

# Feature Selection Using Support Vector Machines and Independent Component Analysis for Wound Infection Detection by Electronic Nose

Jingwei Feng\*, Fengchun Tian, Pengfei Jia, Qinghua He<sup>1</sup>, Yue Shen<sup>1</sup> and Tao Liu

College of Communication Engineering, Chongqing University, Chongqing 400044, China  
<sup>1</sup>Department of Orthopedic and Traumatic Surgery, Center for War Wound and Trauma of PLA, Institute of Surgery Research, Daping Hospital, Third Military Medical University, Chongqing 400042, China

(Received June 13, 2012; accepted January 21, 2013)

**Key words:** feature selection, electronic nose, support vector machine, independent component analysis, wound infection detection

When mice are used as experimental subjects in the detection of wound infection based on electronic nose (Enose), the background, i.e., the smell of the mice themselves, is very strong, and most useful information is buried in it. A new feature selection technique specifically designed to work with support vector machine (SVM) and independent component analysis (ICA) is introduced. The features that represent background and noise are eliminated to improve classification accuracy. To assess this new method, two other datasets are used as validation, and four other feature selection methods are compared. The result shows that this method is effective and practical for feature selection in the detection of wound infection. Besides, this method is also useful in dimensionality reduction.

## 1. Introduction

Enose has been extensively studied and it is now successfully used in many fields. In this paper, Enose is used for wound infection detection. The measurement using Enose represents one way of realizing a cheap and sensitive method of detecting gaseous components.<sup>(1)</sup> Moreover, the type and growth phase of bacteria in wound infection can be monitored by examining the volatile compound concentration around the wound.<sup>(2)</sup> Thus, it is possible to use Enose to detect wound infection. Compared with traditional test methods, such as gas chromatography-mass spectrometry (GC/MS), the Enose has

---

\*Corresponding author: e-mail: cqfjw@cqu.edu.cn

the characteristics of being noninvasive, convenient, highly efficient, can function in real time and potentially superior in the detection of wound infection.

Mice whose wounds are infected with one of three bacteria, *P. aeruginosa*, *E. coli*, and *S. aureus*, are used as experimental subjects, but the background, i.e., the smell of the mice themselves, is very strong. Most useful information is buried in this background, and it constitutes a serious impediment in obtaining good discrimination results.

ICA is widely recognized as a useful tool for analyzing the data structure. In ICA, the data are linearly transformed such that the resulting coefficients are statistically as independent as possible.<sup>(3)</sup> It has been generalized for feature extraction.<sup>(3-7)</sup> ICA can be used to separate useful information and background and noise in wound infection detection signals. However, according to the ambiguities of ICA,<sup>(8)</sup> it cannot be determined which independent components are background and noise.

Variable selection is aimed at getting rid of those sensors or response variables that are redundant, noisy or irrelevant for the classification or quantification tasks envisaged, in such a way that the dimensionality of input space can be reduced without loss of useful information.<sup>(9)</sup> The algorithms of feature selection are divided into three categories: filters, wrappers, and embedded.<sup>(10)</sup> Filter methods are independent of the inductive algorithm, whereas wrapper methods use the inductive algorithm as evaluation function.<sup>(11)</sup> Embedded methods incorporate feature selection as a part of inductive algorithm.

The key and difficult task for wound infection detection by Enose is how to select good ICA features without background and noise to aid in fast processing and pattern classification.

The technique using the support vector machine (SVM), which was developed by Vapnik,<sup>(12)</sup> is a promising classification technique, in which the formulation of a learning problem leads to quadratic programming with linear constraints. SVM is not only good at classification<sup>(13-23)</sup> but also widely used in feature selection.<sup>(9,24-33)</sup>

In this paper, a new method such as the embedded method for feature selection, which is specifically designed to work with SVM and ICA, is introduced. This method selects features that are least affected by background smell or least noisy for constructing the classification model. The performance of this method is compared with those of other methods that have been proposed in refs. 9, 25, 34, and 35. In addition to the wound infection detection dataset, two other datasets that correspond to gas-sensor-based Enose are used as validation datasets to assess this new method. The datasets that are used in this paper are all small in size and it is easy to encounter the problem of overfitting. SVM with good generalization ability is less prone to overfitting than other classifiers, because of structural risk minimization. The feature selection can also help avoid overfitting by getting rid of the features that are redundant, noisy or irrelevant.

