Journal of Sound and Vibration (2001) **243**(1), 3–22
doi:10.1006/jsvi.2000.3398, available online at http://www.idealibrary.com on IDEAL®

THE AUTO-SYNCHRONIZED WAVELET TRANSFORM ANALYSIS FOR AUTOMATIC ACOUSTIC QUALITY CONTROL

S. GÜTTLER† AND H. KANTZ

Max-Planck-Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 *Dresden, Germany. E-mail: guettler@mpipks-dresden.mpg.de*

(*Received* 7 *December* 1999, *and in final form* 18 *May* 2000)

A new feature vector for automatic acoustic quality control is presented and applied to the classification of electric sliding sunroofs at the quality control point by analyzing the sliding noise. The specific features of the sound signals (signatures) which distinguish products of different quality are not explicitly sought, because these signatures are highly specific for each application. Instead, the property is used that the relevant information about the sound signals can be resolved by the ears of experts. The time-frequency resolution of the ear is approximated by a wavelet transform of the signals. As a new approach to the important problem of noise reduction the concept of auto-synchronized wavelet transforms is introduced which allows wavelet transforms (and more general time–frequency representations) to be averaged in the time domain without losing the time-resolved information in the signals. By this averaging process, statistical fluctuations (noise, parameter drifts) can be reduced significantly to reveal the characteristic features of the signals. The classification can then be performed by a next neighbour search on a training set. The concept of auto-synchronized wavelet transforms is developed in a mathematically formal way and the properties of the feature vectors obtained are studied by using artificial noisy signals before applying this method to the experimental data.

1. INTRODUCTION

Since the beginning of the 1950s, the analysis of sound signals for the quality control of products and manufacturing processes has undergone significant development and is now widely applied [1]. The sound signals emitted from machines, manufacturing processes (e.g., cutting and deformation mechanisms, laser welding) or products caused to vibrate (e.g., porcelain, roof tiles) often contain relevant information about the quality of the product. This information is usually evaluated by human experts. The automation of acoustic quality control is desirable not only because of cost reasons but, more importantly, in order to make quality control more objective, i.e., to reduce the dependence of quality standards on individual decisions which may vary largely and often cannot be reproduced. Despite much interest in automatic acoustic quality control there has been only little progress in this direction due to two main difficulties: the first problem is to (automatically) extract the features of sound signals which are relevant for the quality condition of a product for which no general solution is known, the second difficulty is the high variability of the sound signals of good quality products, i.e., the acoustic data of different units being classified as good by

†Present address: Fraunhofer Institute IPA Nobelstr. 12, 70569 Stuttgart, Germany.

0022-460X/01/210003 + 20 $35.00/0

humans may differ considerably due to production tolerances which makes the automatic classification difficult.

Until now several solutions have been proposed for the automatic classification of sound signals (see e.g., references [1–4]) where time–frequency representations (see, e.g., reference [5]) are usually applied. Starting from elementary methods such as the observation of selected bands in the windowed Fourier spectra of the signals [1, 2] more sophisticated methods have been proposed in references [3, 4]. In reference [3], third and fourth order generalizations of the Wigner–Ville distribution are used to detect impulse signals in the acoustic or vibration data of rotating machinery. In reference [4], a method to automate quality control for small electric motors with an attached gear is introduced. The emitted sound signals are transformed into time–frequency representations of Cohen's class [5], where a Gaussian-shaped kernel function can be automatically adapted to the specific type of signals on a training set.

In this paper, a new method for the automatic classification of sound signals which contain information in time and in frequency space is presented where the focus is on signals whose information can (at least in principle) be resolved by the human ear. The starting point is to describe approximately the time–frequency resolution of the ear by a wavelet transform. An important problem which arises with the application of time–frequency representations is the reduction of noise. This is usually performed by pre-processing of the time series. A widely applied method is time domain averaging [6], where a periodic signal is averaged over successive cycles in order to attenuate non-periodic components. For this purpose, a pulse signal synchronized to the signal (in most cases) or the background noise is necessary. If no synchronous signal is available, Lee and White [7] proposed an adaptive filtering process where a delayed feed-back signal is used to separate periodic from non-periodic components of the signal. However, sound signals in general are non-periodic. Here the more general situation is addressed in which the sound signal is stationary above some time scale while the relevant information is contained (everywhere in time space) below that time scale (which is not the case for impulsive signals). This kind of sound signal is emitted by many types of machinery and is usually evaluated by human experts.

In this paper, the concept of auto-synchronized wavelet transforms is introduced which allows wavelet transforms (and more general time–frequency representations) to be averaged in the time domain without losing the time-resolved information of the signals. By this averaging process statistical fluctuations (noise, parameter drifts) can be significantly reduced to reveal the characteristic features of the signals. This method is applied to the classification of electric sliding sunroofs at the quality control point which is performed by a next neighbour search on a training set. For this purpose, the structure-borne sound signals of the sliding noise during closing of the sunroofs were recorded and pre-classified by human experts.

The paper is organized as follows. In the next section, an analysis of the structure-borne sound signals of the noise of sliding sunroofs is given. This introduces the concept of feature selection from auto-synchronized wavelet transforms which is developed in section 3 in a mathematically formal way. The properties of the feature vectors obtained from auto-synchronized wavelet transforms are studied by using artificial noisy signals which are simple but non-trivial for the purpose of numerically testing the averaging method. In section 4, this concept is applied to the classification of the sliding sunroof data where the performance is compared to the standard method of calculating windowed Fourier spectra.

2. PHYSICAL PROPERTIES OF THE SOUND SIGNALS

In order to explore the concept of feature selection from auto-synchronized wavelet transforms, the properties of the recorded sound signals will first be analyzed.

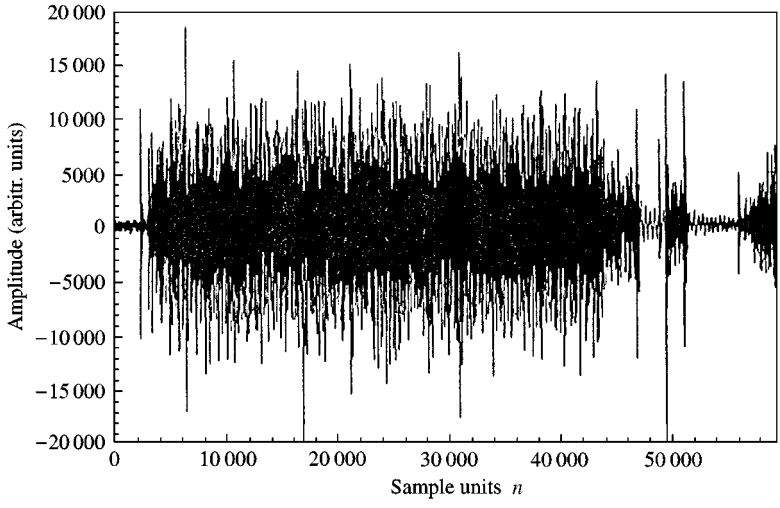

Figure 1. The time series of the structure-borne sound signal of the sliding noise during closing of an electric sunroof. The sampling rate is 10 kHz.

A set of 108 time series of the structure-borne sound signals of the sliding noise were recorded from 13 sunroofs pre-classified as good and from 14 units considered as of poor quality by human testers. From each sunroof two recordings were made in each two different sensor positions. The products denoted by "of poor quality" are usually not defective, but are just expected to have a reduced lifetime due to too high a level of abrasion. Thus, a simple classification into two quality classes A and B is the aim of this analysis. An example of the time series is shown in Figure 1. The sampling rate of the analogue-digital converter is 10 kHz and since only the (intrinsic) time scales of the signals will be of interest (but not their total length) as units in time space the "number of sample units" n are chosen here for simplicity.

Information was provided from the manufacturer that the information content about the quality condition of a sunroof is contained in the middle, approximately stationary part of the sound signal. This is roughly the section between the sample units $n = 10{,}000$ and $n = 40{,}000$ in Figure 1. Therefore, only this middle section of the time series will be examined further. More detailed information about the sound signals or the mechanism by which they were generated is not available.

