ANALYSIS/SYNTHESIS OF THE ANDEAN QUENA VIA HARMONIC BAND WAVELET TRANSFORM

Aldo Díaz, * *Rafael Mendes*

Department of Computer Engineering and Industrial Automation - DCA,
University of Campinas
Campinas, Brazil
aldodiaz64@gmail.com rafael@dca.fee.unicamp.br

ABSTRACT

It is well known that one of the challenges in musical instruments analysis is to obtain relevant signal characteristics and information for sound description and classification. In this paper we study the Peruvian quena flute by means of the Harmonic Band Wavelet Transform (HBWT), a convenient representation for the sound content based on its $1/f$ fractal characteristics. In order to identify a relationship between fractal characteristics of musical sounds, we developed two sound transformations and establish a comparison between quena, a recorder and melodica wind instruments. The sound transformations implemented were noise filtering and pitch-shifting while the sound classification was focused on the γ_p fractal attribute. Our work led us to the conclusion that the HBWT quena representation favored the implementation of sound transformations and that the γ_p fractal feature had great potential in musical instruments recognition and classification applications.

Keywords: sound fractal analysis; quena; $1/f$ noise; noise filtering; pitch-shifting.

1. INTRODUCTION

Fractal sound analysis/synthesis are techniques that aim to develop sound decomposition and reconstruction using a minimum set of relevant fractal attributes. Recently, fractal sound analysis techniques have obtained improved results in musical instruments recognition and classification compared to traditional state of the art methods [1, 2, 3].

Despite most of the fractal techniques use the *box counting* method to analyze the fractal characteristics of musical sound [4], a method called the Harmonic Band Wavelet Transform (HBWT) proposed a different approach inspired on Multirate Filter Banks (MFB) and Perfect Reconstruction (PR) digital signal processing techniques [5]. The HBWT provides a formal representation of musical sound in terms of harmonic noise sources with $1/f$ fractal properties that accurately follow the spectral behavior of a sound signal in the frequency domain. Moreover, the HBWT is based in solid PR and MFB techniques that use overlapping filters in order to achieve distortion-free signal reconstruction [6].

Despite the HBWT was applied as a relevant synthesis tool in audio coding and compression, our purpose was to contribute with the study of the Peruvian *quena* flute (Fig. 1) in the areas of fractal sound analysis and sound transformations. The implementation of two sound transformation techniques were described and discussed: the Harmonic band $1/f$ noise filtering and the pitch-shifting. Additionally, we performed a sound comparison between the quena, recorder and melodica wind instruments based on γ_p fractal parameter in order to illustrate the potential of our methodology in sound content description, musical instruments recognition and classification problems.

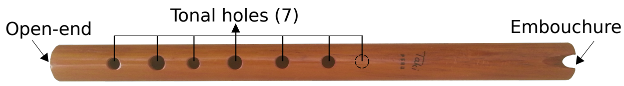

Figure 1: Traditional Peruvian quena with tuning in G major.

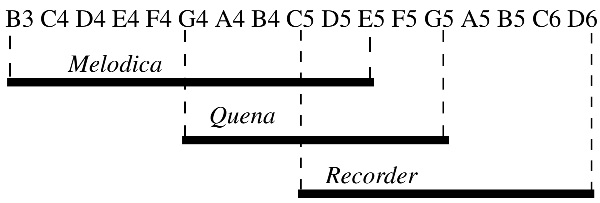

Figure 2: Pitch ranges of the analysis instruments where the overlap between them can be seen.

2. METHOD

It was generated a database with sound of quena, recorder, and melodica. The dataset was conformed by individual notes of one second duration played with *mezzo-forte* intensity, mono channel, sample rate $f_s = 44100$ Hz and 16 bits of resolution. The dataset characteristics are depicted in Fig. 2. A total of 27 sounds samples were collected: 8 for quena, 9 for recorder and 10 for melodica.

