FIRST-ORDER GLOBAL AM-FM DECOMPOSITION AND APPLICATION TO MUSIC ANALYSIS AND TRANSFORMATION

*Mahdi Triki** and *Dirk T.M. Slock***

*Philips Research Laboratories, Eindhoven, The Netherlands
** Eurecom, Sophia Antipolis, France
Email: mahdi.triki@philips.com, dirk.slock@eurecom.fr

ABSTRACT

A refined estimation and tracking of the instantaneous frequency variations is desirable for a variety of audio applications (audio coding, singer segregation, music transcription and transformations, etc). In the present paper, we extend the periodic modeling with global amplitude and frequency modulation approach [16]. We introduce a first order approximation producing an additive term involving the derivative of the 'normalized waveform' multiplied by the instantaneous FM signal. The variations of the global FM get expressed through a subsampled representation and estimated using a simple least-squares scheme.

Index Terms— AM-FM decomposition, global modulation, frequency-selective, first-order approximation, audio transformation

1. INTRODUCTION

The problem of decomposing a signal into amplitude and frequency modulated sinusoids (AM-FM) is encountered in many different applications, such as audio coding, transformation, and segregation. Indeed, AM-FM sinusoidal models have been demonstrated to provide high-quality audio coding, and offers perceptually significant improvement in critical transient signals [1]. Moreover, the AM-FM representation leads to a flexible signal parameterization, often desirable for effect transformations such as time-scaling, pitch-shifting, and timbre-modification [2, 3]. More recently, it has been shown that a singing voice might be identified using the vibrato and tremolo characteristics: segments containing the singing voice could be localized through the analysis of the AM and FM signals [4]. In all these applications, the accuracy of both the instantaneous AM and FM signals estimation and tracking is crucial.

A variety of approaches has been proposed in the literature to perform AM-FM signal decomposition. Perhaps the most successful and ubiquitous is the family for time-frequency representations derived from the sinusoidal modeling paradigm [5]. These generally employ frame-based non-parametric spectral analysis techniques to detect peaks corresponding to sinusoidal-like components. Various techniques have been proposed for accurate peak localization based on non-linear interpolation (e.g. [6]), dichotomy (e.g. [7]) and/or high-resolution analysis (e.g. [8, 9]). Subsequently, these peaks are linked across consecutive time frames [12, 13] and/or coherent frequency bands [10, 11]. A second class of approaches address the AM-FM decomposition using multiband filtering and demodulation [14, 15, 1]. The basic idea is to first locate local spectral formants. Next, the instantaneous AM and FM signals are individually tracked for the different formants. In [16], we have introduced an alternative approach for AM-FM audio signal decomposition. Instead of addressing individually frames and/or formants, the harmonic structure and temporal consistency are both exploited to identify modulations that are common to all partials of a given sound. We have considered a periodic model with non-integer period and global AM and FM modulation (i.e., global amplitude variation and time-warping). The proposed scheme does not treat the harmonics of an audio signal separately as a simple filter bank approach would do. Rather, the energy in all harmonics is exploited jointly through the treatment of the complete periodic signal, in order to robustify the estimation of its modulation characteristics. The Global Modulation (GM) assumptions help the separation of audio signals that have harmonics in common. Furthermore, valuable information could be obtained by individually analyzing the model parameters. Indeed, global amplitude variation reflects mostly attack, sustain, and decay of the whole note signal, whereas global time-warping allows for the detection of musical effects (e.g. vibrato, glissando, etc). In [17] the GM representation was further developed by introduction a frequency-selective global amplitude modulation. The amplitude variations of the various harmonics are modeled using a short FIR filter that introduces a frequency-selective attenuation (allowing for different attack/decay modes), and this in a time-varying fashion to reflect the time-varying amplitude. Simulations show that the proposed scheme is suitable for the analysis of several string and wind instruments [17], and shows good potential for music transcription applications [18].