## 2. Materials and Methods

### 2.1 Wound infection detection

There are five types of mouse, namely, wounded but uninfected, infected with *P. aeruginosa*, infected with *E. coli*, infected with *S. aureus*, as well as no wounds and infection (used as background). Each type has four mice, and each wounded mouse

has one wound in its hind leg. The mice for use in the experiment are provided by the Animal Experiment Center of the Third Military Medical University.

In constructing a gas sensor array, fourteen metal oxide sensors and one electrochemical sensor are selected. They were nine TGS sensors from Figaro Engineering Inc., TGS-826, TGS-813, TGS-825, TGS-800, TGS-816, TGS-2620, TGS-822, TGS-2602, TGS-2600, one XSC sensor from New Creators Electronic Technology Co., Ltd., WSP-2111, two MQ sensors from Winsen Electronics Technology Co., Ltd., MQ-138, MQ-135, one QS sensor from Bluemoon Technology Co., Ltd., QS-01, one FIS sensor from FIS Inc., SP3S-AQ2, and one electrochemical sensor from Dart Sensors Ltd., AQ (air quality) sensor. The gas sensor array is placed in a stainless steel test chamber, the volume of which is 0.24 l (see Fig. 1). The sensors are mounted on a custom designed printed circuit board (PCB) (see Fig. 2), and associated electrical components are mounted on another PCB. There are seventeen sensors in this sensor array, but GSBT11 (Ogam Technology Co., Ltd.) and 4ETO (City Technology Ltd.) are broken, so we just use fifteen sensors that are described above. A 32-channel and 14-bit high-precision data acquisition system (DAS) is employed for the fifteen gas sensors. The heater voltage of each sensor is  $5 \pm 0.05$  V, and the voltage of the amplifying chip is  $5 \pm 0.01$  V.

Figure 3 shows the practical electronic nose system for wound infection detection. From Fig. 3, our Enose system is composed of Enose, DAS, pump, rotor flowmeter, 3-way valve, filter, glass bottle, and computer. The filter is used to obtain clean air. Figure 4 shows the connection of our Enose system. Each mouse was put in the glass bottle with a rubber stopper. Two holes were made in the rubber stopper with two thin glass tubes inserted. One glass tube was placed over the wound as close as possible. The output of the glass tube contains the volatile organic compounds (VOCs) of the wound in the mouse. Another glass tube was used for the input of clean air.

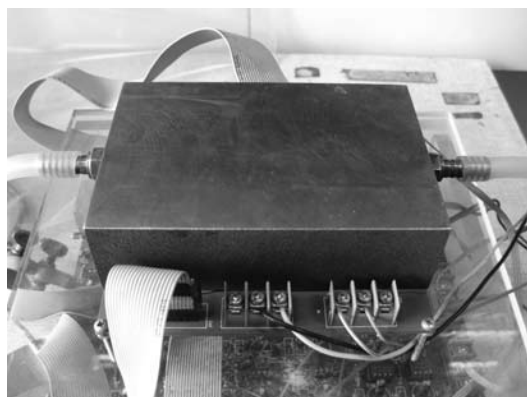


Fig. 1 (left). Test chamber of electronic nose system.

Fig. 2 (right). Sensor array.



Fig. 3. Practical electronic nose system for wound infection detection.

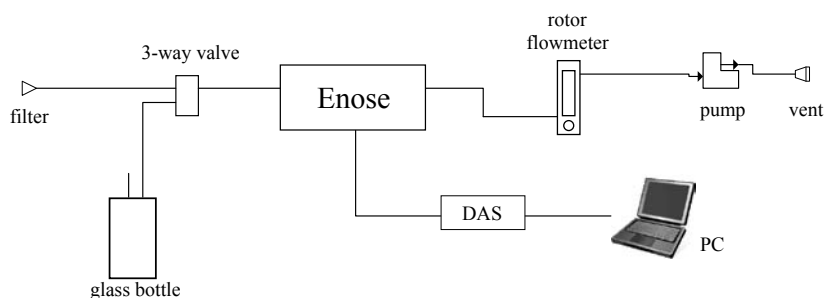


Fig. 4. Connection diagram of the electronic nose system for wound infection.

