

Spectrum Skeletonization: A New Method for Acoustic Signal Feature Extraction

Tibor Dobján* and Gábor Németh†

Abstract

Vibration Analysis Tests (VAT) and Acoustic Emission tests (AE) are used in several industrial applications. Many of them perform analysis in the frequency domain. Peaks in the power density spectrum hold relevant information about acoustic events. In this paper we propose a novel method for feature extraction of vibration samples by analyzing the shape of their auto power spectrum density function. The approach uses skeletonization techniques in order to find the hierarchical structure of the spectral peaks. The proposed method can be applied as a preprocessing step for spectrum analysis of vibration signals.

Keywords: spectrum analysis, skeletonization, spectrum segmentation, feature extraction

1 Introduction

Acoustic events play an important role in several fields of industries [10, 25, 27, 30, 32]. Nowadays, vibration analysis tests are one of the applied methods for non-destructive testing. Vibration analysis tests are used to diagnose rotating machines [31] and for detection of hit and leakage [8]. With acoustic emission burst analysis we can draw conclusions about existing cracks and crack propagation [26].

These methods investigate acoustic samples by extracting some features of the vibration signal. There are three major methods for vibration signal feature extraction:

1. In time domain some statistical parameters (e.g., root mean square, mean, variance, skewness, kurtosis) can be computed [17, 27]. There are also time synchronous average signal (TSA) based methods, filter based methods, and stochastic methods [8, 28].

*College of Dunaujváros, Hungarian Acoustic and Industrial Diagnostic Testing Laboratory, E-mail: dobjan.tibor@email.duf.hu

†University of Szeged, Institute of Informatics, E-mail: gnemeth@inf.u-szeged.hu

2. Frequency domain analysis is used to investigate the frequency components of periodic signal samples [17,32,36]. Applications of frequency-domain analysis use Fast Fourier Transformation (FFT) [14].
3. Time-frequency domain analysis can be used for non-periodic and non-stationary signal samples, like bursts [17,36]. Applications of time-frequency domain analysis use continuous wavelet transforms (CWT) [23].

In this paper we propose a novel method for feature extraction from the auto power spectrum density (APSD) function. Our approach produces an automatic partitioning of a given APSD function along the frequency axis, where each range contains a dominant frequency component. Our method ensures that the bounds of frequency ranges, referred to as *separators* in the following, are defined within the local minima of the APSD function: by this way no relevant peak in the power spectrum density is split to disjunct frequency ranges. The peaks of the APSD function are detected by the skeletonization method [12,29]. Skeleton is a shape descriptor that summarizes the general form of objects. Since skeletons provide also structural properties of the power spectrum, each separator can be located between two skeletal branches.

The rest of this paper is organized as follows: Section 2 gives a short description about our used data. Section 3 details how to determine the relevant frequency ranges of the APSD function. Section 4 contains discussion about some possible improvements and connecting problems of our basic methods. The main conclusions of this document are summarized in Section 5.

2 Collection of recorded signals

Two datasets of acoustic events are examined in this paper. First, we have collected some acoustic emission event from Gleeble-measurements, however, the count of recorded signals was not large enough for a reliable evaluation, hence we also created a dataset of noises from knocks. Each event was registered manually.

2.1 Gleeble-measurements

In the Gleeble-laboratory of College of Dunaújváros we carried out a measurement by Gleeble 3800 thermo-mechanical physical simulator, which is able to perform tensile tests and hot deformation on steel specimens.

In the test we put a standard tensile test specimen into the jaws of Gleeble (see Figure 1). We placed two acoustic emission sensors at the end of the specimen behind the jaws. The specimen was then heated according to the heating profile (see Figure 2). It was heated up at 10 °C/s to 900 °C, and then held at temperature for 120 seconds. The specimen was subsequently controlled cool at 10 °C/s to approximately 400 °C and free cooled the rest of the way to room temperature.