The frequency-selective global modulation leads to a parsimonious representation that efficiently models the different modes of instantaneous amplitude variation with a limited parameter rate (the average number of parameters that appear in the description of one second of the signal). This fact leads to a good estimation vs. modeling noise tradeoff and an effective signal decomposition. The proposed representation, however, allows only for piece-wise constant variation of the instantaneous FM. Indeed, the scheme assumes that the instantaneous fundamental frequency is constant within a frame of length T_f. In practice, the value of the instantaneous FM (in a given frame) is estimated via a local search optimization, and the computational complexity of the whole decomposition depends mainly on FM signal estimation, constraining T_f to be large. Hence, the piece-wise constant FM (expressed through a global time-warping) tracks mainly the very slow variations and the deviation of the fundamental frequency from its a priori value.

On the other hand, as discussed above, a refined estimation and tracking of the FM variations is desirable for a variety of audio applications. In this paper, we introduce a first-order approximation producing an additive term involving the derivative of the 'normal-

** Eurecom research is partially supported by its industrial members: BMW, Bouygues Télécom, Cisco Systems, France Télécom, Hitachi Europe, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The herein disclosed information is secret until March 14^{th} 2010; eyes only for the appointed reviewers of this conference.

ized waveform' multiplied by the instantaneous FM signal. The variations of the global FM get expressed through a subsampled representation and estimated using a simple least-squares scheme.

Notations: upper- and lower-case boldface letters denote matrices and vectors, respectively. As the quantities considered herein are real, $(.)^H$ represents the transpose as well as the complex-conjugate (Hermitian) transpose operators. The symbol T is reserved to denote the assumed period of the audio signal.

2. GLOBAL AM-FM SIGNAL REPRESENTATION

The sinusoidal transform, originally developed by Quatieri and McAulay [5], represents a signal as a sum of P discrete time-varying sinusoids or partials:

$$s(n) = \sum_{k=0}^{P} a_k(n) \cos\left(\psi_k(n)\right) \quad . \tag{1}$$

where $\psi_k(n)$ represents the instantaneous phase of the k^{th} partial. Since the energy of the audio signal is concentrated around the multiples of the fundamental frequency f_0, $\psi_k(n)$ can be decomposed into

$$\psi_k(n) = 2\pi k n f_0 \ + \ 2\pi \varphi_k(n) \tag{2}$$

where $\varphi_k(n)$ characterizes the evolution of the instantaneous phases around the k^{th} harmonic, and can be assumed to slowly vary over time. In [17], we have assumed that the time variations of the instantaneous amplitudes and frequencies of the different harmonics are correlated, and we have expressed the audio signal as a superposition of harmonic components with global frequency selective amplitude modulation and global time-warping, i.e.,

$$s(n) = a_n(q)\, \theta\left(n + \frac{\varphi(n)}{f_0}\right) \tag{3}$$

where:

- $a_n(q) = a_{n,L} q^L + \cdots + a_{n,0} + \cdots + a_{n,L} q^{-L}$ is a symmetric zero-phase FIR filter, and $2L+1$ denotes the amplitude modulating filter length. The introduction of q, where q^{-1} is the one sample time delay operator: $q^{-1}\theta(n) = \theta(n-1)$, allows to introduce a compact notation of transfer functions in the time domain.
- $\theta(n) = \sum_k a_k \cos\left(2\pi k f_0 n + \Phi_k\right)$ is a $T = \frac{1}{f_0}$ periodic signal (normalized waveshape), having a constant spectrum over the whole signal duration. $\theta(n)$ characterizes the spectral envelope of the audio source, and may be considered as a signature for the source (e.g., musical instrument) identification and recognition applications.
- $\varphi(n)$ denotes the global phase modulating signal. In [17], we have assumed that $\varphi(n)$ is piecewise linear, i.e. $\exists T_f$

$$\varphi_{wl}(n) = n\left(f_i - f_0\right) + \Phi_i \ \ \forall n \in [iT_f \ \ (i+1)T_f]$$

where $f_{wl}(n) + f_0 = f_i + f_0$ is the instantaneous FM signal assumed to be piece-wise constant. In such a case, the global phase modulation can be interpreted in term of global time-warping, and parameterized via a time-varying interpolation matrix (see [16] for a detailed description).

