ITERATIVE METHOD FOR HARMONIC AND EXPONENTIALLY DAMPED SINUSOIDAL MODELS

Rémy Boyer and Julie Rosier

ENST, Department of Signal and Image Processing
46, rue Barrault, 75634 Paris Cedex 13
boyer,rosier@tsi.enst.fr

ABSTRACT

We present two evolutions of the well known Exponentially Damped Sinusoidal model, named HDS (Harmonic Exponentially Damped Sinusoidal) model for single pitch signals and HDSM for multi-pitch signals modeling. Additionally, we propose to use an iterative high-accuracy algorithm to estimate the model parameters. Compared with full or fast implementations of high-resolution methods (Matrix Pencil, ESPRIT, Kung's algorithm, ...), this approach has a lower computational complexity and is well adapted to the Harmonic Damped Sinusoidal models since it takes into account the harmonic relations between the angular-frequencies. In the context of low bit-rate audio coding, we show the validity of this approach on several audio signals.

1. INTRODUCTION

Audio modeling with non-stationary models knows a growing interest in the audio community. The EDS (Exponentially Damped Sinusoidal) model is, probably, the best example of these non-stationary models [1, 2, 3]. Moreover, we introduce the closely related Harmonic Damped Sinusoidal models for single-pitch (HDS) and for multi-pitch (HDSM) signals modeling. These models can be seen as a generalization of McAulay & Quatieri's sinusoidal model [4] in the sense that fast time variations of their amplitudes are allowed.

Traditionally, the EDS model is associated to a High-Resolution (HR) parameter estimation method, such as Matrix Pencil, ESPRIT or Kung's algorithm. These methods grouped under the generic designation of "subspace methods" [5] are very efficient in the context of audio modeling [6]. Unfortunately, they do not exploit the harmonic relations between the angular-frequencies and have a high complexity cost of orders. Consequently, these methods become ineffective when the length of the analysis segment is large. In this work, we present an alternative high-accuracy algorithm to estimate the H/EDS model parameters with a lower complexity. This approach is based on an iterative scheme, inspired by the RELAX algorithm. The latter was, in a first time, proposed in the context of the stationary sinusoidal modeling of general signals [7], SAR imaging [8] or audio signals [9]. Then, we extend this approach to the non-stationary H/EDS model case. The final method is a high-accuracy iterative algorithm of complexity order of $O(N \log_2 N)$ or $O(N)$ per iteration.

The paper is organized as follows : in section 2, we define the partial parametric model and the Analysis by Synthesis (A/S) architecture. In section 3, we give the definition of the EDS, HDS, and HDSM models. In sections 4 and 5, we describe our iterative algorithm. In section 6, we begin by showing an example of multi-pitch signals modeling with the harmonic model. After that, we compare the performance of the proposed method with the more costly Matrix Pencil (MP) algorithm by performing a "time-frequency" analysis based on a pseudo-QMF filter-bank architecture. In sections 7 and 8, we compare the algorithmic complexity of this new algorithm with the one of the MP algorithm and we point out the advantages and the drawbacks of the proposed method. Finally, we conclude in section 9.

2. PARAMETRIC MODELS AND ANALYSIS BY SYNTHESIS ARCHITECTURE

Let r be the analysis segment index. We define the general expression for the k-th *partial* parametric model according to

$$\hat{x}_k(n,r) \triangleq \sum_{m=1}^{M_k} A_{m,k}(n,r) \cos\left(\Phi_{m,k}(n,r)\right) \tag{1}$$

where $A_{m,k}(n,r)$ is the generalized time-varying amplitude, M_k is the modeling order, *i.e.*, the number of sinusoids of the k-th partial model. $\Phi_{m,k}(n,r)$ is the generalized time-varying phase. After that, we give the expression of a parametric model

$$\hat{x}(n,r) \triangleq \sum_{k=0}^{K-1} \hat{x}_k(n,r) \tag{2}$$

where K is the number[1] of partial parametric models. The modeling total order is $M = \sum_k M_k$. The analysis process is done through a uniform time segmentation of the audio signal according to $x(n,r) = x(n)h_a(n - rD)$ where $h_a(n)$ is a rectangular analysis window, D is the stride length and the synthesis is made by using a Hanning window $h_s(n)$ as $\hat{x}(n) = \sum_{r=0}^{R-1} \hat{x}(n,r)h_s(n - rD)$ where R is the number of analysis segment. Note that this Overlap and Add (OLA) technique is an Analysis by Synthesis (A/S) architecture which verifies the perfect reconstruction conditions, *i.e.*, $\sum_{r=0}^{R-1} h_s(n - rD) = 1$.