During the test, we listened to acoustic emission events in the 1MHz frequency range ($F_s = 2\text{MHz}$). The measurement systems stored the whole time-series during

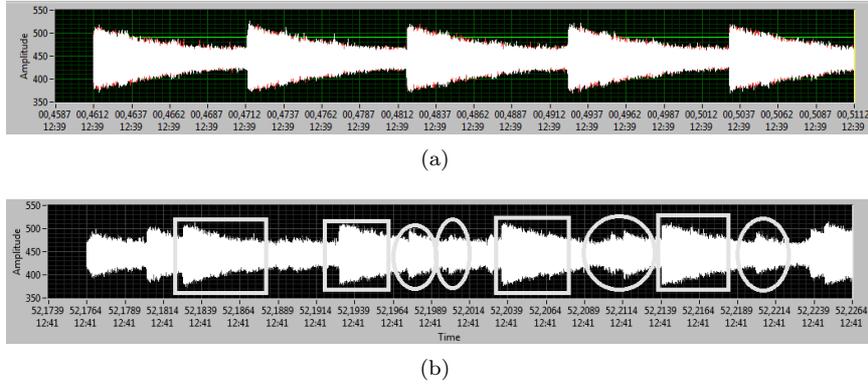


Figure 3: Time series of artificial periodic noise (a). Acoustic events on the cooling period (b). Signal segments marked by squares are periodic artificial, while the segments marked by ellipses are real acoustic events.

2.2 Knocking noises

The acoustic events were recorded by a general headset microphone to a single channel WAV file with 44100Hz sampling frequency ($F_s = 44100\text{Hz}$). As for gathering well-identified short time acoustic signal, we have dropped 6 objects having different shapes and materials into a china and a plastic dishes: toothpick (wood), match (wood), staple (metal), safety pin (metal), screw (metal), and nail (metal). Knocking these objects in the dishes indicates 12 strict classes of noises, but as for considering the material of objects, there are four classes taken into the account. Each of the 12 types of noises contains 100 recorded signals.

The recorded signals have various lengths, hence we had to cut the investigated signals for the same size (i.e., 1024 samples (see Figure 4)). These parts are cutted out around the first location where the signals reach their maximum amplitudes (see Figure 4). One may think that there is no difference between recorded signals, however there were various distances between the dish and the microphone, moreover the dropping height of the objects also differed during the recording.

3 Feature extraction from vibration signals

In this section, we will explain our proposed method for feature extraction.

Let $x(n)$ be a vibration signal containing n sampled data and the i^{th} value of $x(n)$ is denoted by x_i ($i = 1, \dots, n$).

Let us denote f_{max} the maximal frequency that is $\frac{F_s}{2}$, and let Δf the frequency resolution, where $\Delta f = \frac{f_{max}}{\frac{n}{2}}$. Furthermore, we denote the set of investigated frequencies by $\mathcal{F} = \{\Delta f, 2\Delta f, \dots, f_{max}\}$.

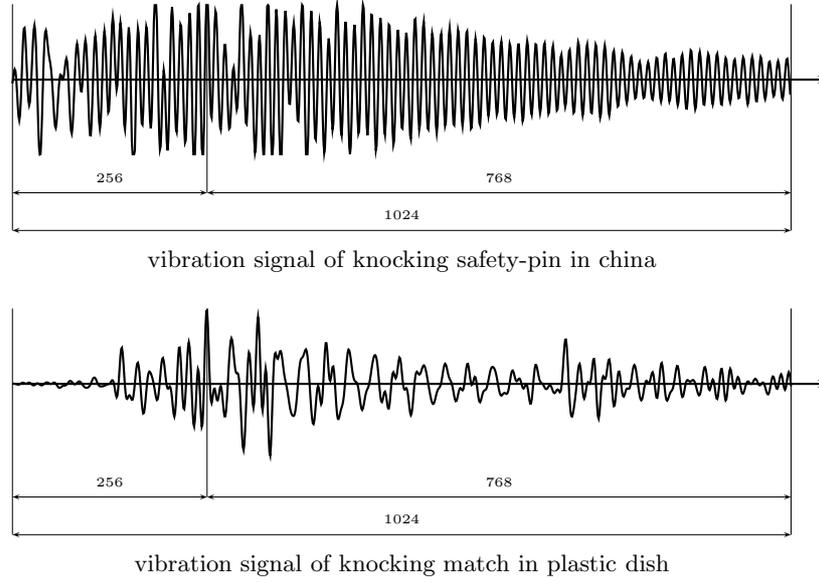


Figure 4: A various vibration signals using 1024 samples. The sampled data begins $1/4n$ before and ends $3/4n$ after location of the maximum peak amplitude (where n is the number of samples).

