THE MELLIN PIZZICATOR

Antonio De Sena, Davide Rocchesso

Dipartimento di Informatica
Università di Verona, Italia
desena@sci.univr.it, davide.rocchesso@univr.it

ABSTRACT

In this paper an application of the Mellin transform to the digital audio effects will be presented. Namely, low-pass and band-pass like filtering in the Mellin domain will be described and used for obtaining some kind of *pizzicato* effect on audio samples (musical instruments, but not only). The pluck and damp effects will be obtained using filtering in Mellin domain only. The algorithm used for implementing the Mellin (scale) transform has been presented in DAFx'04 [1].

1. INTRODUCTION

The effect showed in this paper is entirely based on filtering in the Mellin (scale) domain. The Mellin transform is the mathematical tool that allows to pass from time domain to the so called Mellin domain. The scale transform, a restriction of the Mellin transform, introduced by Cohen [2], can represent a signal in terms of *scale*. The scale can be interpreted, similarly to frequency, as a physical attribute of signals [2]. Thus, we can conceive digital audio effects that work by handling the signal in the scale domain, with transformation of the magnitude and/or phase of the scale image. This is technically feasible as long as fast and accurate realizations of these transforms are available.

The effect presented aims at simulating a *pizzicato* on a generic audio sample. So, for example, a flute can be "pizzicated" for obtaining some kind of plucked string instrument with spectral characteristic of the flute sample. This effect can be realized using two filters in the scale domain. The first filter dynamically cuts the frequencies (from the higher to the lower) as time increases, and the second filter simulates a pluck at the beginning of the sample. Both the filtering processes are realized in the scale domain, the first is a low-pass filter and the second is a band-scale enhancer.

It's important to emphasize the fact that this is an experiment in exotic domains (Mellin, scale) and the first objective is the exploration of these domains. So this paper doesn't want to introduce the best way to do this kind of effects, but an alternative to classical Fourier approaches.

In section 2 an introduction to Mellin and scale domains will be given with a description of the Fast Mellin Transform (FMT). In section 3.1 the scale magnitude low-pass, high-pass and band-pass filtering will be discussed and their behavior will be described. In sections 3.2 a description on how to obtain damp and pluck effects will be given and a discussion about a more classic implementation of the effect will be introduced in section 3.3. Finally, in section 4 experiments and results will be described and shown, and a comparison with a real pizzicato will be described.

2. MELLIN AND SCALE TRANSFORMS

The Mellin transform of a function f is defined as:

$$M_f(p) = \int_0^\infty f(t)\, t^{p-1}\, \mathrm{d}t\ , \tag{1}$$

where $p \in \mathbb{C}$ is the Mellin parameter. The scale transform [2] is a particular restriction of the Mellin transform on the vertical line $p = -jc + \frac{1}{2}$, with $c \in \mathbb{R}$. Thus, the scale transform is defined as:

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(t)\, \mathrm{e}^{(-jc-\frac{1}{2})\ln t}\, \mathrm{d}t. \tag{2}$$

The scale inverse transform is given by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty D_f(c)\, \mathrm{e}^{(jc-\frac{1}{2})\ln t}\, \mathrm{d}c. \tag{3}$$

The key property of the scale transform is the scale invariance. This means that if f is a function and g is a scaled version of f, the transform magnitude of both functions is the same. A scale modification is a compression or expansion of the time axis of the original function that preserves signal energy. Thus, a function $g(t)$ can be obtained with a scale modification from a function $f(t)$, if $g(t) = \sqrt{\alpha} f(\alpha t)$, with $\alpha \in \mathbb{R}^+$. When $\alpha < 1$ we get a scale expansion, when $\alpha > 1$ we get a scale compression. Given a scale modification with parameter α, the scale transforms of the original and scaled signals are related by

$$D_g(c) = \alpha^{jc} D_f(c). \tag{4}$$

This property derives from a similar property of the Mellin transform. In fact, if $h(t) = f(\alpha t)$, then

$$M_h(p) = \alpha^{-p} M_f(p). \tag{5}$$

In both (4) and (5), scaling is reflected by a multiplicative factor for the transforms, and for (4) such factor reduces to a phase difference. So, the scale transform magnitudes of the original signal and the scaled signal are the same.