First, the signals are tested for a possibly deterministic behaviour. This is done by comparing the time series (i.e., the indicated centre section) with surrogate data. The surrogates are generated by the method of phase randomization and subsequent "polishing" of the time series obtained [8]. The "polishing" of the surrogates is done by alternating scaling of the probability distribution and Wiener-filtering of the Fourier transform of the surrogate time series in order to adapt it to the original data. If this scheme converges, (as it usually does) this yields a surrogate time series having the same empirical spectrum and distribution as the original one. The multiple-step forecast error of a next neighbour search forecast algorithm is then used in a reconstructed phase space to yield a discriminating statistics [8]. It turns out that the null-hypothesis that the time series are generated by a linearily stochastic process cannot be rejected. From this it follows that there is no deterministic part of the sound signals which is recognizable. However, later results will show that the time series are not *linearly* stochastic, i.e., they do not *only* contain spectral information but also phase information in frequency space or, equivalently,

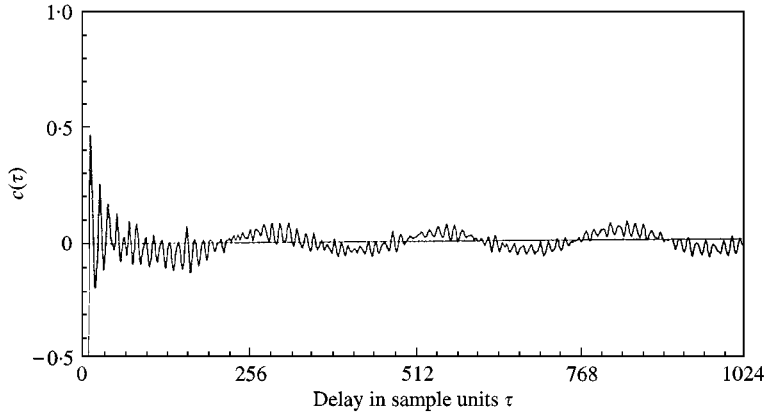

Figure 2. The autocorrelation function $c(\tau)$ of the middle section of the time series shown in Figure 1 (i.e., $n = 10{\cdot}000, \ldots, 40{\cdot}000$).

information in time space. This difference could not be resolved by the discriminating statistics.

To test (at least qualitatively) the stationarity of the indicated middle section of the time series order-10 autoregressive (AR(10)-) models are fitted to subsections of fixed length of the time series and then the parameters estimated are compared. As a fit method the minimization of the squared one-step forecast errors is used [8]. It turns out that (at least) the linear properties of the centre part of the time series are to a good approximation stationary over length scales $N \gtrsim 500$, i.e., on time scales $T \gtrsim \frac{1}{20}$ s.

The information relevant for the classification of the sunroofs can be resolved by the human ear, i.e., by experts. Because the lowest audible frequency is $f \approx 20$ Hz, the sound information is contained on time scales $T \lesssim \frac{1}{20}$ s. Oscillations on larger time scales are perceived as sound fluctuations [9] which do not occur because of the stationarity of the linear properties of the sound signals on time scales $T \gtrsim \frac{1}{20}$ s. This property can also be followed from the autocorrelation function shown in Figure 2. It can be concluded that the information for the classification of the sound signals is contained (everywhere in time space) on length scales $N \lesssim 500$.

The properties listed above may be typical for many kinds of technical signals. Because the method of feature selection from auto-synchronized wavelet transforms does not require detailed information about the sound signals it may be hoped that this concept is also useful for applications other than that studied in this paper.

3. THE AUTO-SYNCHRONIZED WAVELET TRANSFORM

Before going into details, the general idea of feature selection from auto-synchronized wavelet transforms should first be described.

3.1. THE PRINCIPLE IDEAS

Because the relevant information of the sound signals can be resolved by the human ear efforts are made to imitate the time–frequency resolution of the ear. For frequencies $f \gtrsim 500$ Hz this can, to a good approximation be done by a wavelet transform of the sound

signals [10, 11]. This is because the perception of tone pitches takes place according to octaves and this behaviour is simulated by a wavelet transform which divides the frequency space into octaves in a natural manner. This property is also called a *hyperbolic* time–frequency resolution [5]. More detailed psycho-acoustic properties of the sound signals will not be used here, although in general this may be an interesting starting point for feature extraction of sound signals.

The next step is to reduce statistical fluctuations (noise, parameter drifts) of the wavelet transforms of the time series in order to enhance the desired properties of the sound signals. It is assumed that a sufficient reduction of statistical fluctuations will reveal the characteristic features of the signals needed for the classification. If relevant information is contained in the high-frequency bands (as in fact is the case for the sunroof data) this cannot be achieved by high-pass filtering. Instead, the concept in this paper is to apply an appropriate averaging process. It is not possible simply to average over time–frequency representations (as, e.g., wavelet transforms) which are calculated from subsections of a stationary time series, because the information contained in the subspace of the time–frequency plane over which the averaging is performed would be erased. As a new approach to this problem, the class of auto-synchronized wavelet transforms is introduced. The auto-synchronized wavelet transform is defined as the *set* of all cyclically in time space permuted wavelet transforms of a section of a time series. These quantities can be averaged in the time domain without losing the time-resolved information of the signal. It will be shown that the signal-to-noise ratio of feature vectors constructed from auto-synchronized wavelet transforms can be significantly improved through this averaging process.

In the final step, the classification of the sound signals has to be performed according to those features which are relevant for the quality condition of the products. No information about this is known. It is not desirable to search explicitly for the features of the sound signals which distinguish products of different quality, because these signatures are highly specific for each application. Instead, the idea is to construct feature vectors in a general manner from auto-synchronized wavelet transforms and then to perform the classification by a next-neighbour search on a training set. This will be possible due to a sufficient low-noise level of the feature vectors. The training vectors have to be calculated from (reliably) pre-classified items and the training set should contain examples of products being classified as of good quality as well as of poor quality, representing all typical failures. During the testing stage a test vector is compared independently with each of a neighbourhood of training vectors corresponding to items of good and of poor quality. For a sufficiently large training set the difference of the weighted distances to both neighbourhoods is expected to represent the quality condition of the unit under test.

3.2. THE WAVELET TRANSFORM

The relevant information of the sound signals is contained on length scales $N \lesssim 500$, as mentioned above. Therefore, discrete wavelet transforms of sections of the time series of length $N = 512$ are calculated. The wavelet filters are not especially designed to imitate the time–frequency resolution of the human ear; only the hyperbolic time–frequency resolution of a wavelet transform should be used here. Instead, the wavelet filters are designed to obtain high resolution of a signal in the time–frequency plane and to minimize the effect of the periodic boundary conditions because of a finite signal length and the aliasing effect which occurs due to the scaling of discrete time series. These requirements are fulfilled if the wavelet filters are well localized in the time–frequency plane i.e., if they have a small time–frequency uncertainty. Additionally, for practical reasons wavelet filters with compact

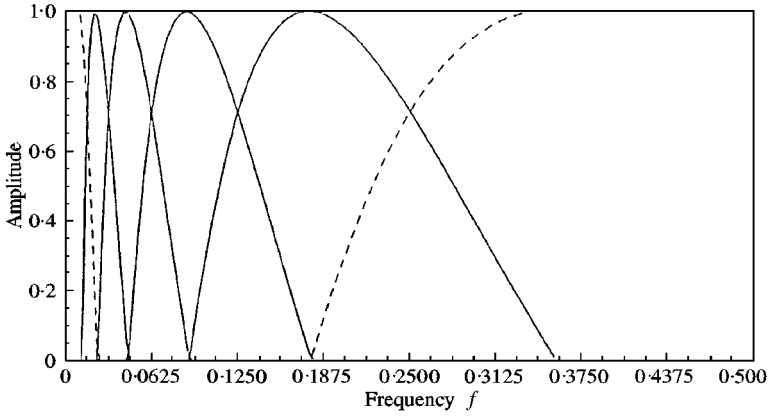

Figure 3. The wavelet filters (band-pass filters) $\{\hat{h}_{k,2}, \ldots, \hat{h}_{k,5}; k = 0, \ldots, 256\}$ in discrete frequency space (solid lines). Additionally, the high-pass filter $\{\widehat{hp}_{k,1}\}$ and the low-pass filter $\{\widehat{lp}_{k,5}\}$ are drawn with dashed lines.

support in frequency space are used. A discussion of these details can be found in reference [12].