To compute the frequency spectrum we use Fast Fourier Transform (FFT). The APSD function of vibration signal $x(n)$ is computed by Equations (1) and (2) [14].

$$apsd = \log \left(\frac{1}{n} |FFT(x(n))| \right) \quad (1)$$

$$APSD(f) = \min(apsd) + apsd(i), \quad (2)$$

$$f = \Delta f \cdot i, \quad (i = 1, \dots, n)$$

We use only the positive half of the APSD function. Figure 5 shows an example for an APSD function.

Next, the APSD function of signal $x(n)$ is mapped into a quadratic binary image I_{APSD} :

$$I_{APSD}(u, v) = \begin{cases} 1 & \text{if } v \leq \left\lfloor \frac{n \cdot APSD(u \cdot \Delta f)}{2 \max_{i \in \mathcal{F}}(APSD(i))} \right\rfloor \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In I_{APSD} the background is white and black points represent the quantized values below the APSD function.

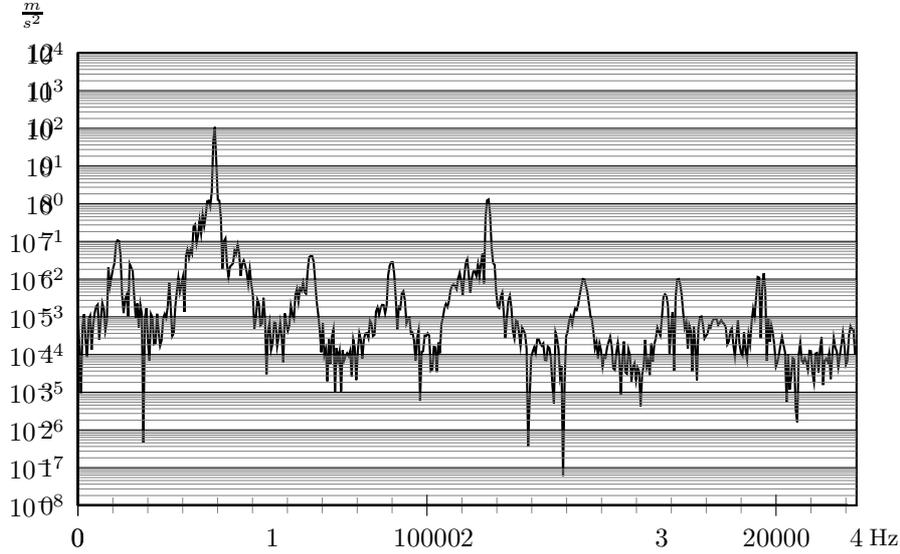
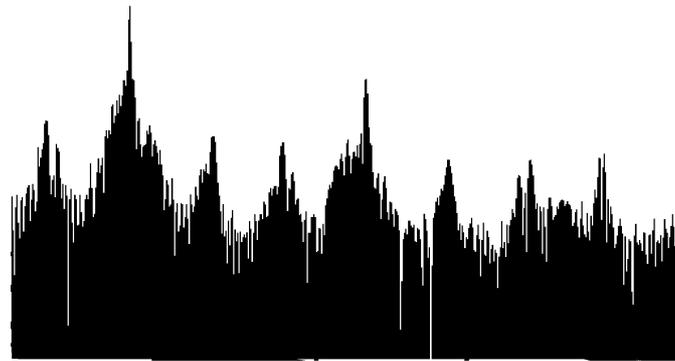


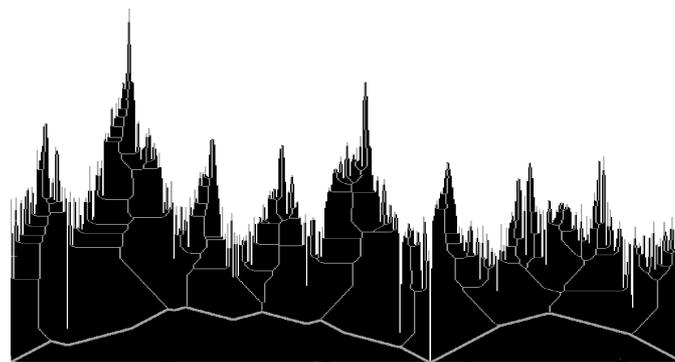
Figure 5: APSD function of a vibration signal

After we created the APSD function image I_{APSD} , the dominant peaks and frequency ranges have to be determined. First we extract the centerline of the APSD function image by a topology preserving skeletonization algorithm [12, 29]. Here we used a thinning algorithm [3, 6, 12, 13, 18, 19, 20] to extract the centerline, since it is the most efficient skeletonization technique for digital pictures. Skeletal branches are growing into dominant peaks of function. Furthermore, the centerline indicates the hierarchical structure of peaks as well.

