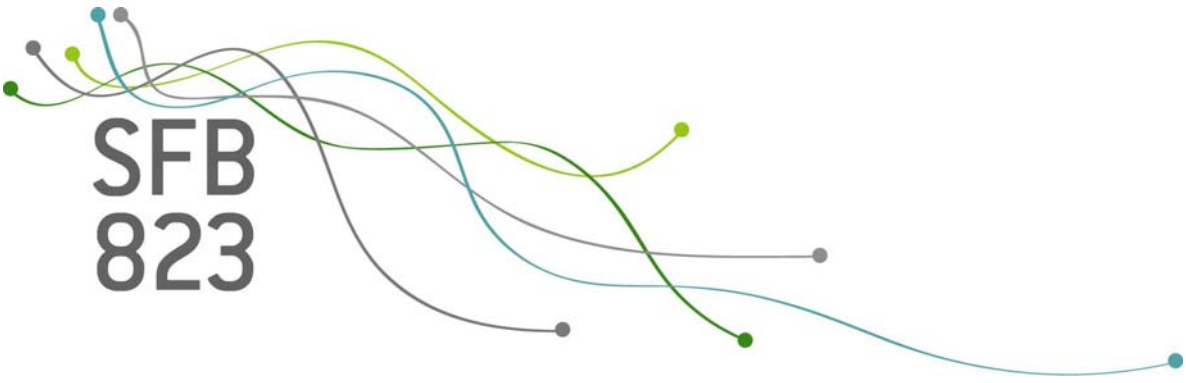
SFB 823

Discussion Paper

A new method for adaptive spectral complexity reduction of music signals

Ekaterina Krymova, Anil Nagathil,
Denis Belomestny, Rainer Martin

Nr. 69/2016



A new method for adaptive spectral complexity reduction of music signals

Ekaterina Krymova[1], Anil Nagathil[2],

Denis Belomestny[3], and Rainer Martin[4*]

[1,3]Universität Duisburg-Essen, Essen, Germany
[2,4]Institute of Communication Acoustics, Ruhr-Universität Bochum, Bochum, Germany
[1]ekaterina.krymova@uni-due.de,[2]anil.nagathil@rub.de,
[3]denis.belomestny@uni-due.de,[4]rainer.martin@rub.de

Abstract

In this discussion paper we present a novel unsupervised segmentation procedure for music signals which relies on an explained variance criterion in the eigenspace of the constant-Q spectral domain. The procedure is used in the context of a spectral complexity reduction method which mitigates effects of cochlear hearing loss. It is compared to a segmentation based on equidistant boundaries. The results demonstrate that the proposed segmentation procedure gives an improvement in terms of signal-to-artefacts ratio in comparison to a segmentation based on equidistant boundaries.

1 Introduction

Recently, in [1] a method was proposed which reduces the spectral complexity of music signals using principal component analysis (PCA) for listeners with cochlear hearing loss. This procedure requires a segmentation step which subdivides the signal into meaningful blocks in the time-frequency domain before PCA is performed. In [1] the segmentation was performed either using blocks of fixed length or by means of note onsets obtained from MIDI files. Although the former procedure does not require any prior information and can be applied in a real-world scenario, it does not take the underlying signal structure into account and thus is prone to smearing temporally important properties like note onsets. Hence, in this discussion paper we present a novel unsupervised

*This work is funded by the German Research Foundation (DFG), Collaborative Research Center 823, Subprojects B3 and C5.

segmentation procedure which detects change points in the eigenspace structure of the constant-Q spectral representation of music signals considering the degree of explained variance. Although statistical change point detection is a well-studied topic in the literature and numerous related methods have been proposed for general multivariate time series both in the time domain [2, 3] and spectral domain [4, 5, 6, 7, 8, 9], change point detection in the eigenspace of the time-frequency domain has not been studied before. The proposed method is evaluated in terms of signal quality measures such as the signal-to-interference ratio (SIR) and the signal-to-artefacts ratio (SAR) as proposed in [10].

