ANNALI DI GEOFISICA, VOL. 42, N. 3, June 1999

Best-basis analysis of broadband tremor signals

Thorsten Bartosch and Peter Steffen
Lehrstuhl für Nachrichtentechnik I, Universität Erlangen-Nürnberg, Erlangen, Germany

Abstract
Active volcanoes usually generate highly non-stationary broadband tremor signals. Short-time shock events with a frequency content of several decades are superimposed on a stationary narrow band continuous tremor. Tremor signals of this type can be observed in the near field of many active volcanoes. In this paper we will demonstrate the analysis of such signals using a specific tremor signal of Mt. Stromboli (Sicily). We used the Best-Basis Algorithm (BBA) in order to compute a spectrogram which is adapted to signal properties on highly different scales. It turns out that the BBA can reveal better fitting properties of the tremor in the time-frequency plane compared to standard methods like Short-Time Fourier Transformation (STFT). Moreover, this very effective algorithm can be used for real time monitoring in the time-frequency plane, for data compression or for denoising of the tremor signals.

Key words *adaptive spectrogram analysis – broadband tremor*

1. Introduction

A conventional approach to reveal the properties of non-stationary signals consists of splitting up the signal into adjacent unit length time intervals. When quasi-stationary behaviour of the signal within an interval can be assumed a transformation for the interval can be applied to study the signal properties. The temporal variations are due to the subsequent time intervals. A standard method like STFT is based on this concept.

Problems arise, *e.g.*, when due to information on different time scales the signal cannot be divided into quasi-stationary blocks. In this case a more flexible concept, which is provided by the adaptive waveform analysis as introduced, *e.g.*, in Coifman and Wickerhauser (1992), Mallat and Zhang (1993), Wickerhauser (1991) or Wickerhauser (1996), has to be used.

The adapted waveform analysis describes how to decompose a signal with respect to a dictionary of basis functions. The properties of these basis functions mainly influence the signal features which can be extracted. Assuming a dictionary with an inherent binary tree structure (see fig. 1), efficient algorithms like the BBA can be applied.

Spectrograms related to adapted waveform analysis result in a more suitable resolution of time and frequency compared to STFT or Wavelet Transformation (WT). Therefore, a frequency range of several decades can be resolved properly. Due to the wide frequency range of seismic broadband records these spectrograms proved to be very useful for obtaining further insight into the nature of tremor signals.

Mailing address: Dr. Thorsten Bartosch, Lehrstuhl für Nachrichtentechnik, Universität Erlangen-Nürnberg, Erlangen, Cauerstraße 7, D-91058 Erlangen, Germany; e-mail: thorsten@nt.e-technik.uni-erlangen.de

2. Wavelet packets

In principle, basis dictionaries may consist of any type or family of orthogonal or non-