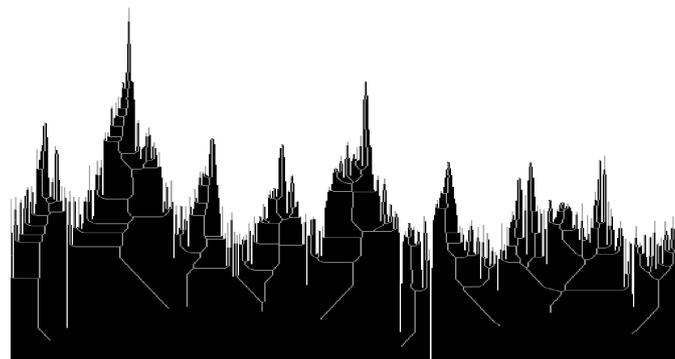
As a next step, the lowest connected skeletal curve segments (i.e., that are depicted with thick gray curve in Figure 6(b)) are removed. This curve segment is characterized as the set of skeletal points having the maximum y coordinate in each column in image I_{APSD} (if the $(0,0)$ point of I_{APSD} is its upper-left corner, and y coordinates increasing vertically down). This provides that the skeletal branches growing into the peaks of APSD function will be disconnected at their roots. Removal of the lowest curve segments disconnects only the skeletal branches into isolated *skeletal trees*, but does not affect the hierarchy of the side branches in the trees (see Figure 6(b) and (c)). The *width of a tree* can be determined by measuring the horizontal distance between the endpoints of the most left and most right branch of the tree. It is not hard to see that, the ranges (i.e., the widths) of the skeletal trees indicate disjunct intervals in the whole frequency range, however the union of these disjunct intervals does not form to whole frequency range. Our aim is to partition the whole frequency range into intervals by separators such that the union of the intervals matches the length of the whole range, with no peaks split by any of the separators.



(a) Spectrum image



(b) Spectrum skeleton (centerline)



(c) Spectrum skeletal trees

Figure 6: Spectrum is mapped to a binary image (a), centerline (b), and the skeletal trees (c) of the spectrum. Note that in (b) the lower skeletal branches (that ensures the connectedness) of the skeleton is removed. Hence only the skeletal trees are left.

Let $\mathcal{T}(1)$ and $\mathcal{T}(2)$ two adjacent skeletal trees. Furthermore, let us denote the most left and most right upper endpoint of tree \mathcal{T} by \mathcal{T}_l and \mathcal{T}_r , respectively. We propose that, the separator line between $\mathcal{T}(1)$ and $\mathcal{T}(2)$ is chosen at the frequency, where the APSD function reaches its minimum between $\mathcal{T}(1)_r$ and $\mathcal{T}(2)_l$ (see Figure 7). Our method is sketched in Algorithm 1.

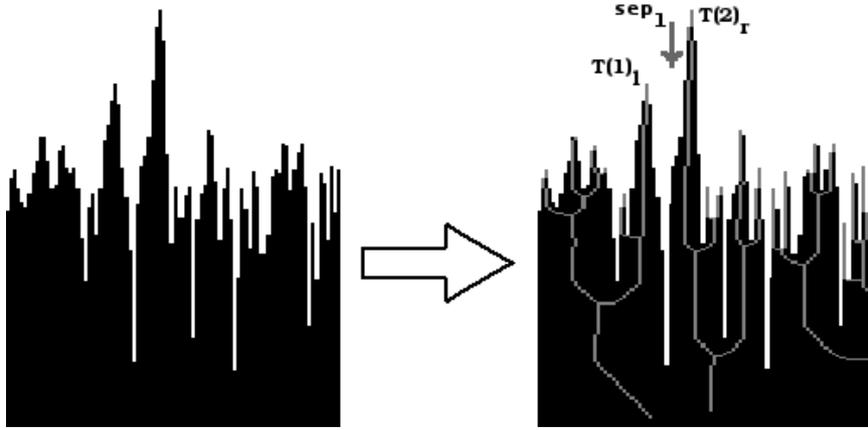


Figure 7: A part of the spectrum image is enlarged (left), and the location between the skeletal trees (right). The separator is drawn where the APSD function reaches its minimum between skeletal trees $\mathcal{T}(1)$ and $\mathcal{T}(2)$.

