Calculation of an entropy-constrained quantizer for exponentially damped sinudoids parameters
Olivier Derrien, Roland Badeau, Gaël Richard

To cite this version:
Olivier Derrien, Roland Badeau, Gaël Richard. Calculation of an entropy-constrained quantizer for exponentially damped sinudoids parameters. OR 20. Technical report, 5 pages. 2010. <hal-00490864>

HAL Id: hal-00490864
https://hal.archives-ouvertes.fr/hal-00490864
Submitted on 9 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Calculation of an Entropy-Constrained Quantizer for Exponentially Damped Sinusoids Parameters

Olivier Derrien, Roland Badeau and Gaël Richard

Abstract

The Exponentially Damped Sinusoids (EDS) model can efficiently represent real-world audio signals. In the context of low bit rate parametric audio coding, the EDS model could bring a significant improvement over classical sinusoidal models. The inclusion of an additional damping parameter calls for a specific quantization scheme. In this report, we describe a new joint-scalar quantization scheme for EDS parameters in high resolution hypothesis, which is much easier to implement than a vector quantization scheme. A performance evaluation of this quantizer in comparison with a 3-dimensional vector quantizer is proposed in a paper submitted to IEEE Signal Processing Letters named "Entropy-Constrained Quantization of Exponentially Damped Sinusoids Parameters".

Index Terms

Parametric audio coding, Exponentially Damped Sinusoids, High resolution, Quantization, Entropy

A. INTRODUCTION

For low bit rate music coding applications, parametric coders are an efficient alternative to transform coders. As a consequence, the interest in parametric audio coding has grown during the last years. Sinusoidal modeling is very popular because most real-world audio signals are dominated by tonal components. In most sinusoidal analysis/synthesis schemes used for parametric coding, sinusoids have constant amplitude over each analysis/synthesis time segment. Sinusoid parameters (amplitude, frequency and phase) are quantized and binary coded. However, some studies have proved that an exponentially damped sinusoidal model (EDS) combined with a variable-length time segmentation is more efficient than a constant-amplitude model. In this report, we describe a joint-scalar quantization method for amplitude, damping and phase parameters. Optimizing the quantizer consists of minimizing the mean distortion under a bit rate constraint. As modern communication techniques commonly use variable-length binary codes, we choose to formulate the bit rate constraint in terms of entropy of quantization indexes. In high resolution hypothesis, i.e. assuming a large number of quantization cells, quantizers are usually defined by their quantization cell density (QCD). The calculation of the optimal QCD is divided in three parts: first, we calculate the distortion between two exponentially damped sinusoids. Then, assuming dependencies between the quantization of amplitudes, damping and phases, we calculate the mean distortion generated by the quantization process. Finally, we obtain the QCD that minimizes the mean distortion under the entropy constraint.

B. THE EDS MODEL AND THE MEAN SQUARE ERROR DISTORTION MEASURE

The EDS modeling of a signal \( x(t), t \in [0, T] \) can be written as

\[
x(t) = \sum_{k=0}^{K-1} s_k(t) + \varepsilon(t)
\]

where \( K \) is the model order, \( \varepsilon(t) \) is a white noise and \( s_k \) is an exponentially damped sinusoid defined as

\[
s_k(t) = a_k e^{\delta_k t} e^{i(\omega_k t + \phi_k)}, \quad \delta_k \geq 0
\]

\[
s_k(t) = a_k e^{\delta_k t} e^{i(\omega_k t + \phi_k)}, \quad \delta_k < 0.
\]

Each EDS is characterized by a set of 4 parameters: amplitude \( a_k \), damping \( \delta_k \), pulsation \( \omega_k \) and phase \( \phi_k \). Note that damping can be positive (increasing envelope) or negative (decreasing envelope). Using different expressions for positive and negative damping avoids numerical errors while estimating amplitudes for high dampings.