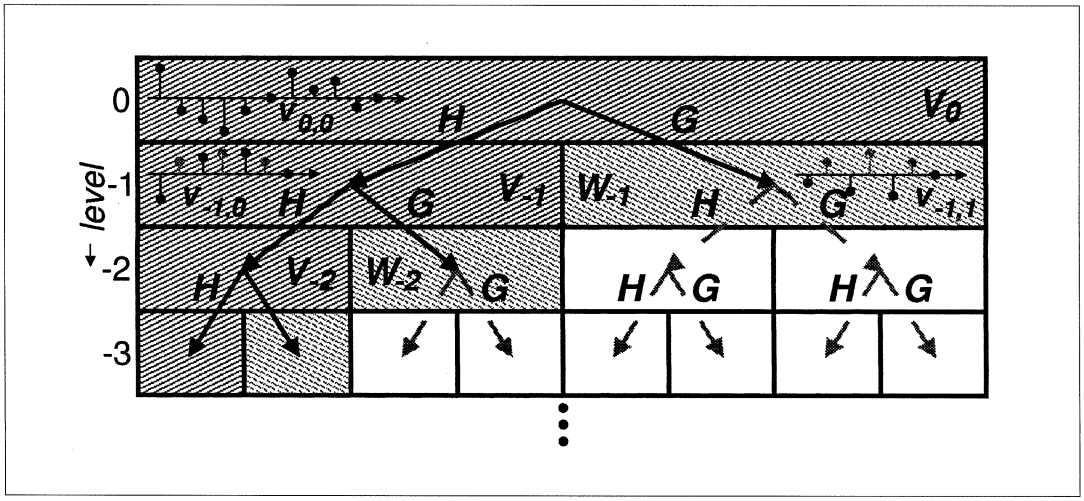

Fig. 1. A nested sequence of subspaces (approximation spaces) V_0, V_{-1}, V_{-2}, ... and their related orthogonal complements (wavelet or detail spaces) W_{-1}, W_{-2}, ... span an MRA. The projection of a discrete signal $f(n)$ onto these spaces is computed by filtering and down-sampling (operators G and H). By additionally decomposing all the projections onto wavelet spaces an over-complete collection of projections called a wavelet packet table is generate (gray arrows).

orthogonal basis functions. Examples are impulses, harmonic signals, modulated windows at different time-scales, wavelets, or wavelet packets. A generalization of the discrete WT (Daubechies, 1992; Jawerth and Sweldens, 1993; Louis *et al.,* 1994; Strang and Nguyen, 1996) leads to the Wavelet Packet Transformation (WPT) (Wickerhauser, 1991, 1996; Coifman and Wickerhauser, 1992; Mallat and Zhang, 1993; which has been used in this approach since the resulting algorithm is very efficient. In the sequel, a short introductory description is given. For further details see Wickerhauser (1996).

To generate wavelets, the scheme of multiresolution analysis (MRA) can be used to provide a wide variety of different wavelets with different properties. An MRA consists of a sequence of nested subspaces V_j, called approximation spaces,

$$\{0\} \subset \cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \cdots \subset \boldsymbol{L}^2(\boldsymbol{R}) \quad (2.1)$$

as depicted in fig. 1. These spaces possess the following properties:

$$\bigcap_{j=-\infty}^{j=\infty} V_j = \{0\} \quad (2.2a)$$

$$\bigcup_{j=-\infty}^{j=\infty} V_j = \boldsymbol{L}^2(\boldsymbol{R}) \quad (2.2b)$$

$$\exists \Phi(t) \text{ with } V_0 = \text{span}\{\Phi(t-k) | k \in \boldsymbol{Z}\} \quad (2.2c)$$

$$f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}, \quad j \in \boldsymbol{Z} \quad (2.2d)$$

$$f(t) \in V_j \Leftrightarrow f(t-k) \in V_j, \quad k, j \in \boldsymbol{Z}. \quad (2.2e)$$

$\Phi(t)$ is called scaling function. These properties induce a refinement equation

$$\Phi(t) = \sum_k a(k)\Phi(2t-k) \quad (2.3)$$

which represents a ladder between spaces at different levels. $\Phi(t) \in V_0$ can be decomposed by the basis functions $\Phi(2t-k) \in V_1$ as $V_0 \subset V_1$. The orthogonal complement of each approximation space V_j with respect to V_{j+1} is called a detail space W_j. It is spanned by translates of the func-

tion $\Psi(t)$ termed wavelet. It can be stated

$$V_j \perp W_j, W_i \perp W_j \; \forall i \neq j$$

$$V_{j+1} = V_j \oplus W_j.$$

Therefore, it is obvious that $\Psi(t)$ can be written as

$$\Psi(t) = \sum_k b(k)\Phi(2t-k). \qquad (2.4)$$

Normalizing $\Phi(t)$ and using the notation

$$\Phi_{j,k}(t) = 2^{j/2}\Phi(2^j t - k), \qquad (2.5a)$$

$$\Psi_{j,k}(t) = 2^{j/2}\Psi(2^j t - k), \qquad (2.5b)$$

the sets of $\Phi_{j,k}(t)$, $k \in \mathbf{Z}$ and $\Psi_{j,k}(t)$, $k \in \mathbf{Z}$ constitute orthonormal bases of V_j and W_j, respectively.