2.1. Sound decomposition

The HBWT belongs to the family of Spectral Modeling Synthesis (SMS) techniques [7]. Related to the representation of relevant sound components, the sound noise is indispensable in the perception of musical sound and inherent to the sound of musical instruments (see Fig. 4). Nevertheless, SMS contributions showing an explicit model for the stochastic component of sound (noise) are scarce [8]. In contrast, the HBWT provides a specific *pseudo-periodic* $1/f$ *noise model* for noise characterization [9]. A block

* The author is grateful to the founding of the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brazil (grant number 160186/2013-7) for funding this project.

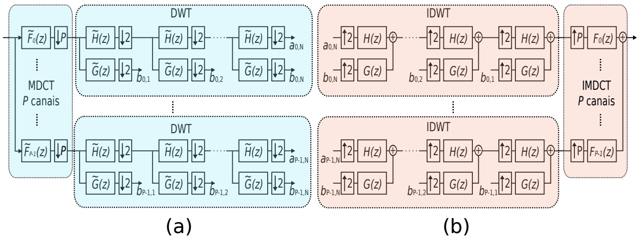

Figure 3: Structure of filters implementing the Harmonic Band Wavelet Transform. (a) Analysis bank. (b) Synthesis bank.

diagram of HWBT implementation is depicted in Fig. 3. As illustrated in Fig. 4, the HBWT model matches the sound signal fractal characteristics by decomposing the frequency harmonics in separated sidebands.

2.2. HBWT Analysis

The HBWT analysis structure is composed by perfect reconstruction filter banks of Modified Discrete Cosine Transform (MDCT) [6] and Discrete Wavelet Transform (DWT) [10]. In Fig. 5 it is shown the frequency response characteristic of an 8-channel MDCT and DWT with Daubechies order 11 filters.

The purpose of MDCT filters is to decompose the signal harmonics into two sidebands, each one containing half of a band that approximates the $1/f$ noise behavior. For MDCT implementation it was used cosine modulated type IV bases [11]:

$$g_{p,r}(k) = g_{p,0}(k - rP) \tag{1}$$

$$g_{p,0}(k) = w(k)\cos\left[\left(k - \frac{P+1}{2}\right)\left(p + \frac{1}{2}\right)\frac{\pi}{P}\right] \tag{2}$$

$$w(k) = \sin\left[\left(k + \frac{1}{2}\right)\frac{\pi}{2P}\right], \tag{3}$$

where $p = 0, \ldots, P-1$; $r \in \mathbb{Z}$; $k = 0, \ldots, 2P-1$. Considering an input signal $x(k)$ with fundamental frequency f_0 and average period P measured in number of samples then, the number of MDCT channels is equal to $P = f_s/f_0$. For instance, in Fig. 4 it is represented the spectrum of the first three harmonics of an A4 quena sound with $f_0 = 444.8$ Hz and the MDCT channels synchronized to the average period of the signal $P = 99$.

Each sideband is immediately decomposed by a bank of the Discrete Wavelet Transform (DWT). In wavelet theory, the wavelet representation bases execute shift and scaling operations of a wavelet function $\psi(t)$ defined as:

$$\psi_{\alpha,\tau}(t) = \frac{1}{\sqrt{\alpha}}\psi\left(\frac{t-\tau}{\alpha}\right), \tag{4}$$

where τ is called the shifting factor and $\alpha > 0$ is the scaling factor. Following the recommendations in [9], the DWT filter bank was implemented using Daubechies wavelets of order 11. The number of wavelet analysis levels N varies according to the instrument. In case of the quena, recorder and melodica, it was found $N = 4$ to be a proper number of levels for analysis.