In this paper, we relax further our assumptions on the global phase/frequency modulating signal:

$$\varphi(n) = n f_0 + \varphi_{wl}(n) + \widetilde{\varphi}(n). \tag{4}$$

$\widetilde{\varphi}(n)$ allows a refined modeling of the global FM variations around the fundamental frequency, and it is assumed having a small relative magnitude $\widetilde{\varphi}(n)/f_0 \ll 1$. The audio signal can be approximated by:

$$\begin{aligned} s(n) &= a_n(q)\, \theta\left(n + \frac{\varphi_{wl}(n) + \widetilde{\varphi}(n)}{f_0}\right) \\ &\approx a_n(q)\left(\theta\left(n + \frac{\varphi_{wl}(n)}{f_0}\right) + \frac{\widetilde{\varphi}(n)}{f_0}\theta'\left(n + \frac{\varphi_{wl}(n)}{f_0}\right)\right). \end{aligned}$$

The first-order Taylor approximation of the phase dependence produces an additive term involving the derivative of the periodic signal $\theta'(n)$ multiplied by the fluctuating term $\dfrac{\widetilde{\varphi}(n)}{f_0}$.

Moreover, if one assumes that the first-order FM signal $\widetilde{\varphi}(n)$ varies slowly over time (compared to the order of FIR AM filter), the audio signal gets expressed as:

$$s(n) \approx a_n(q)\theta\left(n + \frac{\varphi_{wl}(n)}{f_0}\right) + \frac{\widetilde{\varphi}(n)}{f_0}\left(a_n(q)\theta'\left(n + \frac{\varphi_{wl}(n)}{f_0}\right)\right). \tag{5}$$

3. QUASI-PERIODIC SIGNAL EXTRACTION SCHEME

The audio signal is observed in presence of additive white Gaussian background noise, i.e.,

$$y(n) = s(n) + v(n) \tag{6}$$

In the following section, we will investigate the AM-FM tracking scheme and the extraction of the desired signal. First, we comment on the normalized waveshape derivative $\theta'(n)$, and the estimation of the first-order FM signal $\widetilde{\varphi}(n)$. Then, the extraction scheme of the quasi-periodic signal (following (5)) will be described.

3.1. FIR Derivative Approximation

The derivative $\theta'(n)$ denotes a sampled version of the derivative of the continuous-time signal of which $\theta(n)$ is the sampled version. If the sampling satisfies Nyquist's criterion, $\theta'(n)$ can be obtained from $\theta(n)$ by filtering with the transfer function $j2\pi f$, $f \in (-\frac{1}{2}, \frac{1}{2})$ which we shall approximate with a non-causal FIR filter $h(q) = \sum_{i=-L_h}^{L_h} h_i q^{-i}$ optimized as:

$$\min_{h_i} \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{yy}(f) \left| j2\pi f - h\left(e^{j2\pi f}\right)\right|^2 df \tag{7}$$

where L_h is the order of the FIR derivative filter approximation, and $S_{yy}(f)$ denotes the power spectrum of the observed signal $y(n)$.

The cost function (7) leads to a simple quadratic criterion. The optimal derivative filter coefficients $\mathbf{h} = [h_{-L_h} \cdots h_{L_h}]^H$ are given by:

$$\widehat{\mathbf{h}} = \mathbf{R}^{-1}\mathbf{p} \tag{8}$$

The elements of the matrix $\mathbf{R}$ and the vector $\mathbf{p}$ are expressed as $p_i = 2\pi \int_{-\frac{1}{2}}^{\frac{1}{2}} f S_{yy}(f) \sin(2\pi f i)\, df$ and $R_{ij} = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{yy}(f) e^{2\pi f(i-j)} df$ $= r_{yy}(i-j)$ (the covariance of $y(n)$ at the time-lag $i-j$); and get estimated by approaching the continuous by a discrete integration.