Algorithm 1 Partition the APSD function to dominant frequency ranges

Func PartitionAPSD(APSD(\mathcal{F}))

- 1: Map the APSD(\mathcal{F}) function to the image I . *Getting a binary picture about the APSD function*
 - 2: Extracting centerlines of image I .
 - 3: **for** $i = 1$ **to** $n/2$ **do** *Split centerlines into skeletal trees.*
 - 4: Delete (i.e., change it to white) the skeletal points with maximum y coordinate in column i .
 - 5: **end for**
 - 6: Let c be the count of skeletal tree. *Labeling is used.*
 - 7: Let the trees are labeled from 1 to c from left to right (i.e., the left most tree has Label 1, and the right most tree has Label c).
 - 8: **for** $i = 1$ **to** $c - 1$ **do** *Finding separators between two trees.*
 - 9: Find $sep_i \in [\mathcal{T}(i)_r, \mathcal{T}(i+1)_l]$, where $APSD([\mathcal{T}(i)_r, \mathcal{T}(i+1)_l])$ reaches its global minimum.
 - 10: **end for**
 - 11: **return** $\{sep_1, \dots, sep_{c-1}\}$
-

In the case of spectrum skeletons, the challenge is to match the appropriate branches to each other. In acoustic vibration signals, numerous of extra peaks may appear in the spectrum because of various acoustic noises. Thinning algorithms may produce many skeletal branches that are grown from the extremities of boundary points. Several points in a spectrum-image may fulfill the applied geometric constraints of the thinning algorithm, which means each peak may be represented by a skeletal side branch. The count and the position of the endpoints is strongly dependent on the variance of the spectrum. Hence, instead of the endpoints we focus of usage of skeletal components and the hierarchic structure of skeletal trees, which indicate the dominant spectrum segments.

In spectrum segmentation, conventional methods partition the frequency range equidistantly (see Figure 8(a)) [11,33]. Equidistant segmentation does not consider the shape of the spectrum which means that some peaks of the spectrum may be split during the segmentation. According to our concept, dominant peaks hold relevant information about the source of the event.

In our approach, where the frequency ranges are partitioned dynamically (see Figure 8(b)), each relevant peak is represented by a skeletal tree. The width of a frequency range is determined by the difference of local minima around the given skeletal tree. The APSD functions of recorded signals are always differ from each other because of the background noise of the signal. Hence there is no guarantee that skeletons will indicate the same number of branches and the same number of frequency domain partitions for each APSD function from the class of similar acoustic events. One may think that this noise produce several trees and small frequency ranges. Each noisy peak is represented by a skeletal side branch, but this does not dominant affect to the tree widths. With the help the dynamic partitioning can be described a given APSD function.

4 Further works

The proposed method holds some new ideas that point to a new way of spectrum segmentation. These could be a component of a complex acoustic event recognition system. The quality of event registration is important since it is the input of the segmentation algorithm. It may influence the results of feature extraction. Using skeletal structures we could represent the APSD spectra by graphs. We assume that graph (or sub-graph) isomorphism also can be applied in identification. For identification and clustering of acoustic events we use machine learning methods. What follows is an explanation of their planned use.

4.1 Auto event registration

Our feature extraction method is based on Fourier transformation. The Fourier transformation is not suitable for analysis transient acoustic events because it is suitable only for stationary time series. To get the best approximation, we should select the beginning and the end time of the event correctly, with following rules:

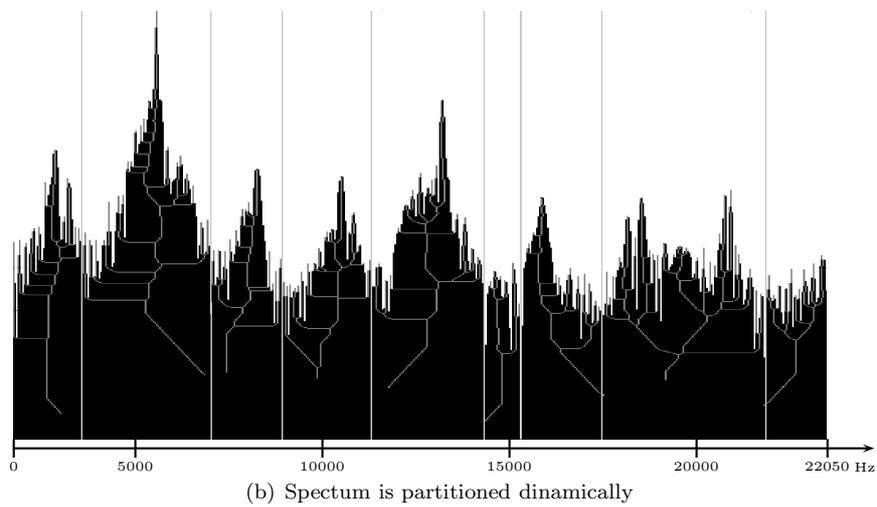
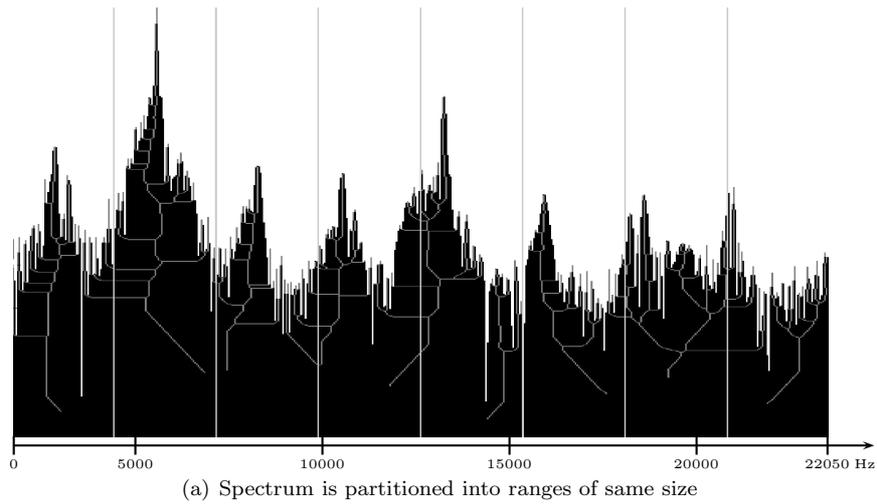


Figure 8: Regular (equidistant) partition of the frequency range (a). Dynamic partitioning using skeletal trees (b). Frequency axis can be computed by product the position of the separator by Δf .

- The selected piece of time series should contain only the event, with the least background noise. (There is always background noise during the event.)
- It should contain the entirety of the event

There are more solutions on the literature for this problem like Auto regression – Akaike Information Criterion (AR-AIC) [21] or Sequential Probability Ratio Test (SPRT) [8]. The frequently used approaches are the threshold-based algorithms.

4.2 Using graph theory

Since skeletonization methods produce reduced structures which represent the dominant peaks in the APSD functions, they are able to build a graphic structure. In future work we would like to build graphs from skeletal branches, and sub-graph isomorphism may help us to find out differences in structure. Since the area below the APSD function does not contain any cavity, this area can be represented by a tree graph derived from centerlines. This concept leads to comparison of graph (or sub-graph) structures for identification. We can assume that similar acoustic events have similar skeletal structures that can be identified by corresponding graph branches. Graph correspondences constructed by centerlines gave a solution for several applications [1, 7, 15, 24, 35].

4.3 Machine learning

In order to recognize the source of the signal, some machine learning algorithms are applied. Machine learning classifies signals based on some features (e.g., maximum amplitude, variance of the signal, or partial RMS rates) [4,5,9,22]. Machine learning is composed of two phases: in the learning phase, a set of sample data is assigned by class labels and their considered features are extracted; while in the second phase some unclassified data must be assigned to a class only based on their features. We investigated of usage the following classifiers in the future work: Linear classifiers determine a hyperplane in the feature space that separates data into two classes correctly. Support vector machines (SVM) [2, 16, 34] are kernel based classifiers that allow non-linear class borders for better accuracy. K -means classifiers are capable of distinguish more than two classes as well. Both of the linear and SVM classifiers are binary classifiers (i.e., they separate data into two disjunct sets), however there exists some extension of them to distinguish more than two classes. One of the concepts is *one-against-all*, while the other one is the *one-against-one* comparison [34].