The overall HBWT analysis procedure was implemented as follows:

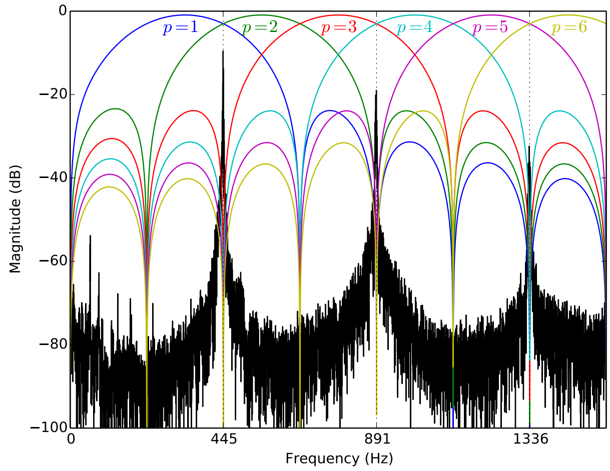

Figure 4: Quena A4 signal and MDCT channels synchronized to the average period P.

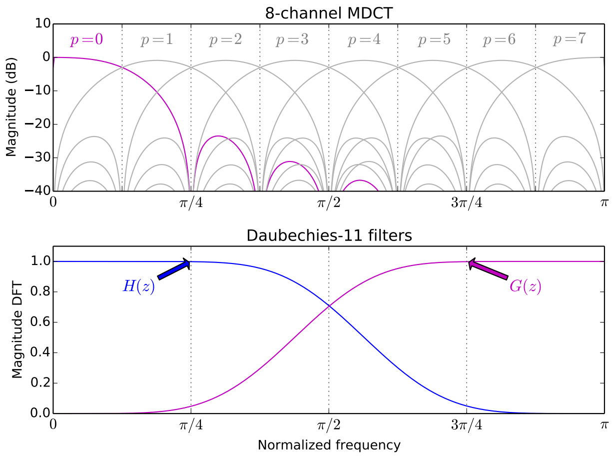

Figure 5: Frequency responses of MDCT and DWT filter banks.

Step 1

Decomposition of input signal $x(k)$ via a P-channel MDCT filter bank. Each MDCT filter $\tilde{F}_p(z)$ is band-pass type with bandwidth equals to π/P.

Step 2

Decimation bank of factor P[1].

Step 3

Wavelet decomposition via a P-channel DWT filter bank at the output of the decimation bank.

As a result, it is obtained the decomposition of the sidebands of the harmonics into two components: A deterministic component, represented by the scale coefficients $a_{p,N}$ and a stochastic component, represented by the wavelet coefficients $b_{p,n}$, where p

[1]The P decimation rate does not cause aliasing since the MDCT channels have a bandwidth equals to π/P.

is the MDCT channel index, n the DWT level index, and N the DWT total number of levels.

2.3. HBWT Synthesis

Step 1
: Wavelet Reconstruction via a P-channel filter bank implementing the inverse of the DWT (IDWT)[2].

Step 2
: Expansion bank of factor P[3].

Step 3
: Output signal $\hat{x}(k)$ reconstruction via a P-channel IMDCT (inverse of MDCT)[4].

By means of HBWT decomposition and reconstruction, a sound signal is modeled by a set of periodic $1/f$ noise-like stochastic processes. A formal definition of $1/f$ noise-like processes can be found in [12]. One of the benefits of using the $1/f$ noise-like model is that only two parameters are needed to describe the complete behavior of a sound signal: The parameter σ_p^2 that controls the amplitude of the $1/f$ spectrum and the parameter γ_p that controls the slope of the pseudo-periodic $1/f$ noise-like sidebands of the spectrum.

The discrete-time harmonic band wavelets (DT-HBW), as defined in [13], are:

$$\xi_{n,m,p}(k) = \sum_{r=-\infty}^{\infty} \psi_{n,m}(r) g_{p,r}(k) \tag{5}$$

$$\zeta_{N,m,p}(k) = \sum_{r=-\infty}^{\infty} \varphi_{N,m}(r) g_{p,r}(k), \tag{6}$$

where $n = 1, 2, \ldots, N$; $m \in \mathbb{Z}$; $p = 0, 1, \ldots, P-1$; $\psi_{n,m}$ and $\varphi_{N,m}$ are the discrete-time ordinary wavelets and the corresponding scale residue function, respectively; $g_{p,r}$ are the MDCT functions of Eq. 2. Therefore, a signal $x(k) \in l^2$ can be expanded on a discrete-time harmonic band wavelet set according to:

$$x(k) = \sum_{p=1}^{P} \left(\sum_{n=1}^{N} \sum_{m=-\infty}^{\infty} b_{p,n}(m) \xi_{n,m,p}(k) + \sum_{m=-\infty}^{\infty} a_{p,N}(m) \zeta_{N,m,p}(k) \right), \tag{7}$$

where the $b_{p,n}(m)$'s and $a_{p,N}(m)$'s are the expansion coefficients and the corresponding harmonic-band scale residue coefficients at scale N, respectively. From *Proposition 3.4* in [13] it was deducted the resultant energy for the $b_{p,n}(m)$ analysis coefficients:

$$Var\{b_{p,n}(m)\} = \sigma_p^2 2^{n\gamma_p} \tag{8}$$

Considering the logarithm of the energies of each n-subband of a single p-sideband it was found a linear relationship for γ_p ($1/f$ noise slope) at each harmonic band wavelet analysis level n:

$$\log_2(Var\{b_{p,n}[m]\}) = \gamma_p n + const, \tag{9}$$

[2] This step performs the reconstruction of the sidebands.

[3] This step returns the properly bandwidth for the reconstructed spectrum of each sideband.

[4] The IMDCT filter bank selects the appropriate frequency range for the sidebands on each channel.

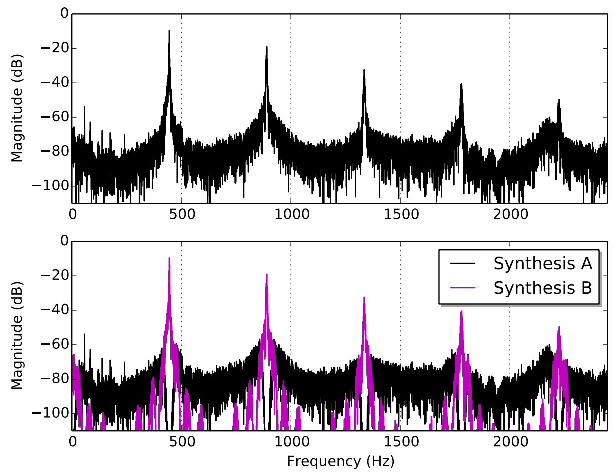

Figure 6: Above: First five harmonics of quena A4 sound. Below: Synthesis results of experiments A and B.

where the spectral component γ is directly related to the self-similarity attribute H_p (Hurst exponent) according to Eq. 10.

$$\gamma_p = 2H_p + 1 \tag{10}$$

In case of $1/f$ noise signals, or fractional Brownian motion processes (fBm), the typical values for the Hurst exponent are in the interval $0 < H < 1$, for a corresponding $1 < \gamma < 3$ [12].

3. RESULTS

3.1. Harmonic band $1/f$ noise filtering

We performed two sets of experiments in order to filter the harmonic band $1/f$ noise. The results are shown in Fig. 6. In the first experiment, called *Synthesis A*, the wavelet expansion coefficients $b_{p,n}$ were used as input to the synthesis bank with up to $N = 3$ scale levels. Separately, in experiment *Synthesis B*, the wavelet scale coefficients $a_{p,N}$ ($N = 3$) were used as input to the synthesis bank. *Synthesis A* resulted in the isolation of the harmonic band $1/f$ noise content associated to the sidebands of harmonics. In *Synthesis B*, the result was the $1/f$ noise-free harmonic signal (purely deterministic content).

These experiments benefited the study of the sound signal fluctuations related to the $1/f$ noise. It was found that the harmonic band $1/f$ noise content provided an important contribution to the sound in terms of perception. For instance, in *Synthesis A*, it was found that the harmonic band $1/f$ noise was related to the action mechanism of the instrument. In case of the quena, from the $1/f$ signal it was the retrieved the sound of the blowing air.