In order to establish an MRA an appropriate set of coefficients $a(k)$ has to be found by means of which a solution $\Phi(t)$ of (2.3) can be obtained. The coefficients $b(k)$ can then be calculated easily as $b(k) = (-1)^k a(1-k)$ (see Daubechies, 1992). Further additional restrictions like *e.g.,* orthogonality, biorthogonality, number of vanishing moments, compact support or smoothness of the basis functions $\Phi(t)$ result in different types of wavelets.

The decomposition starts by projecting a signal $f(t) \in \boldsymbol{L}^2(\boldsymbol{R})$ onto an initial approximation space V_0

$$f(t) \approx \sum_k c_0(k)\Phi_{0,k}(t), c_0(k) = \int f(t)\Phi_{0,k}(t)dt. \qquad (2.6)$$

Then the coefficients of the projection onto the next coarser space V_{-1} are given by

$$c_{-1}(k) = \int f(t)\Phi_{-1,k}(t)dt.$$

Using (2.3), (2.5a) and exploiting the orthogonality of the $\Phi_{j,k}(t)$ it can be shown that

$$c_{-1}(k) = \frac{1}{\sqrt{2}} \sum_n a(n-2k)\, c_0(n).$$

Obviously, the $c_{-1}(k)$ are obtained by convolution of $a(-k)$ and $c_0(k)$ followed by a decimation of factor 2. The operator which produces approximations on different levels acts as a low-pass filter (H), while the operator (G) can be recognized as a high-pass filter. The coefficients of an MRA can efficiently be calculated by successively filtering the coefficients of all approximations followed by a down-sampling operation. This procedure is called Mallat's Algorithm with asymptotically fewer computations than the FFT (Louis *et al.,* 1994).

An appropriate choice of successive approximation and detail spaces results in the representation of signals by wavelet packets. The resulting expansion is called a wavelet packet table, see fig. 1. An over-complete collection of high-pass, low-pass and band-pass basis systems is provided. Because of the inherent binary tree structure of the basis systems efficient algorithms exist. *E.g.,* the Best Basis (Coifman and Wickerhauser, 1992), Matching Pursuit (Mallat and Zhang, 1993) or Atomic Pursuit (Chen *et al.,* 1996) algorithms select an «optimal» decomposition of the signal out of this dictionary. Because of less computational afford, we applied BBA.

3. Coifman wavelets (Coiflets)

Probably the most famous wavelets were introduced by Daubechies (1992). They constitute an orthonormal family and are obtained by the requirements

$$a(k) = 0, \; k < 0, \; k \geq 2M, \; M \in \boldsymbol{N},$$

and

$\Phi(t)$ should be as regular as possible.

Similar to the Daubechies' wavelets the Coifman wavelets, often called «Coiflets», lead to an MRA. They are obtained by the additional requirement of vanishing moments for the scaling function

$$\int t^l \Phi(t)dt = 0, \quad l = 1(1)M-1 \qquad (3.1)$$

thus leading to a sequence $a(k)$ of length $3M$.

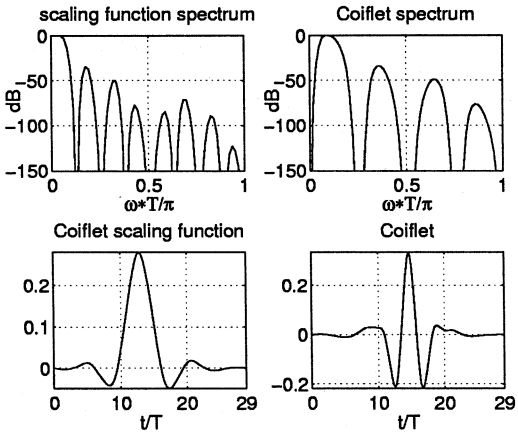

Fig. 2. Coiflet scaling function and wavelet of filter order 30 in time and Fourier domain (ω: frequency).