3.2. Global FM Parameterization and Estimation

If we assume that the normalized waveshape $\theta(n)$, the time-warping signal $\varphi_{wl}(n)$, and the FIR AM filter $a_n(q)$ are given, the signal $r(n) = y(n) - a_n(q)\theta(n + \frac{\varphi_{wl}(n)}{f_0})$ is linear with respect to $\widetilde{\varphi}(n)$. Moreover, the coefficients of $\widetilde{\varphi}(n)$ are assumed to be slowly varying over time. Therefore, $\widetilde{\varphi}(n)$ can be parameterized by a down-sampled version. The remaining samples could be interpolated, i.e.,

$$\widetilde{\boldsymbol{\varphi}} = \begin{bmatrix} \widetilde{\varphi}(1) \\ \widetilde{\varphi}(2) \\ \vdots \\ \vdots \\ \widetilde{\varphi}(N) \end{bmatrix} = \begin{bmatrix} 1 & 0 \cdots 0 \\ P_{21} P_{22} \cdots 0 \\ P_{31} P_{32} \cdots 0 \\ 0 & 1 \cdots 0 \\ \vdots \\ 0 \cdots \cdots 1 \end{bmatrix} \begin{bmatrix} \widetilde{\varphi}(1) \\ \widetilde{\varphi}(T_\varphi) \\ \vdots \\ \vdots \end{bmatrix} = \mathbf{P}_\varphi \, \widetilde{\boldsymbol{\varphi}}_\downarrow$$

where $\widetilde{\boldsymbol{\varphi}}_\downarrow$ characterizes the first-order FM signal $\widetilde{\varphi}(n)$, downsampled by the factor T_φ. $\mathbf{P}_\varphi$ represents the interpolation matrix used to reconstruct $\widetilde{\boldsymbol{\varphi}}$ from its downsampled version $\widetilde{\boldsymbol{\varphi}}_\downarrow$ (see [17] for more details on the design of the interpolation matrices).

Thus, using matrix notations, we have

$$\mathbf{r} \;=\; \mathbf{G}_\varphi \, \widetilde{\boldsymbol{\varphi}}_\downarrow + \mathbf{v} \tag{9}$$

where $\mathbf{G}_\varphi = \text{ diag } \left(\frac{1}{f_0} a_n(q) \theta^{'}(n + \frac{\varphi_{wl}(n)}{f_0}) \right) \mathbf{P}_\varphi$.

Assuming additive white gaussian noise, the ML approach lead to the least-squares solution:

$$\widehat{\widetilde{\boldsymbol{\varphi}}}_\downarrow = \left(\mathbf{G}_\varphi^H \mathbf{G}_\varphi \right)^{-1} \mathbf{G}_\varphi^H \mathbf{r} \tag{10}$$

Remark that $\mathbf{G}_\varphi$ assumes the knowledge of $a_n(q)$ (or an accurate estimate). Alternatively, we consider $\mathbf{G}_b = \text{ diag } \left(\theta^{'}(n + \frac{\varphi_{wl}(n)}{f_0}) \right) \mathbf{P}_\varphi$ and the statistics

$$\mathbf{b}_\downarrow = \left(\mathbf{G}_b^H \mathbf{G}_b \right)^{-1} \mathbf{G}_b^H \mathbf{r} \tag{11}$$

Assuming that $a_n(q)$ has a smooth frequency response and that $a_n(q)$ and $\widetilde{\varphi}(n)$ vary slowly over time, one can show that:

1) $\mathbf{b}_\downarrow$ is a sufficient statistics for $\widetilde{\boldsymbol{\varphi}}_\downarrow$ [19].

2) $\mathbf{b}_\downarrow \approx \frac{\left\langle a_n(q)\theta^{'}(n) \right\rangle_{T_\varphi}}{f_0 \left\langle \theta^{'}(n) \right\rangle_{T_\varphi}} \, \widehat{\widetilde{\boldsymbol{\varphi}}}_\downarrow$; where $\langle . \rangle_{T_\varphi}$ denotes temporal averaging over the T_φ-wide interval centered on n.

Compared to $\widehat{\widetilde{\boldsymbol{\varphi}}}_\downarrow$, $\mathbf{b}_\downarrow$ incorporates 'all the relevant' information on the time varying amplitude modulating filter $a_n(q)$ and does not need prior information on the filter coefficients (only $\theta(n)$ and $\varphi_{wl}(n)$ are assumed known). Figure 1 plots the time evolution of input signal (musical note played by an acoustic guitar and sampled at 20.050 kHz) and the extracted signal $\mathbf{b} = \mathbf{P}_\varphi \mathbf{b}_\downarrow$. One may indeed observe that the $b(n)$ captures both the vibrato effect (FM information) and the decay of the note signal (AM information).