We define \( z_k = e^{\frac{\delta_k}{2} i \omega_k} \), named poles. If \( \delta_k \geq 0 \), we define \( \alpha_k = a_k e^{-\delta_k + i \phi_k} \) and if \( \delta_k < 0 \), we define \( \alpha_k = a_k e^{i \omega_k} \), named complex amplitudes. In both cases, the EDS can be written as

\[
s_k(t) = \alpha_k z_k^t.
\]

Olivier Derrien is with the Laboratory of Mechanics and Acoustics; CNRS, 13402 Marseille cedex 20 and with the University of Toulon, France (e-mail: olivier.derrien@lma.cnrs-mrs.fr). Roland Badeau and Gaël Richard are with the Institut Télécom; Télécom ParisTech; CNRS LTCI, 75634 Paris cedex 13, France (e-mail: roland.badeau@telecom-paristech.fr, gael.richard@telecom-paristech.fr).

The research leading to this paper was supported by the French GIP ANR under contract ANR-06-JCJC-0027-01, Décompositions en Éléments Sonores et Applications Musicales - DESAM.
In this study, we consider only the joint-quantization of amplitude, damping and phase. We assume that the pulsation is not quantized. Practically, the pulsation would be quantized in an independent way. We define \( p_k = \{ a_k, \delta_k, \phi_k \} \) as the set of parameters to be quantized and \( \hat{p}_k = \{ \hat{a}_k, \hat{\delta}_k, \hat{\phi}_k \} \) as the set of reconstructed parameters. We note \( P \) and \( \hat{P} \) the random variables associated with \( p_k \) and \( \hat{p}_k \).

Using the time-continuous signal model, the Mean Square Error distortion measure between an EDS \( s_k(t) \) and the reconstructed EDS \( \hat{s}_k(t) \) is defined as

\[
d(p_k, \hat{p}_k) = \frac{1}{T} \int_0^T |s_k(t) - \hat{s}_k(t)|^2 \, dt.
\]

This equation can be rewritten as

\[
d(p_k, \hat{p}_k) = |\alpha_k|^2 e^{2\delta_k} - 1 + |\hat{\alpha}_k|^2 e^{2\hat{\delta}_k} - 2 \Re \left( \alpha_k \hat{\alpha}_k e^{(\delta_k + \hat{\delta}_k) + iT(\omega_k - \hat{\omega}_k) - 1} \right)
\]

where \( \Re \) is the complex conjugate of \( x \) and \( \Re(x) \) the real part of \( x \).

Since the pulsation is not quantized, we assume \( \omega_k = \hat{\omega}_k \). If \( \delta_k \) and \( \hat{\delta}_k \) have the same sign (a sufficient condition is that the damping quantizer is symmetric around zero), we finally obtain

\[
d(p_k, \hat{p}_k) = a_k^2 h(2\delta_k) + a_k^2 h(2\hat{\delta}_k) - 2a_k \hat{a}_k \cos(\phi_k - \hat{\phi}_k) h(\delta_k + \hat{\delta}_k)
\]

where \( h \) is the real-valued function defined as

\[
h(x) = \frac{1 - \exp(-|x|)}{|x|}, \quad h(0) = 1.
\]

C. Computation of the Mean Distortion

We note \( D = \{ \hat{p}_n \}, n \in \{0 \ldots N - 1 \} \) the reconstruction dictionary. \( C_n \) is the quantization cell associated with the reconstruction value \( \hat{p}_n \). The mean distortion over \( C_n \) can be written as

\[
d_{C_n}(\hat{p}_n) = \frac{\int_{C_n} \rho_P(p) d(p, \hat{p}_n) \, dp}{\int_{C_n} \rho_P(p) \, dp}
\]

where \( \rho_P(p) \) is the probability density function (PDF) of EDS parameters.

The overall mean distortion is

\[
D = \sum_n \rho_n d_{C_n}(\hat{p}_n) \quad (1)
\]

where \( \rho_n = \text{proba} \{ P \in C_n \} = \int_{C_n} \rho_P(p) \, dp \).

The first step in the computation of \( D \) consists in finding an analytic expression for \( d_{C_n}(\hat{p}_n) \).

We assume that amplitude, damping and phase are quantized with scalar quantizers. Thus, the 3-dimensional quantization cell can be seen as the product of 3 scalar quantization cells:

\[
C = \{ [z_a, z_a + \Delta_a], [z \delta, z \delta + \Delta \delta], [z \phi, z \phi + \Delta \phi] \}
\]

where \( \Delta_a, \Delta \delta \) and \( \Delta \phi \) denote the widths of scalar quantization cells respectively for amplitude, damping and phase. As we consider only one quantization cell, note that we omit the cell index \( n \).