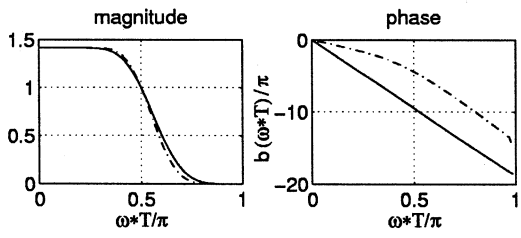

Fig. 3. Frequency responses of the magnitude and phase of a Coiflet (solid) $M = 3$ and a Daubechies (dotted) $M = 5$ filters.

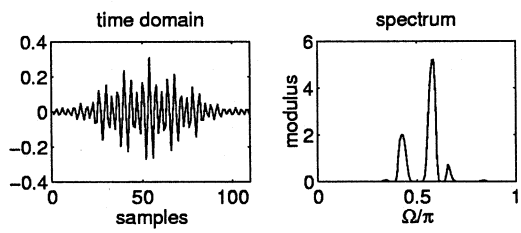

Fig. 4. Time and frequency domain of a Coiflet (30) packet basis function out of a badly frequency localized basis block at level 5. (Ω: normalized frequency).

This implies that $\Phi(t)$ has finite support as well

$$\text{supp } \{\Phi\} = [0, 3M - 1]. \qquad (3.2)$$

The corresponding wavelet is obtained as the solution of (2.4) where $b(k) = (-1)^k a(3M - 1 - k)$. As a consequence we have

$$\int t^l \Psi(t)dt = 0, \quad l = 0(1)M - 1, \qquad (3.3)$$

$$\text{supp}\{\Psi\} = [0, 3M - 1].$$

Their support length (time localization) compared with other MRA wavelets increases. On the other hand, an increasing number of vanishing moments provides increasing symmetry properties of the wavelet and the scaling function and results in a decreasing approximation error in (2.6) when the signal samples are used as the approximation coefficients of the initial space (Louis *et al.*, 1994). In comparison to the Daubechies' filters the Coiflet filters show approximately almost linear phase. For this reason they are very useful since they induce fewer phase distortions and therefore the center of the coefficients when interpreted in the time-frequency plane can be well estimated (Wickerhauser, 1996). Otherwise the stop band damping is slightly better in case of the Daubechies' filter which results in decreasing aliasing effects due to down-sampling, see fig. 3.

A Coiflet filter with length 30 was chosen in order to obtain a good trade-off between the support length of the wavelet and the localization properties in the frequency domain, see fig. 2. The scaling function and the wavelet function are orthogonal for integer translates k, $k \in \mathbf{Z}$. The frequency localization is quite acceptable.

By generating a wavelet packet dictionary out of Coiflets the maximum decomposition level has to be controlled because of the spreading of the packets in the frequency domain (Wickerhauser, 1994). Figure 4 shows a Coiflet packet basis function at level 5 which is badly localized in the frequency domain. Apparently, an interpretation as a single tile in the time-frequency plane is not given.

4. Best-basis algorithm

By decomposing a signal in a wavelet packet table we obtain the representations $v_{d,b}(k)$ of the signal in the over-complete collection of subspaces (sub-bands). d denotes the decomposition level and b counts the basis blocks from left to right (fig. 1). The table $v_{d,b}(k)$ contains a lot of redundancy. Therefore, an algorithm is needed to select an orthonormal basis which describes the whole L^2 and which finds a basis which is best adapted to the signal with respect to a cost functional. Because of the first two requirements, the sum of adjacent sub-bands must not have gaps or overlaps; see fig. 5a,b for two valid selections. Many well-known criteria can be used as information cost functional (Taswell, 1994). In the present approach Shannon's entropy cost functional

$$M(v) = \sum p(k) \log \frac{1}{p(k)}, \quad p(k) = \frac{|v(k)|^2}{\|v(k)\|^2}$$