One choice (certainly not unique) which fulfils these properties are the wavelet filters $\hat{h}_{k,j}$ defined in discrete frequency space as

$$\hat{h}_{k,j} = \widehat{lp}_{k,j-1}\,\widehat{hp}_{k,j}, \quad k = -N/2 + 1, \ldots, N/2, \quad j = 2, \ldots, \log_2(N) - 1, \tag{1}$$

where $N = 512$ is the period of the discrete wavelet transform, and the high- and low-pass filters $\widehat{hp}_{k,j}$ and $\widehat{lp}_{k,j}$ are given by

$$\widehat{hp}_{k,j} = \begin{cases} 0 & \text{if } 0 \leqslant |\frac{k}{N}| < f_j/\sqrt{2}, \\ \sqrt{\frac{1}{2}(1 + \sin[\pi \log_2(f_j^{-1}|\frac{k}{N}|)])} & \text{if } f_j/\sqrt{2} \leqslant |\frac{k}{N}| \leqslant \sqrt{2} f_j, \\ 1 & \text{if } \sqrt{2} f_j < |\frac{k}{N}| \leqslant \frac{1}{2}, \end{cases}$$

$$\widehat{lp}_{k,j} = \begin{cases} 1 & \text{if } 0 \leqslant |\frac{k}{N}| < f_j/\sqrt{2}, \\ \sqrt{\frac{1}{2}(1 - \sin[\pi \log_2(f_j^{-1}|\frac{k}{N}|)])} & \text{if } f_j/\sqrt{2} \leqslant |\frac{k}{N}| \leqslant \sqrt{2} f_j, \\ 0 & \text{if } \sqrt{2} f_j < |\frac{k}{N}| \leqslant \frac{1}{2}, \end{cases} \tag{2}$$

In equations (2) $f_j = 2^{-j} f_c = 2^{-(j+1)}$ denotes the split frequency of the filter $\widehat{hp}_{k,j}$ and $\widehat{lp}_{k,j}$ where $j = 1, \ldots, \log_2 N - 1$ and $f_c = \frac{1}{2}$ is the Nyquist frequency. Figure 3 shows the wavelet filters $\{\hat{h}_{k,2}, \ldots, \hat{h}_{k,5}; k = 0, \ldots, 256\}$, and the high-pass filter $\{\widehat{hp}_{k,1}\}$ and the low-pass filter $\{\widehat{lp}_{k,5}\}$. The discrete wavelet filters are related to each other by scaling and approximately by inverse scaling, i.e., by trigonometrical interpolation with the factor 2.

The wavelet transform of a time series $\{x_n; n = 1, \ldots, N\}$ (where N must be a power of 2) is now defined as the matrix

$$(W T_x^h(n, j)), \quad n = 1, \ldots, N, \quad j = 1, \ldots, J, \tag{3}$$

whose columns are the high-pass, band-pass, and low-pass filtered components of the signal. $J = 2, \ldots, \log_2 N$ denotes the number of filter components.[‡] It should be noted that this definition of a discrete wavelet transform is different from the concept of wavelet transforms related to multi-resolution analysis [11]. The idea here is to construct a discrete approximation of a continuous wavelet transform which yields a high resolution of a signal in the time–frequency plane.

This wavelet transform is by construction an *isometric* mapping between the corresponding vector spaces with respect to the standard scalar products, i.e.,

$$\langle x, y\rangle \equiv \sum_{n=1}^{N} x_n \bar{y}_n = \sum_{j=1}^{J} \sum_{n=1}^{N} WT_x^h(n,j)\,\overline{WT}_y^h(n,j) \equiv \langle WT_x^h, WT_y^h\rangle, \tag{4}$$

where $\bar{\cdot}$ denotes the complex conjugation. A further important property of this transform is that it preserves translations in time space with periodic boundary conditions, i.e., if the time series $\{y_n\}$ and $\{x_n\}$ are related to each other by a translation in time space this property also holds for their wavelet transforms:

$$y_n = x_{n-m} \Leftrightarrow WT_y^h(n,j) = WT_x^h(n-m,j). \tag{5}$$

For this property to hold the redundancy of the wavelet transform is necessary: i.e., the mapping of a vector to a matrix.

Now, a method of averaging over a set of these wavelet transforms which are calculated from subsections of a stationary time series can be introduced.

3.3. DEFINITION OF THE AUTO-SYNCHRONIZED WAVELET TRANSFORMS

Assume a set of wavelet transforms $WT_x^h(n,j)$, equation (3), calculated from subsections of a stationary time series. According to reference [13] a time series $\{x_n, n \in \mathbb{N}\}$ is called stationary (stationary on time scales $\geqslant N$, respectively) if all *joint* probabilities $p(x_1; x_2; \ldots; x_n)$ are invariant under translations in time space (on time scales $\geqslant N$, resp.): i.e.,

$$p(x_1; x_2; \ldots; x_n) = p(x_{1+k}; x_{2+k}; \ldots; x_{n+k}) \quad \text{for } k \in \mathbb{N} \quad (\text{for } k = lN;\ l \in \mathbb{N}, \text{ resp.}).$$

Note that this definition also applies to deterministic periodic time series with period N.

The (random) subsections of a stationary time series differ mainly by their initial conditions which can be described by the realization of a random variable. Since this property also holds for the columns of $WT_x^h(n,j)$ the idea is to permute these matrices cyclically in time space in order to adapt (or synchronize) their initial conditions. It will turn out that this synchronization process allows time–frequency representations to be averaged without losing the time-resolved information.

First, a formal definition of auto-synchronized wavelet transforms is given. The practical implementation will be explained later. A quantity which is independent of the initial conditions is obviously the *set* of all *cyclic permutations* in the time variable n of $WT_x^h(n,j)$. This set is defined as

$$R_x^h(j) := \{P_{cyc}(WT_x^h(n,j)) \,|\, n = 1, \ldots, N\}, \quad j = 1, \ldots, J, \tag{6}$$

[‡]Strictly speaking only the band-pass filters are wavelet filters, but this does not allow a complete decomposition of discrete signals.

where $P_{cyc}(WT_x^h(n,j))$ denotes the cyclic permutations of $WT_x^h(n,j)$ in the variable n. The definition of $R_x^h(j)$ requires $WT_x^h(n,j)$ to preserve translations in time space (equation (5)). Further, this definition does not depend on $J = 2, \ldots, \log_2 N - 1$.

On the *set* $\{R_x^h\}$ of all sets $R_x^h(j)$ a *metric* can be defined by taking the *minimum* of the Euclidean distances

$$\| WT_x^h - WT_y^h \|_2 \equiv \sqrt{\sum_j \sum_n (WT_x^h(n,j) - WT_y^h(n,j))^2} \tag{7}$$

of all cyclic permutations of $WT_y^h(n,j)$ in the variable n with $WT_x^h(n,j)$ fixed. This yields the definition

$$d(R_x^h, R_y^h) := \min_n \| WT_x^h - P_{cyc}(WT_y^h(n,j)) \|_2 \equiv \| WT_x^h - P_{min}(WT_y^h) \|_2 . \tag{8}$$

It has to be confirmed that $d(R_x^h, R_y^h)$ in fact has all properties of a metric. The definition immediately yields

$$d(R_x^h, R_y^h) = d(R_y^h, R_x^h) \geqslant 0 \quad \text{and} \quad d(R_x^h, R_y^h) = 0 \Leftrightarrow R_x^h = R_y^h .$$

The triangle inequality finally follows from this property of the Euclidean metric:

$$\begin{aligned} d(R_x^h, R_y^h) &\leqslant \| WT_x^h - WT_y^h \|_2 \\ &= \| WT_x^h - WT_z^h + WT_z^h - WT_y^h \|_2 \\ &\leqslant \| WT_x^h - WT_z^h \|_2 + \| WT_z^h - WT_y^h \|_2 . \end{aligned} \tag{9}$$

Because inequality (9) holds for each choice of WT_y^h and WT_z^h one can take the permutations of WT_y^h and WT_z^h which minimize both expressions in the last line. This finally yields $d(R_x^h, R_y^h) \leqslant d(R_x^h, R_z^h) + d(R_z^h, R_y^h)$.

As the next step an *addition* on the set $\{R_x^h\}$ can be defined by adding the matrices WT_x^h and $P_{min}(WT_y^h)$ which are synchronized to each other and then all cyclic permutations of the sum are taken. This yields the definition

$$R_x^h \oplus R_y^h := \{ P_{cyc}([WT_x^h + P_{min}(WT_y^h)](n,j)) \,|\, n = 1, \ldots, N \}, \quad j = 1, \ldots, J, \tag{10}$$

where the cyclic permutation $P_{min}(WT_y^h)$ is determined by the condition of minimal Euclidean distance between WT_x^h and $P_{cyc}(WT_y^h)$, (equation (8)). It follows that the addition in equation (10) is *commutative* but *not associative*. By using equation (5) the closure relation, $R_x^h \oplus R_y^h = R^h_{(x + P_{min}(y))} \in \{R_x^h\}$ is obtained. It follows that the set $\{R_x^h | \oplus\}$ has the mathematical structure of a *ring* with a *metric*. Further, one has $\forall a \in \mathbb{R}$, $a(R_x^h \oplus R_y^h) = aR_x^h \oplus aR_y^h$.