2 Spectral Complexity Reduction

2.1 Spectral Analysis and Signal Reconstruction

We consider a discrete-time signal $x(n)$ with time index n, which is sampled at the sampling frequency f_s. It contains a leading voice signal $t(n)$ and an accompaniment signal $i(n)$ such that $x(n) = t(n) + i(n)$. For the spectral analysis we use the constant-Q transform (CQT) [11] which allows to adjust the center frequencies of frequency bins f_κ and their frequency spacing Δf_κ, with $\kappa = \{0, 1, \ldots, K-1\}$, to the geometric frequency distribution of notes $f_\kappa = f_0\, 2^{\frac{\kappa}{12b}}$ in western music. Here, f_0 denotes the frequency of the lowest note to be considered and b determines the number of frequencies describing a semitone. The CQT is defined such that the quality factor $Q = f_\kappa/\Delta f_\kappa = 1/(2^{\frac{1}{12b}} - 1)$ of a frequency bin is constant. Hence, the analysis length becomes frequency-dependent such that $N_\kappa = f_s/\Delta f_\kappa = Q f_s/f_\kappa$. The CQT is then computed by

$$X_{\mathrm{cqt}}(\kappa, \lambda) = \frac{1}{N_\kappa} \sum_{n \in \mathcal{N}_\kappa} x(n, \lambda)\, w_\kappa(n)\, \exp\left(-j \frac{2\pi Q n}{N_\kappa}\right), \tag{1}$$

where $x(n, \lambda) = x(n + \lambda B)$ is a signal frame of length $N_0 = Q f_s/f_0$ with segment index λ and segment shift B and κ denotes the CQT bin index. Furthermore, $w_\kappa(n)$ is an analysis window with a frequency-dependent length. The windows are chosen as Hann windows. Note that the analysis windows are aligned such that they all attain their maximum value for $n = N_0/2$, respectively. Outside of their support the windows attain values of zero.

The CQT in (1) is not directly invertible since the reconstruction of N_0 signal samples from their $K < N_0$ CQT coefficients constitutes an under-determined problem. However, using the method proposed in [12] we can restrict ourselves only to reconstruct a subset of L samples centred around $n = N_0/2$ which transforms the under-determined problem into a segmented overdetermined problem if $L < K$. This problem can then be solved by means of a least-squares procedure. The full-length signal can then be reconstructed using the overlap-add method.

2.2 Reduced-rank Approximations Based on PCA

In order to reduce the spectral complexity of the signal $x(n)$, block-wise reduced-rank approximations of the CQT representation are computed using PCA. To this end, the CQT-based time-frequency representation of the full signal is denoted by a matrix $\mathbf{X} \in \mathbb{C}^{N \times K}$, where N is the total number of frames. This matrix is then segmented into M non-overlapping blocks $\mathbf{U}^{(m)} \in \mathbb{C}^{B_m \times K}$ with $m = \{0, 1, \ldots, M-1\}$ and B_m being the number of frames contained in the m-th block. For notational convenience we will drop the index m in the following whenever possible.

PCA projects the CQT matrix $\mathbf{U}$ on a signal-dependent orthogonal basis such that it represents a high amount of the total variance in the matrix $\mathbf{U}$ by the first few dimensions of the transform space. To this end, PCA solves the eigenvalue problem $\mathbf{U}^{\mathrm{H}}\mathbf{U}\mathbf{w}_k = d_k\mathbf{w}_k$ where $k \in \{1, 2, ..., K\}$ denotes the index of principal components and $[\cdot]^{\mathrm{H}}$ is the Hermitian conjugate operator. It finds the eigenvalues d_k of the covariance matrix $\mathbf{C}_{uu} \sim \mathbf{U}^{\mathrm{H}}\mathbf{U}$, which correspond to the variance represented in the k-th dimension of the principal component space, and the corresponding eigenvectors $\mathbf{w}_k$. It is worth to note, that the eigenvectors describe the most prominent and most covarying spectral bands. These eigenvectors span an orthogonal basis $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_k, ..., \mathbf{w}_K]$ and the eigenvalues are sorted in descending order, i.e. $d_1 \geq d_2 \geq \ldots \geq d_K$, which ensures that the first principal components carry the highest percentage of total variance. Each block $\mathbf{U}$ can then be projected onto its basis $\mathbf{W}$ yielding the score representation