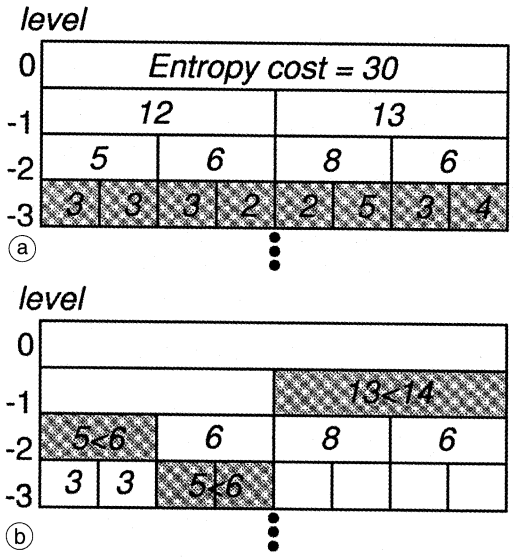

Fig. 5a,b. The cost functional of each basis block is calculated (number). The BBA starts at the bottom of the basis tree (a). From bottom to top the cost of the parent basis is compared with the sum of costs of the associated two children bases. b) The basis with the lower cost is chosen.

was used. For each basis block the entropy cost is calculated. The algorithm starts at the bottom (see fig. 5a,b) and compares the sum of entropy of the two children basis systems with the entropy cost of the parent basis. A minimal cost factor indicates a more peaked and therefore better fitting basis.

5. Time-frequency plane

To compute a time-frequency representation the bandwidth and time duration of a rectangular tile related to a coefficient has to be known. At a constant decomposition level each basis block (sub-band) is spanned by basis functions with constant bandwidth. Therefore, all coefficients in a basis block at level d and block number $b \in \{0, ..., 2^{-d} - 1\}$ are related to tiles which approximately cover the frequency interval

$$\left[\frac{b}{2^{1-d}} f_s, \frac{(b+1)}{2^{1-d}} f_s\right] \quad (f_s \text{ sampling-frequency}).$$

The initial space is spanned by orthonormal integer shifted Dirac impulses. To get a non-overlapping tiling along the time axis an interval of $1/f_s$ is chosen. Using (2.5a,b) we get the time intervals $[k - 2^{-d-1}/f_s, k + 2^{-d-1}/f_s]$ for a tile which corresponds to a coefficient at position k in a basis block at level d. The position k has to be corrected according to the phase distortion which occurred to the $v_{d,b}(k)$ after applying a cascade of operators H and G (Wickerhauser, 1996).

Each tile is mapped with a color related to its amplitude (black is highest) in the mentioned intervals. The resulting representation is called idealized time-frequency plane as we did not calculate the dimensions of the rectangular tiles belonging to the Coiflet used (Wickerhauser, 1996).

However, this interpretation of the coefficients does not take into account the misbehaviour of wavelet packets at higher levels in the frequency domain. Therefore we have to limit the maximum decomposition level to 5-6 for the filters used.

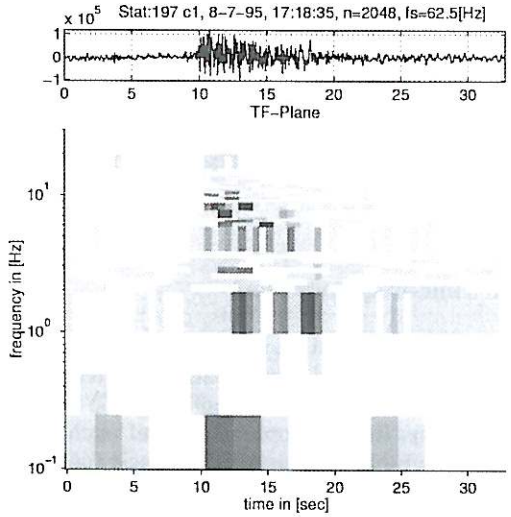

Fig. 6. BBA spectrogram of a shock event from Stromboli volcano. Most of the shock events consist of a deep frequent bump occurring during the onset of the signal.