5 Conclusion and open questions

In this paper we presented a novel method for spectrum segmentation based on skeletons. The proposed algorithm can be applied as a preprocessing step in spectrum description and analysis.

The APSD function computed from acoustic vibration signals are mapped to a binary image, where the background is represented by white pixels, and the area below the APSD function is formed by black pixels. The centerline of the APSD function image is extracted by a topology preserving thinning algorithm. The dominant peaks in the APSD function image are represented by skeletal branches which hold some structural information about the set of local maxima of APSD function. The extracted centerline is used for spectrum segmentation. Removal of the lowest segments of centerline splits the “skeleton” into skeletal trees. These skeletal trees indicate the main frequency components in the APSD function, however their union does not form the whole frequency range. Hence, we found the border of the frequency ranges in the minimum of function value in the frequency intervals between two skeletal trees. This fulfills our requirement to the effect that no peak in the spectrum is split into disjunct frequency segments. According to our concept, peaks in the APSD function hold relevant information about the origin of the acoustic signal.

We are working on some open questions in our further works:

- Thinning methods work on binary images, and quantization the APDS values yield some distortions in our data. Voronoi skeleton [12, 29] works with the original values, hence it can be a more appropriate solution for segmentation based on skeletons.
- In pattern recognition and shape analysis skeletons are converted to graphs, then it can be used for the formal description of the shape. In the further works we are focusing to the graph representation to describe similarity of APSD functions.

The proposed feature extraction by skeletons opens a new window for vibration signal processing and spectrum segmentation.

References

- [1] Bai, X. and Latecki, L. J. Path similarity skeleton graph matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(7):1282–1292, july 2008.
- [2] Ben-Hur, A. and Weston, J. A user’s guide to support vector machines. In Carugo, Oliviero, Eisenhaber, Frank, and Walker, John M., editors, *Data Mining Techniques for the Life Sciences*, volume 609 of *Methods in Molecular Biology*, pages 223–239. Humana Press, 2010.
- [3] Bertrand, G. and Couprie, M. Two-dimensional parallel thinning algorithms based on critical kernels. *Journal Mathematical Imaging and Vision*, 31(1):35–56, May 2008.
- [4] Bishop, C. M. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

- [5] Bishop, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [6] Couprie, M. Note on fifteen 2d parallel thinning algorithms. In *Internal Report, Université de Marne-la Vallée*, 2006.
- [7] Demirci, M. F. and Osmanlioglu, Y. Many-to-many matching under the l_1 norm. In *ICIAP*, pages 787–796, 2009.
- [8] Dobján, T., Pletl, Sz., Deák, T., Doszpod, L., and Pór, G. Identification of the place and materials of knocking objects in flow induced vibration. *Acta Cybernetica*, pages 53–67, 2010.
- [9] Duda, R. O., Hart, P. E., and Stork, D. G. *Pattern Classification*. Wiley, New York, 2. edition, 2001.
- [10] Eltabach, M., Vervaeke, T., Sieg-Zieba, S., Padioleau, E., and Berlingen, S. Features extraction using vibration signals for condition monitoring of lifting cranes. In *Proc. of International conference (Surveillance 6), Compiègne*, 2011.
- [11] Ghaderi, H. and Kabiri, P. Automobile independent fault detection based on acoustic emission using FFT. In *Singapore International NDT Conference & Exhibition (SINCE 2011)*, 2011.
- [12] Gonzalez, R. C. and Woods, R. E. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.
- [13] Guo, Z. and Hall, R. W. Parallel thinning with two-subiteration algorithms. *Commun. ACM*, 32(3):359–373, March 1989.
- [14] Heinzl, G., Rüdiger, A., and Schilling, R. Spectrum and spectral density estimation by the Discrete Fourier Transform (DFT), including a comprehensive list of window functions and some new at-top windows, 2002.
- [15] Hiransakolwong, N., Vu, K., Hua, K. A., and Lang, S.-D. Many-to-many skeletal-graphs matching approach to shape recognition. In *Proc. International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, 2004.
- [16] Hsu, C.-W., Chang, C.-C., and Lin, C.-J. A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University, 2003.
- [17] Jossa, I., Marschner, U., and Fischer, W.-J. *Signal-based feature extraction and SOM based dimension reduction in a vibration monitoring microsystem.*, pages 283–288. London: Springer, 2001.