In high resolution hypothesis, it is reasonable to assume that \( \rho_P(p) \) is constant over the quantization cell. Thus, we get

\[
d_C(\hat{a}, \hat{\delta}, \hat{\phi}) \approx \frac{1}{\Delta_a \Delta \delta \Delta \phi} \int_{z_a}^{z_a + \Delta_a} \int_{z \delta}^{z \delta + \Delta \delta} \int_{z \phi}^{z \phi + \Delta \phi} \left( a^2 h(2\delta) + \hat{a}^2 h(2\hat{\delta}) - 2a \hat{a} \cos(\phi - \hat{\phi}) h(\delta + \hat{\delta}) \right) \, da \, d\delta \, d\phi. \quad (2)
\]

1) Optimal reconstruction values: In order to find the reconstruction values \( \hat{a}, \hat{\delta} \) and \( \hat{\phi} \) which minimize the distortion over each quantization cell, we solve the following system:

\[
\frac{\partial d_C}{\partial \hat{a}}(\hat{a}, \hat{\delta}, \hat{\phi}) = 0 \quad \Rightarrow \quad \hat{a} \approx \left( z_a + \frac{\Delta_a}{2} \right) \cos \left( z \phi + \frac{\Delta \phi}{2} - \phi \right) \sin \left( \frac{\Delta \phi}{2} \right) \frac{\sin \left( \frac{\Delta \phi}{2} \right)}{\Delta \delta} \left[ \hat{h}(z \delta + \Delta \delta + \hat{\delta}) - h(z \delta + \hat{\delta}) \right] \quad (3)
\]

\[
\frac{\partial d_C}{\partial \hat{\delta}}(\hat{a}, \hat{\delta}, \hat{\phi}) = 0 \quad \Rightarrow \quad \hat{\delta} \approx \left( z_a + \frac{\Delta_a}{2} \right) \cos \left( z \phi + \frac{\Delta \phi}{2} - \phi \right) \sin \left( \frac{\Delta \phi}{2} \right) \frac{\sin \left( \frac{\Delta \phi}{2} \right)}{\Delta \delta} \left[ h(z \delta + \Delta \delta + \hat{\delta}) - \hat{h}(z \delta + \hat{\delta}) \right] \quad (4)
\]

\[
\frac{\partial d_C}{\partial \hat{\phi}}(\hat{a}, \hat{\delta}, \hat{\phi}) = 0 \quad \Rightarrow \quad \sin \left( z \phi + \frac{\Delta \phi}{2} - \phi \right) \left[ \hat{h}(z \delta + \Delta \delta + \hat{\delta}) - \hat{h}(z \delta + \hat{\delta}) \right] \approx 0 \quad (5)
\]
where \( h(x) \) is a primitive of \( h(x) \) and \( h'(x) \) is the first order derivative of \( h(x) \). In high resolution, the quantization cells are assumed to be small, so \( \Delta_a \approx 0, \Delta_\delta \approx 0 \) and \( \Delta_\phi \approx 0 \). Taylor series expansions give
\[
\frac{\sin \left( \frac{\Delta_\phi}{2} \right)}{\Delta_\phi} \approx 1
\]
\[
\frac{\tilde{h}(z_\delta + \Delta_\delta + \tilde{\delta}) - \tilde{h}(z_\delta + \tilde{\delta})}{\Delta_\delta} \approx h \left( z_\delta + \frac{\Delta_\delta}{2} + \tilde{\delta} \right)
\]
\[
\frac{h(z_\delta + \Delta_\delta + \tilde{\delta}) - h(z_\delta + \tilde{\delta})}{\Delta_\delta} \approx h' \left( z_\delta + \frac{\Delta_\delta}{2} + \tilde{\delta} \right).
\]
Thus, equations (3)-(5) can be simplified as
\[
\hat{a} \approx \left( z_a + \frac{\Delta_a}{2} \right) \cos \left( z_\phi + \frac{\Delta_\phi}{2} - \tilde{\phi} \right) \frac{\tilde{h} \left( z_\delta + \frac{\Delta_\delta}{2} + \tilde{\delta} \right)}{h(2\tilde{\delta})}
\]
\[
\sin \left( z_\phi + \frac{\Delta_\phi}{2} - \tilde{\phi} \right) h \left( z_\delta + \frac{\Delta_\delta}{2} + \tilde{\delta} \right) \approx 0
\]
\[
\hat{a} \approx \left( z_a + \frac{\Delta_a}{2} \right) \cos \left( z_\phi + \frac{\Delta_\phi}{2} - \tilde{\phi} \right) \frac{h' \left( z_\delta + \frac{\Delta_\delta}{2} + \tilde{\delta} \right)}{h'(2\tilde{\delta})}.
\]
The only solution is:
\[
\hat{a} \approx z_a + \frac{\Delta_a}{2}, \quad \tilde{\delta} \approx z_\delta + \frac{\Delta_\delta}{2}, \quad \tilde{\phi} \approx z_\phi + \frac{\Delta_\phi}{2} \quad (6)
\]
which means that, in high resolution, reconstruction values are approximately in the middle of quantization cells.

