Available online at www.sciencedirect.com
ScienceDirect
Applied Soft Computing 7 (2007) 246–256
Applied Soft Computing
www.elsevier.com/locate/asoc

Wavelet filter for noise reduction and signal compression in an artificial nose

C. Zanchettin *, T.B. Ludermir

Federal University of Pernambuco, Center of Informatics, P.O. Box 7851-50.732-970, Cidade Universitária, Recife-PE, Brazil

Received 29 April 2004; received in revised form 14 December 2004; accepted 12 June 2005

Abstract

This work presents results of the use of a wavelet filter for noise reduction and data compression of signals generated by artificial nose sensors. To verify the performance of the wavelet analysis in the treatment of odor patterns, we compare two widely used artificial nose classifiers, multi-layer perceptron neural network and time delay neural network in the analysis of signals generated by eight conducting polymer sensors exposed to gases derived from the petroliferous industry.

Keywords: Artificial neural networks; Wavelet transform; Artificial noses

1. Introduction

For a number of decades, scientists have recognized the power of incorporating biological principles into the design of artificial devices or systems. One contemporary example of this approach is the development of artificial noses [11]. Artificial noses are systems developed for the automatic detection and classification of odors, vapors and gases. These electronic devices have two main components, the sensor system and the automated pattern recognition system. In the artificial nose, the odor recognition process begins in the sensor system, which is responsible for the capturing or measurement of the odorant stimulus through the sensitivity of its sensors. The sensor system can be made up of either a set of distinct sensors, where each element measures a different property of the odorant composition or a single device that produces a set of measurements and/or characteristics. Each odorant substance is presented to the sensor system, which generates a pattern of resistance values that characterize the odor. This pattern is presented to the recognition system, which, in turn, classifies the odorant stimulus.

In this process, data preprocessing is an important step before classification can be performed. The phase is important because a number of different problems can compromise the performance of sensor system. These include: (1) the odor signal measurer can present disturbance or noise; (2) the data acquisition process tends to be unstable (e.g. the sensors present variations during the acquisition phase, which can be

* Corresponding author.
E-mail addresses: cz@cin.ufpe.br (C. Zanchettin), tbl@cin.ufpe.br (T.B. Ludermir).

1568-4946/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2005.06.004

transformed into ‘outliers’ in the database); and (3) the propagation of the signal through the communication channel between the sensor and pattern recognition systems can be contaminated by interference signals coming from the surrounding environment. The treatment of odor patterns is necessary before they are presented to the pattern recognition system to compensate for the concentration drifts in the response arrays, eliminate noise and normalize data.

Considering the odorant signals obtained by the artificial nose are signals acquired by a sensor in a particular time space, it seems natural that techniques used for digital signal processing are the appropriate tools for the treatment of these data. For this purpose, the wavelet transform [6] was used as the preprocessing method for the odorant signals.

The wavelet transform is a time-scale representation that has been used successfully in a broad range of applications. Its main feature is its time–frequency analysis capability using a single transformation, which makes it useful in applications such as signal de-noising, wave detection, data compression, feature extraction, etc. [2,19,27] and [12]. This work proposes the utilization of the wavelet transform for preprocessing odor patterns as a filter for noise reduction and compressing the signal generated by the sensor system. The results of this wavelet filter are compared in two widely used artificial nose classifiers, the multilayer perceptron neural network (MLP) [23] and the time delay neural network (TDNN) [15]. The remainder of this paper is divided into five sections. Section 2 presents the wavelet method and the preprocessing results. Section 3 shows details of the experiments. Section 4 describes the strategy used for performance assessment. In Section 5, the results of the experiments are presented. Section 6 contains a summary and conclusions.

2. Wavelet transform

The wavelet transform is a signal processing technique that represents a transient or non-stationary signal in terms of time and scale distribution. Due to its light computational complexity, the wavelet transform is an excellent tool for on-line data compression, analysis and de-noising.

Unlike more traditional filtering methodologies, wavelet transforms have the ability to preserve the temporal locality of sharp transitions within time domain signals. This property is important in a fault detection context, as sharp transitions are likely indicators of fault conditions and, hence, any utilized filtering methodology should not disturb the location of their occurrence.

The basic idea behind signal processing with wavelets is that, as in Fourier analysis [3], a signal can be decomposed into its component elements through the use of basis functions. In Fourier analysis, the basis functions are sine and cosine waves. In the case of wavelet analysis, the basis functions consist of the wavelet scale function, as well as scaled and shifted versions of the mother wavelet function. The scale function in wavelets is used to capture the general (or low detail) information on the signal, whereas different mother wavelet scales are used to capture the details of the signal, with each successive scale capturing (describing) finer and finer levels of detail.

