REVISITING QUANTIZATION THEOREM THROUGH AUDIOWATERMARKING

Houssem Halalchi(1), *Gaël Mahé*(1), *Mériem Jaïdane*(2)

(1) Centre de Recherche en Informatique de Paris 5 (CRIP5), Université Paris Descartes, France
{ halalchi, gael.mahe}@math-info.univ-paris5.fr
(2) Unité Signaux et Systèmes, Dpt. TIC, Ecole Nationale d'Ingénieurs de Tunis, Tunisia
meriem.jaidane@enit.rnu.tn

ABSTRACT

In this paper, we propose the concept of "doping watermarking", whose principle is to add an imperceptible noise to an host signal in order to improve its properties. Especially, our aim is to reduce the spectral support of the probability density function (PDF) of an audio signal in order to match the conditions of the quantization theorem. In this context, we develop a specific audiowatermarking algorithm and test its performance on real audio signals. This watermark allows to recover the PDF of a digital signal from a sub-quantized version of the signal, with very low error.

***Index Terms*—** quantization theorem, sub-quantization, audiowatermarking, speech and audio processing.

1. INTRODUCTION

The quantization theorem [1] states that, if the probability density function (PDF) of a sampled signal x is spectrally band-limited, it is possible to recover it from the PDF of the quantized signal. Like in the sampling theorem, a sufficient condition for the recovery is that the quantization "frequency" (inverse of the quantization step q) of the signal is two times greater than the maximal frequency (ν_{max}) of the characteristic function of x: $\frac{1}{q} \geq 2\nu_{max}$

If the conditions of the theorem are not met, the support of the characteristic function should be artificially reduced. In this purpose, we propose to transform the distribution of an audio signal by the way of "doping watermarking". The principle is that an inaudible noise added to the original signal makes the watermarked signal match a target histogram that meets the condition of the quantization theorem.

The idea of doping watermarking was inspired by [2] and [3] in a context of acoustic echo cancellation. In [2], an inaudible noise was added to the signal to reduce the ill-conditioning of the covariance matrix of the signals in the case of multiple loudspeakers. [3] showed that a watermark (actually any piecewise stationary signal) added to the audio signal stationnarizes the latter, which leads to better performance in echo cancellation.

[4] showed the interest of the doping watermarking for non-linear system identification. Adding a noise that makes the PDF of the audio signal Gaussian enhances the conditionning of the matrix to be inverted in the optimal identification.

In the following, we will adapt the quantization theorem to the case of the sub-quantization of an already quantized signal. In the third section, we will propose a specific watermarking algorithm to meet the conditions of the quantization theorem. The PDF recovery with or without watermarking will be compared for real audio signals in section 4.

2. SUB-QUANTIZATION OF QUANTIZED SIGNALS

Widrow showed that quantization is equivalent to an area sampling of the original continuous PDF [1]. Area sampling means : first convolving the original PDF f_x with a uniform pulse function of width q and then multiplying the result of convolution with a Dirac impulse carrier (whose delays are multiples of q), which gives the discrete PDF $f_{x'}$.

We adapted the operations of the classical area sampling to the specific case of digital original signals that are already quantized (discrete original PDF). We define the *sub-quantization* of the digital signal x as the increase of its quantization step q_0. The new quantization step q_1 is a multiple of q_0. The sub-quantized signal is denoted x_Q and the *sub-quantization rate* K is given by $K = \frac{q_1}{q_0}$

In the following, according to the discrete formalism, the index i in the expression $f(i)$, where f is a PDF, stands for iq_0.

Sub-quantization with a factor K means rounding the values belonging to the discrete interval $[nK - \frac{K}{2}, nK + \frac{K}{2}]$ to the value nK (n integer). The PDF of the sub-quantized signal $f_{xQ}(i)$ equals 0 for all $i \neq nK$ and :

$$f_{xQ}(nK) = \sum_{i=\lfloor nK-\frac{K}{2}+1\rfloor}^{\lfloor nK+\frac{K}{2}\rfloor} f_x(i) \quad (1)$$

This work is part of the project WaRRIS granted by the French National Research Agency (project n ANR-06-JCJC-0009)

