Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Sound and Vibration 266 (2003) 667–675
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

# Considering the influence of artificial environmental noise to study cough time–frequency features

A. Van Hirtum, D. Berckmans*

*Department of Agro-Engineering and -Economics, Laboratory for Agricultural Buildings Research, Catholic University Leuven, Kasteelpark Arenberg 30, 3001 Leuven, Belgium*

Received 13 January 2003

**Abstract**

In general the study of the cough mechanism and sound in both animal and human is performed by eliciting coughing in a reproducible way by nebulization of an irritating substance. Due to ventilation the controlled evaporation-protocol causes artificial noises from a mechanical origin. The resulting environmental low-frequency noises complicate cough time–frequency features. In order to optimize the study of the cough-sound the research described in this paper attempts on the one hand to characterize and model the environmental noises and on the other hand to evaluate the influence of the noise on the time–frequency representation for the intended cough sounds by comparing different de-noising approaches. Free field acoustic sound is continuously registered during 30 min citric acid cough-challenges on individual Belgian Landrace piglets and during respiratory infection experiments, with a duration of about 10 days, where room-ventilation was present.


## 1. Introduction

The modelling and optimal representation of naturally occurring phenomena like the free field acoustical cough-sound under study is often attempted with the time–frequency harmonic analysis [1–3]. In practise however there is commonly found a discrepancy between the 'acoustic' model from the noisy signal and the clean signal leading to a robustness problem see e.g., Refs. [4,5]. Therefore this paper intends a quantitative comparison of different de-noising approaches for several signal-to-noise ratios (SNR) of realistic environmental noises by means of optimized error

*Corresponding author. Tel.: +32-16-32-17-26; fax: +32-16-32-14-80.
*E-mail addresses:* annemie.vanhirtum@agr.kuleuven.ac.be (A. Van Hirtum), daniel.berckmans@agr.kuleuven.ac.be (D. Berckmans).

0022-460X/03/$ - see front matter 
doi:10.1016/S0022-460X(03)00592-3

measures and the error signal between the original ($x$) and filtered ($\hat{x}_\delta$) signal. In general the de-noising approach is as follows:

(1) transform time-features $x$: $y = Wx$;
(2) select threshold $\delta$ for de-noising in the transform domain;
(3) de-noise by shrinking ($D$) of the coefficient $y$ in the transform domain: $\hat{y}_\delta = D_\delta y$;
(4) inverse transform to time domain: $\hat{x}_\delta = W^{-1}\hat{y}_\delta$,
where the symbols are defined in Appendix A.

## 2. Methods

### 2.1. High-pass filter

The signal $x(t)$, corrupted by acoustical environment noises, can be written as follows, with $s(t)$ the original signal and $n(t)$ Gaussian noise. The signal $s(t)$ is recovered by inverting the convolution integral.

$$x(t) = \int_a^b l(t-\tau)s(\tau)\,\mathrm{d}\tau + n(t). \tag{1}$$

A comparison with more advanced techniques is described further in this section, and $x(t)$ is assumed to be approximated as in Eq. (1) with the convolution kernel $l(t)$ being a high-pass filter (HPF).

### 2.2. Discrete wavelet decomposition

The basic theory of wavelet representation can be found in many papers (see e.g., Ref. [6]). In discrete (multiresolution) wavelet analysis the discrete signal $x(t) \in l^2$ is determined by

$$x(t) = \sum_{k=-\infty}^{\infty} a_1(k)\phi_{1,k}(t) \tag{2}$$

which is decomposed on different scales as in the following equation, where $\psi_{j,k}(t)$ are discrete analysis wavelets and $\phi_{K,k}(t)$ are discrete scaling functions which are all translations of one father function $\phi(t)$, $d_j(k)$ are the detailed signals (wavelets coefficients) at scale $2^j$, and $a_K(k)$ is the approximated signal (scaling coefficients) at scale $2^K$:

$$x(t) = \sum_{j=1}^{K}\sum_{k=-\infty}^{\infty} d_j(k)\psi_{j,k}(t) + \sum_{k=-\infty}^{\infty} a_K(k)\phi_{K,k}(t). \tag{3}$$

For the signal $x(t)$ the wavelet basis functions $\psi_{j,k}(t)$ are localized in time and translate dilations of one mother wavelet $\psi(t)$. The signal's discrete wavelet transformation (DWT), presented in Eq. (3), allows a sparse signal-representation within the detailed wavelet coefficients indicating signal-singularities at scale $2^j$. So a one-scale representation is decomposed into a sparse and multiscale representation and wavelet coefficients return both time and frequency (or inverse scale) information resulting in a multi-resolution analysis from a coarse to a higher resolution

approximation. At each scale localization in time and frequency space is restricted by the Heisenberg uncertainty principle.