3. DEFINITION OF THE H/EDS MODELS

The choice of the terms $\{A_{m,k}(n,r), \Phi_{m,k}(n,r)\}$ characterizes the model. By omitting the segment index, we can give the following definitions of the EDS model, the HDS model in case of single-pitch signals and the HDSM model in case of multi-pitch signals, according to

[1] possibly $K = 1$.

model label	K	$A_{m,k}(n)$	$\Phi_{m,k}(n)$
EDS	$=1$	$a_m e^{d_m n}$	$\omega_m n + \phi_m$
HDS	$=1$	$a_m e^{d_m n}$	$\omega_0 mn + \phi_m$
HDSM	>1	$a_{m,k} e^{d_{m,k} n}$	$\omega_k mn + \phi_{m,k}$

where $\{\omega_k\}_{0\leq k\leq K-1}$ is the set of fundamental angular-frequencies and we have to satisfy $M_k\omega_k \leq \pi$. We, also, denote, respectively, $\{a_{m,k}, \phi_{m,k}, d_{m,k}\}$ the m-th amplitude, phase and real damping-factor parameters of the k-th partial model. The EDS, HDS and HDSM models have all exponentially time-varying amplitude and the relations bending the angular-frequencies determine the choice of the model. Note that for the simplicity of the notations, we systematically omitted the index k when $K = 1$, *i.e.*, for the EDS and HDS models.
The EDS model is a more general model since it allows any kind of relation between the angular-frequencies. The HDS model fixes a harmonic relation for a single fundamental angular-frequency such as $\omega_m = m\omega_0$. Finally, the HDSM extends this approach by considering multiple fundamental angular-frequencies, according to $\omega_{m,k} = m\omega_k$. This model can be seen as the sum of K HDS models.
The EDS model can be used in the context where no information on the audio signal $x(n)$ is known. The HDS model is well adapted to the single-pitch speech modeling and finally, the HDSM is dedicated to the multi-pitch speech (multi-speaker) signals modeling and to the representation of multiple harmonic musical instruments.
The complex plane for HDS poles (see figure 1-a) and for HDSM poles (see figure 1-b) are represented on figure 1. The pole is defined by $z_{m,k} = e^{d_{m,k}+i\omega_{m,k}}$. Note that due to the varying-amplitude, the modulus of the pole is not limited to the unitary circle. Indeed, we have $|z_{m,k}| = e^{2d_{m,k}}$ and $\arg(z_{m,k}) = \omega_k m$. For the EDS model, the poles can take any value in the complex plane.

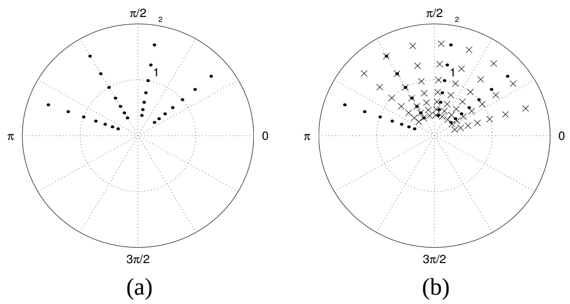


Figure 1: *complex plane (for several damping-factor values) (a) HDS poles with $\omega_0 = 0.7$ (rad) and $M_0 = 4$, (b) HDSM poles with $K = 2$, $\omega_0 = 0.7$, $\omega_1 = 0.3$, $M_0 = 4$ and $M_1 = 8$.*

Note that our approach is quite different of the one in [10] since we consider for each harmonic angular-frequency a damping-factor parameter. This assumption is more consistent with the real structure of speech and audio signals. Moreover, in [10], only the single-pitch case is considered.

