ESTIMATION OF THE PARAMETERS OF SINUSOIDAL SIGNAL COMPONENTS AND OF ASSOCIATED PERCEPTUAL MUSICAL INTERPRETATION EFFECTS

Antony Schutz, Dirk T.M Slock

EURECOM Institute, 2229 Route des Crêtes 06901 Sophia Antipolis, France
firstname.lastname@eurecom.fr

ABSTRACT

The estimation of the parameters of a sinusoidal signal is an old problem and has been extensively studied in the literature. We propose here to use a well known method ; the estimation of the parameters by using the Fast Fourier Transform (FFT) and by taking into account the mutual influence of a part of the frequential peaks. When we extract the sinusoidal part of the signal, we obtain a noise composed, in the best case, by the background noise and the instrumental noise. In this situation estimating the SNR can allow us to detect some interpretation effect.

1. INTRODUCTION

The estimation of the parameters of a sinusoidal signal has been dealt with extensively in the literature [1, 2, 3]. The noise is obtained by subtracting the sinusoidal part of the signal, these leads to the conservation of residual noise like true noise and instrumental noise which has received some interest in the musical domain [4]. The estimate of the SNR will allow us, in the mono-instrumental case, to judge the quality of play of a student on certain acoustics instruments or detecting certain effects of interpretation.

2. ALGORITHM

The musical signal, which is by nature non-stationary, is piece-wise analyzed. The synthesis method consists of estimating the parameters of each frame, generating each partial signal by using the purely sinusoidal model and then reforming the complete signal by using an overlap and add method like in [5]. The noise is extracted by subtracting the synthesized signal from the original noisy signal. The method that we used is similar to the algorithm presented in [6] but here we use constant time windows and the parabolic interpolation made on the short term spectra. However the parabolic interpolation is very sensitive to the presence of some peaks and for reduce the interference we have to clean the spectrum peak by peak [7] in a certain frequency range. After this correction step, we can more correctly estimate the parameters and the result obtained are similar to those obtained by using an ESPRIT method [8].

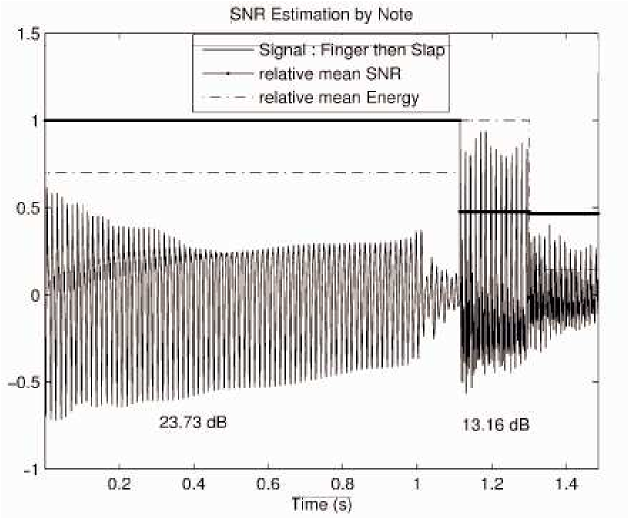

Figure 1: Estimation of the SNR : bass sequence of two notes (finger/slap).

3. APPLICATION TO ACOUSTIC INSTRUMENT

The audio signals are modeled as a sum of sinusoids with time varying parameters. However, in this model the nature of the instruments is completely ignored [9]. The sound of a wind instrument is composed of blow, for a violin the sound is generated by the friction of the bow on the string and more generally for the touch or string instrument the sound is due to striking. For example a bass player can play a note with the finger or by hitting the string with the thumb: the slap interpretation effect.

3.1. Bass: Slap Detection

The slap is a very common technique in bass playing. The strike consists of hitting the strings with the thumb, in the beginning of the fretboard, like a hammer. The resulting sound is almost completely percussive upon the strike and after the sinusoidal regime appears. Note that a note played by slap has a small duration compared to a note played with the finger.