The discrete wavelet transform can be implemented by the scaling and wavelet filters, the following equations, being quadrature mirror filters (QMF) [7]:

$$h(n) = \frac{1}{\sqrt{2}} \langle \phi(t), \phi(2t - n) \rangle, \tag{4}$$

$$g(n) = \frac{1}{\sqrt{2}} \langle \psi(t), \phi(2t - n) \rangle = (-1)^n h(1 - n). \tag{5}$$

The estimation of the detail signal at level $j$ will be done by convolving the approximate signal at level $(j - 1)$ with the coefficients $g(n)$. Convolving the approximate signal at level $(j - 1)$ with the coefficients $h(n)$ gives an estimate for the approximate signal at level $j$. The decomposition scheme involves retaining every other sample of the filter output or a dyadic decomposition structure which is constant for all signals.

### 2.3. De-noising the signal

Assuming that every wavelet coefficient over all scales contains noise, non-linear de-noising by soft-thresholding is performed, i.e., discarding the details exceeding a certain limit [8]. The soft-thresholded wavelet coefficients are defined in the following equations where $\delta$ is the applied soft threshold:

$$\eta(d_j(k)) = \text{sign}(d_j(k))(|d_j(k)| - \delta), \quad \text{if } |d_j(k)| > \delta, \tag{6}$$

$$\eta(d_j(k)) = 0, \quad \text{if } |d_j(k)| \leqslant \delta. \tag{7}$$

The wavelet coefficients whose absolute values are lower than the threshold are first set to zero, and then the remaining non-zero coefficients are shrunk towards zero.

The assumed model for a noisy signal is $x(t) = s(t) + e(t)$, where $s(t)$ is the noise-free signal and $e(t)$ the white or non-white noise. The performance of the de-noising methods is evaluated from the simulations with $l_2$-norm given by the following equation with $\hat{x}(t)$ the de-noised signal $x(t)$.

$$\|s - \hat{x}\|_2 = \left( \sum_t |s(t) - \hat{x}(t)|^2 \right)^{1/2}. \tag{8}$$

### 2.4. Threshold selection

The soft threshold $\delta$ is selected for each signal using four threshold estimating procedures: SURE threshold ($\delta_S$), universal threshold ($\delta_U$), minimax principle ($\delta_M$) and the generalized cross validation (GCV) ($\delta_G$). In the ideal case the optimal threshold should minimize $R(\delta)$ or the mean squared error (MSE) given by the following equation, with $\hat{x}_\delta$ the de-noised signal $x = s + e$ and

taking the wavelet transform $Wx = Ws + We$ or $y = z + v$:

$$R(\delta) = \frac{1}{N}\sum_{k=1}^{N}(\hat{x}_{\delta k} - s_k)^2. \tag{9}$$

The Stein unbiased risk estimate (SURE) [9] is an adaptive threshold selection rule defined below as

$$SURE(\delta_S) = \frac{1}{N}\sum_{k=1}^{N}(\hat{y}_{\delta k} - y_k)^2 + 2\sigma^2\frac{N_1}{N} - \sigma^2, \tag{10}$$

where $y$ denotes the DWT of $x$, $\hat{y}_\delta$ the wavelet transform after thresholding, $N_1$ the number of samples with magnitude above the threshold and $N$ is the number of samples in the signal vector.

In case of non-white correlated noise, $\sigma$ is estimated as the standard variation of the noise estimated from the scale with the highest resolution in the decomposition of each signal. With this approach obtaining risks and minimizing them with respect to $\delta_U$ values give a threshold selection. The method is adaptive through searching a threshold level for each wavelet decomposition level. The universal threshold approach (uni) calculates a fixed threshold with respect to the length of the signal, and the estimated threshold is given by $\delta_U = \sqrt{2\log_e(N)}\sigma$ [10]. The minimax principle applies a fixed threshold $\delta_M = (0.3936 + 0.1829\log(N))\sigma$ to produce a so-called minimax performance for mean square error against an ideal case [10]. The GCV is another approximation of the MSE and defined as in Eq. (11) with $N_0$ the number of wavelet coefficients that is replaced by zero [11]. Note that no estimation of $\sigma$ is necessary:

$$GCV(\delta_G) = \frac{(1/N)\sum_{k=1}^{N}(\hat{y}_{\delta k} - y_k)}{(N_0/N)^2}. \tag{11}$$