$$|D_g(c)| = |D_f(c)|. \tag{6}$$

2.1. The Scale Interpretation

A parallel can be drawn between the properties of the Fourier and scale transforms. In particular, we can define a *scale periodicity* as follows: a function $f(t)$ is said to be scale periodic with period $\mathcal{T}$ if it satisfies $f(t) = \sqrt{\mathcal{T}} f(t\mathcal{T})$, where $\mathcal{T} = b/a$, with a and b starting and ending point of the scale period. $C_0 = 2\pi/\ln \mathcal{T}$ is the "fundamental scale" associated with the periodic function. By analogy with the Fourier theory, we can define a "scale series" and Parseval theorem [3].

For this work it is important to introduce an interpretation of the scale transform based on the scale decomposition of a signal. Like for the Fourier theory in which we can see signals like infinite sums of sines and cosines, for the scale theory we can interpret a signal as an infinite sum of time and frequency damped periodic functions (figure 1). So instead of sines and cosine we have the following functions[1]:

$$\mathrm{ddsin}(t) = \frac{\sin(c \ln t)}{\sqrt{t}} \tag{7}$$

$$\mathrm{ddcos}(t) = \frac{\cos(c \ln t)}{\sqrt{t}}, \tag{8}$$

where c is the scale.

So, the components of a signal at low scales are functions (all components start from time values grater then zero, and are damped in time by the coefficient $t^{-\frac{1}{2}}$) in which their frequencies go rapidly near low values (heavy damped in frequency as time increases). This fact will be used for explaining the filtering in 3.2.

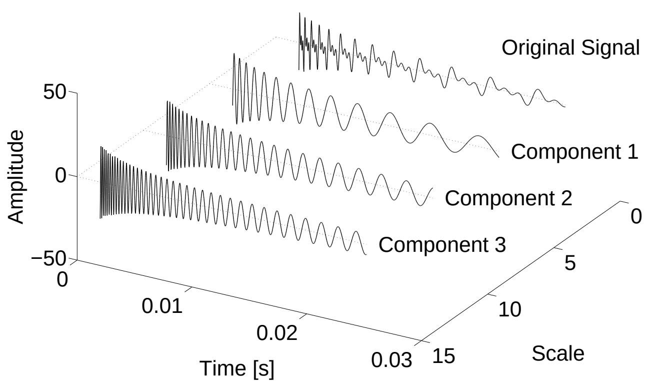

Figure 1: *Example of scale decomposition (Phase, energy and other concepts are not taken into account. Energies are normalized.).*

2.2. Relation with the Fourier transform

From its definition and interpretation, the Mellin transform provides a tight correspondence with the Fourier transform. More precisely, the Mellin transform with the parameter $p = -jc$ can be interpreted as a logarithmic-time Fourier transform. Similarly, we can define the scale transform of a function $f(t)$ using the Fourier transform of a function $g(t)$, with $g(t)$ obtained from $f(t)$ by time-warping f and multiplying the result by an exponential function. This result can be generalized for any p defined as $p = -jc + \beta$, with $\beta \in \mathbb{R}$. So, if $g(t) = \mathrm{e}^{t\beta} f(\mathrm{e}^t)$ with $\beta = \frac{1}{2}$ then:

$$D[f(t)] = F[g(t)], \tag{9}$$

where $F[\cdot]$ and $D[\cdot]$ refer to the Fourier transform and scale transform, respectively.

2.3. A Fast Mellin Transform

Practical modifications of signals in the Mellin domain can be achieved only if an accurate and fast discrete realization of the Mellin transform is available. We have realized[2] a Fast Mellin Transform (FMT [1]) by exploiting the analogy between the Mellin and Fourier transforms (section 2.2), as a sequence of exponential time-warping, multiplication by an exponential, and Fast Fourier Transform, as represented in figure 2. So the algorithm performs