Fig. 1 illustrates the process where the resolution of the time-domain signal $x(k)$, $k = 1, \ldots, N$, is changed by low/high pass filtering operations and the scale is changed by downsampling/upsampling operations. The parameters of the wavelet transform are the type of the wavelet filter used and the number of decomposition levels ($l = 1, \ldots, L$).

At low scale levels, time resolution is traded for better frequency resolution, thereby allowing low frequency events to be analyzed very accurately with respect to their frequency content, but not with respect to their location in time. At high scale levels, frequency resolution is traded for time resolution; the location of high frequency events is accurately marked in time, but their actual frequency content is poorly resolved.

The wavelet transform can be given by:

$$d_{0,0} = \langle g(t), \phi(t) \rangle \tag{1}$$

$$d_{j,k} = \langle g(t), \psi_{j,k}(t) \rangle, \quad j = 1, \ldots, N, \ \ k = 1, \ldots, 2^{j-1} \tag{2}$$

where $d_{j,k}$ are the wavelet coefficients, $g(t)$ is the signal to be transformed, $\phi(t)$ the scale function, $\psi_{j,k}(t)$ the scaled and shifted version of the mother

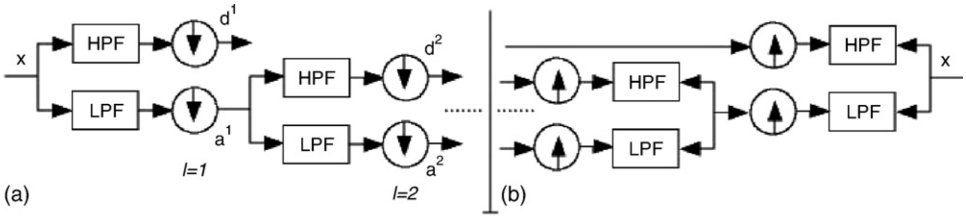

Fig. 1. (A) Forward and (B) inverse wavelet transform.

wavelet function $\psi(t)$ given by:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \quad (3)$$

and, N is the number of wavelet scales over which the wavelet transform is generated. Following this notation, the inverse wavelet transform can therefore be given by:

$$g(t) = d_{0,0}\phi(t) + \sum_{j,k} d_{j,k}\psi_{j,k}(t) \quad (4)$$

2.1. Compressing data and de-noising processing

A possible application of the discrete wavelet analysis is to remove undesired components (noise) from the signal through a de-noising approach. Basically, the procedure includes decomposing the signal into the detail components described above, identifying the noise components, and reconstructing the signal without these components.

The wavelet de-noising approach is based on the observation that random errors in a signal are present in all coefficients, while deterministic changes get captured in a small number of relatively large coefficients. As a result, a non-linear thresholding (shrinking) function in the wavelet domain will tend to keep a few larger coefficients representing the underlying signal, while the noise coefficients will tend to reduce to zero. The advantage of the wavelet de-noising method over classical linear filtering methods is that it attempts to remove whatever noise is present and retain whatever signal is present regardless of the frequency content of the signal.

The compression features of a given wavelet basis are primarily linked to the relative scarceness of the wavelet domain representation for the signal. The notion behind compression is based on the concept that a regular signal component can be accurately approximated using the following elements: a small number of approximation coefficients (at a suitably chosen level) and some of the detail coefficients. For the de-noising and compression process, we consider the following model of a discrete noisy signal:

$$y(n) = f(n) + \sigma e(n) \quad n = 1, \ldots, N \quad (5)$$

The vector y represents a noisy signal and f is an unknown, deterministic signal, where it is assumed that e is Gaussian white noise $N(\mu\sigma^2)$ = M(0, 1). For filtering out white noise, we use a method proposed by Donoho [8].

The method can be carried out through the following three steps:

(1) Transform the signal $f(k)$ corrupted by noise into wavelet domain, and get a group of wavelets coefficients $w_{j,k}$.
(2) Disposal $w_{j,k}$ using soft- or hard-thresholding function, thereby suppressing those coefficients smaller than a certain amplitude, then obtain a group of estimate coefficients $\widehat{w_{j,k}}$.
(3) Recontruct signal using $\widehat{w_{j,k}}$, then obtained the estimated signal $\widehat{f_k}$, which is the de-noised signal.

The choice of the mother wavelet plays a significant role in de-noising and data compression process. Wavelet analysis is a measure of similarity between the basis functions and the signal itself. Here the similarity is in the sense of having similar frequency content. Therefore, in this case, the mother wavelets must be highly localized in time and frequency.