4. ITERATIVE METHOD FOR THE EDS, HDS AND HDSM MODELS

Our approach is based on an iterative M-order decomposition of the audio signal $x(n)$. We introduce the residual audio signal $x_\ell(n)$ at the ℓ-th iteration, according to

$$\begin{aligned} x_\ell(n) &\triangleq x_{\ell-1}(n) - a_\ell e^{d_\ell n}\cos(\omega_\ell n + \phi_\ell) \\ &= x(n) - \sum_{j=1}^{\ell} a_j e^{d_j n}\cos(\omega_j n + \phi_j). \end{aligned} \quad (3)$$

The last expression is obtained by initializing the first residual to the audio signal. Note that there exists a mapping between ℓ and the couple (m, k) according to $\ell = kM_k + m$. Consequently, we have $1 \leq \ell \leq M$.
We want to minimize with respect to the model parameters, the following Non-Linear Least-Square (NLLS) criterion of the residual audio signal, given by

$$\varepsilon_\ell = \sum_{n=0}^{N-1} |x_\ell(n)|^2. \quad (4)$$

The previous criterion can be seen as one ℓ-dimensional NLLS criterion or as ℓ 1-dimensional NLLS criterions. Then, we give the algorithmic description of the method :

Initialization : $x_0(n) \triangleq x(n)$ [audio signal]
iterate ℓ

(1) **Analysis** of the residual audio signal $x_\ell(n)$ according to expression (3).

(2) According to the methodology of section 5, solve the following criterion :

$$\arg\min_{a_\ell, d_\ell, \phi_\ell, \omega_\ell} \varepsilon_\ell \quad (5)$$

(3) **Synthesis** of the modeled signal $\hat{x}_\ell(n)$ according to

$$\hat{x}_\ell(n) = \hat{x}_{\ell-1}(n) + a_\ell e^{d_\ell n}\cos(\omega_\ell n + \phi_\ell). \quad (6)$$

For the EDS model[2] the method is exactly summed up in the previous algorithm. For the HDS model, we perform an estimation of the fundamental angular-frequency at the first iteration, after that, we deduce the set of harmonic angular-frequencies for a fixed modeled order. Consequently, in the subsequent iterations, we estimate only the model parameters $\{a_\ell, d_\ell, \phi_\ell\}$. In the case where $K > 1$, *i.e.*, for the HDSM model, we iterate K times the algorithm dedicated to the HDS model.

5. MODEL PARAMETERS ESTIMATION

5.1. Angular-frequency estimation

5.1.1. EDS model case

In the case of the EDS model, we use the following strategy to solve the one dimensional NLLS criterion (4). First, we begin by estimating the angular-frequency ω_m according to

[2] *i.e.*, $K = 1$ and $\ell = m$.

$$\omega_\ell = \arg \max_{\lambda \in [0,\pi]} X_\ell(\lambda) \tag{7}$$

where $X_\ell(\lambda)$ is the Short-Time Fourier Transform (STFT) of the ℓ-th zero-padded residual audio signal. One should make sure that the selected component is a valid spectral peak [11].

5.1.2. HDS and HDSM models case

For HDS and HDSM models, we exploit the harmonic relations between angular-frequencies : the k-th partial model is featured by its fundamental angular-frequency ω_k. A first estimation stage is then introduced to estimate the fundamental angular-frequencies.
A consequent work on pitch estimation methods has been developed over three decades and one can find several effective single-pitch estimation methods for HDS modeling in [12].
When $K > 1$, multi-pitch estimation methods can be used to obtain the fundamental angular-frequencies of each partial model. Two approaches can be considered : first, one can use single-pitch estimation methods, dedicated to the HDS model and iterated K times [14]. Second, multi-pitch estimation method can be exploited to jointly estimate the fundamental angular-frequencies set [13]. However, we can point out an important defect of pitch estimation methods, *i.e.*, pitch doubling or pitch halving values can occur. Thus, one can use an approach based on the MDL (Minimum Description Length) as described in [15].

5.2. Damping-factor estimation

After that, we estimate the damping-factor d_ℓ by the shifted-STFT method. This method uses the ratio of the modulus of two STFT segments of the same length but one shifted from the other. Consequently, we have :

$$d_\ell = \frac{1}{P} \ln \frac{\left|X_\ell^{(1)}(\omega_\ell)\right|}{\left|X_\ell^{(0)}(\omega_\ell)\right|} \tag{8}$$

where $X_\ell^{(0)}(\omega)$ and $X_\ell^{(1)}(\omega)$ are the respective STFTs of the signals :

$$\begin{array}{rcll} x_\ell^{(0)}(n) & = & x_\ell(n)w(n) & \text{for} \quad n = 0, \ldots, N-P-1 \\ x_\ell^{(1)}(n) & = & x_\ell(n)w(n-P) & \text{for} \quad n = P, \ldots, N-1 \end{array} \tag{9}$$

where P is a time offset chosen small with respect to the analysis duration. Note that we only calculate the Fourier transforms for the considered ω_ℓ. Consequently, the complexity is not the STFT one but only the product of two $(N-P)$ dimensional vectors.
$w(n)$ is a weighting window which is designed for isolating the pole from its conjugate [16]. We choose here a Blackman's window [17]. We sum up the method in figure 2.