## 3. Signals and acoustical environment

Coughing in individual Belgian Landrace piglets (10 kg) is elicited either by nebulization of an irritating substance ([1], and the references therein), or by induction of a controlled and reproducible respiratory infection protocol [12]. Example acoustical free field cough-signals, $s(n)$, are registered (22 kHz) immediately after nebulization of the irritating substance in the absence of disturbing noises. In the following, these example sounds are referred to as a chemical cough, $s_{chem}$. Fig. 1 shows respectively the time and fast Fourier transforms giving the power spectral density (PSD) for three examples of $s_{chem}$. Due to the presence/need for ventilation a controlled evaporation-protocol involves artificial noises from a mechanical origin. Three occurring distinct types of experimental ventilation noises, $e(n)$ with $n = 1, \ldots, N$, are individually recorded. Firstly the noise accompanying the nebulization, $e_{nebu}$, of the irritating substance into the closed laboratory installation of dimensions $2000 \times 800 \times 950$ ($l \times w \times h$). Secondly, the noise due to preserve air-circulation, $e_{circ}$, in the laboratory installation and thirdly the ventilation present in the room, $e_{room}$, containing the laboratory installation is recorded. Fig. 2 shows, respectively, the time and the fast Fourier transforms giving the power spectral density (PSD) for each of the

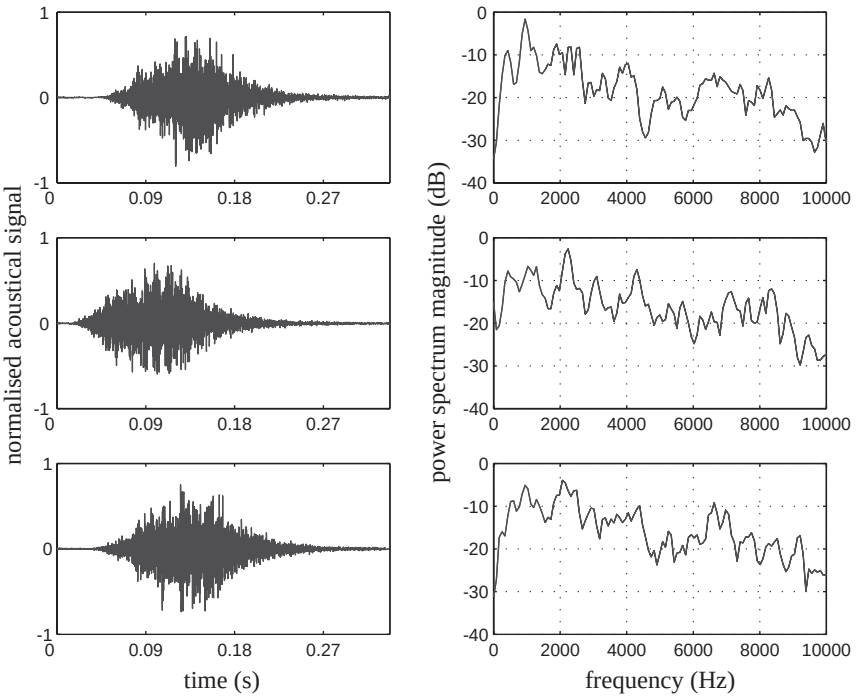


Fig. 1. Time-signal-amplitude versus number of samples and corresponding power spectral magnitude (dB) frequency features (Hz) for three examples of $s_{chem}$.