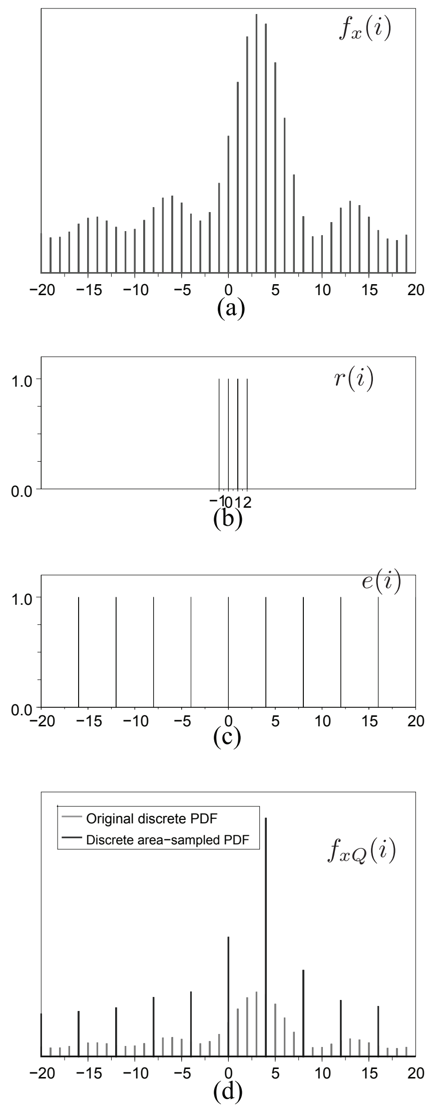

Fig. 1. Discrete area sampling of PDF of x: (a) original PDF $f_x(i)$; (b) discrete rectangular filter $r(i)$;(c) impulse train $e(i)$; (d) final PDF $f_{xQ}(i)$

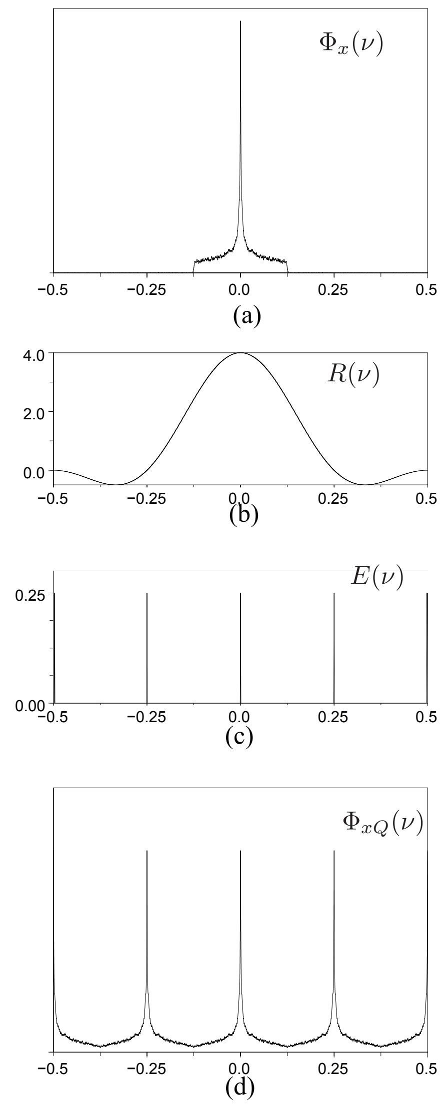

Fig. 2. Discrete area sampling in the spectral domain

As for the quantization theorem, in terms of PDF, sub-quantization of the original signal leads to a discrete area sampling, as illustrated in Figure 1 for $K = 4$:

1. Convolution of the original PDF $f_x(i)$ (Fig. 1a) by a discrete rectangular filter r (Fig. 1b) :

$$r(i) = \sum_{l=\lfloor -\frac{K}{2}+1 \rfloor}^{\lfloor \frac{K}{2} \rfloor} \delta(i-l) \quad (2)$$

2. Multiplication of the result of convolution by a uniform impulse train e (Fig. 1c) :

$$e(i) = \sum_{n=-\infty}^{+\infty} \delta(i-Kn) \quad (3)$$

The PDF $f_{xQ}(i)$ of the sub-quantized signal (Fig 1d) is :

$$f_{xQ}(i) = [f_x(i) * r(i)].e(i) \quad (4)$$

The discrete time Fourier transform (DTFT) of the PDF is known as the "characteristic function" $\Phi(\nu)$. From (4), we obtain :

$$\Phi_{xQ}(\nu) = [\Phi_x(\nu)\, R(\nu)] * E(\nu) \quad (5)$$

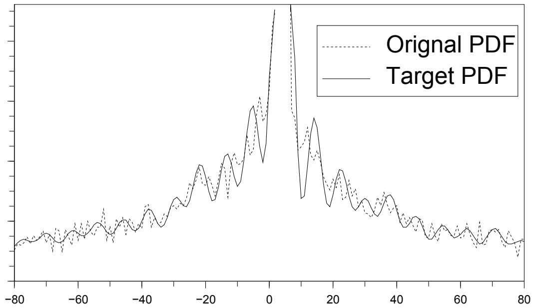

Fig. 3. Original and target distributions (zoom)

where, from relations (2) and (3) :

$$R(\nu) = \begin{cases} K & \text{if } \nu \in \mathbb{Z} \\ \frac{sin(\pi K \nu)}{sin(\pi \nu)} \exp(j\pi\nu) & \nu \notin \mathbb{Z} \text{ and } K \text{ even} \\ \frac{sin(\pi K \nu)}{sin(\pi \nu)} & \nu \notin \mathbb{Z} \text{ and } K \text{ odd} \end{cases} \quad (6)$$

$$E(\nu) = \frac{1}{K} \sum_{n=-\propto}^{\propto} \delta(\nu - \frac{n}{K}) \quad (7)$$