Similarly, the coefficient of the amplitude modulating filter $a_n(q)$ could be parameterized via a downsampled representation $\{\mathbf{a}_{\downarrow i}\}_{i=0:L}$ (possibly with a different downsampling factor T_a) [17]. The received signal $\mathbf{y}$ could be expressed as a linear combination of the AM and FM downsampled parameter

$$\mathbf{y} \cong [\mathbf{G}_{a0} \cdots \mathbf{G}_{aL} \; \mathbf{G}_b] \begin{bmatrix} \mathbf{a}_{\downarrow 0} \\ \vdots \\ \mathbf{a}_{\downarrow L} \\ \mathbf{b}_\downarrow \end{bmatrix} + \mathbf{v} \tag{12}$$

The matrices $\mathbf{G}_{-}$ share the same structure, a diagonal matrix right multiplying an interpolation matrix. The estimation of these parameters could be jointly performed using a least-squares scheme. Remark that from a computational complexity, the tracking of the first-order FM signal is equivalent to raising the order of the AM filter by 1.

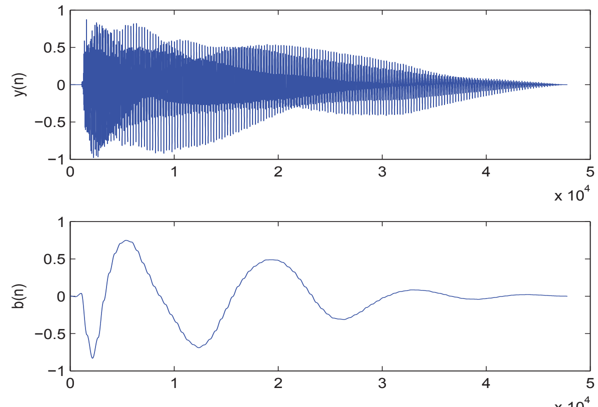

Fig. 1. Time evolution of of input signal (top) and the extracted first-order FM signal (bottom).

3.3. Quasi-Periodic Signal Extraction

Following [16, 17], the model parameters are estimated in an iterative (cyclic) fashion:

1. Assuming that $\widehat{a}_n(q)$, $\widehat{\varphi}_{wl}(n)$, $\widehat{\widetilde{\varphi}}(n)$ are given and using the derivative filter approximation, the received signal could be expressed as a linear function of the normalized waveshape $\theta(n)$. Hence, the periodic signature is recovered via least-squares estimation.
2. Using the current estimates of $\widehat{a}_n(q)$, $\widehat{\theta}(n)$ and $\widehat{\widetilde{\varphi}}(n)$, The piece-wise FM signal $\varphi_{wl}(n)$ gets estimated on a frame-by-frame basis using a local search scheme (as described in [17]).
3. Using the current estimate of $\widehat{\theta}(n)$ and $\widehat{\varphi}_{wl}(n)$, the amplitude modulating filter $a_n(q)$ and the first-order FM signal $\widetilde{\varphi}(n)$ get estimated through the joint estimation of $\{\mathbf{a}_\downarrow \, , \, \mathbf{b}_\downarrow\}$.

4. EXPERIMENTAL RESULTS

We validate the proposed extraction approach using real musical signals. The audio signals were played by an acoustic guitar, recorded at 44.100 kHz, then downsampled to 22.050 kHz.

We have compared the extraction Signal-to-Noise Ratio $SNR_{out} = \frac{\sum_n s(n)^2}{\sum_n (s(n) - \widehat{s}(n))^2}$ using the proposed global modulation representations, with and without first order approximation (denoted $QPSE$ and $QPSE_{mod}$, respectively). The smoothing AM and FM modulation factors were set to $T_a = T_\varphi = 3T$ ($T = ceil(1/f_0)$ is the period of the harmonic component, assumed known). The order of the derivative filter was selected $L_h = 5$. For a fair comparison, the order of the amplitude modulation filters of $QPSE_{mod}$ and $QPSE$ were set to L and $L+1$ respectively (such as the two models have the same total degree of freedom).