3.2. Pitch-shifting via HBWT

In sound synthesis, pitch-shifting is a technique that consists in modification (shift) of the fundamental frequency f_0 (pitch) of a sound. In terms of HBWT decomposition, pitch-shifting was interpreted as a frequency modulation process: The frequency content extracted in the initial signal analysis was transferred to new frequency bands in the synthesis.

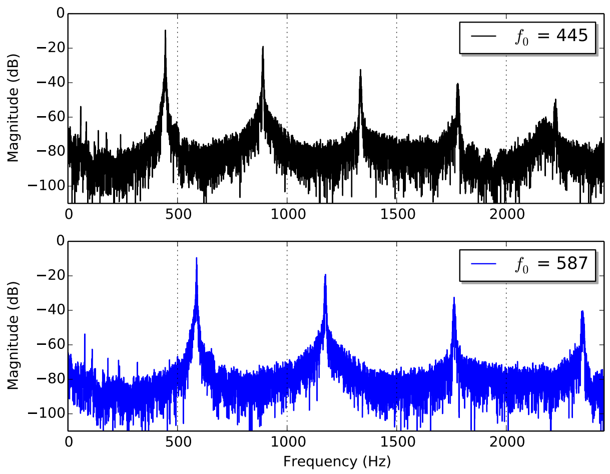

Figure 7: Pitch-shifting of A4 note (above) to D5 (below).

The pitch-shifting implementation via HBWT consisted of two main steps: **Step 1:** Design of the MDCT synthesis filter bank (Eq. 2) with an appropriate number of channels equals to $P_2 = \left\lfloor \frac{f_s}{f_2} \right\rfloor$. This step allowed the modulation of the original signal pitch to a new $f_0^{'}$ frequency. **Step 2:** Apply the original HWBT analysis coefficients as input to the synthesis bank. In Fig. 7, it can be appreciated the modification of the A4 quena note ($f_0 = 445$ Hz) to D5 ($f_0^{'} = 587$ Hz). This technique resulted in high-quality pitch-shifted signals that preserved the acoustical characteristics of the original quena sound.

3.3. Quena fractal analysis and comparison

The comparison between wind instruments quena, melodica and recorder was established by analysis of the fractal parameter γ_p ($1/f$ slope). In the experiments we compared the fractal characteristics of the frequency harmonics ($1/f$ noise) for all instruments playing the same note. The parameters γ_p were computed by linear regression according to Eq. 9.

In Fig. 8 are shown the results for the f_0 harmonic of C5 note ($f_0 = 523.25$ Hz). The results showed a strong $1/f$ fractal behavior for the main harmonics of all instruments. This fact was corroborated by a strong Pearson correlation coefficient (greater than the 80%) between the variables of the linear regression.

By including the analysis of the Hurst exponent, we found between 7 to 12 fractal harmonics on the quena, 12 to 19 fractal harmonics on the recorder and 20 to 40 fractal harmonics on the melodica. Additionally, it was found that the quena fractal slopes γ_p were closer to the recorder rather than the melodica. The latter was supported by the fact that the quena and recorder shared similar physical characteristics like the typical resonator tube of the flutes. In contrast, the melodica analysis returned greater differences with respect to the quena that were probably caused due to the differentiated free-reeds action mechanism and other distinctions in execution like the blowing air pressure or articulations. Based on the $1/f$ fractal analysis, these findings supported our initial hypotheses that γ_p fractal parameter was a useful attribute for the description of musical sound with potential as a feature for

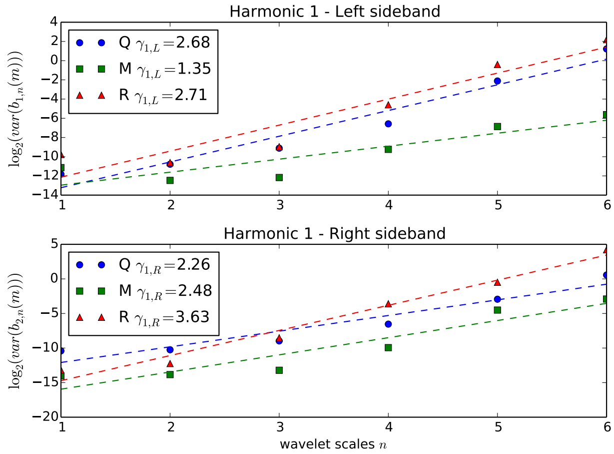

Figure 8: Quena (Q), melodica (M), and recorder (R) analysis of the left and right sidebands (channels $p = 1, 2$) for the first harmonic of C5 note with six subbands (wavelet scales $n = 1, \ldots, 6$).

musical instrument recognition and classification applications.