The quantities $R_x^h(j)$, $R_y^h(j)$ are called *auto-synchronized* wavelet transforms (in fact they are *sets* of time-permuted wavelet transforms), because the initial conditions of the matrices WT_x^h and WT_y^h calculated from subsections of a time series are adapted to each other (or auto-synchronized) by the permutation process. This property follows from the condition of minimal Euclidean distance between the permuted matrices, (equation (8)). Because it suffices to regard *one representative* (i.e., one element) of each set $R_x^h(j)$, the recipe to add auto-synchronized wavelet transforms is simply to take the matrices WT_x^h and WT_y^h, to permute one of them until the distance (7) is minimal, and then to add them. The result is a representative of $R_x^h \oplus R_y^h$. It follows that the addition is commutative, but not associative if more than two elements should be added.

3.4. CONSTRUCTION OF THE FEATURE VECTORS

As the next step a feature vector should be constructed from the time series $\{x_n; n = 1, \ldots, KN\}$ by averaging over the elements $\{R^h_{x^i};\ i = 1, \ldots, K\}$ which are calculated from the K sections of this series $\{x^i_n; n = 1, \ldots, N\}$. Because the addition (10) of more than two elements is not uniquely defined, an order of addition has to be determined. This is done by a *clustering algorithm.*

Initially each element $R^h_{x^i}$ defines its own cluster which is described by what can be called its "centre-of-mass vector" $c_{i,1} \equiv R^h_{x^i}$. The index (i, k) of $c_{i,k}$ indicates that cluster i consists of k elements. In a first step the elements with the smallest distance $d(c_{i,1}, c_{j,1})$, (equation (8)) to each other are unified and a common "centre-of-mass vector" is defined by $c_{i,2} = (c_{i,1} \oplus c_{j,1})/2$. In this manner, all clusters are successively unified in the order of smallest distances $d(c_{i,k}, c_{j,l})$ where the common "centre-of-mass vector" of two clusters is given by $c_{i,(k+l)} = (kc_{i,k} \oplus lc_{j,l})/(k + l)$. The clustering process ends with the element $c_{1,K}$ which defines an average of the set $\{R^h_{x^i};\ i = 1, \ldots, K\}$. In the next section, the properties of the feature vectors thus obtained will be studied.

Of course, this clustering scheme is not unique. A further method, for example, to obtain an ordering of the set $\{R^h_{x^i};\ i = 1, \ldots, K\}$ is to take the first two-element cluster which is created as described above and then to unify all remaining clusters only with this multiple-element cluster in the order of smallest distances $d(\tilde{c}_{1,k}, c_{j,1})$ of the one-element cluster $c_{j,1}$ to the "centre-of-mass vector" $\tilde{c}_{1,k}$ of the large one. However, for the data studied the constructed feature vectors $\tilde{c}_{1,K}$ show inferior properties compared to $c_{1,K}$.

For the auto-synchronization process it suffices to consider those columns of the matrices WT^h_x which contain relevant time information. These are the high-pass and the $J-1$ following band-pass components, where J is determined by the lowest dominant frequency of the signal. For the sunroof of data $J = 5$ is obtained.

At this point, an initial comment about the computational costs of the averaging method should be made. The auto-synchronization process in fact does not require the matrices $WT^h_x(n, j)$ to be cyclically permuted, but just one-dimensional vectors. Let $\widehat{WT}^h_x(n, j)$ denote the matrix which consists of the first J columns of $WT^h_x(n, j)$. By using the unitarity relation of the Fourier transform (Parseval's theorem) it follows from equation (4) and the definition of the wavelet transform that

$$\begin{aligned}
\| \widehat{WT}^h_x - \widehat{WT}^h_y \|_2^2 &= \sum_{j=1}^{J} \sum_{k=-N/2+1}^{N/2} (\hat{x}_k - \hat{y}_k)\hat{h}_{k,j}(\bar{\hat{x}}_k - \bar{\hat{y}}_k)\bar{\hat{h}}_{k,j} \\
&= \sum_k (\hat{x}_k - \hat{y}_k)(\bar{\hat{x}}_k - \bar{\hat{y}}_k) \sum_{j=1}^{J} |\hat{h}_{k,j}|^2 \\
&= \sum_k (\hat{x}_k - \hat{y}_k)(\bar{\hat{x}}_k - \bar{\hat{y}}_k) |\widehat{hp}_{k,J}|^2 \\
&= \| \widehat{hp}_J \hat{x} - \widehat{hp}_J \hat{y} \|_2^2 = \| x * hp_J - y * hp_J \|_2^2. \qquad (11)
\end{aligned}$$

In equation (11), $\hat{}$ denotes the discrete Fourier transform and $*$ the discrete convolution in time space. In the third line, $\widehat{hp}_{k,J} := \sqrt{\sum_{j=1}^{J} |\hat{h}_{k,j}|^2}$ defines a high-pass filter in frequency space. From equation (11) it follows that the synchronization of two wavelet transforms $\widehat{WT}^h_x(n, j)$ and $\widehat{WT}^h_y(n, j)$ is equivalent to the synchronization of the corresponding high-pass filtered sections $x * hp_J$ and $y * hp_J$ of the time series. It is therefore not necessary

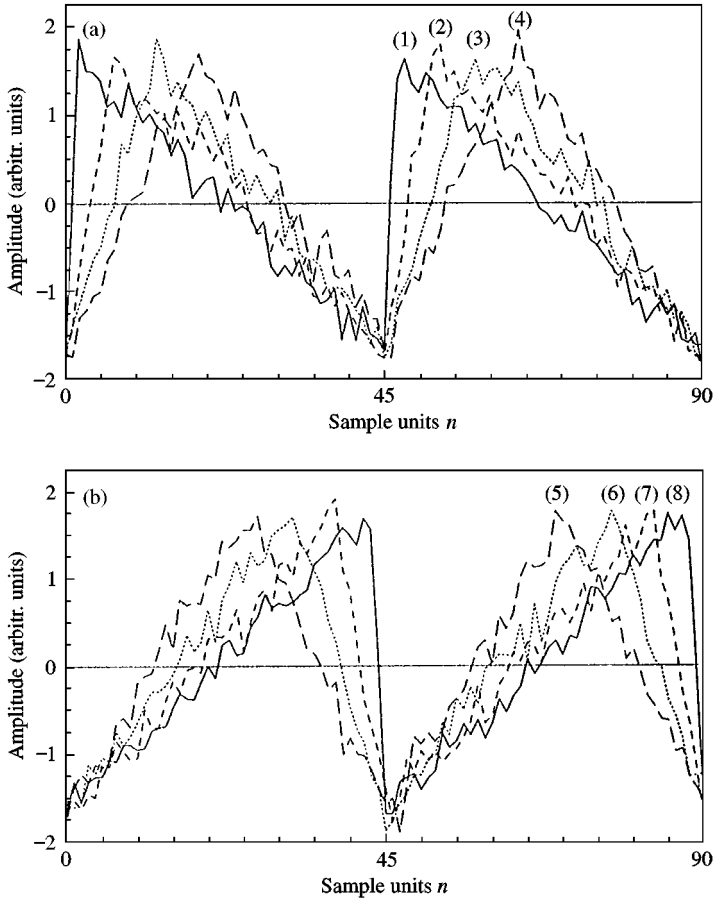

Figure 4. Sawtooth signals with 15% white noise added. The clean signals in (a) and (b) are pairwise related by reflection symmetry.

to calculate all cyclical permutations of $\widetilde{WT}_x^h(n, j)$ to obtain the feature vector $c_{1,K}$, but only those of $x * hp_J$. This significantly decreases the computational effort.

3.5. A NUMERICAL TEST OF THE FEATURE VECTORS

The properties of the feature vectors $c_{1,K}$ obtained from auto-synchronized wavelet transforms will now be studied. Therefore, artificial noisy signals are considered which are simple but non-trivial for the purpose of numerically testing the averaging method. Because the principle applicability of this averaging scheme should be shown here and since this concept is not expected to be limited to a narrow class of signals it is not intended to simulate the sliding noise data of the sunroofs. Apart from that, there is no detailed information available about the mechanism of the sound signals generation.