In the present work, we are interested in detecting low amplitude, short duration, fast decaying and oscillating types of signals. One of the most popular orthonormal wavelet is Daubechies' wavelet, which has been shown to meet the requirements. Daube-

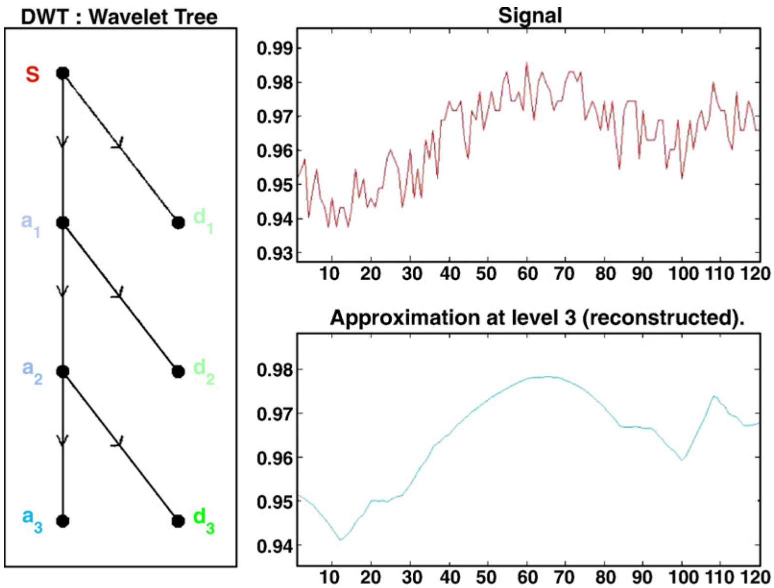

Fig. 2. Tree of decomposition used to de-noise and compress the odor signal.

chies' wavelet with various filter coefficients was studied. It was been found that larger filter coefficients generate more localized wavelets in both time and frequency dimensions.

In experiments, we employed the order three wavelet transform of the *Daubechies*' family [5]. The wavelet filter was implemented using the *wavelets toolbox* [18] in the Matlab 6.5.1 simulation software. The wavelet filter is implemented through the multi-resolution pyramid decomposition technique [16]. In principle, signal s is decomposed into two sets: a d_{m} set of wavelet coefficients (also known as 'details') and an a_{m} set of coefficients known as 'approximation', using filters g and h for the decomposition, respectively. Digital filter g is a pass band filter. Thus, the filtered d_1 signal is a detailed version of the s signal and possesses high frequency components in comparison to the a_1 signal approach, which uses a pass low filter. The a_m set later received a new decomposition to generate the a_{m+1} and $d_{\mathrm{m}+1}$ coefficients, and so on.

The decomposition of the odorant signal s can be described as:

$$s = s + n = a_m + \sum_{i=1}^{m} d_i \tag{6}$$

where m is the greatest decomposition level and n the noise.

The data acquisition method in the artificial nose induces low-frequency noises. The signals free of noise were found in the component of lowest frequency of a tree with three decomposition levels (a_3). Fig. 2 displays the decomposition tree of the analyzed signal. From this decomposition, the original odor signal was reconstructed from the third decomposition level without the noise components that could hinder the classifier performance.

3. Experiments

This study sought to classify five gases from the petroliferous industry (Petro-bras, Brazil). A prototype of an artificial nose was used to acquire the data [24]. Sensor systems are often built with polypyrrol-based gas sensors. Some advantages of using this kind of sensor include [21]: (1) rapid absorption kinetics at room temperature; (2) low power consumption (in terms of microwatts), as no heating element is required; (3) resistance to poisoning; and (4) the possibility of building sensors tailored to particular classes of chemical compounds.

The prototype was composed of eight distinct gas sensors built by electro-chemical deposition of polypyrrol using different types of dopants. Tests were carried out with the petroleum gas derivates methane, ethane, propane and butane. Sensitivity was verified with carbon monoxide gas.

The data sets were obtained with nine data acquisitions for each of the gases by recording the resistance value of each sensor every twenty seconds over a period of 40 min. In the acquisition phase, each sensor obtained 120 resistance values for each gas. A pattern is a vector of eight elements representing the resistance values recorded by the sensor array. Thus, each acquisition contains 600 patterns formed by 960 values for each of the gases.