2) Mean distortion for one cell: Using the results of equation (6), we calculate the integral defined by equation (2):
\[
d_C(\hat{a}, \tilde{\delta}, \tilde{\phi}) \approx \left( \frac{\Delta^2}{12} + \hat{a}^2 \right) \left[ \frac{\tilde{h}(2\tilde{\delta} + \Delta_\delta) - \tilde{h}(2\tilde{\delta} - \Delta_\delta)}{2\Delta_\delta} + \hat{a}^2 h(2\tilde{\delta}) - 2 \hat{a}^2 \sin \left( \frac{\Delta_\phi}{2} \right) \left[ \frac{\tilde{h} \left( 2\tilde{\delta} + \frac{\Delta_\delta}{2} \right) - \tilde{h} \left( 2\tilde{\delta} - \frac{\Delta_\delta}{2} \right)}{\Delta_\delta} \right] \right]. \quad (7)
\]
Taylor series expansions give
\[
\frac{\sin \left( \frac{\Delta_\phi}{2} \right)}{\Delta_\phi} = 1 - \frac{\Delta_\phi^2}{24} + O(\Delta_\phi^4)
\]
\[
\frac{\tilde{h}(2\tilde{\delta} + \Delta_\delta) - \tilde{h}(2\tilde{\delta} - \Delta_\delta)}{2\Delta_\delta} = h(2\tilde{\delta}) + \frac{\Delta_\delta^2}{6} h''(2\tilde{\delta}) + O(\Delta_\delta^4)
\]
\[
\frac{\tilde{h} \left( 2\tilde{\delta} + \frac{\Delta_\delta}{2} \right) - \tilde{h} \left( 2\tilde{\delta} - \frac{\Delta_\delta}{2} \right)}{\Delta_\delta} = h(2\tilde{\delta}) + \frac{\Delta_\delta^2}{24} h''(2\tilde{\delta}) + O(\Delta_\delta^4)
\]
where \( h''(x) \) is the second order derivative of \( h(x) \). Thus, keeping only terms in \( O(\Delta_\delta^2), O(\Delta_\delta^4) \) and \( O(\Delta_\delta^2) \), equation (7) can be simplified as
\[
d_C(\hat{a}, \tilde{\delta}, \tilde{\phi}) \approx \frac{1}{12} \left[ h(2\tilde{\delta})\Delta^2_a + \hat{a}^2 h''(2\tilde{\delta})\Delta^2_\delta + \hat{a}^2 h(2\tilde{\delta})\Delta^2_\phi \right]. \quad (8)
\]
3) Overall mean distortion: We assume that the amplitude quantizer depends on damping, that the damping quantizer depends on amplitude, and that the phase quantizer depends both on amplitude and damping. As amplitude, damping and phase are quantized with scalar quantizers, the 3-dimensional QCD can be split in 3 scalar functions: \( g_A(a, \delta) \), \( g_\Delta(a, \delta) \) and \( g_\phi(a, \delta, \phi) \) are respectively the QCD on amplitude, damping and phase. The scalar QCDs at reconstruction points are defined as follows:
\[
g_A(\hat{a}, \delta) = \frac{1}{\Delta_a}, \quad g_\Delta(\hat{a}, \delta) = \frac{1}{\Delta_\delta}, \quad g_\phi(\hat{a}, \delta, \phi) = \frac{1}{\Delta_\phi}.
\]
Thus, using the result of equation (8), the mean distortion defined by equation (1) can be written as
\[
D \approx \frac{1}{12} \sum_n \rho_n \left[ \frac{h(2\tilde{\delta}_n)}{g_A^2(\hat{a}_n, \delta_n)} + \frac{\hat{a}^2 h''(2\tilde{\delta}_n)}{g_\Delta^2(\hat{a}_n, \delta_n)} + \frac{\hat{a}^2 h(2\tilde{\delta}_n)}{g_\phi^2(\hat{a}_n, \delta_n, \phi_n)} \right].
\]
Assuming that the PDF of EDS parameters $\rho_P(p)$ is constant over each quantization cell leads to the following approximation:

$$\rho_n \approx \rho_P(\hat{p}_n) \Delta_n$$  \hspace{1cm} (9)

where $\Delta_n$ is the volume of quantization cell $C_n$, yielding

$$D \approx \frac{1}{12} \sum_n \rho_P(\hat{a}_n, \hat{\delta}_n, \hat{\phi}_n) \left[ \frac{h(2\hat{\delta}_n)}{g^2_A(\hat{a}_n, \hat{\delta}_n)} + \frac{\hat{a}^2 h''(2\hat{\delta}_n)}{g^3_A(\hat{a}_n, \hat{\delta}_n)} + \frac{\hat{a}^2 h(2\hat{\delta}_n)}{g^4_A(\hat{a}_n, \hat{\delta}_n, \hat{\phi}_n)} \right] \Delta_n.$$  

The sum can be approximated by an integral:

$$D \approx \frac{1}{12} \int \int \int \rho_P(a, \delta, \phi) \left[ \frac{h(2\delta)}{g^2_A(a, \delta)} + \frac{a^2 h''(2\delta)}{g^3_A(a, \delta)} + \frac{a^2 h(2\delta)}{g^4_A(a, \delta, \phi)} \right] da \, d\delta \, d\phi.$$  \hspace{1cm} (10)

### D. Computation of the optimal quantization cell density

The optimal quantizer minimizes the overall mean distortion under a bit rate constraint. We write the constraint as a condition on the entropy of quantization indexes. Thus, the optimal quantizer minimizes $D = E[d(P, \hat{P})]$ under the constraint $H(I) \leq R$, where $H(I)$ denotes the entropy of quantization indexes and $R$ the average number of coding bits.

We denote $i_n$ the quantization index associated with the quantization cell $C_n$. We get

$$\text{proba}\{I = i_n\} = \text{proba}\{P \in C_n\} = \rho_n.$$  

Thus, the entropy of quantization indexes can be written as

$$H(I) = - \sum_n \rho_n \log_2(\rho_n).$$

According to equation (9), we get

$$H(I) \approx - \sum_n \rho_P(\hat{p}_n) \Delta_n \log_2(\rho_P(\hat{p}_n)) \Delta_n.$$

$\Delta_n$ is related to the QCD at reconstruction points:

$$g_P(\hat{p}_n) = \frac{1}{\Delta_n}.$$  

The expression of the entropy is then

$$H(I) \approx - \sum_n \rho_P(\hat{p}_n) \log_2(\rho_P(\hat{p}_n)) \Delta_n + \sum_n \rho_P(\hat{p}_n) \log_2(g_P(\hat{p}_n)) \Delta_n.$$  

Sums can be approximated by integrals, yielding

$$H(I) \approx - \int \rho_P(p) \log_2(\rho_P(p)) \, dp + \int \rho_P(p) \log_2(g_P(p)) \, dp.$$  

The first integral is the differential entropy of EDS parameters named $H(P)$. The 3-dimensional QCD $g_P(p)$ can be written as the product of 3 scalar QCD on amplitude, damping and phase. Finally, the entropy constraint can be written as

$$H(I) \approx H(P) + \int \int \int \rho_P(a, \delta, \phi) \log_2(g_A(a, \delta) g_\Delta(a, \delta) g_\Phi(a, \delta, \phi)) \, da \, d\delta \, d\phi \leq R.$$  \hspace{1cm} (11)

Thus, we look for the expressions of $g_A$, $g_\Delta$ and $g_\Phi$ which minimize the distortion defined by equation (10) under the entropy constraint defined by equation (11).