As illustrated in Figure 2, sub-quantizing with a factor K is equivalent, in the spectral domain, to replicating K times the characteristic function in $[-\frac{1}{2}; \frac{1}{2}]$, with a periodicity of $1/K$. This leads to the:

Sub-quantization theorem: if the characteristic function of a quantized signal x is equal to zero for $|\nu| > \frac{1}{2K}$ in $[-\frac{1}{2}; \frac{1}{2}]$, then the PDF of x can be derived from that of the signal x_Q resulting from the sub-quantization of x with a factor K.

The original PDF is recovered through low-pass filtering of f_{xQ} with a cut-off frequency $\frac{1}{2K}$ and division by $R(\nu)$ on $[-\frac{1}{2K}; \frac{1}{2K}]$ (note that $R(\nu) \neq 0 \quad \forall |\nu| < \frac{1}{K}$). Thus,

$$\Phi_x(\nu) = \Phi_{xQ}(\nu) G(\nu) \quad (8)$$

where $G(\nu) = 1/R(\nu)$ for $|\nu| < \frac{1}{2K}$, 0 otherwise.

3. REDUCING THE SUPPORT OF THE CHARACTERISTIC FUNCTION

Unlike to the previous example, the characteristic function of a digital audio signal generally spreads over the whole normalized frequencies intervall ($\nu_{max} = \frac{1}{2}$). The sub-quantization implies the reduction of the spectral support, in order to meet the Widrow condition.

The reduction of the spectral support means a low-pass filtering of $f_x(i)$, with a cut-off frequency $\nu_c = \frac{1}{2K}$. Figure 3 shows the target low-pass PDF for a speech signal of duration 3 s, sampled at $Fs = 8000$ Hz and coded with 16 bit/sample, for $K = 4$.

In order to make the signal follow the target filtered PDF, we need to move the values of the samples, by addition of a noise that we will call *doping watermark*: $z = x + w$, where w is the watermark and z the watermarked signal. The noise w should be inaudible. One may expect w to be all the louder as the target PDF is far from the original, so that ν_c cannot be freely chosen.

Preliminary histogram normalization : the proposed algorithm works with the histogram h_x of a sequence x of length N. The target histogram (filtered) h_{target} is rounded to integer values so that its total number of samples equals to N. Initially, $z = x$. The goal is that the histogram of z, h_z, equals to h_{target}.

Iterative histogram adjustment: then, for $i = -2^{b-1} \rightarrow 2^{b-1} - 1$, where b is the number of coding bits :

- If $h_z(i) = h_{target}(i)$, the samples with value i do not need to be moved.

- If $h_z(i) - h_{target}(i) = n > 0$, we select randomly n samples of z of value i. Each of those samples gets the value $i+1$, so that $h_z(i) = h_{target}(i)$ and $h_z(i+1) = h_z(i+1) + n$.

- If $h_z(i) - h_{target}(i) = -n < 0$, we select randomly n samples of z of value $i+1$. Each of those samples gets the value i. If $h_z(i+1) < n$, we select randomly the missing samples among those of value $i+2$, and so on, until $h_z(i) = h_{target}(i)$. $h_z(j > i)$ decreases according to the number of samples moved.

At the end of this algorithm, $h_z = h_{target}$.

The noise w generated by the transformation of x into z has to be inaudible, in other words masked by the audio signal. This constraint of frequency masking is not integrated in the proposed algorithm, which works in the PDF domain, but is checked *a posteriori*, by comparing frame by frame the spectrum of w to the masking threshold. The latter was computed according to the Johnston model [5] for 32 ms frames. The masking depends on the transformation imposed to the PDF : the watermark is all the more audible as the sub-quantization rate K is high. According to our experiments on various signals, choosing $K < 8$ leads to an acceptable masking. Figure 4 illustrates the masking in the case $K = 4$, for one frame of the previous speech signal.

The audibility of the watermark was also assessed through Signal to Watermark Ratio (SWR) and perceptual measures : mean opinion score (MOS) predicted by PESQ [6] for speech and objective difference grade (ODG) from PEAQ [7] for music. See results in Table 1 for the same speech example and a violin signal of duration 4 s, sampled at 44.1 kHz, coded with 16 bit/sample. For the speech sequence, the watermark is inaudible for $K = 4$. For the violin sequence, it is slightly audible for both values of K.