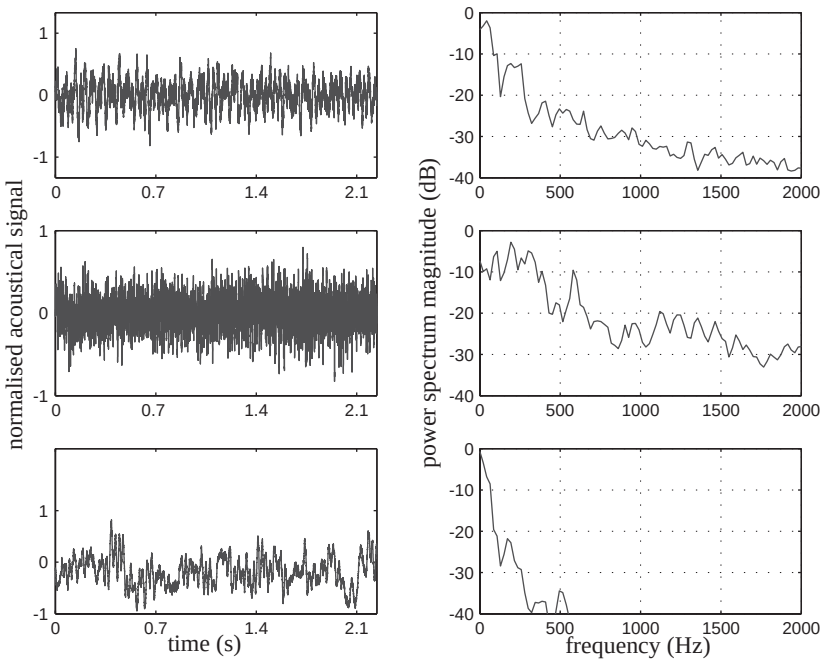


Fig. 2. Time-signal-amplitude versus number of samples and corresponding power spectral magnitude (dB) frequency features (Hz) for $e_{nebu}$, $e_{circ}$ and $e_{room}$.

distinct noises. As expected for noises of mechanical origin, the distinct ventilation noises are clearly low-frequency coloured noises.

In order to study the influence of realistic artificial environmental noises on the example cough-sounds the cough-signals, $s(n)$, are additively corrupted with environmental noise, $e(n)$. Several SNR between 2 and 20 dB are attempted. As already mentioned in Section 2.4 the noisy signal, $x(n)$, is assumed to be modelled as in the following equation where $\|\cdot\|$ denotes the $l^2$ norm.

$$x(n) = s(n) + re(n) \Rightarrow \mathrm{SNR} = 10\log_{10}\frac{\|s\|^2}{\|x-s\|^2}. \tag{12}$$

## 4. Results and discussion

Different de-noising approaches are compared on 10 cough-sounds $x_{chem,i} = s_{chem,i} + r_i e$ for SNR respectively 2, 5, 10, 15 and 20 dB in accordance with Eq. (12). De-noising is assessed firstly with DWT soft thresholding described in Sections 2.2–2.4 and secondly by applying a well-known HPF to the data $x_{chem,i}$ as described in Section 2.1. The results of both approaches are compared. Evaluation of the MSE for the applied HPF as a function of the high-pass cut-off frequency $\delta_{HPF}$ is illustrated in Fig. 3. $R(\delta_{HPF})$ is computed for the noise-signals $e_{room}$, $e_{nebu}$ and $e_{circ}$ for SNR 2 and 20 dB. Note the typical form of $R(\delta_{HPF})$. For small values of the threshold $\delta_{HPF}$, the input noise $e$ still has an important contribution to the MSE of the result. If the threshold is large, the shrinking operation induced by the de-noising deforms the original signal and causes a bias. Both the left- and right-hand side of this figure show the best de-noising performance when $\delta$ or the cut-off frequency equals $\pm 100$ Hz. Since only for very corrupted signals (low SNR) the influence of $e_{circ}$ and $e_{nebu}$ differs from $e_{room}$ and only the last one is present during infection protocols, in the rest of the paper only $e_{room}$ is considered.

For wavelet-based de-noising the Coiflet wavelet function of order 5 [7] was chosen out of the compactly supported orthogonal wavelet families for discrete wavelet analysis following from the good de-noising performance mentioned in Ref. [4] for electrocardiogram-bio-signals. Extensive discussion of the optimal wavelet choice for use in acoustic analysis is beyond the scope