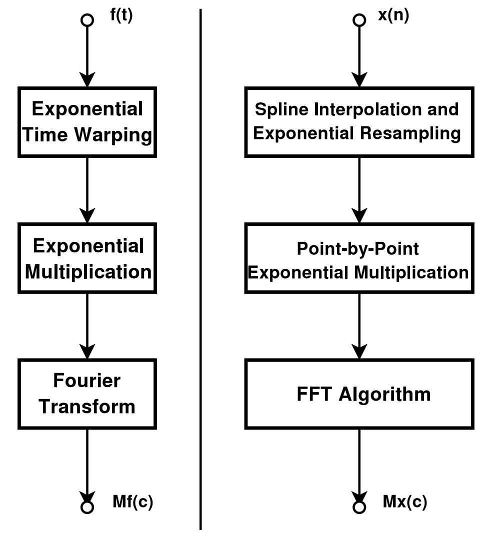

Figure 2: *Implementation of the Fast Mellin Transform: theory (left) and practice (right).*

an exponential resampling using a spline interpolator (cardinal or natural cubic [4]) for achieving the exponential time warping required, a point by point multiplication with an exponential function and a FFT (Fast Fourier Transform). The asymptotic complexity of the algorithm is $\mathcal{O}(n \ln^2 n)$, where n is the number of samples.

3. THE PIZZICATO EFFECT AND FILTERING IN SCALE DOMAIN

When a string (for example a guitar string [5]) is plucked we hear a sound. If analyzed from a time-frequency point of view (using a spectrogram for example) this signal presents a damping in time and in frequency. The signal looses energy (due to air resistance, losses at string termination, internal losses in the material due to viscoelastic losses, etc.), but the losses are not uniformly distributed in the spectrum. The attenuation works first (from a temporal point of view) on high frequencies and then on low frequencies.

This behavior can be simulated using filtering in Mellin (scale) domain. In the next sections we will explain how.

[1] Due to our purposes of magnitude filtering, the phase is not taken into account.

[2] The MATLAB code is available on http://profs.sci.univr.it/~desena/FMT/

3.1. Filtering in Scale Domain

In this section low-pass, high-pass and band-pass filters in scale domain will be introduced. Other types of filtering in scale domain can be found in [1] and a sample with water drops filtered using the scale transform has been submitted to the *Freesound Project*[3].

The low-pass filter can be performed by multiplying the transform magnitude by a window (like filtering in Fourier domain). The simplest window could be a rectangular window, that sets to zero all magnitude components that are found between a cutoff scale and the the signal maximum scale. Observing the results (the original signal spectrogram can be seen in figure 3, top), we can interpret this filter like a time-varying low-pass filter. The cutoff frequency approaches zero (hyperbolically, cf. [6]) as time increases and, at the same time, the amplitude of the signal is damped (see figure 4, top). The speed of frequencies cutting and the amplitude damping depend on the cutoff scale.

The high-pass filter behaves symmetrically, gradually moving the cutoff frequency toward zero (see figure 4, bottom).

The band-pass works in a midway, allowing to pass only certain frequencies at precise time instants. Viewing a band-pass in a spectrogram plot one could see only frequencies between two (time vs frequency) hyperbolic curves (see figure 3, bottom).

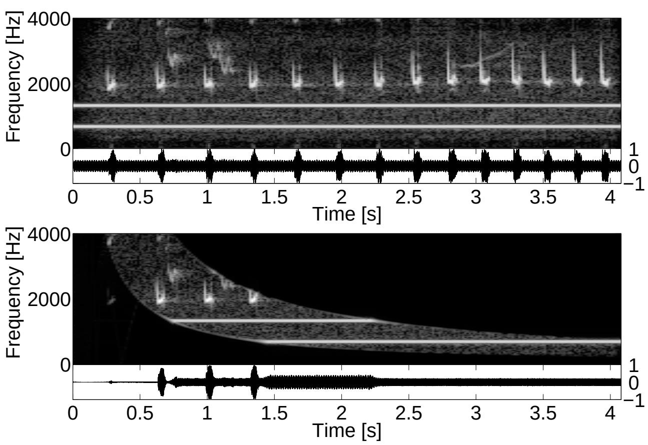

Figure 3: *Original signal (top) and band-pass filtered signal (bottom) spectrograms.*

3.2. The Damping Effect and the Plucking Effect

After the previous two sections we are now in the position to explain how the damping effect and the plucking effect have been obtained with the scale transform.