Fig. 2 and 3 plot the extraction results for two different notes (B3 and D3). As a reference, we have compare the GM-based schemes to time-frequency representation. The desired signals were retrieved using an ABSOLA analysis/synthesis algorithm (with peaks interpolation and tracking) [3, Ch10]. In the time-frequency processing, the block size, zero-padding factor, and maximum number of sinusoids were set to 512, 2, and 32 respectively (the signals were segmented using a Hamming window with 50% overlap). The extracted first-order FM signal $b(n)$ is also shown in the bottom of the figures.

Curves show that first-order FM representation allows better extraction accuracy. This enhancement cannot be achieved by further increasing the order of the AM modulating filter: the first-order modulating signal $b(n)$ models and tracks the instantaneous frequency

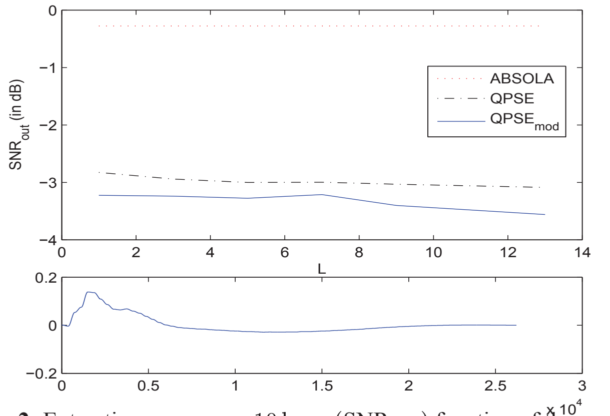

Fig. 2. Extraction accuracy $-10\log_{10}(\mathrm{SNR_{out}})$ function of the amplitude modulating order L (top) and the extracted first-order FM signal $b(n)$(bottom).

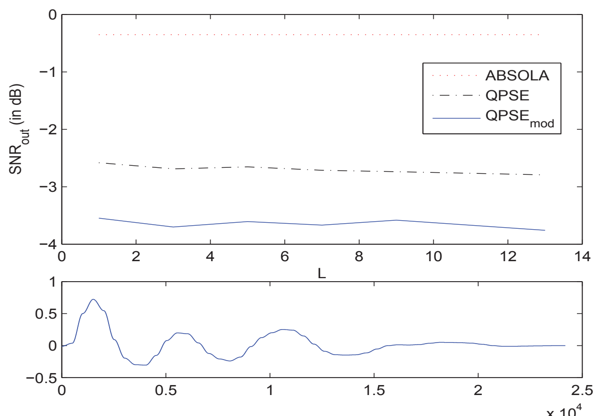

Fig. 3. Extraction accuracy $-10\log_{10}(\mathrm{SNR_{out}})$ function of the amplitude modulating order L (top) and the extracted first-order FM signal $b(n)$(bottom).

fluctuations, while the amplitude modulating filter $a_n(q)$ focuses on amplitude variations. The enhancement, however, depends on the frequency effects present in the signal. Indeed, larger improvement is achieved when a vibrato effect is played (Fig. 3).

Accurate FM variation tracking is a key building block in several sound analysis and transformation systems. With this respect, the representation presented herein offers an intuitive interpretation of the model parameters and gives direct access to the perceptual attributes that are used to control a large variety of audio effects (loudness, pitch and timbre). The analysis/synthesis scheme also offers the possibility to apply several effects with at least similar flexibility as the classic techniques. To validate the effectiveness of the proposed representation, we have implemented a vibrato effect. The vibrato is a common effect for various acoustical instruments. It is used for emphasis and timbral variety, and is considered as a low frequency modulation applied to the frequency of the partials (with a constant spectral shape), i.e.,