3.1. Multi-layer perceptron neural networks

Artificial neural networks have been widely applied to pattern recognition systems of artificial noses [4]. This approach has many advantages, such as the ability to handle non-linear signals from the sensor array, adaptability, fault tolerance, noise tolerance and inherent paralelism generating a high speed of operation subsequent to training. Among the several existing models of artificial neural networks (ANNs), the MLP is the most widely used, partly for its ease of implementation, and partly for its simplicity. These characteristics also make the MLP one of the most employed ANNs for odor classification in artificial noses [28,7,11,14,25] and [24].

The data set for training and testing the network are divided into three distinct sets: a training set, containing 50% of the total amount of patterns; a validation set, containing 25% of the patterns; and test set, which contains the remaining 25%. This division of data is suggested by [22]. The five gases to be classified are represented by the same amount of patterns in each of the three data sets.

The pattern set was normalized so as to be within a range of values between −1.0 and +1.0. A MLP neural network was used, containing only one hidden layer. The input layer had eight units, corresponding to the values of the eight sensors. The output layer had five units, corresponding to the five gases to be classified (in this experiment the 1-of-*m* output coding was used, where *m* is the number of classes). All network-processing units were implemented *hyperbolic tangent activation functions* [22]. The neural network contains all possible feedforward connections between adjacent layers, without having any connection between non-adjacent layers. In this work, five distinct topologies were trained (4, 8, 12, 16 and 20 units in the hidden layer).

The training algorithm used is a version of the Levenberg–Maquardt method described in [9]. For each topology, 30 different and random weight initializations were performed. In all cases, the maximum quantity of iterations was 5000. The training stops if the criterion GL_5 of Proben1 [22] is satisfied twice (to avoid initial oscillations in validation errors). The GL_5 criterion provides an idea of the generalization loss during training and it is sufficiently useful in avoiding overfitting. It is defined as the increase in the overall validation error in relation to the minimum validation error. Training also stops if the training progress criterion defined in Proben1 [22], with $P_5(t) < 0.1$, is satisfied, or if the maximum quantity of 5000 iterations is reached.

The error measure used in the analysis of results was the squared error percentage, presented in Eq. (7).

$$E = 100\frac{o_{\max} - o_{\min}}{NP}\sum_{p=1}^{P}\sum_{i=1}^{N}(o_{pi} - t_{pi})^2 \qquad (7)$$

where $o_{\min}$ and $o_{\max}$ are the minimum and maximum values of output coefficients in the problem representation, N is the number of output nodes of the network, P the number of patterns (examples) in the data set considered, and t the desired output of the network.

Another error measure used is the classification error of the test set, which corresponds to the number of incorrectly classified patterns divided by the total quantity of patterns. Thus, the aspects observed at the end of training were the squared error percentage of the training, validation and test sets and the classification error of the test set.

3.2. Time-delay neural networks

Time-delay neural networks (TDNN) were originally developed for speech recognition by [26]. The main goal in development of TDNN was to have neural network architecture for non-linear feature classification invariant under translation in time or space. TDNN uses built-in time-delay steps to represent temporal relationships. The translation invariant classification is accomplished by sharing the connection units of the time delay steps. The activation of each TDNN processing unit is computed by the weighted summation of all activations of predecessor processing units in an input window over time and applying a non-linear function (i.e. a sigmoid function) to the sum.

The TDNN proposed for odor recognition consists of a pattern recognition system capable of analyzing the temporal features of the signals generated by the sensors of the artificial nose. This system considers the variation of the signals generated by the sensors along the time interval in which the data acquisitions were accomplished. Thus, the TDNN has 16 inputs, corresponding to the curves generated by eight sensors along time, and five outputs representing the classification of the odor signal. The TDNN is a type of ANN that has presented promising results in odor classification [28,31,29] and [30].

As described in [28], the architecture of the pattern recognition system needs to receive complete curves generated by the sensors during the data acquisition. Thus, only complete data acquisitions can be used as training, validation and test sets. One of the data acquisitions (960 patterns of each gas) was used as the training set. Similarly, two further acquisitions (with the same amount of data) were used as the validation and test sets. The choice of the acquisitions was made arbitrarily by using the three first data acquisitions.

The same data set in the previous section was used, allowing a comparison between approaches. The same normalization, topologies, activation function, output coding, training algorithm and stopping criteria adopted for the MLP experiments were used. The architecture used was a TDNN containing only one hidden layer. The input layer had 16 units, as the input was formed by the current pattern and the previous pattern (one delay for each sensor). For each topology, 30 runs were performed with different and random weight initializations. In all cases, the maximum number of epochs allowed was 100.

It is important to emphasize that the analyzed errors were computed separately for each gas. In other words, for each sensor the network evaluated the five curves corresponding to the five gases independently, so that no single curve contained the values of the five gases in sequence. This was adopted so that the presentation order of the gases to the network would not influence the results.

