Noise Robust AM-FM Demodulation using Least-Squares Truncated Power Series Approximation

Wooi-Boon Goh
School of Computer Engineering,
Nanyang Technological University, Singapore 639798
aswbgoh@ntu.edu.sg

ABSTRACT

Abstract - **This paper describes a parametric approach for demodulating multicomponent AM-FM sinusoidal signals. The parametric model consist of several truncated power series whose coefficients are estimated using linear least-square minimisation cast within an iterative parameter-substitution framework. It is shown that the proposed technique for AM-FM estimation is relatively robust to background noise compared to other techniques such as the smoothed energy operator separation algorithm and the Hilbert transform method.**

Keywords – **Signal Processing, AM/FM Demodulation**

I. INTRODUCTION

The analysis of an AM-FM signal usually involves the estimation of the signal's time-varying amplitude envelope and instantaneous frequency. AM-FM demodulation techniques can be classified into parametric and non-parametric methods. AM-FM parametric models are generally not popular because the parameter estimation process is non-linear in nature and difficult to solve [1]. Grenier [2], Mukhopadhyay and Sircar [3] have proposed using time-dependent auto-regressive moving average (TARMA) models and attempted to linearised the parameter estimation process by expressing the time-dependent AR coefficients in terms of some suitable basis functions such as the Fourier and Chebyshev basis functions. Problematic issues such as the optimum choice of basis functions and the estimation of the time-varying AR coefficients under noisy conditions still remain [3]. On the other hand, non-parametric methods such as the Hilbert transform method [4] and the energy separation algorithm [5]-[7] are popular and have been applied to speech resonance modeling and musical instruments analysis [8]-[10]. Quatieri *et al.* proposed a technique for AM-FM separation based on the relative amplitude envelope of the output of two linear transduction filters [6]. Non-parametric methods are generally simpler and computationally more efficient but are not as noise robust compared to parametric methods [3]. Additionally, many parametric methods [1], [11] can be readily extended to signals containing multiple AM-FM sinusoids. Non-parametric methods work-around this problem by applying appropriate bandpass filtering (e.g. Gabor filter) to the signal in order to extract each of the sinusoids before the demodulation algorithm is applied to each extracted component in turn [5], [8]-[9]. But such filters introduce amplitude distortions since not all frequencies with the bandwidth of the filter are equally attenuated.

This paper proposes a parametric method for demodulating multicomponent AM-FM sinusoidal signals. The proposed least-squares truncated power series approximation (L-STPSA) model is developed in section 2. In section 3, the performance of the proposed parametric model is evaluated for different orders of L-STPSA model and under various noise conditions. An application of the L-STPSA model to singing voice (multicomponent) analysis is also presented.

II. THE L-STPSA AM-FM MODEL

An AM-FM signal $x(n)$ can represented by

$$x(n) = A(n)\cos[w_c n + \theta(n)] \tag{1}$$

where w_c is a fixed carrier frequency with time-varying amplitude $A(n)$. The instantaneous frequency $w(n)$ is the derivative of the time-varying phase $\theta(n)$ given by

$$w(n) = w_c + [d\theta(n)/dn] \tag{2}$$

Given a signal $x(n)$ and the carrier frequency w_c, the classical AM-FM demodulation problem is to estimate the amplitude envelope $|A(n)|$ and the instantaneous frequency $w(n)$.

A. The AM Model

We start with a simple AM signal model given by

$$\tilde{x}(n) = A(n)\cos[w_c n + \varphi] \tag{3}$$

where the carrier frequency w_c is known to be present in the AM-FM signal $x(n)$. $A(n)$ is the time-varying amplitude and φ is some phase value. We expand (3) into its odd and even quadrature components given by

$$\tilde{x}(n) = a(n)\cos(w_c n) + b(n)\sin(w_c n) \tag{4}$$

where $a(n)$ and $b(n)$ are given by

$$a(n) = A(n)\cos(\varphi) \quad \text{and} \quad b(n) = -A(n)\sin(\varphi) \tag{5}$$

Here we assume the functions describing the time-varying amplitude $A(n)$ and phase $\theta(n)$ of the signal are *analytic*, which means a derivative exist at all points on the function [12]. Analytic functions can be described by a power series. For example $\cos x$ is given by

$$\cos x = 1 - x^2/2! + x^4/4! - x^6/6! + - \ldots \tag{6}$$

With this assumption, it is proposed that the quadrature components of $\tilde{x}(n)$ be modeled as a general truncated power series (with centre at 0) of order P given by

$$a(n) = \sum_{k=0}^{P} n^k a_k \quad , \quad b(n) = \sum_{k=0}^{P} n^k b_k \qquad (7)$$