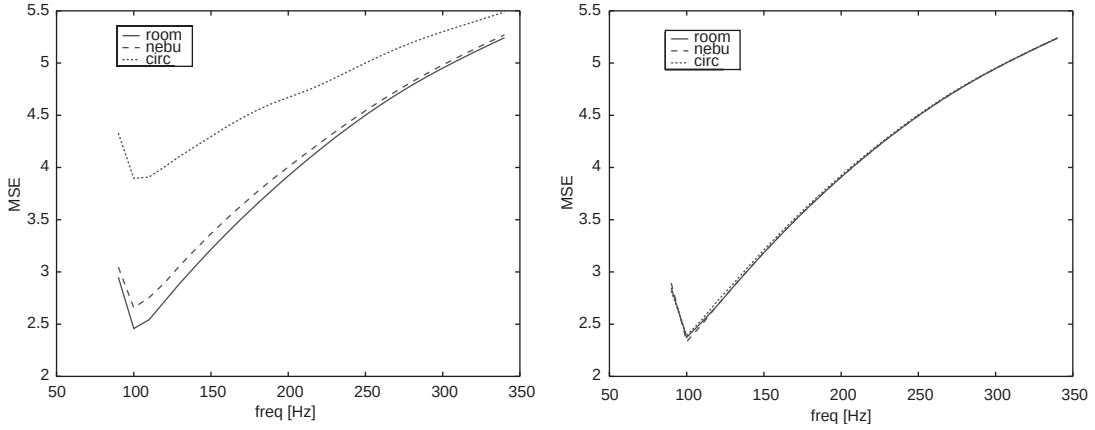


Fig. 3. $R(\delta_{HPF})$ with threshold $\delta_{HPF}$ corresponding to the high-pass frequency for SNR = 2 (left) and SNR = 20 (right).

of this paper. The choice of a particular wavelet is mainly determined by its number of vanishing moments, its smoothness properties and its compact support, as well as its localization in the frequency domain. In general, matching characteristics of the used wavelet analysis function with expected signal-characteristics yields improved performance. The principle of threshold-selection and de-noising presented in Section 2.4 is adequate for every wavelet choice. Table 1 shows the $R$ obtained for distinct SNRs for discrete wavelet decomposition with depth $L = 3$ and HPF. Increasing the wavelet decomposition depth to $L = 5$ or 7 did not yield a better de-noising performance.

Since noise is coloured, a level-dependent scaling of the thresholds was used for adjusting to the non-white noise structure. In general, the GCV yields the best performance. However, both the small performance differences and the best de-noising performance of HPF with a cut-off frequency of 100 Hz indicate the absence of much signal-energy for the cough-sound in the low-frequency domain. As a consequence, although very disturbing for human hearing, the presence of ventilation or more general low-frequency ventilation disturbing sounds, does not influence the cough-spectra to a high degree. This finding is very important to the analysis and time–frequency characterization of the cough-sounds due to infection where $e_{room}$ is always present and no noise-free cough-signal is available. Moreover Table 1 shows the good performance of high-pass filtering for acoustical environments with low SNR. Therefore, for cough-processing involving the total frequency-range, e.g., real-time cough-recognition, it is appropriate to discard the low-frequency range with linear HPF or to apply classical processing techniques as adaptive filtering in case of well-situated specific sources. For higher signal-to-noise ratios the DWT approach using generalized cross validation for threshold selection yields the best de-noising performance. Therefore in less noisy acoustic environments aiming specific feature-extraction mostly present in the non-white noise-corrupted part of the spectrum, e.g., pitch present in parts of the cough-sound [2], de-noising with DWT multi-resolution decomposition in time and frequency can be motivated because of its scale-adaptivity, level-dependent thresholding and non-linearity.

De-noising established by maximum SNR solution (or minimum mean square error solution) as incorporated in the wavelet and GCV methods are well accepted in image-processing [6], but questionable for speech-processing where sound-quality is commonly assessed with auditive quality being the most important de-noising criterion for speech-synthesis. However to study the cough-sound the ‘source of sound’ or sound production is of much more interest and importance than ‘sound-perception’ and therefore a more objective quality measure is preferred above the sensitivity of human hearing.

Table 1
$R$ for distinct $e_{room}$ SNR for a Coiflet wavelet DWT (order 5, depth $L = 3$) and HPF

| SNR (dB) | SURE | uni | minimax | GCV | HPF |
|---|---|---|---|---|---|
| 2 | 0.55 | 0.76 | 0.69 | 0.53 | 0.48 |
| 5 | 0.40 | 0.65 | 0.57 | 0.37 | 0.48 |
| 10 | 0.25 | 0.55 | 0.46 | 0.21 | 0.48 |
| 15 | 0.18 | 0.51 | 0.41 | 0.12 | 0.48 |
| 20 | 0.14 | 0.50 | 0.39 | 0.07 | 0.48 |