Finally, we can note that in the reference [18], the authors propose to determine each damping-factor by using a 1-D optimization method. Then, the global algorithmic complexity is generally higher than the STFT one.

5.3. Complex amplitude estimation

The $N \times 2$ Vandermonde matrix can then be deduced from the pole z_ℓ :

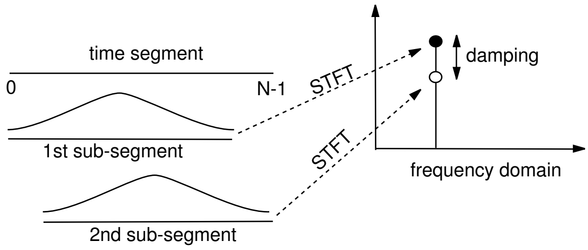


Figure 2: *Principle of the damping-factor estimation.*

$$\boldsymbol{V}_\ell = \begin{pmatrix} 1 & 1 \\ z_\ell & z_\ell^* \\ \vdots & \vdots \\ z_\ell^{N-1} & z_\ell^{(N-1)*} \end{pmatrix} \tag{10}$$

and the following criterion can be solved :

$$\arg \min_{\boldsymbol{\alpha}_\ell} \|\boldsymbol{x}_\ell - \boldsymbol{V}_\ell \boldsymbol{\alpha}_\ell\|_2^2 \iff \boldsymbol{\alpha}_\ell = \boldsymbol{V}_\ell^\dagger \boldsymbol{x}_\ell \tag{11}$$

where † denotes the pseudo-inverse, $\boldsymbol{x}_\ell = (x_\ell(0) \ldots x_\ell(N-1))^T$ and $\boldsymbol{\alpha}_\ell = (a_\ell e^{i\phi_\ell}/2 \;\; a_\ell e^{-i\phi_\ell}/2)^T$. Note that the pseudo-inverse of $\boldsymbol{V}_\ell$ exists if and only if the Hermitian matrix $\boldsymbol{V}_\ell^H \boldsymbol{V}_\ell$ is full rank. If we make the assumption that $z_\ell \in \mathbb{C}$, *i.e.*, $\omega_\ell \neq 0$, then the rank of $\boldsymbol{V}_\ell$ is 2. Thanks to rank conservation, $\boldsymbol{V}_\ell^H \boldsymbol{V}_\ell$ is also a 2-rank matrix and is square of dimension 2. We conclude that the pseudo-inverse of $\boldsymbol{V}_\ell$ exists under the assumption that the pole is complex. In the audio compression context, the zero angular-frequency component is not important and is systematically eliminated at the synthesis stage. Let us indicate that [19] points out that estimating $\boldsymbol{\alpha}_\ell$ according to the expression (11) is asymptotically consistent (for N great enough).

5.4. Modified RELAX algorithm for the EDS model

The previous algorithm can be improved by using the RELAX method [7]. The basic three step RELAX procedure for the first component estimation could be briefly described in the following manner :

step 1 Assume $\ell = 1$. Estimate from $x(n)$, the one-component signal $\hat{x}_1(n)$, *i.e.*, the model parameters : $\{\omega_1, d_1, a_1, \phi_1\}$.

step 2 Build the first residual $x_1(n)$ and estimate the one-component signal $\hat{x}_2(n)$, *i.e.*, the model parameters : $\{\omega_2, d_2, a_2, \phi_2\}$.
From the signal $x(n) - \hat{x}_2(n)$, redetermine the following model parameter set $\{\omega_1, d_1, a_1, \phi_1\}$. Iterate the previous operations until "practical convergence" [7].

step 3 Compute $\hat{x}_3(n)$ from the residual signal $x_2(n)$, *i.e.*, estimate $\{\omega_3, d_3, a_3, \phi_3\}$.
From the signal $x(n) - \hat{x}_2(n) - \hat{x}_3(n)$, redetermine the following model parameter set $\{\omega_1, d_1, a_1, \phi_1\}$. Iterate the previous operations until "practical convergence".

$\longrightarrow$ Estimate the second component ($\ell = 2$), *i.e.*, the model parameter set $\{\omega_2, d_2, a_2, \phi_2\}$.

The remaining operations consist in continuing this three step process until the model order is reached. It is well known, that this method improves the angular-frequency resolution of the Fourier-like methods [7, 18, 20] and, thus, improves also the estimation of the other model parameters.