Given a signal $x(n)$ of sample length N, the modeled signal $\tilde{x}(n)$ in (4) can be estimated by minimising the mean square-error ε in (8) with respect to the (P+1) pairs of amplitude coefficients (a_0 to a_k, b_0 to b_k) and a predetermined carrier frequency of ($w = w_c$), as detailed in [13].

$$\varepsilon_A = \sum_{n=1}^{N} \{\tilde{x}(n) - x(n)\}^2 \qquad (8)$$

B. The AM-FM Model

If the signal being modelled is a pure AM sinusoidal signal, the phase component φ of $\tilde{x}(n)$ in (3) would be a constant. Since our concern is with signals that are modulated both in amplitude and frequency, then the instantaneous phase φ will be time-varying and can be described by a time-varying function $\hat{\theta}(n)$ given by

$$\hat{\theta}(n) = \arctan[-b(n)/a(n)] \qquad (9)$$

where $a(n)$ and $b(n)$ are the signal's quadrature amplitude components given in (7). Starting with an AM model, the time-varying functions $a(n)$ and $b(n)$ are computed for a fixed carrier frequency w_c using a least-square truncated power series approximation (L-STPSA) model of order P. These values are then substituted into (9) to obtain a preliminary estimate of the time-varying instantaneous phase function $\hat{\theta}(n)$. This unwrapped phase function $\hat{\theta}(n)$ is incorporated into the AM-FM signal model given in (1) by tracking the 2π jumps in the phase values obtained in (9). The unwrapped phase is then modeled as an analytic function using a L-STPSA of order Q in the same way the L-STPSA quadrature amplitude functions were computed earlier. The Q-order phase function $\hat{\theta}(n)$ is then substituted into (1) as $\theta(n)$. Like before, the AM-FM signal model in (1) can be expanded into its odd and even components, given by

$$\tilde{x}(n) = a(n)\cos[w_c n + \theta(n)] + b(n)\sin[w_c n + \theta(n)] \qquad (10)$$

Similar to the AM model, the time-varying functions $a(n)$ and $b(n)$ of the AM-FM model are obtained from the (P+1) pairs of amplitude coefficients (a_0 to a_k, b_0 to b_k) using the least-square minimization. However, this time around, instead of using a fixed carrier frequency of (w=w_c), a time-varying frequency of (w=[$w_c n$+ $\theta(n)$]) is used. To obtain a new expression for the instantaneous phase of the AM-FM signal model, (10) is expanded using suitable trigonometric identities to the form

$$\begin{aligned}\tilde{x}(n) = \; & a(n)[\cos(w_c n)\cos(\theta(n)) - \sin(w_c n)\sin(\theta(n))] + \\ & b(n)[\sin(w_c n)\cos(\theta(n)) + \cos(w_c n)\sin(\theta(n))] \qquad (11)\end{aligned}$$

Rearranging (11), the quadrature components of $\tilde{x}(n)$ are

$$\begin{aligned}\tilde{x}(n) = \; & [a(n)\cos(\theta(n)) + b(n)\sin(\theta(n))]\cos(w_c n) + \\ & [b(n)\cos(\theta(n)) - a(n)\sin(\theta(n))]\sin(w_c n) \qquad (12)\end{aligned}$$

A new time-varying phase function $\hat{\theta}(n)$ can be obtained like in (8) and is given by

$$\hat{\theta}(n) = \arctan\left[\frac{-b(n)\cos(\theta(n)) + a(n)\sin(\theta(n))}{a(n)\cos(\theta(n)) + b(n)\sin(\theta(n))}\right] \qquad (13)$$

This newly estimated phase function $\hat{\theta}(n)$ is then unwrapped, smoothened using a L-STPSA model and substituted back into the signal model in (1) to obtain new values of $a(n)$ and $b(n)$. These new values of $a(n)$ and $b(n)$ in turn, allow the computation of another time-varying phase function $\hat{\theta}(n)$ using (13). This iterative parameter-substitution process is repeated until the AM-FM signal model $\tilde{x}_M(n)$ obtained in the Mth iteration closely resembles that in the (M+1)th iteration. The termination criterion for this iterative process could be a simple mean-square error criterion given by

$$\frac{1}{N}\sum_{n=1}^{N}\left[\tilde{x}_M(n) - \tilde{x}_{M+1}(n)\right]^2 \le \varepsilon_M \qquad (14)$$

where the error threshold ε_M is some suitable small value that is obtained empirically. It is dependent on the average signal magnitude. Fixed iteration counts could also be used. For the signals presented in this paper, reasonable convergence was obtained in ≤ 20 iterations. The L-STPSA parametric model completely described an AM-FM sinusoid as