As already introduced, when plucked, a string starts to vibrate until the energy is dissipated. From a time and frequency viewpoint we can see the frequencies components approaches zero, with high frequencies converging faster. A similar effect can be obtained using a scale low-pass filter. So, the scale transform can be used to mimic the natural damping effect of a string. The effect can be driven using three parameters: the damping velocity (how fast frequencies are attenuated) that can be controlled modifying the bandwidth of the window (larger values meaning slower decays); the floor value of the window that affords an attenuation

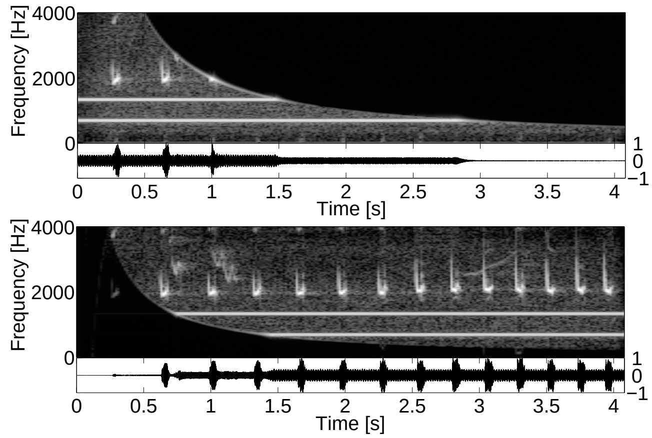

Figure 4: *Low-pass (top) and high-pass filtered signal (bottom) spectrograms.*

instead of a full cutting of out-of-band frequencies, and the window type (rectangular, Hann, etc.) that affects the way in which the frequencies are cut or attenuated over the time.

The plucking effect has been simulated using a band-enhancer that is like a band-pass but, instead of cutting, amplifies a range of scale components. This band-enhancer works at very low scale range with a very narrow band. So, only few scale components are enhanced and their frequencies go rapidly near zero (because they are ultra low scale components). Thus, an attack/pluck effect at the beginning of the signal will be produced. The parameters for this effect are similar to the previous one, but are more constrained. In particular the bandwidth of the window must be small and the central scale of the band-enhancer must be very small. There is also an enhancing factor that can be tuned to obtain heavier or softer attack/pluck effect.

3.3. A Comparison with STFT Approach

Even if the primary objective of the paper is the exploration of scale and Mellin domains, we can make a comparison of our framework with a classical Fourier approach. One way to obtain similar effects using a more classical approach is the use of the short time FFT/ IFFT-based analysis-synthesis scheme [7, 8]. The STFT gives obvious advantages from a computational viewpoint, but the filtering framework is more complex. In fact using the STFT one should build a low-pass filter sequence with the cutoff frequency that varies with time (high in the first applied filters, and low in the last filters). Hop size and windows types that allow a correct re-synthesis must be taken into account. So, the entire structure becomes bigger then the proposed one. Another advantage of the STFT approach is that it is a more general approach that allows a control over the "shape" of the time-frequency damping that in the Mellin case can only be hyperbolic, but the latter can produce a continuous variation of the frequency cut over time. This is something difficult to achieve with the STFT.

4. EXPERIMENTS AND RESULTS

The experiments have been mainly conducted on samples of musical instruments. The samples come from the University of Iowa

[3] http://freesound.iua.upf.edu/samplesViewSingle.php?id=17531

(http://theremin.music.uiowa.edu/MIS.html) and from the "McGill University Master Samp" cds (http://www.mcgill.ca/).
Other experiments can be done on non instruments samples, like vowels for instance, to try to "pizzicate" the vowels.
All the samples are WAV files with a sampling frequency of 44.1 kHz and with a quantization of 16 bit. A normalization has been performed after processing.
Let us introduce some definitions for general readers. The term *pizzicato* means an audio event similar to the sound produced by a string plucked (sounded) by fingers or a plectrum. *Percussive* sound indicates a sound similar to the hit of a body on a surface. On the contrary of the previous definition this sound does not involve a string instrument. Finally, *vibrato* means a tremulous or pulsating effect produced by minute and rapid variations in pitch.