$$\mathbf{T} = \mathbf{U}\mathbf{W} \tag{2}$$

where $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_k, ..., \mathbf{t}_K]$ and $\mathbf{t}_k \in \mathbb{C}^{B_m \times 1}$ denotes the k-th coefficient vector. A dimensionality reduction can be performed by retaining a selected number of eigenvectors $\mathbf{w}_k$ with $k \in \widehat{\mathcal{K}} = \{1, 2, ..., \hat{k}\}$ and $\hat{k} \leq K$, which span a subspace $\widehat{\mathbf{W}} \subset \mathbf{W}$ resulting in the reduced coefficient representation

$$\widehat{\mathbf{T}} = \mathbf{U}\widehat{\mathbf{W}}. \tag{3}$$

Since $\mathbf{W}$ is a unitary matrix, a reduced-rank approximation of the original spectrogram block can be obtained by

$$\widehat{\mathbf{U}} = \widehat{\mathbf{T}}\widehat{\mathbf{W}}^{\mathrm{H}} = \mathbf{U}\widehat{\mathbf{W}}\widehat{\mathbf{W}}^{\mathrm{H}}. \tag{4}$$

Depending on the number of retained components such a reduced-rank approximation retrieves the most prominent and most temporally correlated harmonics of both the leading voice and the accompaniment and attenuates the low-variance spectral contributions. As demonstrated in [1] this reduces the spectral complexity of music signals which leads to a reduction of auditory distortion in the presence of cochlear hearing loss.

3 Proposed Segmentation Procedure

3.1 Explained variance ratio statistics

In applications the number of principal component to retain $\hat{k}$ is usually selected as the minimal number of the first principal components, for which the explained variance ratio is big enough (i.e. not less than preset value). The explained variance ratio for the $\hat{k}$ components is usually defined as

$$R(\hat{k}, \mathbf{U}) = \sum_{i=1}^{\hat{k}} d_i \Big/ \sum_{j=1}^{K} d_j. \tag{5}$$

The ratio (5) takes values in $[0, 1]$ and represents the proportion between explained variance of the data along the first $\hat{k}$ principal components and total variance in the data. Assume the number of principal component to retain is fixed and U contains observations from some probability distribution. Then one might consider $R(\hat{k}, \mathbf{U})$ as a measure of quality of data compression with $\hat{k}$ first principal components and use it for structural change point detection.

Suppose we are given a new sample data matrix $Z \in \mathbb{R}^{N \times K}$. In this case we will use the definition of explained variance ratio for ratio between the variance explained by $\hat{k}$ first eigenvectors of $\mathbf{U}$

$$R(\hat{k}, \mathbf{W}, Z) = \frac{\sum_{i=1}^{\hat{k}} \mathbf{w}_i^{\mathrm{H}} Z^{\mathrm{H}} Z \mathbf{w}_i}{\sum_{j=1}^{K} \mathbf{w}_j^{\mathrm{H}} Z^{\mathrm{H}} Z \mathbf{w}_j}. \tag{6}$$

To our knowledge explained variance ratio was not previously used for the change point detection in the spectral domain. It includes the information about the structure preserved in principal components and may be used for a segmentation procedure combined with dimension reduction by PCA.