6. SIMULATIONS

6.1. HDS and HDSM models

The HDSM model is used on a mixing of two voices (see figure 3-a). On figure 4-α, the log-spectra of the two previous signals and their modeled versions (see figures 3-b) are represented. The autocorrelation method [12] has been used to estimate the fundamental angular-frequency in the single-pitch case. For the two-pitch case, we have exploited the same method in an iterative scheme context. The log-spectra of the two partial models are represented on figures 4-β-a,b. The fundamental angular-frequency trajectories are plotted on figure 3-c. We choose the autocorrelation method for two reasons. Firstly, it is a high-resolution method and secondly, its computational cost is quite moderate.

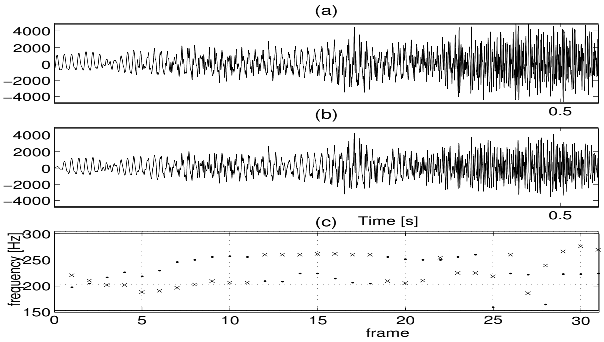
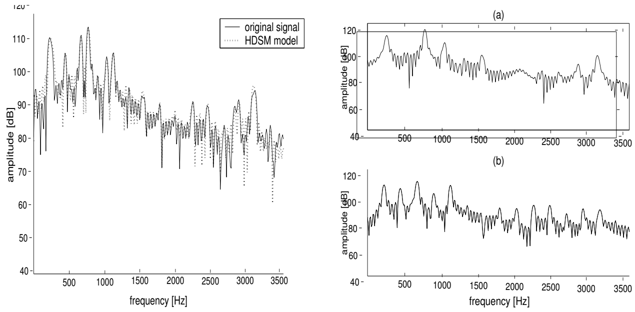


Figure 3: *HDSM model, (a) original two-pitch signal, (b) modeled time-waveform $\hat{x}(n)$ with $M_0 = M_1 = 20$, (c) fundamental angular-frequency trajectories.*

Finally, one can see the harmonic structure of the test signals and the great spectral and temporal matching of the modeled signals.

6.2. Time-frequency analysis by filter bank approach

In this section, we show through two examples of Glockenspiel and Harpsichord signals (see figures 5-a and 6-a) the effectiveness of the proposed method by comparison with the Matrix Pencil (MP) algorithm. The classic SNR criterion is consistent with exact temporal modeling of signal but it does not reflect the spectral behavior of the considered models. Introducing a spectral aspect in the analysis, we use the polyphase 32-band pseudo-QMF filterbank $\{h_b(n)\}_{1\leq b\leq 32}$ of MPEG1-audio [21] providing a uniform partition of the frequency axis. The bandwidth of each subband is 500 Hz with a 32 kHz sampling frequency. In each subband, we use the criterion SNR to characterize the temporal modeling performance. The final criterion can thus be reformulated as



(α) (β)

Figure 4: *HDSM model,* (α) *log-spectrum of the original* $x(n)$ *and of the modeled signal* $\hat{x}(n)$*,* (β) *(a) log-spectrum of the first partial model* $\hat{x}_0(n)$*, (b) log-spectrum of the second partial model* $\hat{x}_1(n)$*.*

$$\text{SNR}_{TF}^{(r,b)} = 10\log_{10}\sum_{n=0}^{N-1}\frac{|h_b(n)*x(n,r)|^2}{|h_b(n)*(x(n,r)-\hat{x}(n,r))|^2}\ [\text{dB}] \tag{12}$$

Note that this representation is to be related to power levels computation in each subband. In fact, $\text{SNR}_{TF}^{(r,b)}$ interpretation in weak power subband b is not significant as $\text{SNR}_{TF}^{(r,b)}$ derived from a high power subband in terms of considered model performance analysis.

According to figures 5-b,c and for the Glockenspiel signal of figure 5-a, we can note the great modeling obtained by the EDS model with the two tested estimation methods. Due to the use of short windows (8 ms), we can note the absence of pre-echo phenomenon[3] (see figures 5-b,c).