4. Performance assessment

Empirically evaluating the accuracy of hypotheses is fundamental to machine learning. The statistical methods for estimating hypothesis accuracy focus on three basic questions. First, given the observed accuracy of a hypothesis over a limited sample of data, how well does it estimate its accuracy over additional examples? Second, given that one hypothesis outperforms another over some sample of data, how probable is it that this hypothesis is more accurate in general? Third, when data is limited, what is the best way to use this data both to learn a hypothesis and estimate its accuracy? As limited samples of data might misrepresent the general distribution of data, estimating true accuracy from such samples can be misleading. Statistical methods, together with assumptions on the underlying distributions of data, allow one to draw boundaries between the observed accuracy over the sample of available data and the true accuracy over the entire distribution of data [20].

To verify the classifier performance, all continuous variables were expressed as mean standard deviation in the analysis of the experiments. The 2-tailed paired Student's *t*-test was used to compare the related samples [13]. The Student's *t*-test is a statistical method that deals with problems associated to inference based on '*small*' samples. The test is used to verify the hypothesis that a given variable differs between two groups, but the paired test is specifically used when each data point in one group corresponds to a matching data point in another group.

The degree of agreement between the methods, as well as the variabilities of the wavelet filter method, were assessed by use of the analysis method reported by Bland and Altman [1]. This technique compares two methods by calculating the mean and 95% range of the differences between the data points of the methods. In the Bland–Altman test, the data analysis may be displayed in the form of a graph on which the *x*-axis shows the mean of the results of the two methods ([A + B]/2), while the *y*-axis represents the difference between the two methods ([B − A]), which may be plotted on either an absolute, percentage, or logarithmic scale. The graph includes a line for the mean difference between the methods and lines for the 2-s limits of the mean difference.

The linear regression analysis and Pearson correlation were also used for assessing the relation between both methods (multi-layer perceptron versus multi-layer perceptron with wavelet filter). The results were considered statistically significant when $P < 0.05$.

5. Results

The experimental results are presented in this section. To measure the performance of wavelet analysis, the MLP was trained with and without the preprocessing data by the wavelet filter. The results are displayed in both Table 1 (original data) and Table 2 (preprocessing data).

With the use of the wavelet analysis, a considerable improvement in the classification errors of odor patterns was achieved. The best classification result was in the topology with four nodes in the hidden layer, classification error of 14.39% (Table 1). With the use of the wavelet analysis the error was reduced to 11.50% (Table 2).

The same comparison was performed for the TDNN neural network. The results are presented in Table 3 (original data) and Table 4 (preprocessing data). The use of the wavelet filter improved the results of odor classification. The results with the TDNN are better than those presented for the MLP network. Without the use of wavelet analysis, the TDNN presented a classification error of 11.34% (Table 3,

Table 1
Mean and standard deviation for the results of MLP experiments with original data

Hidden nodes	Training error		Validation error		Test error		Classification error of test set	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4	2091.89	3500.98	1067.58	1783.34	1090.91	1773.15	0.1439	0.1882
8	1849.77	2681.81	967.09	1379.39	968.29	1371.35	0.1462	0.913
12	3060.68	4440.45	1549.07	2217.48	1554.59	2214.10	0.2024	0.572
16	3690.66	4174.59	1868.96	2093.64	1873.48	2090.15	0.2622	0.626
20	5636.20	5182.06	2845.62	2588.31	2864.12	2588.75	0.3521	0.3088

Table 2
Mean and standard deviation for the results of MLP using wavelet filter

Hidden nodes	Training error		Validation error		Test error		Classification error of test set	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4	2067.22	3401.50	1063.94	1711.02	1053.51	1710.92	0.1378	0.2039
8	1425.67	2373.96	737.11	1209.04	731.34	1194.09	0.1150	0.1735
12	2640.24	4278.83	1329.16	2149.28	1326.96	2141.86	0.1729	0.2547
16	3761.41	4944.03	1899.51	2483.61	16250.16	2504.78	0.2365	0.2861
20	5456.01	4913.29	2751.25	2454.55	2761.49	2488.42	0.3446	0.2834

Table 3
Mean and standard deviation for the results of TDNN, with original data

Hidden nodes	Training error		Validation error		Test error		Classification error of test set	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4	6.4900	4.4133	85.0300	25.6867	61.9833	19.2867	0.1259	0.1024
8	2.9167	2.9167	73.3800	30.1867	64.0167	16.3033	0.1134	0.0873
12	4.7700	9.9867	72.7833	33.9867	69.2333	15.2767	0.2052	0.1053
16	46.8000	49.7900	106.1500	47.5533	87.8567	38.3400	0.3151	0.1545
20	52.1033	35.4200	106.7600	43.6933	100.8667	50.4200	0.3238	0.1872