In figure 5 a bowed cello sample has been processed using a low-pass window with minimum value set to zero. This implies that the scale components after the scale-cutoff value are deleted and not simply attenuated. The spectrogram shows the hyperbolic decay in time of the frequencies, emulating the damp effect. The attack/pluck effect is not clearly visible in the graphical representation, but can be seen in the time domain. The obtained sample sounds different from a true pizzicato cello string, but, from a perceptual point of view, the effect resembles a true string like behavior.

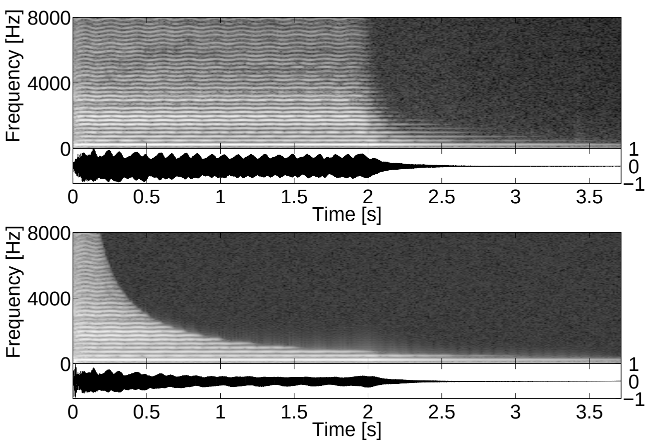

Figure 5: *Cello sample. Original sample (top) and filtered (bottom).*

In figure 6 a flute sample has been processed using a low-pass window with minimum value set to a value greater then zero. The spectrogram shows the hyperbolic decay in time of the frequencies, but, since the filter makes an heavy attenuation and not a full cut, the harmonics are still present and visible. This sound is perceived more as percussive rather then "pizzicato".

The third example (figure 7) is again a flute sample, but played with a vibrato effect, processed with a low-pass window with minimum value set to a value greater then zero. The same considerations of the previous example can be done. In this case more harmonics can be seen along their "wave" motion due to the vibrato.

In figure 8 a comparison between a true pizzicato and a synthetic pizzicato is shown. The two signals sound different. The synthetic hold all the effects bound to the original sound, like the bow and string interaction that does not exist in the true pizzicato.

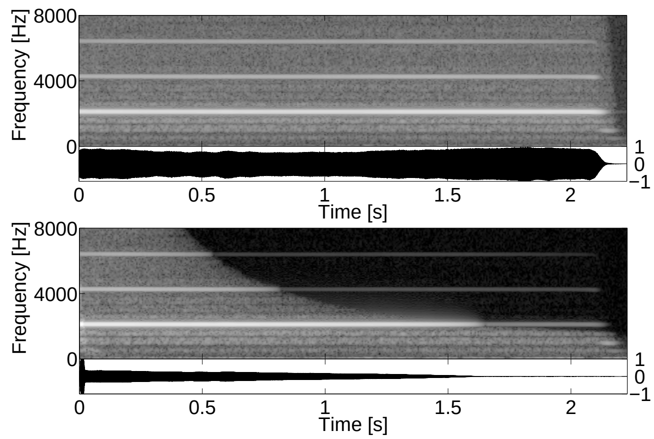

Figure 6: *Flute sample. Original sample (top) and filtered (bottom).*

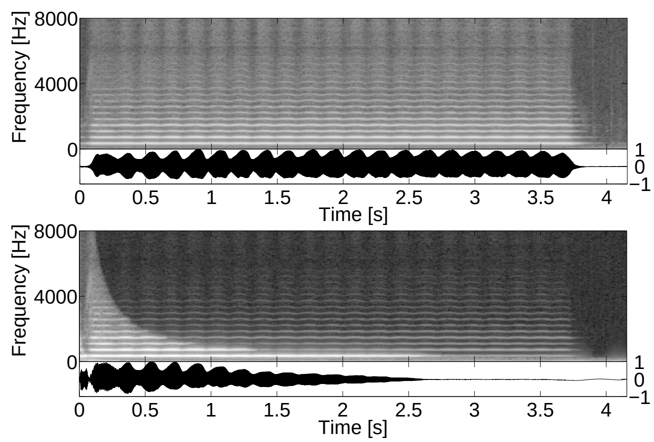

Figure 7: *Flute sample played with vibrato effect. Original sample (top) and filtered (bottom).*

The spectrogram shows that the actual pizzicato does not have a perfect hyperbolic decay in the time-frequency plane, even though a hyperbolic envelope can be traced over it.