In the light of figures 7-b, we can say that the modified RELAX algorithm presents better performance in the regions where we have the higher subband powers (figure 7-a) *i.e.* for $b \leq 5$. On the other hand, the MP method presents a better modeling in higher subband, especially for $b = \{11; 12; 17\}$. In other subbands, the power can be considered negligible, in the context of low bit-rate audio modeling.

In the second example, on figure 6, we choose a Harpsichord signal. Contrary to the Glockenspiel signal, the two methods show their defects in the high-frequency domain (see figure 7-c for $b \geq 22$ *i.e.* for frequencies higher than 11 kHz) and the modeled signal spectrograms on figures 6-b,c. However, we can see on figure 7-d, that the two methods present very close performances. Indeed, we have less than 2dB between the two approaches. These conclusions are confirmed by informal listenings.

7. ALGORITHMIC COMPLEXITY ANALYSIS

7.1. Complexity order of the proposed method

In the context of the proposed method and for the EDS model, we have a complexity order of $O(N\log_2 N)$ per iteration for the determination of the angular-frequency by the STFT. The complexity order for the estimation of the damping-factor is $O(N)$.

[3] additional energy before the onset.

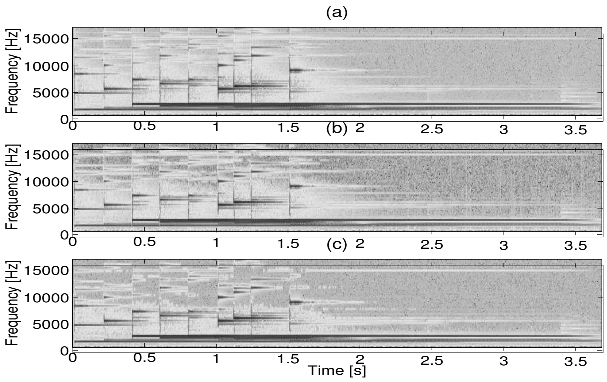


Figure 5: *EDS modeling, Spectrograms, (a) Glockenspiel signal, (b) modeled signal with the MP algorithm ($M = 30$) (c) modeled signal with the modified RELAX algorithm ($M = 30$).*

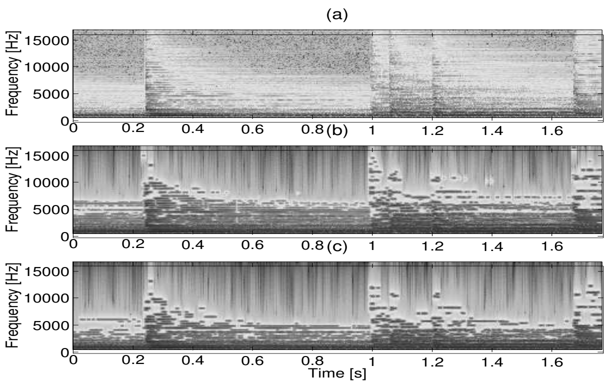


Figure 6: *EDS modeling, Spectrograms, (a) Harpsichord signal, (b) modeled signal with the MP algorithm ($M = 30$) (c) modeled signal with the modified RELAX algorithm ($M = 30$).*

The pseudo-inversion of a thin matrix for the complex amplitude estimation is $O(N)$. Then, the complexity per iteration is dominated by the STFT, *i.e.*, $O(N \log_2 N)$. In the case of using the RELAX approach, we reiterate this process few times. Consequently, we can neglect it. For M iterations, the total complexity is $O(MN \log_2 N)$.
For the HDS model, we perform a fundamental angular-frequency estimation based on the autocorrelation method which presents a complexity of $O(N)$ and, finally, M damping-factors and complex amplitudes estimations. Then, the total complexity is $O(N)$ plus $O(MN)$ which is dominated by $O(MN)$.
The HDSM model makes the assumption that $K > 1$. In this case, the modeling order is defined by $\sum_{k>1} M_k$ and the total complexity is $O(N \sum_{k>1} M_k)$.

7.2. Full implementation of a HR method

The Matrix Pencil algorithm has a complexity dominated by the cost of the full implementation of the SVD (Singular Value Decomposition) [26], *i.e.*, $O(N^3)$. In relation with subsection 7.1, we can conclude that the proposed algorithm presents a strongly reduced computational cost.

