Adam, Y. Qu, J. W. Davis, M. D. Ward, M. A. Clements, L. H. Cazares, O. J. Semmes, P. F. Schellhammer, Y. Yasui, Z. Feng and G. L. Wright, Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from begin prostate hyperplasia and healthy men, *Cancer Res.*, vol.62, pp.3609-3614, 2002.
[4] J. Quinn, L.-H. Loo, J. Armitage, H. Cordingley, S. Roberts, P. J. Bugelski, M. Kam and L. Hrebien, Classification of pharmaceutical toxicity by feature analysis, *Proc. of the 28th IEEE Annual Bio. Eng.*, Philadelphia, PA, pp.211-212, 2002.
[5] M. S. Mohamad, S. Omatu, M. Yoshioka and S. Deris, A three-stage method to select informative genes for cancer classification, *International Journal of Innovative Computing, Information and Control*, vol.6, no.1, pp.117-125, 2010.
[6] D. Woetzel, D. Driesch, M. Pfaff, F. V. Eggeling, K. Junker and R. Guthke, Applying data mining methods to SELDI-TOF analysed renal cell carcinoma samples to identify relevant tumor markers, *Jena Center for Bioinformatics*, 2002.
[7] K. Junker, J. Gneist, C. Melle, D. Driesch, J. Schubert, U. Claussen and F. V. Eggeling, Identification of protein pattern in kidney cancer using proteinchip arrays and bioinformatics, *Int. J. of Molecular Medicine*, vol.15, pp.285-290, 2005.
[8] L. Dong, H. Zhang, X. Ren and Y. Li, Classifier learning algorithm based on genetic algorithms, *International Journal of Innovative Computing, Information and Control*, vol.6, no.4, pp.1973-1981, 2010.
[9] E. G. Mansoori, M. J. Zolghadri and S. D. Katebi, Protein superfamily classification using fuzzy rule-based classifier, *IEEE Trans. Nanobioscience*, vol.8, no.1, pp.92-99, 2009.
[10] E. G. Mansoori, M. J. Zolghadir and S. D. Katebi, SGERD: A steady-state genetic algorithm for extracting fuzzy classification rules from data, *IEEE Trans. Fuzzy Sys.*, vol.16, no.4, pp.1061-1071, 2008.
[11] L. I. Kuncheva and L. C. Jain, Designing classifier fusion systems by genetic algorithms, *IEEE Trans. on Evolutionary Computation*, vol.4, no.4, pp.327-336, 2000.
[12] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Kuhn and A. K. Jain, Dimensionality reduction using genetic algorithms, *IEEE Trans. on Evolutionary Computation*, vol.4, no.2, pp.164-171, 2000.
[13] J. H. Yang and V. Honavar, Feature subset selection using a genetic algorithm, *IEEE Intelligent Systems and Their Applications*, vol.13, no.2, pp.44-49, 1998.
[14] L. Yao, Genetic algorithm based identification of nonlinear systems by sparse Volterra filters, *IEEE Trans. Signal Processing*, vol.47, no.12, pp.3433-3435, 1999.
[15] C.-R. Huang, P.-C. Chung, B.-S. Sheu, H.-J. Kuo and M. Popper, Helicobacter pylori-related gastric histology classification using support-vector-machine-based feature selection, *IEEE Trans. Information Tech. in Biomedicine*, vol.12, no.4, pp.523-531, 2008.
[16] Y. Wang, S. Wan and K. K. Lai, A new fuzzy support vector machine to evaluate credit risk, *IEEE Trans. Fuzzy Syst.*, vol.13, no.6, pp.820-831, 2005.
[17] B. Gholami, W. M. Haddad and A. R. Tannenbaum, Relevance vector machine learning for neonate pain intensity assessment using digital imaging, *IEEE Trans. Biomedical Engr.*, vol.57, no.6, pp.1457-1466, 2010.
[18] X. Song, S. K. Halgamuge, D. Chen, S. Hu and B. Jiang, The optimized support vector machine with correlative features for classification of natural spearmint essence, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.1089-1099, 2010.
[19] C. Wu, Y. Guo, X. Zhang and H. Xia, Study of personal credit risk assessment based on support vector machine ensemble, *International Journal of Innovative Computing, Information and Control*, vol.6, no.5, pp.2353-2360, 2010.
[20] E. F. Petricoin, A. M. Ardekani, B. A. Hitt, P. J. Levine, V. A. Fusaro, S. M. Steinberg, G. B. Mills, C. Simone, D. A. Fishman, E. C. Kohn and L. A. Liotta, Use of proteomic patterns in serum to identify ovarian cancer, *Lancet*, vol.259, pp.572-577, 2002.
[21] L. H. Loo, J. Quinn, H. Cordingley, S. Roberts, L. Hrebien and M. Kam, Classification of SELDI-TOF mass spectra of ovarian cancer serum samples using a proteomic pattern recognizer, *Proc. of the 29th IEEE Annual Bio. Eng.*, pp.130-131, 2003.
[22] L. Yao and C.-C. Lin, Identification of nonlinear systems by the genetic programming based Volterra filter, *IET Signal Processing*, vol.3, no.2, pp.93-105, 2009.