All the examples shown that the artificial nature of the "Mellin pizzicato" can be immediately recognized if compared to a real pizzicato, especially with the help of time-frequency analysis. Is rather obvious that the real pizzicato decays in a more complex way. However, the artificial pizzicato gives the listener a strong sensation of pizzicato (or percussive sound). Other experiments have been done but are not presented in this paper, like an application of the pizzicator to vowels, thus producing an effect similar to plucking a string in the vocal tract.

5. CONCLUSIONS

The results obtained in this paper show how digital audio effects can be built in the Mellin/scale domain. In particular, this work shows how to simulate a "pizzicato" effect using only the scale domain. The results are satisfactory, in fact the artificial pizzicato gives to the listener an heavy sensation of pizzicato (or percussive

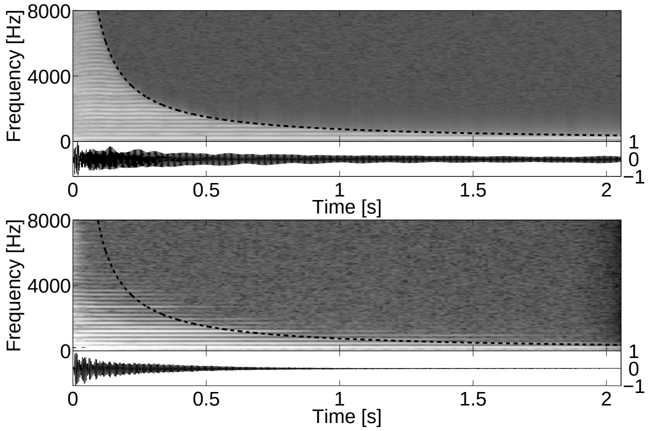

Figure 8: *Synthetic pizzicato (top) compared to real pizzicato (bottom). The idealized hyperbolic envelope of cutoff frequencies is drawn in dashed line.*

sound). A simple low-pass coupled with a band-enhancer in scale domain have been used to generate this effects. The entire process needs only few parameters that can be chosen with respect to the kind of damping and attack to simulate. Again, the first aim of this experiment is the exploration of the scale domain, in particular how to work using scale instead of more classical approaches, thinking at the scale like a physical attribute of the signal, interpreted like a joint time-frequency dimension.

6. REFERENCES

[1] A. De Sena and D. Rocchesso, "A fast Mellin transform with applications in DAFx," in *Proc. Int. Conf. on Digital Audio Effects (DAFx-04),* Naples, Italy, Oct. 2004, pp. 65–69.

[2] L. Cohen, "The scale representation," *IEEE Trans. Sig. Proc.*, vol. 41, no. 12, pp. 3275–3291, Dec. 1993.

[3] H. Sundaram, S. Joshi, and R. Bhatt, "Scale periodicity and its sampling theorem," *IEEE Trans. Sig. Proc.*, vol. 45, no. 7, pp. 1862–1864, July 1997.

[4] M. Unser, "Splines: A perfect fit for signal and image processing," *Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[5] V. Välimäki, J. Pakarinen, C. Erkut, and M. Karjalainen, "Discrete-time modelling of musical instruments," *Report on Progress in Physics*, vol. 69, no. 1, pp. 1–78, Nov. 2006.

[6] J. Bertrand, P. Bertrand, and J. Ovarlez, "The Mellin transform," in *The Transforms and Applications Handbook*, A. Poularikas, Ed. CRC Press LLC, Boca Raton, Florida, U.S. in cooperation with IEEE Press, 2000, pp. (11–53)–(11–54).

[7] D. Arfib, F. Keiler, and U. Zölzer, "Source-filter processing," in *DAFX – Digital Audio Effects*, U. Zölzer, Ed. J. Wiley & Sons, 2002, pp. 361–362.

[8] X. Amatrian, J. Bonada, A. Loscos, and X. Serra, "Spectral processing," in *DAFX – Digital Audio Effects*, U. Zölzer, Ed. J. Wiley & Sons, 2002, pp. 429–431.