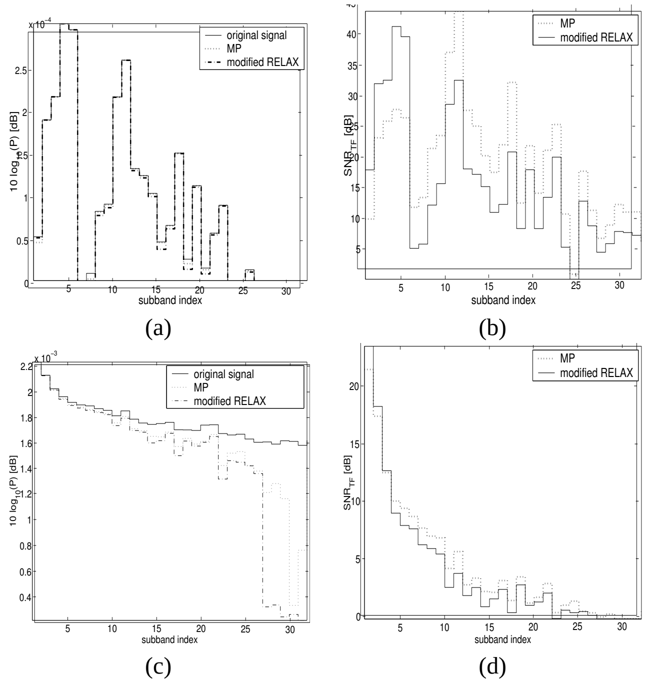


Figure 7: *Glockenspiel signal (a) Power by subband, (b) SNR$_{TF}$, Harpsichord signal , (c) Power by subband, (d) SNR$_{TF}$.*

7.3. Fast implementation of a HR method

In [22, 23], we present a fast and a very fast processing of the "signal part" of the SVD. This approach, in its fast version, has a total complexity order of $O(N^2M)$ if we use only the Orthogonal Iteration algorithm. A very fast version is obtained by adding a fast technique to compute the products between a vector and a structured (Hankel or Toeplitz) matrix. The complexity order is, then, $O(NM^2)$.
We conclude that the total complexity of the proposed new method is attractive in comparison with fast implementation of the SVD.
In the case of very fast implementation, the proposed method remains more economic if the model order verifies $M \gg \log_2 N$ in the EDS model case. Note that this constraint is realistic in the context of low bit-rate audio coding. In the case of HDS and HDSM models, the new method has always a lower computational cost.

8. ADVANTAGES AND DRAWBACKS OF THE PROPOSED METHOD

We have to point out the limitations of the proposed method. Firstly, in the sharp transient audio modeling context, *i.e.*, when we have to deal with very localized time events (*e.g.* castanet onset), it remains preferable to use a HR method (see [24, 25]). Secondly, in the frequency tracking context (with short analysis segments) a HR method has proven its efficiency [22]. Finally, note that HR methods remain interesting in the context of high-quality audio analysis [6].

However, in the context of compact audio representation (low bit-rate coding), *i.e.*, for a given order M, we choose a large analysis range N, the proposed method is well adapted and provides very satisfactory results for a reduced computational cost. For this reason, we have chosen this approach for the quasi-stationary signals and ”soft” transients representation in the context of a parametric audio (speech and music) coder, designed for a target bit-rate lower than 30 kbits and for a 32 kHz sampling frequency [3].

9. CONCLUSION

In this communication, we have presented two modifications of the EDS model which are dedicated to harmonic signals. These models are named HDS for Harmonic Damped Sinusoidal model for single-pitch signals and HDSM for Harmonic Damped Sinusoidal model for multiple-pitch signals. In addition, we propose to use for the three models a high-accuracy iterative algorithm to estimate the model parameters. After that, we present in the simulation part and in the context of compact audio representation (low bit-rate audio coding), a ”time-frequency” analysis by means of a filter-bank architecture which tends to show that this method provides similar performance to the more costly Matrix Pencil method.

10. REFERENCES

[1] P. Lemmerling, I. Dologlou and S. Van Huffel, ”Speech Compression based on exact modeling and Structured Total Least Norm optimization”, *Proc. of IEEE Int. Conf. on Acoustic, Speech and Signal Processing (ICASSP 98)*, May 1998.

[2] J. Nieuwenhuijse, R. Heusdens and E.F. Deprettere, ”Robust Exponential Modeling of Audio Signal”, *Proc. of IEEE Int. Conf. on Acoustic, Speech and Signal Processing (ICASSP 98)*, May 1998.

[3] R. Boyer and S. Essid, ”Decomposition, Representation and Coding of Audio Signals based on the Exponentially Damped Sinuoidal (EDS) Model”, EURASIP Journal on Applied Signal Processing (JASP), *submitted*.