Statistical properties of explained variance ratio could be derived based on the distribution of the eigenvalues. For example, following [13] the density of distribution of the eigenvalues $\bar{d}_1 \geq \cdots \geq \bar{d}_K$ of sample covariance for the matrix Z with i.i.d. rows from $\mathcal{N}(0, \Sigma)$ is

$$p_{(\bar{d}_1, \cdots, \bar{d}_K)}(l_1, \ldots, l_K) = \frac{\prod_{i=1}^{K} l_i^{(N-K-1)/2} \prod_{i<j} (l_i - l_j)}{C_{\Sigma, K, l_1, \ldots, l_K} \prod_{k=1}^{K} d_k^{K/2}}, \tag{7}$$

where for the case $\Sigma = \rho I_K$:

$$C_{\Sigma,K,l_1,\ldots,l_K} = \frac{2^{NK/2}\Gamma_K(K/2)\Gamma_K(N/2)}{\prod_{i=1}^{K}\exp(-l_i/2\rho)}. \tag{8}$$

3.2 Segmentation via explained variance ratio

Suppose the assumption that $\hat{k}$ principal components within the blocks contain the main information about the signal is fulfilled. Then if one breaks the block $\mathbf{U}$ into two sub-blocks $U_1 \in \mathbb{R}^{B_m^1 \times K}$, $U_2 \in \mathbb{R}^{B_m^2 \times K}$, $B_m^1 + B_m^2 = B_m$, it can be expected that

$$R(\hat{k}, \mathbf{W}_1, \mathbf{U}_1) - R(\hat{k}, \mathbf{W}_1, \mathbf{U}_2) < \delta, \tag{9}$$

where matrix $\mathbf{W}_1 \in \mathbb{R}^{K \times K}$ contains the principal component vectors of $\mathbf{U}_1$, and δ is a constant. The sub-block $\mathbf{U}_1$ should be big enough, namely $B^1 > \text{rank}(\mathbf{U}_1)$.

Denote by X_I, $I \subset \{1, \ldots, N\}$ a sub-matrix of X, which contain rows of X with indices in the set I. Inequality (9) defines a segmentation Algorithm 1, where parameter δ controls the number of resulting blocks. The procedure results in the set S of change points. The number of change points depends on the value of δ: increase of δ leads to smaller number of change points. There is no rule to select δ at the moment. Intuitively, the choice of δ also determines the delay in change point detection and it shouldn't be neither very small, nor too big. Also it should take into account the number of principal components to keep: for the fixed δ if one increases the number of principal components, then detection of change points becomes less frequent. The distribution of the left-hand side statistics (9) could be derived from the distribution of the data. For example, for a sample from the standard normal distribution one can use the common distribution of eigenvalues of covariance matrix (7), (8) to derive the distribution of (5) and (6). Then, from the derived distribution it could be possible to estimate δ, and to construct the confidence intervals for the reliability of estimation. In the Section 5 below we will show results of simulations for different values of δ. For a benchmark we use the segmentation with equidistant boundaries and a number of blocks equal to the number of change points in the proposed segmentation method.

3.3 An idea for δ selection

Suppose that $\mathbf{U}_1$ consists of N_1 Gaussian vectors from the distribution $\mathcal{N}(\mathbf{0}, C_1)$ and $\mathbf{U}_2$ consists of N_2 Gaussian vectors from the distribution $\mathcal{N}(\mathbf{0}, C_2)$. The covariance matrices C_1 and C_2 are supposed to be known. Apply SVD to the covariance matrices: $C_i = \mathbf{W}_i^\top D_i \mathbf{W}_i$, $i \in \{1, 2\}$, where D_i is a diagonal matrix with eigenvalues $d_{i,1}, \ldots, d_{i,K}$, $\mathbf{W}_i$ is a matrix with eigenvectors of C_i. Denote by $\mathbf{W}_{1,\hat{k}}$ a sub-matrix of $\mathbf{W}_1$, which consists of the first $\hat{k}$ columns. We are