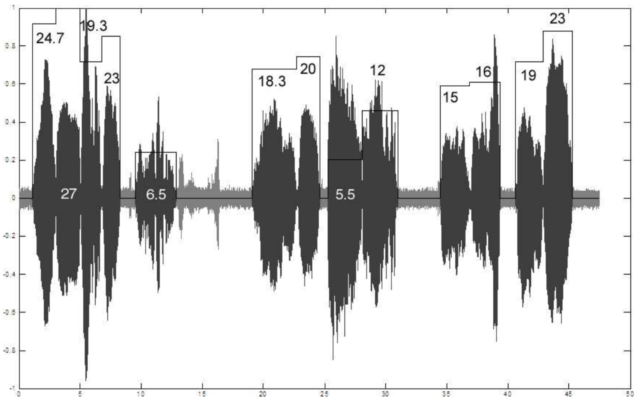

Figure 2: *A Violin piece played by a student*

Fig.1 shows the result of the SNR estimation on a bass sequence composed of two single notes. The first note is plucked with the finger and the second is played by slap. As we expected, the SNR of the slap note is small compared to the note played with the finger.

3.2. Violin: The Practice of the Bow

When a violinist plays, he moves the bow upon the strings and the sound is generated by the friction of this movement. A well played note is constrained by (at least) four parameters:

- The speed of the displacement of the bow.
- The pressure exerted on the string.
- And the two orientations of the bow (on itself and on the string).

A good player exercises constant speed and pressure, keeps the bow completely parallel on the string and the displacement orthogonal to the string. When one or more of these constraints are not respected, the sound becomes noisier. In the worst case, we only heard the displacement of the bow. Fig. 2 shows a succession of note play by a student. The black part corresponds to the detected note, the lines correspond to the relative SNR and the SNR is given by the value. We can observe the difference of the estimation of each note, so we can qualify a good note and a bad note. So, here, a good note has a SNR bigger than 18dB and a very bad note have a SNR smaller than 8dB. In practice all the thresholds have to be adjusting in accordance with the background noise.

4. DISCUSSION FOR SOME OTHER INSTRUMENT

Studying the noise is an interesting concept for musical applications. [4] has used the noise decomposition for estimating the tempo. This is due to the fact that in the noise, the attack is easier to detect. For a wind instrument, we can think of some applications as well. In fact, when a player blows into a wind instrument, the quality of the note is defined by two aspects: The constancy of the blow, which can be determined by the analysis of the evolution of the energy in the noise and the constancy of the frequency during the note.

With a clarinet, if the player blows strongly, the note becomes lower and it is easy to detect this. For an oboe, one can blow with all his strength without obtaining a note. For a string instrument, the noise will contain the attack which allows us to separate note with attack and note without attack (legato).

5. CONCLUSION

The work reported in this paper is by no means complete. Many more methods have to be including like a non-stationary model, noise whitening, a multi-scale analysis, an adaptive estimation of the parameters during the time evolution and more efficient parameter estimation. But we have chosen the low complexity method to allow the simulator to do some quick post treatment. However, this work opens some possibility for estimating some perceptual aspects of musical interpretation. The result of the analysis has to be adapted for each instrument due to the different thresholds but, the application can be implemented in a real time simulator and can be a pedagogic simulator for students who want to improve their playing skills.

10. REFERENCES

[1] F.Keiler, S.Marchand. Survey on extraction of sinusoids in stationary sounds. Proc. Of the 5th Int. Conf. on Digital Audio Effects (DAFx – 02).

[2] R.Althoff, F.Keiler and U .Zölzer. Extracting sinusoids from harmonic signals. (DAFx-99), NTNU, Trondheim 1999.

[3] E.Aboutanios. A modified dichotomous Search Frequency estimator. IEEE Signal Processing Letters, vol 11, N°2 2004.

[5] X.Serra. Musical sound modeling with sinusoids and plus noise. Musical signal processing. Editor C.Roads et al., Swets and Zeitlinger Publisher 1997.

[4] M.Alonzo, R.Badeau, B.David and G.Richard. Musical tempo estimation using noise subspace projection. IEEE Workshop on Application of Signal Processing to Audio and Acoustics 2003.

[6] G. Peeters, X. Rodet. Signal Characterization in terms of Sinusoidal and Non-Sinusoidal Components. Proceedings DAFX98 - Barcelona, Spain, November 19-21, 1998

[7] J.S.Marques, L.B Almeida. A bacground for sinusoid based representation of voiced speech. In Proc. ICASSP 1986.

[8] R.Badeau, G.Richard and B.David. Fast adaptive ESPRIT Algorithm. IEEE Workshop on Statistical Signal Processing 2005.

[9] N.H.Fletcher and T.D. Rossing. The Physics of Musical Instrument. Second Edition, Springer-Verlag 1998.