Each experiment comprises three stages: baseline stage, response stage, and recovery stage. In the first stage, the sensors were exposed to clean air for 3 min. In the second stage, the gas stream that contains VOCs of the wound was passed over the sensors for 5 min. In the last stage, the sensors were exposed to clean air again for 15 min. In all the three stages, the flow rate was maintained at 50 ml/min and the DAS sampled the data every 100 ms. The overall experiment was repeated five times for every mouse. The interval between every experiment was 5 min for the cleaning of the chamber with clean air.

## 2.2 Validation datasets

The first validation dataset comes from breast cancer detection. Enose is used to detect the volatile markers of breast cancer in the breath, which are nonane, heptanal, and 1-phenylethanol.

The second comes from wound pathogen detection. Enose is used to detect seven species of pathogen most common in wound infection. The seven species of pathogen are *P. aeruginosa*, *E. coli*, *Acinetobacter sp.*, *S. aureus*, *S. epidermidis*, *K. pneumoniae*, and *S. pyogenes*.

The Enose in breast cancer detection is the same as that in wound pathogen detection, which consists of six metal oxide gas sensors and one electrochemical gas sensor. More details on these gas sensors and wound pathogen detection experiment are given in a previously published paper.<sup>(36)</sup>

Because of strong noise, the electrochemical gas sensor is discarded. The first dataset has 27 samples and each gas has 9 samples. The second dataset has 70 samples and each species of pathogen has 10 samples.

### 3. SVM-Based Feature Selection

ICA is a new multivariate data analysis technique for blind source separation. In wound infection detection, the signals from the wound of mice become mixed with background smell or other noise before the sensors receive them. ICA, which extracts statistically independent components from the obtained dataset, could help in eliminating the noise from the obtained signals and retain useful information.<sup>(4)</sup> The method that we proposed is specifically designed to work with SVM<sup>(37)</sup> and ICA. This method uses ICA to separate background smell and useful information, and uses SVM to eliminate background smell.

Figure 5 summarizes the main steps and concepts of the feature selection process that we proposed. The details on how the different steps are implemented are given as follows.

**Step 1:** The preprocessing step and pattern recognition system are often integrated parts of the Enose and, thus, a fast evaluation of data is possible.<sup>(38)</sup> Preprocessing is very important, which can directly affect the discrimination. First, a moving average filter whose span is 5 is used to smooth the original response data of sensors. Then, the relative method<sup>(38)</sup> is used for baseline correction, and maximum values of each sensor response are extracted as features. Scaling provides that each variable is scaled to an equal variation and, thus, each variable will have the same opportunity to affect the classifier.<sup>(39)</sup> Thus, array autoscaling is also used in this step.

**Step 2:** The FastICA package provided in ref. 3 is used in the Matlab environment to conduct the ICA analysis. This FastICA algorithm was chosen, because when compared with the other methods of estimating the independent components, in this method, the convergence is cubic in nature.<sup>(8)</sup> The nonlinearity function 'tanh' used in the fixed-point algorithm is chosen. The FastICA package is run using the symmetric approach that estimates all the independent components in parallel. The number of independent components to be estimated equals the dimension of data, i.e., the number of IC scores equals the number of sensors. Array autoscaling is also used on the new data that are composed of IC scores.

**Step 3:**  $w_j$  denotes the optimal weights of kernel function in the SVM classifier, which is obtained by using all the features except the  $j$ th one, and  $j = 1, \dots, M$ , where  $M$