Table 4
Mean and standard deviation for the results of TDNN using wavelet filter

Hidden nodes	Training error		Validation error		Test error		Classification error of test set	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4	1.1800	1.7400	30.6900	24.2867	71.4233	12.4700	0.0316	0.0678
8	1.2533	1.9700	26.6500	22.0067	66.7467	7.3867	0.0075	0.0371
12	3.8100	9.6367	34.3900	23.7067	70.5833	11.1700	0.0200	0.0610
16	41.6367	48.9900	76.7333	52.6367	87.3267	38.2867	0.1895	0.1892
20	49.9000	44.3400	82.2567	51.8800	92.0033	38.1067	0.2372	0.1940

topology with eight hidden nodes); using wavelet filter the error is less than 1%, more accurately at about 0.75% (Table 4, topology with eight hidden nodes). Another point to consider was the low standard derivation presented for the network using the filter, which suggests a small variation in the results of the 30 runs. The excellent results obtained by the TDNN were possible due to its capacity for using the temporal characteristics of odor patterns in the neural network learning, which guarantees better adaptability regarding the odor recognition problem in artificial noses.

In order to verify the statistic relevance of the results, the hypothesis *t*-test was carried out on the performance of each one of the classifiers. The test is accomplished from the best average results obtained in experiments.

The classification error of ANNs was slightly greater when the classifier did not use wavelet filter in the data ($14.39 \pm 17.96\%$ and $11.34 \pm 8.50\%$ for MLP and TDNN, respectively, Tables 1 and 3) than when wavelet filter was used ($11.50 \pm 15.88\%$ and $0.75 \pm 3.6\%$, respectively, Tables 2 and 4).

In the multi-layer perceptron results, the mean difference between the measurements obtained using both methods was -2.88% (95% confidence interval from -0.1167 to 0.059%). Despite the better classification error of multi-layer perceptron using wavelet filter, statistically both methods present the same performance (P = 0.5073). However, the 95% agreement limit between the two methods was -4.3 to +4.9% (Fig. 3(A)).

In time delay neural network results, the mean difference between the measurements obtained using both methods was -10.59% (95% confidence interval from -0.1406 to -0.072%). The results were statistically significant ($P < 0.001$), confirming the hypothesis that the average performance of the time delay neural networks making use of wavelet analysis is superior to the same ANN without this filter. More importantly, the 95% agreement limit between the two methods was -7.6 to +2.88% (Fig. 3B).

The significant correlation between the two methods was also demonstrated through the linear regression analysis, considering the classifier without wavelet filter as the dependent variable and the classifier using the filter as the independent variable. For the multi-layer perceptron experiments, the

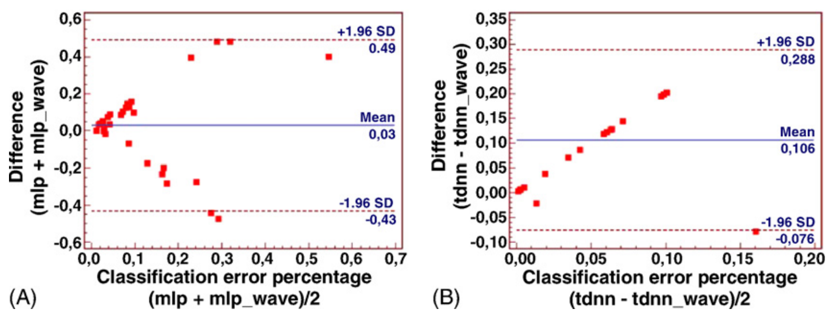

Fig. 3. (A) Result of the Bland–Altman plot of the data obtained by use of the multi-layer perceptron with and without wavelet filter and (B) result of the Bland–Altman plot of the data obtained by use of the time delay neural network with and without wavelet filter. The mean difference (bias) is represented by the solid horizontal line, and the agreement limit is represented by the dotted horizontal lines.