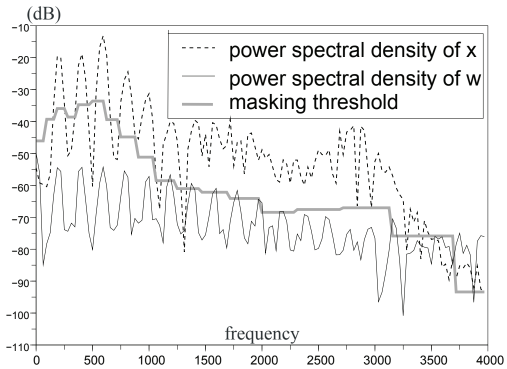

Fig. 4. Frequency masking of the watermark

signal	K	SWR	quality	$d_{KS}(f_x, \hat{f}_x)$ $\times 10^4$	$d_{KS}(f_z, \hat{f}_z)$ $\times 10^4$
speech	4	19.2	MOS: 4.14	69	4.3
	8	15.7	MOS: 3.85	197	5.9
violin	4	30.0	ODG: -0.49	7.3	1.4
	8	15.7	ODG: -0.52	13	1.3

Table 1. Audibility of the watermark and PDF recovery error, according to the sub-quantization rate K

4. SUB-QUANTIZATION AND PDF RECOVERY

The original signal x and the watermarked signal z are sub-quantized into x_Q and z_Q, respectively.

According to the sub-quantization theorem, the characteristic function of z, $\Phi_z(\nu)$, can be recovered without error from that of z_Q, $\Phi_{zQ}(\nu)$, using Eq. (8).

A recovery method in the PDF domain was proposed by [8] in the case of continuous PDF, but it is based on a differentiator, which is delicate to implement in the discrete case. We propose the following process in the frequency domain to avoid the temporal aliasing which may result from the Discrete Fourier Transform (DFT).

- computing the DFT of G on 2^b samples, then the inverse DFT on 2^{b+1} samples (zero-padding);
- computing the inverse DFT of f_{zQ} on 2^{b+1} samples (zero-padding);
- computing the estimation of the distribution of z :

$$\hat{f}_z = DFT[\Phi_{zQ} G] \tag{9}$$

We measured the dissimilarity between two PDF f and $\hat{f}$ through the Kolmogorov-Smirnov distance $d_{KS}(f, \hat{f})$. As shown in Table 1, the watermark reduces significantly the recovery error. The improvement is lower for violin than for the speech signal, because the original characteristic function is less spread, so that the frequency aliasing caused by the sub-quantization of the original signal is less critical.

5. CONCLUSION

We have formulated a discrete-to-discrete version of the quantization theorem, adapted to the sub-quantization of quantized signals. Since digital audio signals often not meet the condition of the sub-quantization theorem, we have proposed a doping-watermarking algorithm that reduces the spectral band of the characteristic function. The inaudibility of the watermark needs the targeted sub-quantization rate to be reasonably low.

The proposed method allows to sub-quantize digital audio signals and recover the original PDF with a very reduced error. The watermarking algorithm may be improved through integrating the inaudibility constraint in the process, instead of checking it *a posteriori*. This may be achieved through spectral shaping of the watermark, taking into account temporal dependencies between samples of the audio signal.

6. REFERENCES

[1] B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," *IRE Transactions on Circuit Theory*, vol. 3, no. 4, pp. 266–276, December 1956.

[2] A. Gilloire and V. Turbin, "Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, May 1998, vol. 6, pp. 3681–3684.

[3] S. D. Larbi and M. Jaidane, "Audio Watermarking : A Way to Stationnarize Audio Signals," *IEEE Trans. Signal Processing, Supplement on Secure Media*, vol. 53, no. 2, pp. 816–823, February 2005.

[4] I. Marrakchi, G. Mahé, M. Jaidane-Saidane, S. Djaziri-Larbi, and M. Turki-Hadj Alouane, "Gaussianisation method for identification of memoryless nonlinear audio systems," in *Proceedings of the European Signal Processing Conference*, 2007, pp. 2316–2320.

[5] J.D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. on selected areas in communications*, vol. 6, no. 2, pp. 314–323, Feb. 1988.

[6] "Perceptual evaluation of speech quality (pesq), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," 2001, ITU-T Rec. P.862.

[7] "Method for objective measurement of perceived audio quality," 1998, ITU-R Rec. BS.1387.

[8] B. Widrow and I. Kollar, *Quantization Noise: Round Off Error in Digital Computation, Signal Processing, Control and Communications*, Cambridge University Press, 2008.