Algorithm 1 Structural changes via explained variance ratio

Require: $\hat{k}$, δ;
  return S
  $i_{\text{break}} = 1$, $j = 1$, $S = \emptyset$,
  while $i_{\text{break}} + K/2 - 1 < N$ **do**
    $I = \{i_{\text{break}}, \ldots, i_{\text{break}} + K/2 - 1\}$,
    $\mathbf{U}_1 := X_I$, $\mathbf{U}_2 := \mathbf{U}_1$,
    Compute principal components $\mathbf{W}_1$ for $\mathbf{U}_1$,
    Compute $R(\hat{k}, \mathbf{W}_1, \mathbf{U}_1)$,
    $R(\hat{k}, \mathbf{W}_1, \mathbf{U}_2) = R(\hat{k}, \mathbf{W}_1, \mathbf{U}_1)$,
    while $R(\hat{k}, \mathbf{W}_1, \mathbf{U}_1) < R(\hat{k}, \mathbf{W}_1, \mathbf{U}_2) + \delta$ **do**
      $I := \{i_{\text{break}} + K/2 - 1 + j\}$,
      $\mathbf{U}_2 := X_I$,
      Compute $R(\hat{k}, \mathbf{W}_1, \mathbf{U}_2)$,
      $j = j + 1$,
    end while
    $i_{\text{break}} := i_{\text{break}} + K/2 - 1 + j - 1$, $j := 1$,
    $S = \{S \bigcup i_{\text{break}}\}$,
  end while

interested in solving a hypothesis testing problem

$$H_0 : \mathbf{U}_2 \text{ has columns from } \mathcal{N}(\mathbf{0}, C_1), \tag{10}$$
$$H_1 : \mathbf{U}_2 \text{ has columns from } \mathcal{N}(\mathbf{0}, C_2), \tag{11}$$

where

$$\frac{\text{tr}(W_{1,p}^\top C_2 W_{1,\hat{k}})}{\text{tr}(C_1)} - \frac{\text{tr}(W_{1,\hat{k}}^\top C_1 W_{1,\hat{k}})}{\text{tr}(C_1)} \leq -\delta. \tag{12}$$

To this end consider a statistics

$$S(\mathbf{U}_2) = \frac{\text{tr}(\mathbf{W}_{1,\hat{k}}^\top \mathbf{U}_2^\top \mathbf{U}_2 \mathbf{W}_{1,\hat{k}})}{N\text{tr}(C_1)} - \frac{\text{tr}(\mathbf{W}_{1,\hat{k}}^\top C_1 \mathbf{W}_{1,\hat{k}})}{\text{tr}(C_1)}. \tag{13}$$

One might use $S(\mathbf{U}_2)$ to check whether hypothesis H_0 is satisfied to estimate δ. Under H_0

$$S(\mathbf{U}_2) = \frac{1}{N_2} \sum_{i=1}^{\hat{k}} (\eta_i - \mathbf{E}\eta_i), \tag{14}$$

where η_i are $\Gamma\left(\frac{N_2}{2}, 2\frac{d_{1,i}}{\sum_{j=1}^K d_{1,j}}\right)$ random variables. Thus under H_0 the statistics $S(\mathbf{U}_2)$ has centered weighted χ^2-distribution. Then an estimate of δ might be found from the confidence bound with fixed confidence level α.

$$\mathbf{P}\big(S(\mathbf{U}_2) < -\delta \big| H_0\big) < \alpha. \tag{15}$$

The covariance matrix C_1 might be estimated from the first data matrix $\mathbf{U}_1$: $\hat{C}_1 = \frac{1}{N_1}\mathbf{U}_1^\top \mathbf{U}_1$.

The more correct way is to consider also the distribution of statistics $S(\mathbf{U}_2)$ under H_1. One can obtain

$$S(\mathbf{U}_2) = \frac{1}{N_2}\sum_{i=1}^{\hat{k}}(\zeta_i - \mathbf{E}\zeta_i) - \delta \tag{16}$$

where ζ_i, $i = 1, \ldots, \hat{k}$ are independent $\Gamma\left(\frac{N_2}{2}, 2\frac{d_{Z,i}}{\sum_{j=1}^{n} d_{1,j}}\right)$ random variables and $d_{Z,i}$ are the eigenvalues of the matrix $Z = D_2^{1/2}\mathbf{W}_2^\top \mathbf{W}_{1,\hat{k}}\mathbf{W}_{1,\hat{k}}^\top \mathbf{W}_2 D_2^{1/2}$. Thus under H_1 the statistics $S(\mathbf{U}_2)$ has weighted χ^2-distribution with a mathematical expectation $-\delta$. The estimate of δ could be obtained from the solution of hypothesis problem:

$$H_0: \ S(\mathbf{U}_2) \text{ has distribution (14) with mean 0,} \tag{17}$$

$$H_1: \ S(\mathbf{U}_2) \text{ has distribution (16) with mean } -\delta,\ \delta > 0. \tag{18}$$