- [18] Kardos, P., Németh, G., and Palágyi, K. An order-independent sequential thinning algorithm. In Wiederhold, P. and Barneva, R. P., editors, *IWCIA*, volume 5852 of *Lecture Notes in Computer Science*, pages 162–175. Springer, 2009.
- [19] Kardos, P. and Palágyi, K. Order-independent sequential thinning in arbitrary dimensions. In *Proc. IASTED International Conference on Signal and Image Processing and Applications, SIPA 2011*, pages 129–134, Crete, Greece, 2011.
- [20] Kong, T. Yung and Rosenfeld, A., editors. *Topological Algorithms for Digital Image Processing*. Elsevier Science Inc., New York, NY, USA, 1996.
- [21] Küperkoch, L. *Automated recognition, phase arrival time estimation and location of local and regional earthquakes*. PhD thesis, Ruhr-Universität, Bochum, Germany, 2010.
- [22] MacKay, D. J. C. *Information Theory, Inference & Learning Algorithms*. Cambridge University Press, New York, NY, USA, 2002.
- [23] Mallat, S. *A Wavelet Tour of Signal Processing, 2nd ed.* Academic, San Diego, CA, 1999.
- [24] Palágyi, K., Tschirren, J., Hoffman, E. A., and Sonka, M. Quantitative analysis of pulmonary airway tree structures. *Computers in Biology and Medicine*, 36(9):974–996, 2006.
- [25] Peng, Z., Chu, F., and He, Y. Vibration signal analysis and feature extraction based on reassigned wavelet scalogram. *Journal of Sound and Vibration*, 253(5):1087 – 1100, 2002.
- [26] Pór, G., Doszpod, L., and Dobján, T. Developing an acoustic emission measuring system based on modular high speed data acquisition devices. In *30th European Conference on Acoustic Emission Testing and 7th International Conference in Acoustic Emission*, September 2012.
- [27] Samanta, B. and Al-Balushi, K.R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical Systems and Signal Processing*, 17(2):317 – 328, 2003.
- [28] Schoonewelle, H., Hagen, T. H. J. J. Van Der, and Hoogenboom, J. E. A comparison of three time-domain anomaly detection methods. *Annals of Nuclear Energy*, 23(2):159–170, 1996.
- [29] Siddiqi, K. and Pizer, S. *Medial Representations: Mathematics, Algorithms and Applications*. Springer Publishing Company, Incorporated, 1st edition, 2008.
- [30] Su, H., Chong, K. T., and Ravi Kumar, R. Vibration signal analysis for electrical fault detection of induction machine using neural networks. *Neural Computing and Applications*, 20(2):183–194, March 2011.

- [31] Szántó, J. and Nagy, I. Integration of fault diagnostic technologies into a complex condition monitoring system and its practical results. In *World Congress on Maintenance (WCM-2008)*, pages 24–26, November 2008.
- [32] Tang, J., Chai, T.Y., Zhao, L.J., and Wen, Y. Feature extraction and selection based on vibration spectrum with application to mill load modeling. In *Advanced Control of Industrial Processes (ADCONIP), 2011 International Symposium on*, pages 266–271, may 2011.
- [33] Tsunodaa, T., Kato, T., Hirata, K., Sekido, Y., Sendai, K., Segawa, M., Yamatoku, S., Morioka, T., Sanoa, K., and Tsuneokaa, O. Studies on the loose part evaluation technique. *Progress in Nuclear Energy*, 15:569–576, 1985.
- [34] Weston, J. and Watkins, C. Support vector machines for multi-class pattern recognition. In *ESANN*, pages 219–224, 1999.
- [35] Xu, Y., Wang, B., Liu, W., and Bai, X. Skeleton graph matching based on critical points using path similarity. In Zha, Hongbin, Taniguchi, Rin-ichiro, and Maybank, Stephen, editors, *Computer Vision – ACCV 2009*, volume 5996 of *Lecture Notes in Computer Science*, pages 456–465. Springer Berlin / Heidelberg, 2010.
- [36] Yang, H., Mathew, J., and Ma, L. Vibration feature extraction techniques for fault diagnosis of rotating machinery : a literature survey. In *Asia-Pacific Vibration Conference*, pages 801–807, Gold Coast, Australia, 2003.