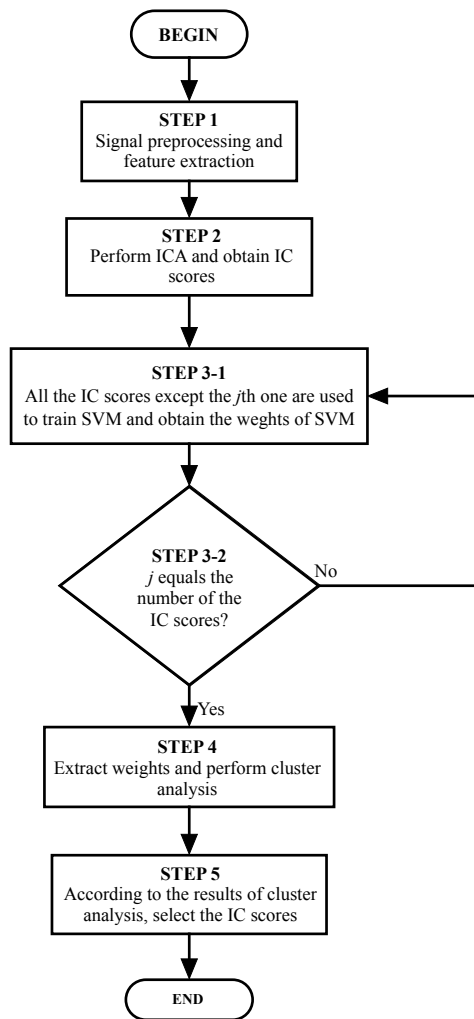


Fig. 5. Conceptual steps of the feature selection process.

is the number of IC scores. In this study, the kernel function is radial basis and a one-against-one strategy is used to build the SVM classifier. Considering an  $N$ -class problem, the one-against-one method builds  $N(N - 1)/2$  classifiers where each one is trained using input patterns from two classes.  $\mathbf{w}_j$  consists of weights of  $N(N - 1)/2$  classifiers. The weights of each classifier have two parts: one is for the class labeled by +1, and the other is for the class labeled by -1. All the maximum and minimum values are extracted in the part that corresponds to the class labeled by +1 and the part that corresponds to the class labeled by -1, respectively, in all the classifiers in each  $\mathbf{w}_j$ .

**Step 4:** After all the maximum and minimum values in all the parts of  $\mathbf{w}$  are extracted, a new feature matrix of  $W$  is formed. Each row of the matrix represents the maximum and minimum values of all the classifiers in each  $\mathbf{w}_j$ . Scaling is used on the matrix. The IC scores form clusters by using cluster analysis on the matrix  $W$ . During cluster analysis, the distance between samples in the matrix  $W$  is computed by using one minus the sample correlation between points. The agglomerative hierarchical cluster tree is created on the basis of this distance by using the unweighted average distance (UPGMA). The distance between the two subnodes merged at a node to measure node height in the agglomerative hierarchical cluster tree is computed to construct clusters.

**Step 5:** The features, IC scores, in one cluster are eliminated and the rest of the features in the other clusters are retained to be used as the inputs of SVM to accomplish the recognition. The kernel function of SVM is radial basis and a one-against-one strategy is used. Particle swarm optimization (PSO)<sup>(40)</sup> is used to determine the parameters in SVM. To evaluate the identification performance, the leave-one-out method<sup>(41)</sup> is used. After trying every cluster, we select the cluster that has the best discrimination results after eliminating this cluster. The features in the cluster that we selected are the background smell or noise. We consider that the background smell or noise has the same effect on the kernel weights of SVM. This method can select the features that have useful information and eliminate the useless ones.

#### 4. Results and Discussion

The dataset of wound infection detection has fifteen features (IC scores), and Fig. 6 shows the clustering results of the fifteen features.

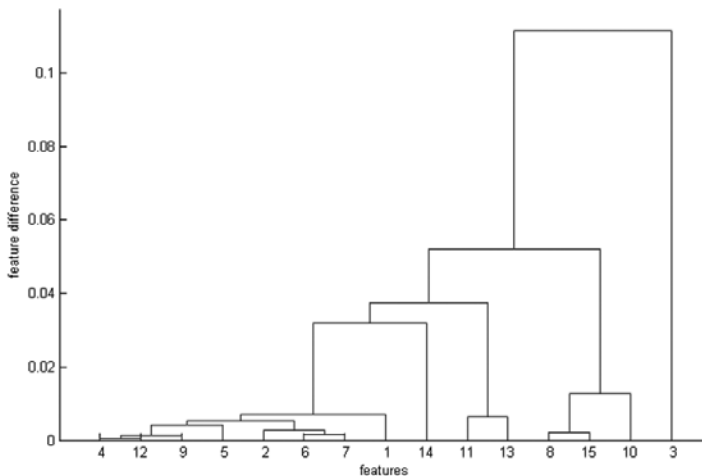


Fig. 6. Clustering results of the fifteen features.