Assuming that the rate-distortion function of any quantizer (i.e. $D$ as a function of $H(I)$) is decreasing, the optimal solution (i.e. the minimum value for $D$) is reached when $H(I) = R$. This constrained optimization problem can be conveniently solved with a Lagrange optimization technique. The Lagrangian functional is defined as

$$\mathcal{L} = D + \mu [H(I) - R]$$

where $\mu$ is the real-valued Lagrange multiplier. The Euler-Lagrange equations give the optimal expressions of $g_A$, $g_\Delta$ and $g_\Phi$ as functions of $\mu$:

$$\frac{\partial \mathcal{L}}{\partial g_A} = 0 \iff g_A(a, \delta) \approx \left( \frac{\ln(2)h(2\delta)}{6\mu} \right)^\frac{1}{2}$$

$$\frac{\partial \mathcal{L}}{\partial g_\Delta} = 0 \iff g_\Delta(a, \delta) \approx a \left( \frac{\ln(2)h''(2\delta)}{6\mu} \right)^\frac{1}{2}$$
\[ \frac{\partial L}{\partial g_{\Phi}} = 0 \Leftrightarrow g_{\Phi}(a, \delta, \phi) \approx a \left( \frac{\ln(2)h(2\delta)}{6\mu} \right)^{\frac{1}{2}}. \]

The optimal value for \( \mu \) can be obtained from the constraint. Equation (11) can be rewritten as

\[ R \approx H(p) + \int \int \int \rho_p(a, \delta, \phi) \log_2 \left( \frac{a^2 \ln(2) h^2(2\delta) h''(2\delta)^{\frac{1}{2}}}{(6\mu)^{\frac{3}{2}}} \right) da d\delta d\phi. \]

Defining the following constant:

\[ \sigma = H(p) + \int \rho_{\Delta}(\delta) \log_2(h(2\delta)h''(2\delta)^{\frac{1}{2}})d\delta + 2 \int \rho_A(a) \log_2(a) da \]

where \( \rho_A(a) \) and \( \rho_{\Delta}(\delta) \) are respectively the marginal PDFs of amplitude and damping, the optimal value for \( \mu \) is

\[ \mu \approx \ln(2) \frac{2^{\frac{1}{2}} \sigma - R}{6 \cdot 2^{\frac{1}{4}} (R - \sigma)} \]

and the optimal QCDs are finally

\[ g_{A}(\delta) \approx h(2\delta)^{\frac{1}{2}} 2^{\frac{1}{4}} (R - \sigma) \]
\[ g_{\Delta}(a, \delta) \approx a h''(2\delta)^{\frac{1}{2}} 2^{\frac{1}{4}} (R - \sigma) \]
\[ g_{\Phi}(a, \delta) \approx a h(2\delta)^{\frac{1}{2}} 2^{\frac{1}{4}} (R - \sigma). \]

E. Conclusion

One can observe that amplitude and phase quantizers are uniform, since the QCDs \( g_{A} \) and \( g_{\Phi} \) do not depend respectively on \( a \) and \( \phi \). In contrast, damping quantization is not uniform: Quantization is more precise for small damping values.

This work can be considered as an extension of the Entropy-Constrained Polar Quantizer proposed by R.Vafin et al, since for \( \delta = 0 \), the resulting QCDs \( g_{A} \) and \( g_{\Phi} \) are identical to the ones described by R.Vafin et al.

For practical implementation, integrating each QCD with respect to the quantized variable leads to a compression/expansion function that can be combined with a unitary and uniform scalar quantizer.

A performance evaluation of this quantizer in comparison with a 3-dimensional vector quantizer is proposed in a paper submitted to IEEE Signal Processing Letters named "Entropy-Constrained Quantization of Exponentially Damped Sinusoids Parameters".