$$f_k(n) = m(n).f_k \tag{13}$$

where $m(n) = 1 + \Delta_a.\sin(2\pi\Delta_f n)$ is the vibrato modulating signal, Δ_f and Δ_a represent the vibrato frequency and depth, respectively. We have compared the proposed scheme with the time-frequency representation based implementation [3, Ch10]. Vibrato effects were synthetically added to monophonic acoustic guitar notes (recorded in real environment and sampled at 22.050 kHz). Δ_f and Δ_a were set to 5Hz and 0.1 respectively. Subjective tests with 10 persons (musical and/or audio processing experts) show a preference for the GM based approach (GM was prefered by 75%, compared to 16% for time-frequency processing). The preference was even more pronounced in case of polyphonic sounds (e.g. chords transformation).

5. REFERENCES

[1] M.G.Christensen, S. van de Par,"Efficient Parametric Coding of Transients," *IEEE Trans. on Audio, Speech, and Language Processing*, Jul. 2006

[2] M. Goodwin and M. Vetterli,"Time-Frequency Signal Models for Music Analysis, Transformation,and Synthesis," *In Proc. of TFTS*, Jun. 1996.

[3] U. Zölzer (Ed.), "DAFX - Digital Audio Effects," *John Wiley & Sons*, 2002.

[4] L. Regnier and G. Peeters, "Singing Voice Detection in Music Tracks using Direct Voice Vibrato Detection," *In Proc. of ICASSP*, Apr. 2009.

[5] R. McAulay and T. Quatieri, "Speech Analysis/Synthesis Based on a Sinusoidal Representation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Aug. 1986.

[6] F. Keiler and S. Marchand,"Survey On Extraction Of Sinusoids In Stationary Sounds," *In Proc. of DAFX*, Sept. 2002.

[7] E. Aboutanios,"A modified Dichotomous Search Frequency Estimator," *Signal Processing Letters*, Feb. 2004.

[8] M.G. Christensen, A. Jakobsson, and S.H. Jensen, "Joint High-Resolution Fundamental Frequency and Order Estimation," *IEEE Trans. on Audio, Speech, and Language Processing*, Jul. 2007.

[9] R. Badeau, B. David, and G. Richard, "High-Resolution Spectral Analysis of Mixtures of Complex Exponentials Modulated by Polynomials," *IEEE Trans. on Signal Processing*, Apr.2006.

[10] X. Wen, "Harmonic Sinusoid Modeling of Tonal Music Events," *PhD Thesis*, University of London, 2007.

[11] R. Gribonval and E. Bacry, "Harmonic Decomposition of Audio Signals with Matching Pursuit," *IEEE Trans. on Signal Processing*, Jan. 2003.

[12] X. Serra and J. Smith, "Spectral Modeling Synthesis:A Sound Analysis/Synthesis Based on a Deterministic plus Stochastic Decomposition," *Computer Music Journal*, 1990.

[13] T. Virtanen and A. Klapuri,"Separation of Harmonic Sound Sources Using Sinusoidal Modeling," *In Proc. of ICASSP*, 2001.

[14] A. Potamianos and P. Maragos, "Speech Analysis and Synthesis Using an AM-FM Modulations Model," *Speech Communication*, Jul. 1999.

[15] R.B. Sussman and M. Kahrs,"Analysis and Resynthesis of Musical Instrument Sounds Using Energy Separation," *In Proc. of ICASSP*, May. 1996.

[16] M. Triki and D.T.M. Slock, "Periodic Signal Extraction with Global Amplitude and Phase Modulation for Music Signal Decomposition," *In Proc. of ICASSP*, March 2005.

[17] M. Triki and D.T.M. Slock, "Periodic Signal Extraction with Frequency-Selective Amplitude Modulation and Global Time-Warping for Music Signal Decomposition," *In Proc. of MMSP*, Oct. 2008.

[18] M. Triki and D.T.M. Slock, "Perceptually Motivated Quasi-Periodic Signal Selection for Polyphonic Music Transcription," *In Proc. of ICASSP*, Apr. 2009.

[19] W. Bischoff, H. Cremers, and W. Fieger, "A Characterization of the Normal Distribution by Sufficiency of the Least Squares Estimation," *Metrika*, Dec. 1987.