From Fig. 6, we agglomerated the features in four clusters. The features of each cluster are set to zero, in turn. The rest of the features are used as the inputs of SVM to accomplish the discrimination of infection type. The recognition probability when each cluster is eliminated is given in Table 1. Table 1 shows that we can obtain the best results if we eliminate features 11 and 13 in cluster 2. Thus, features 11 and 13 in cluster 2 would be the background smell or noise that should be eliminated to improve the discrimination of infection type. We consider that the background smell or noise has the same effect on the kernel weights of SVM. The features that represent background smell or noise will agglomerate in one cluster. The VOCs of the wound contain more different sources than background smell or noise, thus, only one cluster is eliminated. The results that eliminate more than one cluster are shown in Table 2. According to Table 1, cluster 1 is very important; thus, this cluster is not eliminated. From Table 2, we can see that if we eliminate more than one cluster, the result will become worse.

The performance of this method is compared with those of other methods that have been proposed in refs. 9, 25, 34, and 35.

The first method is termed PSO+SVM. This method can simultaneously determine the parameter values of SVM and select features, without reducing the SVM classification accuracy. More details are given in ref. 34.

The second method is termed the L-J method.<sup>(30)</sup> This method ranks the features according to their influence on the decision hyperplane. The influence of the features is evaluated using the angle between the gradient of decision function of SVM and unit vectors that represent the indices of the individual features. More details are given in ref. 25.

The third method is given in ref. 9 and termed SFS+SVM. This method is inspired in sequential forward selection (SFS). The influence of one feature is evaluated by the difference between the squared norm of the optimal weight vectors of two separation

Table 1  
Recognition probability when each cluster is eliminated.

Cluster	Feature	Dimension	Recognition probability
1	1,2,4,5,6,7,9,12,14	6	71.25% (57/80)
2	11,13	13	96.25% (77/80)
3	8,15,10	12	91.25% (73/80)
4	3	14	91.25% (73/80)

Table 2  
Results of eliminating more than one cluster.

Eliminated clusters	Dimension	Recognition probability
1	2,3	95.00% (76/80)
2	2,4	87.50% (70/80)
3	3,4	87.50% (70/80)
4	2,3,4	87.50% (70/80)



hyperplanes. One is obtained by using all the features, and the other is obtained by using all the features except the evaluated one. The optimal weights of kernel function in SVM are used considering the difficulty in getting the weight vectors of the separation hyperplane in SVM when a kernel function is used.

The features in the three methods above are all IC scores. The classifier used in these three methods is also SVM, and PSO is used to determine the parameters in SVM.

The fourth method is given in ref. 35 and termed WT+RBF. It needs background data (20 background samples) that are obtained from the healthy mice. Thus, this method needs much more experiment. The direct multiplication of wavelet transform coefficients at corresponding scales between the response signals from wounded and healthy mice is used to eliminate the background smell. The feature values are extracted from the wavelet coefficients with the background eliminated.

The classifier used in method four is the Radical Basis Function network. According to comparison with SVM, the Radical Basis Function network acquires better results in this method.

The fifth method termed ICA+SVM is the method proposed by us. To assess the usefulness of ICA, the sixth method, which is the fifth method without ICA, is compared. This method is termed SVM. The feature values are maximum response values in this method.

The discrimination capabilities that are leave-one-out errors, dimension, and the features that are selected to be eliminated are given in Table 3.

From Table 3, we can see that the method that we proposed and the first method give the best classification rate, and the dimension is the lowest. However, the result of the first method is not constant, and the proposed method is constant. We run the first method 10 times, and the best result is given in Table 3, and all the ten results are given in Table 4. The proposed method can always give the best result. The reason is

Table 3  
Comparison among six different methods.

Method	Eliminated features	Dimension	Recognition probability
PSO+SVM (method 1)	11,13	13	96.25% (77/80)
L-J (method 2)	0	15	92.50% (74/80)
SFS+SVM (method 3)	0	15	92.50% (74/80)
WT+RBF (method 4)	/	14	95.00% (76/80)
ICA+SVM (method 5)	11,13	13	96.25% (77/80)
SVM (method 6)	1,4,5,7,13	10	83.75% (67/80)

Table 4  
Ten results of PSO+SVM.

