AN IMPROVED CEPSTRAL METHOD FOR DECONVOLUTION
OF SOURCE-FILTER SYSTEMS WITH DISCRETE SPECTRA:
APPLICATION TO MUSICAL SOUND SIGNALS

Thierry GALAS - Xavier RODET
LAFOREST - UA CNRS N°1095 - Laboratoire de Traitement de la Parole
Université Pierre & Marie Curie - 4, place Jussieu - 75222 PARIS CEDEX 05

Abstract: Most of music sound signals exhibit a discrete power spectrum, i.e. a
limited number of values only are of importance, which generally correspond to
harmonic or inharmonic partials. It is well known that source-filter modeling of such
signals leads to difficulties with usual methods such as AR modeling. This article
deals with a new method for modeling this class of signals. Many musical
applications are in view such as additive synthesis, spectral envelope estimation,
source-filter synthesis, etc....

1. Introduction

Source-filter models are quite useful in many domains, particularly when modeling sound
production systems. Identification of the parameters of such models is in fact a deconvolution
problem. There are two main classes of deconvolution methods, AR deconvolution and
homomorphic deconvolution. When applied for modeling sounds this problem has to be solved
using analysis by synthesis with an error criterion pertinent from a perceptual point of view. For
example, given X the power spectrum of the signal to be analyzed, S the power spectrum of the
hypothetical source, d(y,z) a spectral distance and the class C of the power spectra of considered
filters, then we search for the parameters of in C which minimize d(X,S,p).

AR deconvolution has already been examined from this point of view [Iwakura 68], and
improvements for treatment of discrete spectra have been proposed [El-Jamoudi 86][Gallas 89]. Let us
choose S(u)=1 for all u, d(Y,Z) the spectral distance defined by the quadratic error between the log
spectra and the class C of power spectra P(u) defined by:

\[ P(u) = \prod_{i=0}^{p} \cos(u_i) \]

Then we obtain the classical homomorphic deconvolution (cepsrum). When applied to
discrete spectra this method give erroneous results if the order of the model is not negligible in
front of the number of spectra peaks (as classical AR deconvolution also does). An improved
criteria method has already been proposed [Iwakura 79], but the algorithm is iterative and requires 2
FFT in each step. Our algorithm is direct and its computational cost is relatively low.

2. Discrete Cepstrum

We assume that the power spectra S and X are defined on the same discrete set \( \Omega = \{ \omega_n \mid n = 1 \text{ to } n \} \):

\[ S = \sum_{\omega} S_{\omega} \delta_{\omega, \omega_0} \] (2)

\[ X = \sum_{\omega} X_{\omega} \delta_{\omega, \omega_1} \] (3)

2.1 Spectral Envelope Domain

The spectral envelope domain which we consider, is the same as in the homomorphic case,
i.e. defined by expression (1).

2.2 Error Criterion

ICMC GLASGOW 1990 PROCEEDINGS
The error criterion we choose is the quadratic error between the log spectra with a spectral weighting factor \( h_i \) are strictly positive reals:

\[
E = \sum_{i=1}^{n} h_i \left( \log(y_i) - \log(x_i) \right)^2
\]  

(4)

This criterion is rather pertinent from a perceptual point of view.

2.3 Parameter Identification

By zeroing the partial derivatives of the error criterion we found that the identification of the \( c_i \) parameters can be obtained by solving the matrix equation \( A C = B \), where:

- \( A \) is the \( p+1 \) order matrix defined by:
  \[
  a_{k+1} = \sum_{j=1}^{n} h_j \cos(\omega_k j) \cos(\omega_k l)
  \]  
  (5)

- \( C \) is the vector of the searched parameters.

- \( B \) is the vector defined by:
  \[
  b_j = \sum_{k=1}^{n} h_k \log(x_k y_k) \cos(\omega_k l)
  \]  
  (6)

We compute \( A \) very efficiently by using an intermediate vector \( R \) defined as:

\[
r_{i+1} = \sum_{k=1}^{n} h_k \cos(\omega_k k) \]  
  (7)

then

\[
a_{i+1} = r_{i+1} - r_{i+1}
  \]  
  (8)

We solve this equation with the Cholesky algorithm.

2.4 Results

We estimate spectral peaks by some heuristics on a short term Fourier transform of the signal. We obtain better results with discrete cepstrum (Fig 2) than with classical cepstrum (Fig 1). But if we choose a too high order in comparison to the signal pitch then results are badly affected (see Fig 3).

3. Probabilistic Approach

One of the reasons for this phenomenon is that we assume a perfect knowledge of spectral peaks positions. The observation we perform by short term Fourier transform of the signal leads to some uncertainty. In consequence we replace the position \( (\omega_k, x_k) \) of every spectral peak by a probably distribution \( P_{\omega}(x, \omega) \).

3.1 New Error Criterion

We then replace the first error criterion by its mathematical expectation. Assuming \( s_i = h_i = 1 \) for all \( i \), we have:

\[
E = \sum_{i=1}^{n} \int P_{\omega}(x, \omega) \left( \log(y) - \log(x) \right)^2 \, dx \, d\omega
\]  

(10)

3.2 Parameters Identification

If \( P_{\omega}(x, \omega) \) has no particular properties, computation of the parameters can be done by using a sampling of \( P_{\omega}(x, \omega) \). Every spectral peak \( (\omega_k, x_k) \) is replace by a set of peaks \( (\omega_k, x_k) \) weights \( h_k = P_{\omega}(x_k, \omega_k) \). If we choose gaussian distributions for \( P_{\omega}(x, \omega) \):

\[
P_{\omega}(x, \omega) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-x_k)^2}{2\sigma^2}} e^{\frac{-\omega(\omega-a_k)^2}{2c}}
\]  

(11)
it is possible to compute explicitly the corresponding matrix equation $A.C=B$. The added computational cost of this procedure is of only 3.n supplementary products.

3.3 Results
An example of result is given on Fig 4 to be compared with Fig 3. We have used this algorithm to analyse high pitched women voices, with the parameters so obtained a new signal has been synthesized with pitch modifications. The high quality of resulting synthetic speech shows the good adaptation of our new method to the analysis of high pitched voices and sounds.

4. Conclusion
We have showed that homomorphic deconvolution can be applied to discrete spectra such as those of music sounds with high pitch. We have introduced a probabilistic approach which not only overcomes the difficulties which appear with high order models, but also allows for the introduction of more information or constraints on the spectral envelopes. These methods can be successfully applied to analysis and synthesis of high pitched sounds.

Fig 1: Spectrum (Pitch = 240 Hz)
Fig 2: Discrete Spectrum (Pitch = 240 Hz)
Fig 3: Discrete Spectrum (Pitch = 280 Hz)
Fig 4: Probabilistic Discrete Spectrum (Pitch=280Hz)

Bibliography