## 5. Conclusion

In this paper, noise-free cough signals are artificially corrupted with realistic low-frequency ventilation noises. The performances of several variations of de-noising including thresholding rules were compared for soft thresholding. The performed DWT multi-resolution decomposition in time and frequency is motivated because it is well suited for removing specific unwanted signal components that may vary spectrally. For acoustical environments with low SNR due to noises of mechanical origin a simple high-pass filtering yields a good performance. It is concluded that the presence of low-frequency disturbing sounds does not influence the spectral characteristics of the cough-sound to a high degree. For higher signal-to-noise ratios, in general, the DWT approach using generalized cross validation for threshold selection yields the best de-noising performance.

## Acknowledgements

This work was supported by the Belgian Ministry of Agriculture (Projects No. A-5778 and No. S-5899).

## Appendix A. Nomenclature

| | |
|---|---|
| $x, x_i$ | noisy input (general index $i$) |
| $s, s_i$ | original, noise-free signal |
| $e, e_i$ | noise |
| $n, n_i$ | Gaussian noise |
| $y, y_i$ | noisy wavelet coefficients |
| $W$ | forward wavelet transform |
| $W^{-1}$ | inverse wavelet transform |
| $\delta$ | smoothing parameter, threshold |
| $D$ | shrinking transform |
| $\hat{y}_\delta, \hat{y}_{\delta i}$ | noisy wavelet coefficients after thresholding |
| $\hat{x}_\delta, \hat{x}_{\delta i}$ | output of threshold algorithm |
| $N$ | number of data points in a discrete final signal |
| $R$ | mean squared error (MSE) |
| $l$ | convolution kernel |
| $\delta_{HPF}$ | high-pass frequency |
| $\delta_S$ | sure threshold |
| $\delta_U$ | universal threshold |
| $\delta_M$ | minimax threshold |
| $\delta_G$ | generalized cross validation (GCV) threshold |
| $\sigma^2$ | variance |
| $N_0$ | number of noisy wavelet coefficients below threshold |
| $N_1$ | number of noisy wavelet coefficients above threshold |
| $\phi(t)$ | primal scaling (father) function |

| | |
|---|---|
| $\psi(t)$ | primal wavelet (mother) function |
| $g$ | high-pass filter in filter bank, coefficients in wavelet equation (primal) |
| $h$ | low-pass filter in filter bank, coefficients in dilation equation (primal) |
| $l^2$ | Hilbert space of square summable sequences |

## References


[1] A. Van Hirtum, D. Berckmans, The fundamental frequency of cough by autocorrelation analysis, in: *Proceedings of the Seventh European Conference on Speech Communication and Technology*, Aalborg, Denmark, 2000, pp. 2435–2438.

[2] A. Van Hirtum, D. Berckmans, Fuzzy approach for improved recognition of pig-coughing from continuous registration, in: *Proceedings of the 25th International ISMA Conference on Modal Analysis Noise and Vibration Engineering*, Leuven, Belgium, 2001, pp. 1535–1541.

[3] J. Korpas, J. Sadlonona, M. Vrabec, Analysis of the cough sound: an overview, *Pulmonary Pharmacology* 9 (1996) 261–268.

[4] P.E. Tikkanen, Nonlinear wavelet and wavelet packet de-noising of electrocardiogram signal, *Biological Cybernetics* 80 (1999) 259–267.

[5] L.G. Weiss, T.L. Dixon, Wavelet-based de-noising of underwater acoustic signals, *Journal of the Acoustical Society of America* 101 (1) (1997) 377–383.

[6] S.G. Mallat, *A Wavelet Tour to Signal Processing*, Academic Press, San Diego, CA, 1998.

[7] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[8] D. Donoho, De-noising by soft-thresholding, *IEEE Transactions on Information Theory* 41 (1995) 612–627.

[9] D. Donoho, I. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, *Journal of the American Statistical Association* 90 (1995) 1200–1224.

[10] D. Donoho, I. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *Biometrika* 81 (1994) 425–455.

[11] M. Jansen, M. Malfait, A. Butheel, Generalized cross validation for wavelet thresholding, *Signal Processing* 45 (1) (1997) 33–44.

[12] M. Kobisch, T.N.F. Friis, Swine mycoplasmoses, *Revue Scientifique et Technique de l'Office International des Epizooties* 15 (4) (1996) 1569–1605.