Thus the problem consists in estimation of the mean of $S(\mathbf{U}_2)$. Unfortunately in general weighted χ^2-distribution has no known closed-form solution. For the δ estimation based on (15) on might use different approximation schemes (e.g. [14]). For the problem (17)-(18) the straight-forward calculation of the decision rule yet seems to be not feasible.

4 Evaluation

For the evaluation we used extracts of 110 synthesized MIDI files of chamber music pieces which were also used for evaluation in [1]. The resulting leading melody and accompaniment signal waveforms are sampled at $f_s = 16$ kHz, converted to mono signals, and mixed at an input SIR of 0 dB. To compensate for the spectral tilt towards higher frequencies, the music signals are fed to a first order pre-emphasis filter yielding the filtered signal $x_{\mathrm{f}}(n) = x(n) - 0.9x(n-1)$. The CQT (1) of each signal is computed for $f_0 = 55$ Hz, $f_{K-1} = 7040$ Hz, $b = 2$, $B = 32$, $L = 64$ which denote the minimal analysis frequency, the maximal analysis frequency, the number of CQT bins per note, the frame shift and the synthesis window length, respectively. Hann windows are used as analysis windows $w_\kappa(n)$ and as the synthesis window. Hence, in total we obtain 168 CQT bins corresponding to seven octaves. Note that this set of parameters provides a high-quality signal reconstruction from an unmodified CQT spectrum [12]. To reverse the effects of the pre-emphasis filter the reconstructed signal is fed to the corresponding first order de-emphasis filter.

Then we applied explained variance ratio Algorithm 1 and segmentation with equidistant boundaries to CQT-spectrum to obtain sets of non-overlapping blocks $\mathbf{U}^{(m)} \in \mathbb{C}^{B_m \times K}$ for PCA (see 2.2) and computed the reduced rank approximations. Similarly as in [1] we use the signal-to-inference ratio (SIR) and

the signal-to-artefacts ratio (SAR) [10] to assess the degree of accompaniment attenuation and leading voice distortion.

5 Results

5.1 Example of explained variance ratio segmentation

To illustrate explained variance ratio Algorithm 1 consider as an example a piece from Bach Siciliano for Oboe and Piano with spectrogram in Figure 1. In Figures 2 result of change point detection by Algorithm 1 is presented. One can see the correspondence between the stationary and non-stationary regions in the spectrogram and the segmentation result of the proposed algorithm.

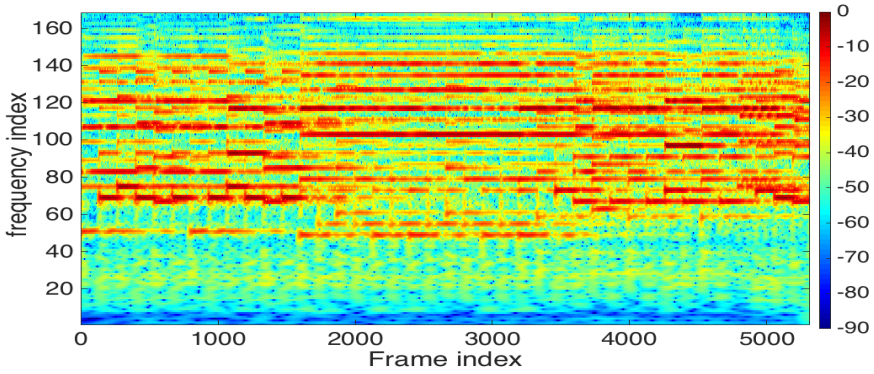


Figure 1: CQT-Spectrogram of Bach Siciliano for Oboe and Piano.