Table 5
Linear regression and Pearson correlation

Equation of the regression line	Coefficient of determination	Standard deviation of the residues	Pearson correlation	P
MLP	$y = 0.0321x + 0.1104$	0.0013	0.1615	0.03630.8489
TDNN	$y = -0.0057x + 0.0081$	0.0002	0.0373	−0.01320.9448

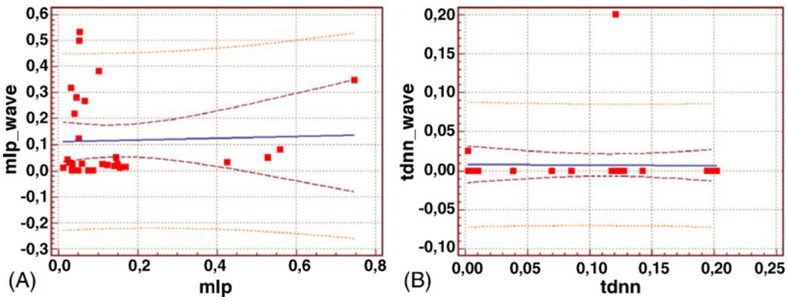

Fig. 4. (A) Multi-layer perceptron and (B) time delay neural network. Correlation between the results obtained by use of the classifier with and without wavelet filter.

equation of the regression line was calculated as $y = 0.0321x + 0.1104$, with a 95% confidence interval of the 'b' coefficient from −0.3283 to 0.3915, and a standard deviation of the residues of 0.1615 (Table 5). The Pearson correlation coefficient showed little relation between both methods ($r = 0.0363$; $P = 0.8489$) (Fig. 4A). In the time delay neural networks experiments, the equation of the regression line was calculated as $y = -0.0057x + 0.0081$, with a 95% confidence interval of the 'b' coefficient from −0.3717 to 0.3487, and a standard deviation of the residues of 0.0373 (Table 5). The Pearson correlation coefficient also showed little relation between both methods ($r = -0.0132$; $P = 0.9448$) (Fig. 4B).

6. Discussion and conclusions

In this work, we presented the results obtained with wavelet analysis in the processing of odor patterns in an artificial nose. This analysis was used as a preprocessing method of input patterns for multi-layer perceptron and the time delay neural networks classifiers.

The MLP presented a better performance with the use of the wavelet filter in the input patterns. The classification error decreased from 14.39 to 11.34%. For the TDNN, the classification error decreased from 11.50 to 0.75%.

The correlation between the two methods (MLP and TDNN, with and without wavelet filter), measured through the Pearson correlation coefficient, was neither expressive nor statistically significant ($r = 0.0363$ and $r = -0.0132$; $P = 0.8489$ and $P = 0.9448$ for multi-layer perceptron and time delay neural networks experiments, respectively). The linear regression analysis also showed very consistent results (Table 5). However, as reported by Bland and Altman in their study on the comparison of the two measurement methods, assessment [1] of the agreement between methods is more important than the correlation or linear regression.

Based on the analysis reported by Bland and Altman, in the multi-layer perceptron experiments, we showed that the mean difference between the values of the classification error using wavelet filter was not significant in practice, with only a −2.88% (95% confidence interval from −0.1167 to 0.059%) of difference and not statistically significant ($P = 0.5073$). In the time delay neural networks experiments, the mean difference was very significant, with a −10.59% (95% confidence interval from −0.1406 to −0.072%) of difference and the results are statistically significant ($P < 0.0001$).

The agreement limit of the two methods (MLP and TDNN, with and without wavelet filter) was −4.3 to +4.9% and −7.6 to +2.88%, respectively. These results show that the percentage of the classification error using the wavelet filter, 95% of the time, will have a difference lower than 4.9 and 2.88% of the value that would be obtained without wavelet filter, for multi-layer perceptron and time delay neural networks, respectively. This characterizes a fully acceptable degree of agreement, showing that the filter has homogeneous results. In addition, as shown in the Bland–Altman plot (Fig. 3), the dispersion of the individual differences around the mean difference (bias) is very homogeneous. In other words, the reproducibility of the wavelet filter results is good for both methods.

In future work, other techniques of signal compression and noise reduction will be combined with wavelet analysis for increasing the compression rate and improving the quality of the signal reconstruction. Other wavelets and their tradeoffs between performance and complexity need further investigation.

Acknowledgments

The authors would like to thank CNPq, CAPES and FINEP (Brazilian research agencies) for their financial support.