For a numerical test eight sawtooth signals are considered which differ by their slope where each pair of them are related by reflection symmetry (see Figure 4). 15% white noise is added to the signals. According to the definition of stationarity stated in section 3.3 these signals are stationary above the time scale $T = 45$ of their period while all information

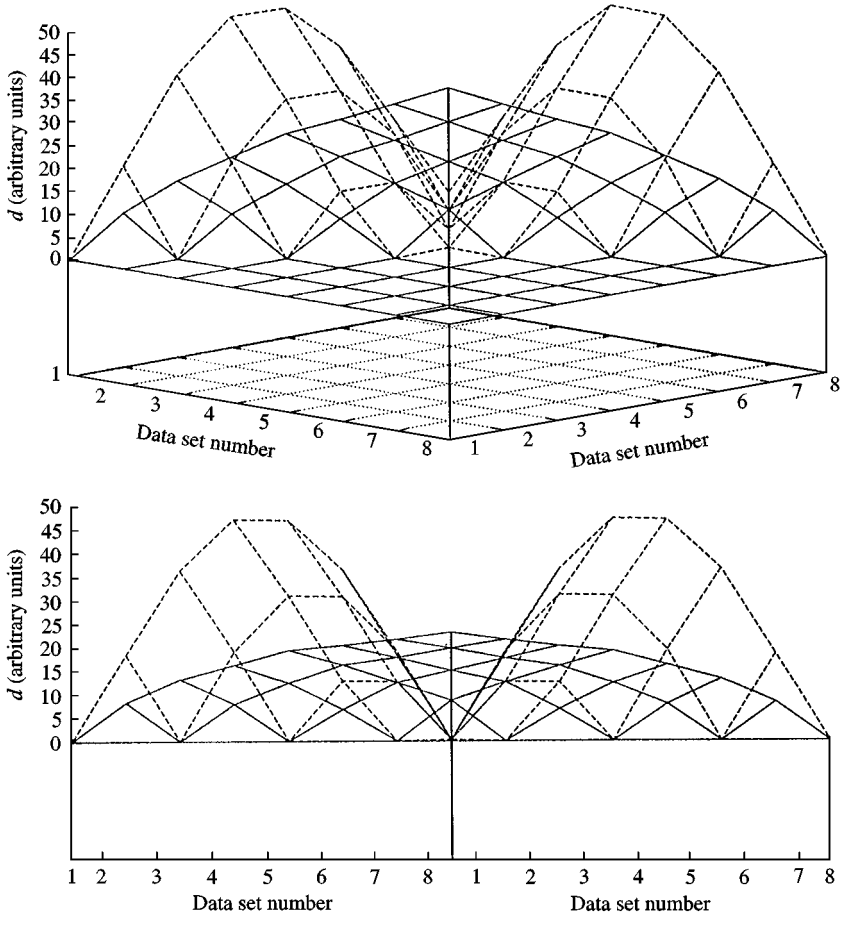

Figure 5. The pairwise distances $\{d(c^i_{1,30}, c^j_{1,30});\ 1 \leqslant i \leqslant j \leqslant 8\}$, equation (8), of the feature vectors $\{c^i_{1,30};\ i = 1, \ldots, 8\}$ (——) and the Euclidean distances of the spectral feature vectors (- - - - - -), both calculated from the signals shown in Figure 4. The numbering of the signals is given in Figure 4, the perspective is 15° in the upper and 0° in the lower panel. The front part of the figure would be symmetric to the rear part and is therefore not plotted.

needed for the classification is contained below that time scale. These properties are similar to the sunroof data.

First, the performance of the feature vectors $c_{1,K}$ is compared to the standard method of calculating windowed Fourier spectra. Therefore, windowed Fourier transforms over 512 sample units of the time series are calculated and the squared magnitudes are averaged over a section of length 15,360. Figure 5 shows the pairwise Euclidean distances of the spectral feature vectors drawn with dashed lines.

As a comparison the feature vectors $\{c^i_{1,K};\ i = 1, \ldots, 8\}$ are calculated from the sawtooth signals where the number of elements $\{R_{x^l};\ l = 1, \ldots, K\}$ which are averaged per time series is chosen as $K = 30$. The wavelet transforms $\{WT^h_{x^l}(n, j);\ l = 1, \ldots, K\}$ have the period $N = 512$ which also yields a signal length of 15,360. It is important to note that the period of the wavelet transforms is *not a multiple* of the signals' period $T = 45$. This in fact makes this numerical example a highly non-trivial test for the averaging method. Since the dominant frequency of the sawtooth signals is $f = \frac{1}{45}$ it suffices to regard the high-pass and the four following band-pass filtered components for the calculation of the feature vectors. Figure 5

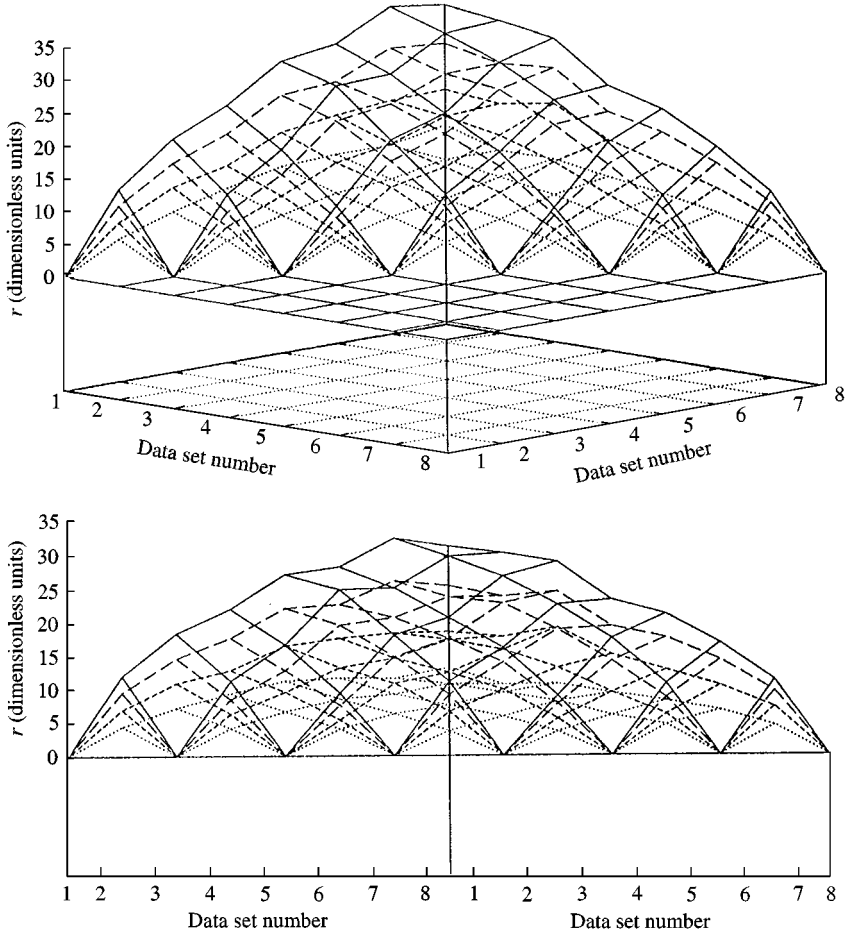

Figure 6. The matrix $\{r_{ij}(K);\ 1 \leqslant i \leqslant j \leqslant 8\}$, equation (12), plotted for $K = 4, 10, 20$ and 30 (dotted (· · · · · ·), dashed (- - - -), long dashed (– – –), and solid lines (——), respectively). The numbering of the signals and the perspectives are the same as in Figure 5.

shows the pairwise distances $\{d(c^i_{1,30}, c^j_{1,30});\ 1 \leqslant i \leqslant j \leqslant 8\}$, (equation (8)) of the feature vectors $\{c^i_{1,30};\ i = 1, \ldots, 8\}$ drawn with solid lines.

As can clearly be seen, the spectral feature vectors are only able to separate the fraction of higher harmonics contained in the sawtooth signals while the signals being symmetric to each other cannot be distinguished (the diagonal in Figure 5). In contrast, the feature vectors $c^i_{1,30}$ resolve the information in time space well. The distance $d(c^i_{1,30}, c^j_{1,30})$ monotonically increases as a function of $|i - j|$, the absolute value of the difference between the signals' numbers given in Figure 4.