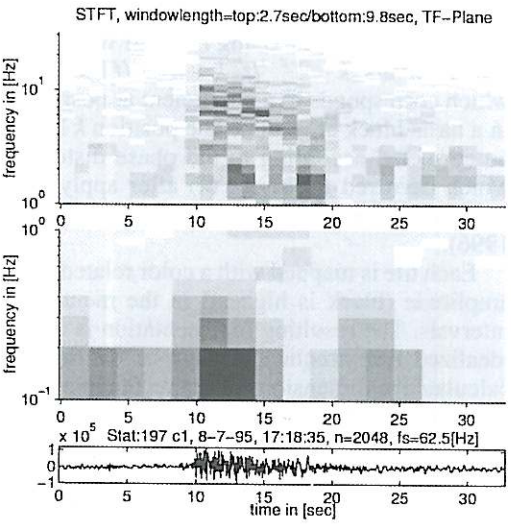

Fig. 7. STFT spectrograms computed with two different window lengths in order to get a good time resolution and to measure long period events.

6. Best-basis analysis of tremor signals

Mt. Stromboli is one of the best investigated volcanos. Its seismic signals possess a typical continuous tremor with spectral energy in several frequency bands within $1 < f < 10$ Hz and superimposed shock events with spectral content from $f << 0.5$ Hz up to Nyquist frequency.

For our investigations we used velocity data recorded by J. Neuberg during an array measurement with Guralp broadband seismometers on the Stromboli volcano in 1995.

Figure 6 shows a spectrogram computed with the BBA and fig. 7 shows two spectrograms computed with the STFT of a seismic signal of a duration of 33 s (2048 samples) which contains a shock event of Mt. Stromboli. The partitioning of the frequency axis in the case of the STFT is necessary since a short window is needed in the upper frequency band in order to get a good time resolution and a long window is needed in order to catch the deepest 10 s periodicity. The BBA inherently computes this kind of adaptation to the signal. Therefore, a better fit partitioning of the frequency axis in adjacent frequency intervals is obtained. When the signal changes its spectral properties in time significantly different algorithms like Matching Pursuit (Mallat and Zhang, 1993) or Atomic Pursuit (Chen *et al.*, 1996) should be used instead of BBA. They decompose the signal by recursively searching for the best fitting basis function instead of choosing a whole sub-band basis system. Consequently they do not fix the frequency resolution along the whole time axis. On the other hand, one can restrict the window length so that only one shock event is under consideration.

The BBA obviously results in a better decorrelated representation. The coarse time-frequency distribution is equal in both representations but the BBA gives a clearer pattern. Above 5 Hz the STFT reveals several spectral lines which appear quite less apparent in case of the BBA.

For studying the dominant sub-bands, individual frequency bands out of the BB spectrogram have been reconstructed in fig. 8. The Low-Pass (LP) trace on the bottom shows a bump shaped event with its maximum just on the very beginning of the shock. The Band-Pass

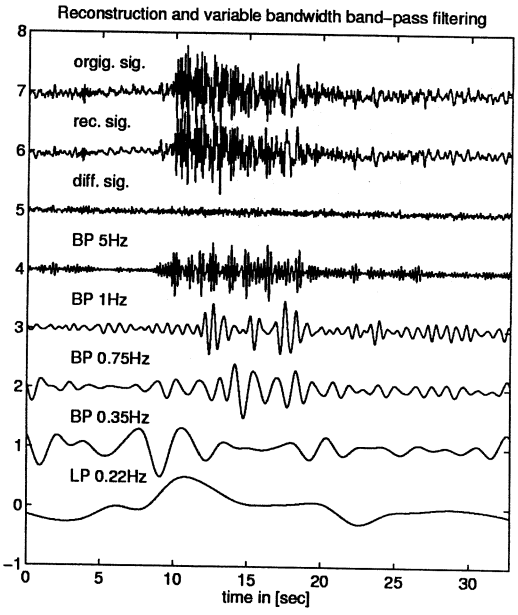

Fig. 8. From top to bottom the first three traces (identically scaled) show the original shock event, the reconstructed signal based on 310 coefficients, and the difference signal. The following four traces (scaled to maximum) show the reconstructed signal (without compression) based on coefficients which correspond to the center frequency of one band.