[4] R.J. McAulay and T.F. Quatieri, ”Speech analysis & synthesis based on a sinusoidal representation”, *IEEE Trans. on ASSP*, Vol. 34, No. 4, August 1986.

[5] A-J. Van Der Veen, ED. F. Deprettere and A. Lee Swindlehurst, “Subspace-Based Signal Analysis Using Singular Value Decomposition”, *Proc. of the IEEE*, Vol. 81, No 9, September 1993.

[6] J. Laroche, ”The use of the matrix pencil method for the spectrum analysis of musical signal”, *J. Acoust. Soc. Am.*, 94(4), October 1994.

[7] J. Li, P. Stoica, ”Efficient mixed-spectrum estimation with applications to target feature extraction”, *IEEE Trans. on Signal Processing*, Vol. 44 Issue: 2, Feb. 1996.

[8] J. Li, P. Stoica, Z. Bi, R. Wu and E. G. Zenio, ”A robust Hybrid Spectral Estimation Algorithm for SAR imaging”, *Proc of IEEE Asilomar conference*, Vol. 2 , 1998.

[9] K. Wang, H. Ge, and Y. Ding, ”Adaptive Parametric Schemes for Analysis and Synthesis of Musical Signals”, *J. Audio Eng. Soc.*, Vol 49, No 5, May 2001.

[10] J. Jensen, S.H. Jensen and E. Hansen, ”Harmonic exponential modeling of transitional speech segment”, *Proc. of IEEE Int. Conf. on Acoustic, Speech and Signal Processing (ICASSP 2000)*, 2000.

[11] X. Serra, *A system for sound analysis/transformation/synthesis based on a detreministic plus stochastic decomposition*, Ph.D. Thesis, Stanford University (CCRMA), October 1989.

[12] W. Hess, *Pitch determination of speech signals*, Springer-verlag, 1983.

[13] R.C. Maher, ”An Approach for the Separation of Voices in Composite Musical Signal”, *J. Acoust. Soc. Am.*, vol. 6, june 1989.

[14] T. Virtanen and A. Klapuri, ”Separation of Harmonic Sounds Using Multipitch Analysis and Iterative Parameter Estimation”, *Proc. of IEEE WASPAA*, 2001.

[15] J. Rosier and Y. Grenier, ”Pitch Estimation for the Separation of Musical Sounds”, *Proc. of the 112th Conv. of the AES*, may 2002.

[16] P. O’Shea, ”The use of sliding spectral windows for parameter estimation in power system disturbance monitoring”, *IEEE Trans. on PES*, 2000.

[17] B. Porat *A course in Digital Signal Proecessing*, John Wiley & Sons, New York, 1997.

[18] Z-S. Liu, J. Li, P. Stoica, ”RELAX-based estimation of damped sinusoidal signal parameters”, *Signal Processing*, 62, 1997.

[19] P. Stoica, H. Li and J. Li, ”Amplitude Estimation of Sinusoidal Signals : Survey, New Results, and an Application”, *IEEE Trans. on SP*, Vol. 48, No. 2, February 2000.

[20] I. Santamaria, C. Pantaleon and Jesus Ibanez, ”A comparative study of high-accuracy frequency estimation methods”, *Mechanical Systems and Signal Processing*, 14(0), Academic Press, 2000.

[21] K. Banderburg and G. Stoll, ”ISO-MPEG-1 Audio : a generic standard for coding of high-quality digital audio”, *JASA*, Vol. 42, October 1994.

[22] R. Badeau, R. Boyer and B. David, ”EDS parametric modeling and tracking of audio signals”, *Proc. of 5th Int. conf. on Digital Audio Effects (DAFx02)*, September 2002.

[23] R. Boyer, ”Fast algorithm and non-stationary model for high-resolution audio signal modeling”, ENST internal report : `http:/www.tsi.enst.fr/∼boyer/`, 2002.

[24] R. Boyer, S. Essid and N. Moreau, ”Dynamic temporal segmentation in parametric non-stationary modeling for percussive musical signals”, *Proc. of IEEE Int. Conf. on Multimedia and Expo (ICME 02)*, August 2002.

[25] R Boyer and K. Abed-Meraim, ”Efficient parametric modeling of audio transients”, *Proc. of 5th Int. conf. on Digital Audio Effects (DAFx02)*, September 2002.

[26] G.H. Golub and C.F. Van Loan, *Matrix Computation*, North Oxford Academic, Oxford, second edition, 1983.