4. CONCLUSION

In this article it was presented an analysis of the quena by means of the *pseudo-periodic* $1/f$ *noise model* by using the *Harmonic Band Wavelet Transform*. The method and analysis described in this paper were performed for the first time on the andean wind instrument *quena*. By means of HBWT, it was obtained a suitable quena sound representation that enabled the development of interesting sound transformations and sound analysis applications.

The two sound transformation techniques we presented were: The *harmonic band* $1/f$ *noise filtering*, applied to the $1/f$ noise isolation or to the pure harmonic content reconstruction and the *pitch-shifting*, a modulation technique used to create new sound signals with a different fundamental frequency based on the analysis of a previous source. Both transformation techniques demonstrated acoustically interesting results. The last experiment was a comparison between the quena, recorder and melodica in terms of the fractal parameter γ_p. The results of the analysis pointed out that the γ_p fractal parameter was a useful attribute for sound description and characterization and a potential tool for musical instruments recognition and classification applications.

The results presented in this paper were part of a research project focused on the quena signal analysis and characterization based on distinctive fractal attributes. The methodology presented in this paper is scalable to the analysis of other types of musical instruments, thus we considered the study of additional andean instruments for the future. The source code of our Python implementation of the sound analysis/synthesis method described in section 2 and the sound examples to the experiments described in section 3 are available at the URL: `https://sites.google.com/site/aldodiazsalazar/`.

5. REFERENCES

[1] A. Zlatintsi and P. Maragos, "Multiscale Fractal Analysis of Musical Instrument Signals With Application to Recogni-

tion," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 4, pp. 737–748, Apr. 2013.

[2] S. Gunasekaran and K. Revathy, "Fractal dimension analysis of audio signals for Indian musical instrument recognition," in *2008 International Conference on Audio, Language and Image Processing*. July 2008, pp. 257–261, IEEE.

[3] A. Das and P. Das, "Fractal Analysis of different eastern and western musical instruments," *Fractals*, vol. 14, no. 3, pp. 165–170, 2006.

[4] P. Maragos, "Fractal signal analysis using mathematical morphology," in *Advances in Electronics and Electron Physics*, vol. 88, pp. 199–246. 1994.

[5] Pietro Polotti and Gianpaolo Evangelista, "Fractal additive synthesis," *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 105–115, Mar. 2007.

[6] P. P. Vaidyanathan, *Multirate Systems And Filter Banks*, Prentice Hall, New Jersey, first edition, 1993.

[7] X. Serra and J. Smith, "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic Plus Stochastic Decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12, Jan. 1990.

[8] G. Evangelista, "Pitch-synchronous wavelet representations of speech and music signals," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3313–3330, 1993.

[9] P. Polotti and G. Evangelista, "Fractal Additive Synthesis via Harmonic-Band Wavelets," *Computer Music Journal*, vol. 25, no. 3, pp. 22–37, Mar. 2001.

[10] S. Mallat, *A Wavelet Tour of Signal Processing, The Sparse Way*, Academic Press, third edition, 2008.

[11] T. Nguyen and R. Koilpillai, "The theory and design of arbitrary-length cosine-modulated filter banks and wavelets, satisfying perfect reconstruction," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 473–483, Mar. 1996.

[12] G. Wornell, "Wavelet-based representations for the 1/f family of fractal processes," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1428–1450, 1993.

[13] P. Polotti, *Fractal additive synthesis: spectral modeling of sound for low rate coding of quality audio*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2003.