Times	1	2	3	4	5	6	7	8	9	10
Results (%)	93.75	93.75	93.75	93.75	96.25	92.50	92.50	92.50	93.75	93.75

that the first method has to simultaneously determine seventeen parameters, and PSO in this method is easily trapped in local optimum; most of the time, it cannot give the best result. After feature reduction using the proposed method, the PSO just determines two parameters; it is much easier to obtain the best result. The features in the fourth method are wavelet coefficients that are different from the features in this paper; thus, the features that are eliminated were not given.

Because the performance changes with the varying  $k$  value in  $k$ -fold cross validation, we give the  $k$ -fold cross validation results of different  $k$  values in Table 5. From Table 5, we can see that this new method can always give the best results with different  $k$  values among six different methods.

In addition to the wound infection detection dataset, two other datasets that correspond to gas-sensor-based Enose are used to verify this new method. The first dataset comes from breast cancer detection. The second dataset comes from wound pathogen detection. The results of two datasets are shown in Table 6. The PSO+SVM method is also compared in Table 6.

One problem in the two datasets is that the dimension that is six is so low, but this method still can effectively select beneficial features. The result of PSO+SVM is also not constant. We also run the method 10 times and choose the best result.

From Tables 3 and 6, we can see that the method that we proposed gives a higher classification accuracy rate and a lower dimension across different datasets.

Table 5  
Recognition probability with different  $k$  values.

Method	$k$ value				
	80	40	20	10	5
PSO+SVM	96.25%	95.00%	95.00%	92.50%	92.50%
L-J	92.50%	92.50%	92.50%	91.25%	90.00%
SFS+SVM	92.50%	92.50%	92.50%	91.25%	90.00%
WT+RBF	95.00%	95.00%	95.00%	90.00%	88.75%
ICA+SVM	96.25%	97.50%	97.50%	95.00%	92.50%
SVM	83.75%	85.00%	85.00%	82.50%	83.75%

Table 6  
Results of the new method on the validation datasets.

Dataset	New method			PSO+SVM		
	Feature	Dimension	Recognition probability	Feature	Dimension	Recognition probability
1	5,6	4	96.30% (26/27)	5,6	4	96.30% (26/27)
2	3,5	4	100% (70/70)	3,5	4	100% (70/70)

## 5. Conclusions

In this study, we present a new method that is based on ICA and SVM, which is capable of selecting beneficial features. These optimal features are then adopted in both training and testing to obtain the optimal outcomes in classification. The classifier used in this new method is SVM, in which parameters are determined by PSO. The new method can be applied to eliminate background or noise features and improve the overall classification results in wound infection detection. Comparison of the obtained results with those of the other methods demonstrates that the new method has better classification accuracy and lower dimension than the others tested. This method can also give a higher classification accuracy rate and a lower dimension across different datasets.

## Acknowledgements

This work was financially supported by the Science Foundation Project of CQ CSTC, P. R. China (CSTC, 2009BA2021) and the Natural Science Foundation Project of CQ CSTC, P. R. China (CSCT, 2012JJA40005).