The question of major interest is *how strongly* the noise can be reduced by this averaging process. The averaging over spectral components which are calculated from K sections of a time series leads to the convergence of the noise term to zero as $1/\sqrt{K}$. To study this behaviour for the quantities $c_{1,K}$ eight identical, approximately isosceles sawtooth signals (No. 4 or 5 in Figure 4) are taken to which 15% (differently realized) white noise is added. A quantity which defines a measure for the *signal-to-noise ratio* of the feature vectors $c^i_{1,K}$ is

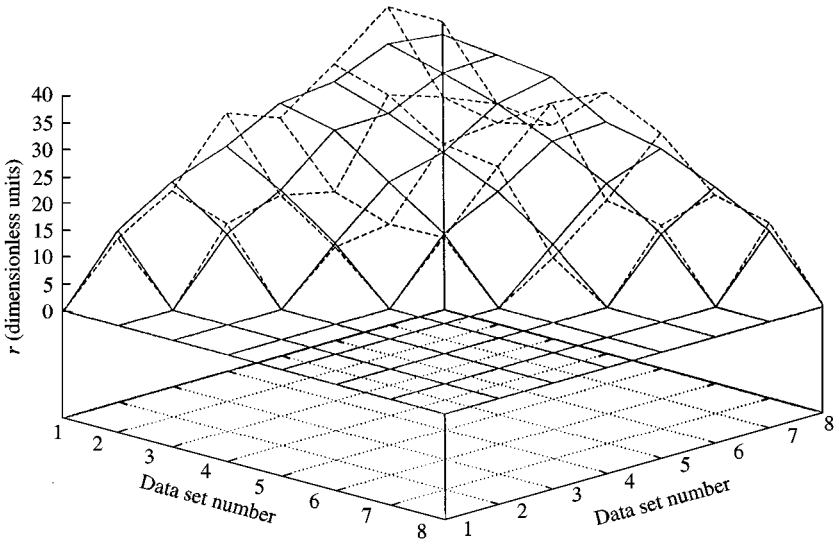

Figure 7. The matrix $\{r_{ij}(50);\ 1 \leqslant i \leqslant j \leqslant 8\}$ (- - - -) to compare with $\{r_{ij}(30);\ 1 \leqslant i \leqslant j \leqslant 8\}$ (——) from Figure 6. The perspective is 25°.

the (symmetric) matrix

$$r_{ij}(K) = \frac{d(c^i_{1,K}, c^j_{1,K})}{d(\tilde{c}^i_{1,K}, \tilde{c}^j_{1,K})}, \quad i,j = 1, \ldots, 8, \tag{12}$$

which is regarded as a function of K, the number of elements $\{R_{x^l};\ l = 1, \ldots, K\}$ over which the average per time series is performed. The feature vectors $\tilde{c}^i_{1,K}$ are calculated form the (except the noise term) identical sawtooth signals while the vectors $c^i_{1,K}$ are defined as above. In Figure 6, the matrix $\{r_{ij}(K);\ 1 \leqslant i \leqslant j \leqslant 8\}$ is plotted for $K = 4$, 10, 20 and 30.

As is clearly visible in Figure 6, for $K \lesssim 30$ the quantities $r_{ij}(K)$ increase for each pair i, j in approximately equal steps when K is doubled. An important criterion for the performance of the noise reduction are the fluctuations of $r_{ij}(K)$ as a function of i, j with K hold fixed, i.e., how "smooth" the surfaces in Figure 6 are.

For larger values of K the averaging process does not converge. First, in the region $30 \lesssim K \lesssim 40$ the signal-to-noise ratios of $c^i_{1,K}$ appear to saturate, i.e., almost stay the same. There is no visible difference between $\{r_{ij}(40)\}$ and $\{r_{ij}(30)\}$; therefore these values are not shown in Figure 6. For $K > 40$ the averaging process becomes unstable, i.e., $r_{ij}(K)$ starts to fluctuate and again decreases for many pairs i, j. This is illustrated in Figure 7. This behaviour appears to be similar to that of *asymptotic* power series expansions which approach the function they expand for small values of the order parameter n, reach the best approximation for a finite value $n = n_{max}$, and diverge thereafter.

The performance of this noise reduction depends on the kind of signals. In the case of the sawtooth signals the signal-to-noise ratios improve for all values of K if less than 15% noise is added. If there is too much noise on the other side ($> 20\%$) this method fails. However, because the sawtooth signals have a very pronounced period which does not divide the period of the wavelet transform this example may be difficult for the averaging mechanism (as was intended here). In the following application to the experimental sound signals the numerical results can be (qualitatively) confirmed. In these tests only the convergence

properties of the clustering algorithm described in section 3.4 that works best are shown. It is easy to understand that any arbitrary clustering scheme will impair or even disable the results.

4. ACOUSTIC QUALITY CONTROL OF SLIDING SUNROOFS

In this section, the application of the auto-synchronized wavelet transform analysis to the classification of structure-borne sound signals of electric sliding sunroofs is shown. Because the pre-classifications of the sunroofs turned out not to be completely reliable (the repetition of the acoustic tests by humans lead to partially deviating results) the following question is considered.

Assume that a classification algorithm is trained with time series recorded from a set of units which represent all quality conditions. Can then the pre-classifications be reproduced in an *out-of-sample* test? It is assumed that only those items which lie near the border of acceptable tolerances may be unreliably pre-classified. This means that the algorithm is mainly trained with features that will also reflect the quality conditions of the units under test (and not arbitrary features). The classification power of different feature vectors can then (at least qualitatively) be compared. The feature vectors $c_{1,K}$ will be compared to the standard method of calculating windowed Fourier spectra. This result appears to be statistically significant for the data base available.

4.1. THE CLASSIFICATION ALGORITHM

A feature vector $c_{1,K}$ is calculated from each sound signal where K is chosen 20, 30 and 40 respectively. For the averaging process only the high-pass and the four following band-pass components of $\{WT_x^h(n, j);\ l = 1, \ldots, K\}$ are considered, as mentioned above.

The data sets recorded in both sensor positions are classified separately in a *leave-one-out* statistic [14]. This means that all but one of the (statistically independent) elements of a data set are used for training of a classification algorithm which then classifies the single test element. The process is repeated until all elements are classified. This yields the classification of the data set in an *out-of-sample* test where the test set contains all elements and each training set all minus one. For the sunroof data a *leave-two-out* statistic has to be considered, because from each sunroof two time series were recorded in each sensor position which have to be removed from the training set since they are not statistically independent.

During the testing stage the test vectors are compared to each of a neighbourhood (of different size) of training samples which represent the items of good and of poor quality, respectively. To both neighbourhoods an average distance is determined and the difference of the weighted distances is then compared to a threshold value to yield a discriminating statistic.

First, the distance of a test vector $c_{1,K}^{\mathrm{T}}$ to a *set* of training vectors $\{c_{1,K}^i;\ i = 1, \ldots, L\}$ has to be uniquely defined. This is done by the definition of a "centre-of-mass vector" of the training neighbourhoods. Because the addition of feature vectors is not associative again several possibilities exist.

The first is to take the same clustering algorithm which is already used in section 3.4 to define the averaging process of the auto-synchronized wavelet transforms. This yields a first "centre-of-mass vector" $\check{C}_L$. But now a natural order for the addition of the training vectors $\{c_{1,K}^i;\ i = 1, \ldots, L\}$ in a neighbourhood exists: the distances of $c_{1,K}^i$ to the test vector $c_{1,K}^{\mathrm{T}}$.

Adding the training vectors in this order yields a second "centre-of-mass vector" C_L. As a third method the clustering algorithm mentioned in section 3.4 which unifies all one-element clusters with only one central cluster defines a third "centre-of-mass vector" $\tilde{C}_L$.

A (qualitative) comparison of these possibilities yields for the sunroof data that calculating the "centre-of-mass vectors" C_L or $\tilde{C}_L$ of the training neighbourhoods leads to similar classification results, while $\check{C}_L$ shows inferior properties. This indicates that for the averaging process of the auto-synchronized wavelet transforms and for the definition of "centre-of-mass vectors", respectively, different clustering schemes are preferable. The important point here is that the calculation of C_L is numerically much more inexpensive than $\check{C}_L$ or $\tilde{C}_L$, because the number of distances $d(c^i_{1,K}, c^j_{1,K})$ which have to be calculated to determine the order of addition grows linearly with the size of the training neighbourhoods instead of quadratically. In order to keep the computational costs of the classification algorithm affordable only the calculation of the "centre-of-mass vectors" C_L is considered further.