5.2 Comparison of segmentation methods

Using the experimental set-up defined in Section 4 we performed numerical comparison of the explained variance Algorithm 1 with segmentation with equidistant boundaries. We set the threshold to $\delta = 0.2$ and take the number of principal components to retain from the set $\{5, 10, 15, 20, 25, 30\}$. For each signal from database (see Section 4) we computed SIR and SAR measures. The results for the mean of SIR and SAR in the database for the explained variance Algorithm 1 (Alg.1) and the segmentation with equidistant boundaries (Equid.) is shown in Figure 3. There's no improvement in terms of SIR, but in the SAR axes one can see that explained variance ratio approach gives better results. For the fixed δ and number of principal components to retain and for different signals the number of resulting segments differ and depend on the variability of the input signal. Sometimes for the fixed number of principal components the selected value of δ is too high, so the delay in change point detection doesn't allow to trace structural changes. Our experiments show that for each signal there exists a certain value of δ which strikes a balance between a moderate number

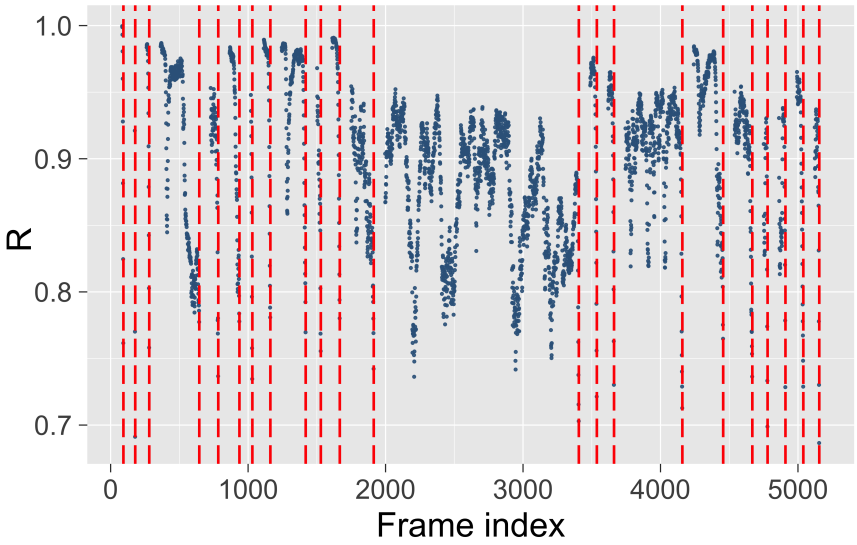


Figure 2: Explained variance ratio $R(\hat{k}, W_1, U_2)$ (blue dotted line) and the result of Algorithm 1 applied to CQT-transform of Bach Siciliano for Oboe and Piano, $\delta = 0.2$, number of principal components to retain equals 10. Red vertical lines denote detected change points.

of blocks and the quality improvement obtained after dimensionality reduction. Nevertheless, for the experiments reported in Figure 3, we have chosen to set δ to a fixed value of 0.2 which we found suitable on average.

6 Conclusions

The results indicate a reduction of artefacts of the leading voice compared to the procedure based on equidistant boundaries while the attenuation of the accompaniment remains unchanged. We argue that the proposed segmentation procedure improves the performance of a spectral complexity reduction scheme, which can further improve the quality of the processed signal and thus make these signals more accessible in the presence of a cochlear hearing loss. The demonstrated improvements were achieved with a fixed threshold δ, however, our experimental results indicate that further improvements are possible when δ is chosen adaptively.