## References

- 1 S. Ehrmann, J. Jüngst, J. Goschnick and D. Everhard: *Sens. Actuators, B* **65** (2000) 247.
- 2 J. Carlsson: *Appl. Microbiol.* **25** (1973) 287.
- 3 O. W. Kwon and T. W. Lee: *Signal Processing* **84** (2004) 1005.
- 4 S. Balasubramanian, S. Panigrahi, C. M. Logue, C. Doetkott, M. Marchello and J. S. Sherwood: *Food Control* **19** (2008) 236.
- 5 L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee and Q. M. Gu: *Neurocomputing* **55** (2003) 321.
- 6 P. O. Hoyer and A. Hyvarinen: *Proc. 2000 Neural Networks Conf. (IEEE, 2000)* p. 369.
- 7 A. J. Bell and T. J. Sejnowski: *Vision Res.* **37** (1997) 3327.
- 8 A. Hyvärinen and E. Oja: *Neural Networks* **13** (2000) 411.
- 9 O. Gualdró, J. Brezmes, E. Llobet, A. Amari, X. Vilanova, B. Bouchikhi and X. Correig: *Sens. Actuators, B* **122** (2007) 259.
- 10 I. Guyon and A. Elisseeff: *J. Mach. Learn. Res.* **3** (2003) 1157.
- 11 M. Dash and H. Liu: *Intell. Data Anal.* **1** (1997) 131.
- 12 V. N. Vapnik: *The Nature of Statistical Learning Theory* (Springer, New York, 1995).
- 13 G. H. Kim, Y. W. Kim, S. J. Lee and G. J. Jeon: *Proc. 2010 SICE Annual Conf. (IEEE, 2010)* p. 1764.
- 14 M. Pardo and G. Sberveglieri: *Sens. Actuators, B* **107** (2005) 730.
- 15 C. Distante, N. Ancona and P. Siciliano: *Sens. Actuators, B* **88** (2003) 30.
- 16 K. Brudzewski, S. Osowski, T. Markiewicz and J. Ulaczyk: *Sens. Actuators, B* **113** (2006) 135.
- 17 K. Brudzewski, S. Osowski and T. Markiewicz: *Sens. Actuators, B* **98** (2004) 291.
- 18 F. J. Acevedo, S. Maldonado, E. Domínguez, A. Narváez and F. López: *Sens. Actuators, B* **122** (2007) 227.
- 19 X. D. Wang, H. R. Zhang and C. J. Zhang: *Proc. 2005 Machine Learning and Cybernetics Conf. (IEEE, 2005)* p. 3394.
- 20 O. Sadik, W. H. Land, A. K. Wanekaya, M. Uematsu, M. J. Embrechts, L. Wong, D. Leibensperger and A. Volykin: *J. Chem. Inf. Comput. Sci.* **44** (2004) 499.

- 21 M. Gaudioso, W. Khalaf and C. Pace: Proc. 2007 IEEE 7th Hybrid Intelligent Systems Int. Conf. (IEEE, 2007) p. 42.
- 22 L. Xie and X. D. Wang: Proc. 2009 Control and Decision Conf. (IEEE, 2009) p. 5148.
- 23 S. J. Dixon and R. G. Brereton: *Chemom. Intell. Lab. Syst.* **95** (2009) 1.
- 24 P. S. Bradley and O. L. Mangasarian: Proc. 1998 Machine Learning Conf. (ICML, 1998) p. 82.
- 25 L. Hermes and J. M. Buhmann: Proc. 2000 Pattern Recognition Int. Conf. (IEEE, 2000) p. 712.
- 26 M. H. Nguyen and F. de la Torre: *Pattern Recognit.* **43** (2010) 584.
- 27 I. Guyon, J. Weston, S. Barnhill and V. Vapnik: *Mach. Learn.* **46** (2002) 389.
- 28 T. Jebara and T. Jaakkola: Proc. 2000 Uncertainty in Artificial Intelligence Conf. (UAI, 2000) p. 291.
- 29 J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio and V. Vapnik: Proc. NIPS **13** (2001) 338.
- 30 C. T. Su and C. H. Yang: *Expert Syst. Appl.* **34** (2008) 754.
- 31 H. Choi, D. Yeo, S. Kwon and Y. Kim: *Comput. Stat. Data Anal.* **55** (2011) 1897.
- 32 S. Maldonado-Bascon, S. Al-Khalifa and F. Lopez-Ferreras: *Artif. Neural Nets Prob. Solving Methods* **2687** (2003) 798.
- 33 E. Phaisangittisagul: Proc. 2010 Knowledge Discovery and Data Mining Conf. (IEEE, 2010) p. 28.
- 34 S. W. Lin, K. C. Ying, S. C. Chen and Z. J. Lee: *Expert Syst. Appl.* **35** (2008) 1817.
- 35 J. Feng, F. Tian, J. Yan, Q. He, Y. Shen and L. Pan: *Sens. Actuators, B* **157** (2011) 395.
- 36 F. Tian, X. Xu, Y. Shen, J. Yan, Q. He, J. Ma and T. Liu: *Sens. Mater.* **21** (2009) 155.
- 37 V. N. Vapnik: *Statistical Learning Theory* (John Wiley and Sons, Hoboken, 1998) p. 375.
- 38 T. Skov and R. Bro: *Sens. Actuators, B* **106** (2005) 719.
- 39 R. Bro and A. K. Smilde: *J. Chemom.* **17** (2003) 16.
- 40 J. Kennedy and R. Eberhart: Proc. Neural Networks **4** (1995) 1942.
- 41 C. M. Bishop: *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).