For each test vector a neighbourhood of n_1 training vectors corresponding to items of good quality and a second one of n_2 vectors which represent the units of poor quality are sought and the corresponding "centre-of-mass vectors" C_L are calculated. Outliers can be removed from the training set in order to train only "typical" quality conditions. Now, the J-dimensional *distance vectors* $\mathbf{d}_g$ and $\mathbf{d}_p$ of a test vector $c^{\mathrm{T}}_{1,K}$ to the "centre-of-mass vectors" C^g_L and C^p_L of the training neighbourhoods which represent units of good and of poor quality, respectively, are defined as

$$\mathbf{d}_{g,p} = (d([c^{\mathrm{T}}_{1,K}]_1, [C^{g,p}_L]_1), \ldots, d([c^{\mathrm{T}}_{1,K}]_J, [C^{g,p}_L]_J)). \tag{13}$$

The metric $d([c^{\mathrm{T}}_{1,K}]_j, [C^{g,p}_L]_j)$ (equation (8)) is calculated from only *one component* of the feature vectors $c^{\mathrm{T}}_{1,K}$, C^g_L or C^p_L (i.e., from one column of the corresponding matrices $WT^h_x(n, j)$. It is now possible to weight individually the frequency bands of the signals. This turns out to be necessary because the information about the quality condition of the sunroofs and other influences is distributed differently in the frequency bands of the sound signals. Finally, the components of the feature vectors are weighted by calculating the *relative distance*

$$d_{rel} = \langle \mathbf{d}_g - \boldsymbol{\alpha}\mathbf{d}_p, \mathbf{n} \rangle \quad \text{with } \|\mathbf{n}\|_2 = 1, \tag{14}$$

where $\langle \cdot \rangle$ denotes the scalar product, $\boldsymbol{\alpha}$ is a vector-valued parameter and $\mathbf{n}$ determines the subspace into which the distance vectors are projected. The relative distance d_{rel} can then be compared to a threshold parameter to obtain the desired classification of the test units into the quality classes A and B.

For the calculation of d_{rel} the parameters $\boldsymbol{\alpha}$, $\mathbf{n}$, n_1 and n_2 (size of training neighbourhoods) have to be estimated. This is done by estimating the *Bayes error* [14] of the classification statistics. This means that the total classification error of the *leave-two-out* statistics is minimized with respect to the parameters. Strictly speaking this yields a mixture of *in-sample* and *out-of-sample* test, because the parameters are not estimated separately on each training set. Therefore, the true minimal classification error (the Bayes error) may be underestimated because of the relatively large number of parameters. However, in this case the (qualitative) comparison of different features vectors is the goal.

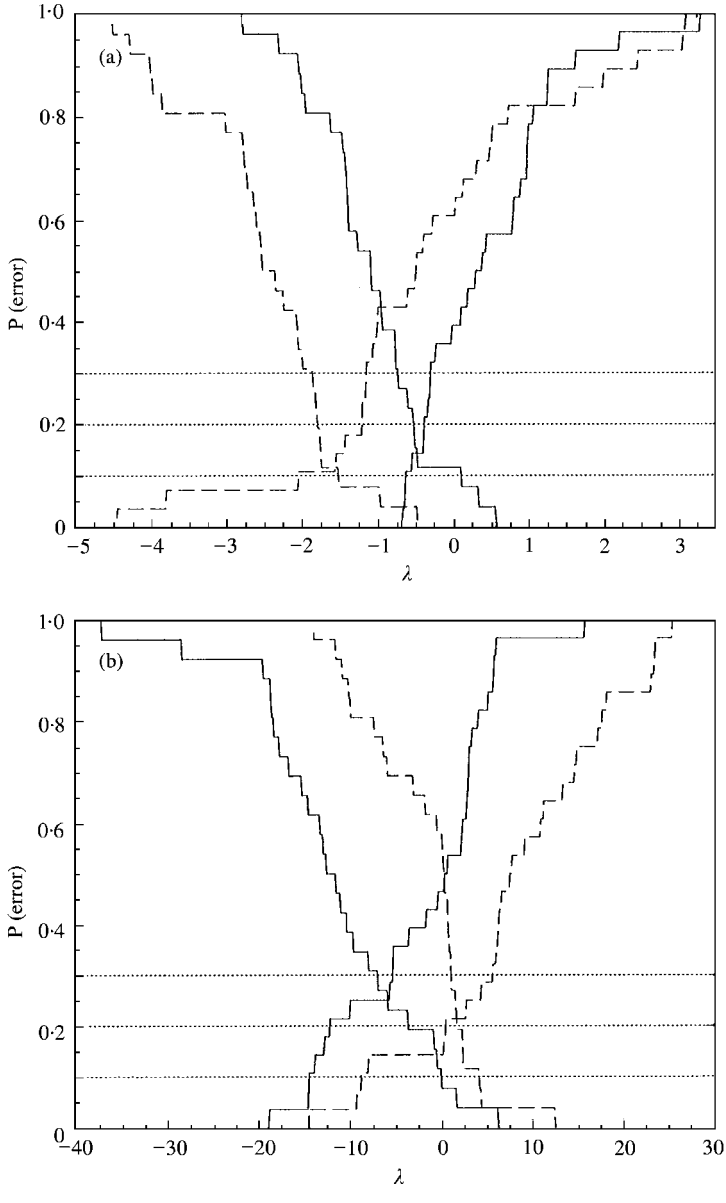

Figure 8. The classification of the sunroof data: probability of error (false rejections (the monotonically decreasing graphs), and false acceptances (the monotonically increasing graphs)) as a function of the sensitivity parameter λ: (——), Data recorded in sensor position 1; (- - - -), Data recorded sensor position 2. (a): Feature vectors $c_{1,40}$ from auto-synchronised wavelet transforms; (b): Feature vectors from spectral components.

4.2. EXPERIMENTAL RESULTS

As in section 3.5 the performance of the feature vectors $c_{1,40}$ is first compared to a spectral method. A spectral feature vector (as described in section 3.5) is calculated from a section of length $40 \times 512 = 20{,}480$ of each sound signal. These vectors are classified analogously to the method described in section 4.1, where now the Euclidean distances

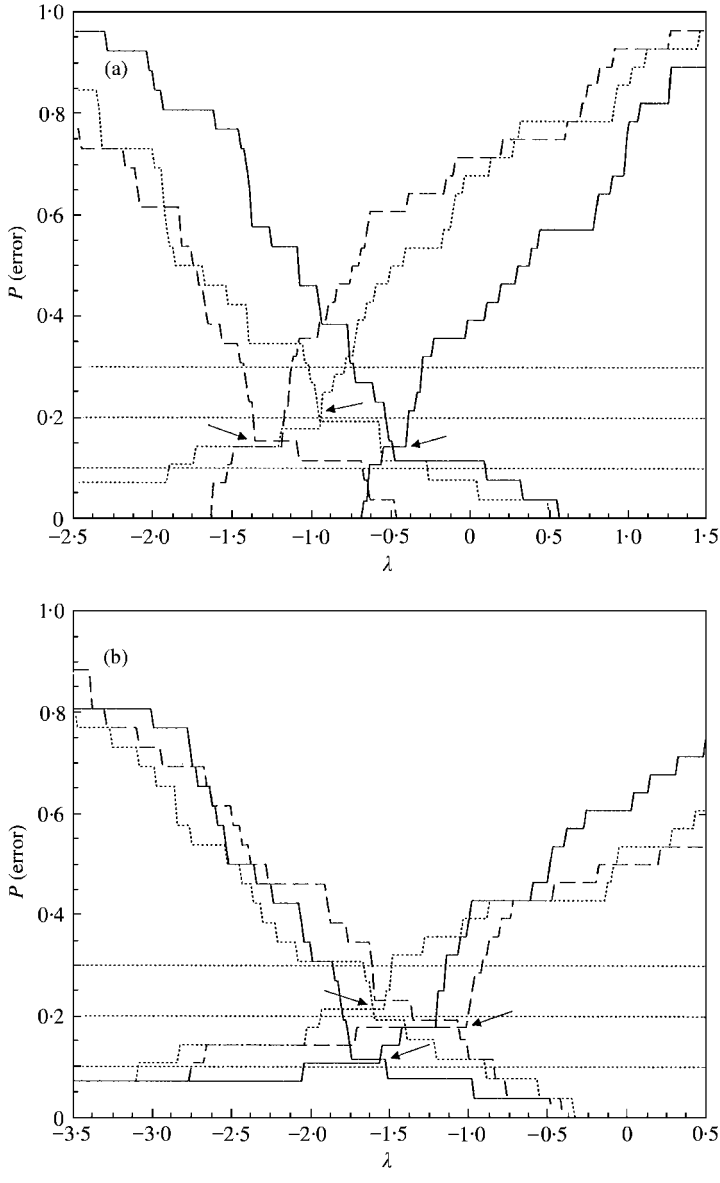

Figure 9. The classification of the sunroof data using the feature vectors $c_{1,K}$ with $K = 20$, 30, and 40 (the dotted (· · · ·), dashed (- - - -) and solid lines (——) respectively). (a): Sensor position 1; (b): Sensor position 2. Compare to Figure 8(a).

d_g and d_p of the test vectors to the centre-of-mass vectors of the training neighbourhoods are calculated and a relative distance is defined as $d_{rel} = d_g - \alpha d_p$ with a scalar weighting constant α.