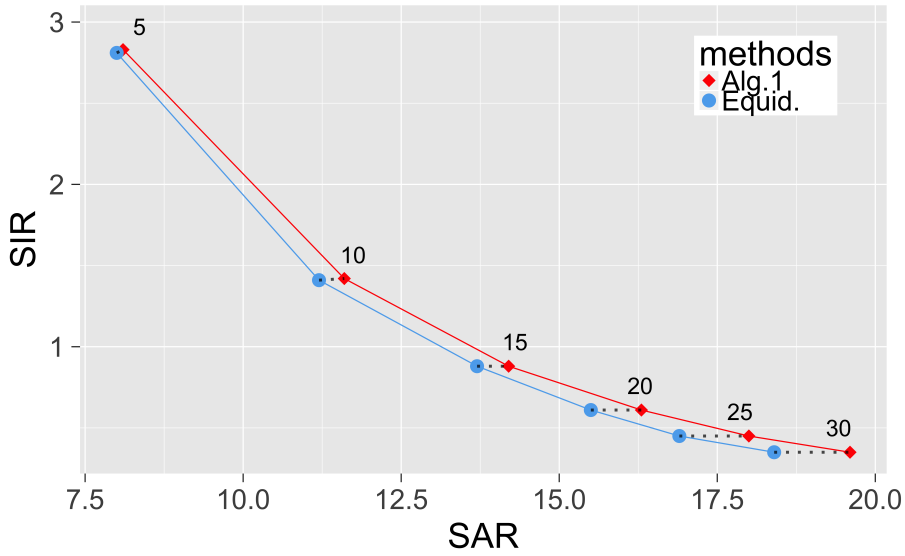


Figure 3: Results in case $\delta = 0.2$ for explained variance ratio method (Alg.1) and segmentation with equidistant boundaries (Equid.). Numbers indicate the number of principal components used in the methods. Dashed lines depict the correspondence between results of the methods: the number of blocks in equidistant boundary segmentation was taken equal to the number of resulting segments of Algorithm 1 for the fixed number of components.

References

[1] A. Nagathil, C. Weihs, and R. Martin, "Spectral Complexity Reduction of Music Signals for Mitigating Effects of Cochlear Hearing Loss," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 445–458, 2016.

[2] A. Aue, S. Hoermann, L. Horvath, and M. Reimherr, "Break detection in the covariance structure of multivariate time series model," *The Annals of Statistics*, vol. 37, pp. 4046–4087, 2009.

[3] C. Kirch and B. Muhsal, "A mosum procedure for the estimation of multiple change points," *Preprint KIT*, 2014.

[4] M. Lavielle and C. Ludena, "The multiple change-points problem for the spectral distribution," *Bernoulli*, vol. 6, no. 5, pp. 845–869, 2000.

[5] L. Giraitis and R. Leipus, "Functional clt for nonparametric estimates of the spectrum and change-point problem for a spectral function," *Lithuanian Mathematical Journal*, vol. 30, no. 4, pp. 302–322, 1990.

[6] E. Paparoditis, "Validating stationarity assumptions in time series analysis by rolling local periodograms," *Journal of the American Statistical Association*, vol. 105, no. 490, pp. 839–851, 2010.

[7] H. C. Ombao, J.A. Raz, R. von Sachs, and B. A. Malow, "Automatic statistical analysis of bivariate nonstationary time series," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 543–560, 2001.

[8] S. Adak, "Time-dependent spectral analysis of nonstationary time series," *Journal of the American Statistical Association*, vol. 93, no. 444, pp. 1488–1501, 1998.

[9] P. Preuss, R. Puchstein, and H. Dette, "Detection of multiple structural breaks in multivariate time series," *Journal of the American Statistical Association*, vol. 110, no. 510, pp. 654–668, 2015.

[10] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1462–1469, 2006.

[11] J. C. Brown, "Calculation of a constant Q spectral transform," *Journal of the Acoustical Society of America (JASA)*, vol. 89, no. 1, pp. 425–434, January 1991.

[12] A. Nagathil and R. Martin, "Optimal signal reconstruction from a constant-Q spectrum," in *Proc. Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012, pp. 349–352.

[13] Robb John Muirhead, *Aspects of multivariate statistical theory*, Wiley series in probability and mathematical statistics. Probability and mathematical statistics. John Wiley & Sons, New York, 1982.

[14] Antonia Castano-Martinez and Fernando Lopez-Blazquez, "Distribution of a sum of weighted central chi-square variables," *Communications in Statistics - Theory and Methods*, vol. 34, no. 3, pp. 515–524, 2005.