(BP) signal with center frequency 0.35 Hz consists of a precursor to the shock. The corresponding tile can clearly be recognized at the shock onset in fig. 6. In the higher frequency bands some short duration wave groups occur temporally during the shock (BP 1 Hz) while others exist for the whole duration (BP 5 Hz).

The analysis of several shock events revealed that most shocks consist of a deep-frequency bump shaped event during the onset of the shock and several dominant sub-bands above 1 Hz which consist of randomly distributed wave groups. The center frequencies of these sub-bands are stable during the occurrence of the shock but vary between different shock events.

As the BB basis is better adapted to the signal properties compared with the STFT basis it results in good compression ratios when only dominant coefficients are used for signal storage. Clipping all coefficients below a threshold such that 96% of the original signal energy is preserved we achieved the following results:

original: 2048 coefficients (100%);
reduced: 310 coefficients ($\hat{=}$ 15.14%).

The top three traces in fig. 8 contain from top to bottom the original signal, the de-noised signal after thresholding the coefficients and inverse transforming, and the difference signal.

7. Conclusions

The time-frequency analysis of broadband tremor by the BBA resulted in a characteristic and clearer representation compared with the standard STFT method. The computation time is similar to the STFT and therefore real-time computation for volcano monitoring and simultaneous compression and de-noising for data storage is possible. An efficient reconstruction algorithm allows the use of this algorithm as an adaptive bandwidth paraunitary filter bank.

One limitation is given by the maximum achievable frequency resolution because of frequency spreading of the wavelet packets. Another limitation is the finite time window for the investigation of the signal whereupon strong changes in the spectral content of the signal cannot be well measured by the BBA. Last but not least, the BB spectrograms are obtained by time-variant operations because of the down-sampling when computing the wavelet packet table. Therefore the sequence of minima and maxima within a sub-band might not correspond to the real energy distribution of the signal.

Acknowledgements

The authors thank J. Neuberg (Department of Earth Science, University of Leeds, U.K.) for the supplement of the Stromboli data as well as for his generous help in many discussions and the reviewers H. Kraus and F. Heinle (Telecommunications Institute, University Erlangen-

Nuremberg, Germany), whose careful efforts resulted in an substantially improved representation. Basic routines of the freely distributed Matlab toolbox WAVELAB .701 which is available at Stanford University, were used to calculate the Wavelet Packet Transformation and the BBA.

REFERENCES

CHEN, S., D. DONOHO and M. SAUNDERS (1996): Atomic decomposition by basis pursuit, *Technical Report,* IBM T.J. Watson Research Center.

COIFMAN, R. and M. WICKERHAUSER (1992): Entropy-based algorithms for Best Basis Selection, *IEEE Trans. Inf. Theory*, **38** (2), 713-718.

DAUBECHIES, I. (1992): *Ten Lectures on Wavelets,* Society for Industrial and Applied Mathematics, Philadelphia.

JAWERTH, B. and W. SWELDENS (1993): An overview of wavelet based multiresolution analysis, *Technical Report*, University of South Carolina.

LOUIS, A., P. MAAß and A. RIEDER (1994): *Wavelets* (Teubner Verlag, Stuttgart).

MALLAT, S. and Z. ZHANG (1993): Matching Pursuits with time-frequency dictionaries, *IEEE Trans. Signal Proces.*, **41**(12), 3397-3415.

STRANG, G. and T. NGUYEN (1996): *Wavelets and Filter Banks* (Wellesley-Cambridge Press), pp. 721.

TASWELL, C. (1994): Near-best basis selection algorithms with non-addaptive information cost functions, *Technical Report*, Scientific Computing and Computational Mathematics.

WICKERHAUSER, M. (1991): Lectures on wavelet packet algorithms, *Technical Report*, Department of Mathematics, Washington University.

WICKERHAUSER, M. (1994): Time localization techniques for wavelet transforms, *Technical Report*, Department of Mathematics, Washington University.

WICKERHAUSER, M. (1996): *Adaptive Wavelet-Analyse* (Vieweg Verlag, Braunschweig/Wiesbaden.