For the calculation of d_{rel} (equation (14)) the components of $c_{1,K}$ which contribute to the classification performance are selected. It turns out that the first band-pass of the wavelet transforms contains no relevant information, therefore only the high-pass and the

band-pass components 2–4 are considered. The classification performance of the spectral feature vectors on the other side could not be improved by the selection of frequency bands.

The classification results of the sound signals are shown in Figures 8 and 9 to be discussed below. On the y-axis the fraction of the feature vectors in a test set being pre-classified as good is plotted whose distances d_{rel} are *larger* than the threshold parameter λ drawn on the x-axis (the monotonically *decreasing* graphs). Equivalently, the fraction of the feature vectors in a test set being pre-classified as of poor quality is plotted whose distances d_{rel} are *smaller* than λ (the monotonically *increasing* graphs). This means that for the units pre-classified as of good quality the probability of *false rejections* of the classification algorithm is shown as a function of λ whereas for the units of poor quality the probability of *false acceptances* is plotted. In other words, the tradeoff between rejecting good items and accepting poor ones as a function of a sensitivity parameter can be read off these figures. This characteristic of a classification algorithm is of major interest for any on-line application. The minimal classification error (the *estimated* Bayes error) is given by the crossing point of the graphs which indicate the probability of false rejections and of false acceptances respectively.

The classification of the acoustic data using both kinds of feature vectors is shown in Figure 8. The data sets recorded in both sensor positions are classified separately; this is indicated by the use of solid and dashed lines respectively. It is clearly seen that for both data sets the application of the feature vectors $c_{1,40}$ leads to smaller classification errors than the use of spectral components. Since the pre-classifications of the sunroofs are not completely reliable these results have to be regarded qualitatively. Nevertheless, the difference in the classification performance of both feature vectors appears to be statistically significant.

The estimated size of the training neighbourhoods turns out to be independent of the particular feature vector, and one obtains $n_1 = 18$ (from 24 to 26 feature vectors which represent the units of good quality) and $n_2 = 8$ (from 26 to 28 vectors corresponding to units of poor quality where 6–8 outliers are removed from the training sets[§]). It is not surprising that n_1 is larger than n_2, because the feature vectors which correspond to units of poor quality may represent different (differently strong) defects. The components of $\boldsymbol{\alpha}$ are in the range between 0 and 2.

Finally, the convergence of the averaging algorithm of the auto-synchronized wavelet transforms should be studied for the acoustic data. Therefore, the feature vectors $c_{1,K}$ are calculated for different values of K, i.e., by averaging a different number of elements $\{R_{x^l};\ l = 1, \ldots, K\}$ per time series. The classification performance of these feature vectors can then be compared. In Figure 9, the classification of the acoustic data is shown again for both sensor positions where now the feature vectors $c_{1,K}$ are applied with $K = 20$, 30 and 40. The arrows in Figure 9 indicate the crossing points of the corresponding curves of false rejections and false acceptances, i.e., the minimal classification errors.

In sensor position 1 the classification power of the feature vectors $c_{1,30}$ and $c_{1,40}$ differs only weakly while there is a visible inferiority of $c_{1,20}$. In sensor position 2 the best result is only obtained using $c_{1,40}$ while $c_{1,30}$ and $c_{1,20}$ show a continuous decrease of the classification performance. Due to the total length of the sound signals the comparison with averages which are calculated over larger sections of the time series cannot be made. This result indicates that the averaging of auto-synchronized wavelet transforms leads

[§]The outliers, of course, remain in the test set.

to a significant reduction of statistical fluctuations also for the experimental data. This confirms (at least qualitatively) the numerical results obtained in section 3.5.

The computational costs of the classification algorithm depend on the number of wavelet transforms which are averaged per time series and on the size of the training set. For $K = 30–40$ and the size of the data set studied, the classification of a single unit needs about 5–10 s of cpu-time on a modern work station. Of course, this is much more than is needed for spectral methods, but for an on-line application it may be sufficient that the data processing time is of the same order as the time needed to generate the sound signal. For the electrical sunroofs this is about 6 s.

5. CONCLUSIONS

This paper has presented the concept of auto-synchronized wavelet transform analysis and applied it to the automatic acoustic quality control of sliding sunroofs. The method aims to classify noisy sound signals which are stationary above some time scale while the relevant information is contained (everywhere in time space) below that time scale and can (at least in principle) be resolved by the human ear.

The concept is to calculate feature vectors from stationary sound signals for, for example, the automatic classification at the quality control point where the particular features of the signals which distinguish units of different quality are not explicitly sought, because these signatures are highly specific for each application. Instead, feature vectors are constructed in a general manner and the classification is then performed by a next neighbour search on a training set. Because the relevant information of the sound signals can be resolved by the human ear, i.e., by experts, the approach is to describe approximately the time–frequency resolution of the human ear by a wavelet transform.

The main subject of this paper is a new method for the reduction of statistical fluctuations (noise, parameter drifts) for (discrete) wavelet transforms or, more general, time–frequency representations. A high noise level turns out to be the main problem for the automatic classification of the data by a next neighbour search algorithm. As a new approach to this problem the class of auto-synchronized wavelet transforms is introduced, which are defined as the *set* of all cyclic permutations of a wavelet transform in the time variable. These quantities, calculated from subsections of a (stationary) time series, can be averaged in the time domain without losing the time-resolved information of the signal. It is shown for numerical data and for experimental sound signals that statistical fluctuations can be significantly reduced by this averaging algorithm in order to reveal the characteristic features of the signals needed for the classification. For the data studied the feature vectors constructed from auto-synchronized wavelet transforms outperformed the standard method of calculating averaged windowed Fourier spectra. But this method of noise reduction can in general not reach the performance of methods which make use of such additional features of a signal (if they are present) as the synchronization of a periodic signal by a pulse signal.

Feature vectors can also be constructed from auto-synchronized time–frequency representations of a different type in an analogous way. If filters with constant width in frequency space are considered instead of the scaled filters used here, a time–frequency representation with a constant instead of a hyperbolic time–frequency resolution is obtained. This would allow the psycho-acoustic hypothesis which motivated the approach to be checked but for a statistically significant result a larger data base is necessary. Nevertheless, the applicability of feature vectors constructed from auto-synchronized wavelet transforms to the classification of stationary sound signals could be shown.

ACKNOWLEDGMENTS

The authors would like to acknowledge the kind support provided by Dr R. Zöller of Carl Schenck AG, Darmstadt, Germany. The experimental data was provided by Carl Schenck AG. The authors want to thank Dr E. Olbrich for many fruitful discussions and comments.

REFERENCES

1. D. DORNFELD 1992 *NDT&E* **25**, 259–269. Application of acoustic emission techniques in manufacturing.
2. D. FARSON, K. HILLSLEY, J. SAMES and R. YOUNG 1996 *Journal of Laser Applications* **8**, 33–42. Frequency–time characteristics of air-borne signals from laser welds.
3. S. LEE and P. WHITE 1997 *Mechanical Systems and Signal Processing* **11**, 637–650. Higher-order time–frequency analysis and its application to fault detection in rotating machinery.
4. C. HEITZ and J. TIMMER 1997 *ACUSTICA* **83**, 1–12. Using optimised time–frequency representations for acoustic quality control of motors.
5. F. HLAWATSCH and G. BOUDREAUX-BARTELS 1992 *IEEE-SP Magazine*, April, 21–67. Linear and quadratic time–frequency representations.
6. P. MCFADDEN 1987 *Mechanical Systems and Signal Processing* **1**, 173–183. Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration.
7. S. LEE and P. WHITE 1998 *Journal of Sound and Vibration* **217**, 485–505. The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery.
8. H. KANTZ and T. SCHREIBER 1997 *Nonlinear Time Series Analysis*. Cambridge, UK: Cambridge University Press.
9. K. HELDMANN 1994 *Wahrnehmung, gehörgerechte Analyse und Merkmalsextraktion technischer Signale*. Fortschritt-Berichte VDI. Düsseldorf, Germany: VDI Verlag.
10. O. RIOUL and M. VETTERLI 1991 *IEEE-SP Magazine*, October, 14–38. Wavelets and signal processing.
11. I. DAUBECHIES 1992 *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics.
12. S. GÜTTLER 1999 *Ph.D. Thesis, On-line library University Wuppertal, Germany. Theoretische Konzepte der Zeitreihenanalyse und Anwendungen auf die Signalklassifikation in Technischen Systemen.*
13. C. GARDINER 1997 *Handbook of Stochastic Methods*. Berlin: Springer.
14. K. FUKUNAGA 1990 *Introduction to Statistical Pattern Recognition*. San Diego: Academic Press.