References

[1] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 1 (1986) 307–310.
[2] J.J. Benedetto, A. Teolis, A wavelet auditory model and data compression, Appl. Comput. Harmon. Anal. 1 (1) (1993) 3–28.
[3] E. Brigham, The Fast Fourier Transform and its Applications, Englewood Cliffs, Prentice Hall, 1988, 448 pp.
[4] M.A. Craven, J.W. Gardner, P.N. Bartlett, Electronic noses development and future prospects, Trend. Anal. Chem. 5 (9) (1996).
[5] I. Daubechies, Orthonormal bases of compactly supported wavelets, Commun. Pur. Appl. Math. 41 (1988) 909–996.
[6] I. Daubechies, The wavelet transform, time–frequency localization and signal analysis, IEEE Trans. Inform. Theory 36 (5) (1990) 961–1005.
[7] C. Di Natale, F.A.M. Davide, A. Damico, G. Sberveglieri, P. Nelli, G. Faglia, C. Perego, Complex chemical pattern recognition with sensor array: the discrimination of vintage years of wine, Sens. Actuators B 24–25 (1995) 801–804.
[8] D.L. Donoho, De-Noising by Soft-thresholding, IEEE Trans. Inform. Theory 41 (3) (1995) 613–627.
[9] R. Fletcher, Pratical Methods of Optimization, Wiley, 1987.
[11] J.W. Gardner, T.C. Pearce, S. Friel, P.N. Bartlett, N. Blair, A multisensor system for beer flavour monitoring using an array of conducting polymers and predictive classifiers, Sens. Actuators B 18–19 (1994) 240–243.
[12] A. Graps, An introduction to wavelets, IEEE Comput. Sci. Eng. 2 (2) (1995) 50–61.
[13] R.A. Johson, D.W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall, 1999.
[14] P.E. Keller, R.T. Kouzes, L.J. Kangas, Three neural network based sensor systems for environmental monitoring, in: Proceedings of the IEEE Electro 94 International Conference, Boston, MA, 1994, pp. 378–382.
[15] K.J. Lang, G.E. Hilton, The Development of the Time-Delay Neural Network Architeture for Speech Recognition, Technical Report CMU-CS-88-152, Carnegie-Mellon University, Pittsburg, PA, 1988.
[16] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693.
[18] M. Misiti, Y. Misiti, G. Oppenheim, J.M. Poggi, Wavelet Toolbox User's Guide for use with Matlab, The MathwWorks Inc., Natick, MA, 1996, 626 pp.
[19] D. Moshou, I. Hostens, G. Papaioannou, H. Ramon, Dynamic muscle fatigue detection using self-organizing maps, Appl. Soft Comput. 5 (4) (2005) 391–398.
[20] T. Mitchell, Machine Learning, McGraw Hill, New York, 1997
[21] K.C. Persaud, P.J. Travers, Arrays of broad specificity films for sensing volatile chemicals, in: E. Kress-Rogers (Ed.), Handbook of Biosensors and Electronic Noses: Medicine, Food and the Environment, 1997, pp. 563–592.
[22] L. Prechelt, Proben1—A Set of Neural Network Benchmark Problems and Benchmarking Rules, Technical Report 21/94, Fakultt fr Informatik, Universitt Karlsruhe, Germany, 1994.
[23] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by backpropagation errors, Nature (323) (1986) 533–536.
[24] M.S. Santos, Construction of an Artificial Nose using Neural Networks (in Portuguese), Ph.D. Thesis, Center of Informatics, Federal University of Pernambuco, Brazil, 2000.
[25] H. Sundgren, F. Winquist, I. Lukkari, I. Lundstrm, Artificial neural networks and gas sensor arrays: quantification of individual components in a gas mixture, Meas. Sci. Technol. 2 (1991) 464–469.
[26] A.H. Waibel, T. Hanazawa, G. Hinton, K. Shikano, K.J. Lang, Phoneme recognition using time-delay neural networks, IEEE Trans. Acoust. Speech Signal Process. 37 (1989) 328–339.
[27] M.V. Wickerhauser, Acoustic signal compression with wavelet packets, in: C.K. Chui (Ed.), Wavelets: A Tutorial in Theory and Applications, vol. 2, Academic Press, 1992, pp. 679–700.

[28] A. Yamazaki, T.B. Ludermir, Classification of vintages of wine by an artificial nose with neural networks, in: Proceedings of the Eighth International Conference on Neural Information Processing, vol. 1, Shanghai, China, 2001, pp. 184–187.

[29] A. Yamazaki, T.B. Ludermir, M.P.C. De Souto, Classification of vintages of wine by an artificial nose using time delay neural networks, IEE Electron. Lett. 37 (24) (2001) 1466–1467.

[30] C. Zanchettin,T.B. Ludermir, Wavelet filter for noise reduction and signal compression in an artificial nose, in: Hybrid intelligent System, Melbourne, Australia, 2003, pp. 907–916.

[31] C. Zanchettin, T.B. Ludermir, A. Yamazaki, Classification of gases from the petroliferous industry by an artificial nose with neural network, in: Proceedings of the Joint 13th International Conference on Artificial Neural Networks, Istanbul, Turkey, 2003, pp. 208–211.