$$\tilde{x}(n) = \left(\sum_{k=0}^{P} n^k a_k\right)\cos\left[w_c n + \sum_{k=0}^{Q} n^k c_k\right] + \left(\sum_{k=0}^{P} n^k b_k\right)\sin\left[w_c n + \sum_{k=0}^{Q} n^k c_k\right] \qquad (15)$$

where (a_0 to a_k, b_0 to b_k) are the (P+1) amplitude coefficients pairs and (c_0 to c_k) are the (Q+1) phase coefficients. The proposed parametric model can describe a single component AM-FM signal segment $x(n)$ compactly using (Q+2P+3) real-valued coefficients and a fixed-carrier frequency component w_c. The amplitude envelope of the AM-FM signal $x(n)$ is given by

$$|A(n)| = \left[\left(\sum_{k=0}^{P} n^k a_k\right)^2 + \left(\sum_{k=0}^{P} n^k b_k\right)^2\right]^{1/2} \qquad (16)$$

The time-varying frequency $w(n)$ in (2) is given by

$$w(n) = w_c + \sum_{k=0}^{Q-1}(k+1)c_{k+1}n^k \qquad (17)$$

where the derivative of the analytic instantaneous phase function $\theta(n)$ can be obtained by differentiating the associated power series term by term [12]. If the phase function $\theta(n)$ is available, backward-difference will provide the instantaneous frequency and it is given by

$$w(n) = w_c + \theta(n) - \theta(n-1) \qquad (18)$$

III. EXPERIMENTAL RESULTS

A. Order of L-STPSA Model

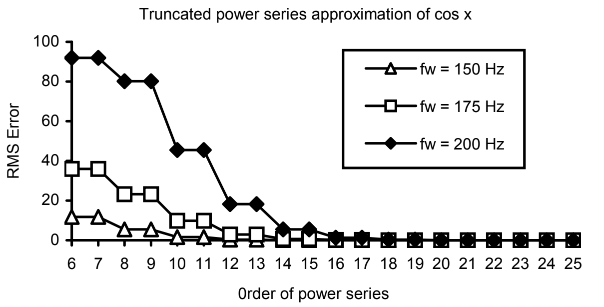

Figure 1. RMS errors in modeling $\cos x$ at different frequencies using a L-STPSA model of various orders.

Fig. 1 shows the root-mean-squared (RMS) errors between the function $f(n) = \cos(2\pi n f_w / f_s)$ and its truncated power series approximated equivalent given in (6) over 300 samples. Function $f(n)$ was computed with only even powers as the order of the power series was varied from 6 to 25. The frequency f_w of the sinusoidal function was set at 150, 175 and 200 Hz, with the sampling frequency f_s = 22050 Hz. As expected, when the order of the truncated power series increases, the accuracy of the estimated signal model improves as can be seen by the rapidly reducing RMS errors. However, it is also observed that for a given order, the RMS error increases almost logarithmically with a linear increase in the signal's frequency.

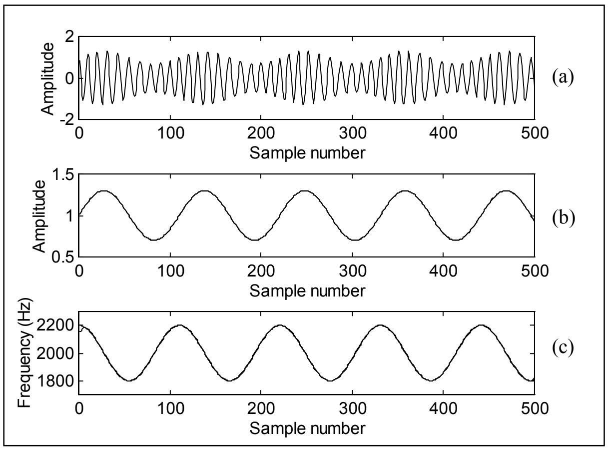

Figure 2. AM-FM estimates with L-STPSA model of orders (P=Q=20). (a) Original test signal. The extracted (b) amplitude modulation and (c) instantaneous frequency. Original signals are in solid lines and estimated signals are in dash lines (not visible as they are overlapping the solid lines).

In order to observe the accuracy of the L-STPSA AM-FM signal model, the RMS errors were measured for different orders of L-STPSA models. These results were obtained for a sinusoidal AM-FM signal with carrier frequency w_c=2kHz, FM and AM frequencies of w_{fm}=200Hz and w_{am}=200Hz respectively (see Fig. 2a). Fig. 3a and 3b show the RMS errors in the amplitude envelope and instantaneous frequency estimates for model orders P = Q from 6 to 25. Except for lower order values, the error characteristics in Fig. 3 are similar to those observed in Fig. 1, exhibiting a sharp drop in RMS error as the order is increased. With AM and FM frequencies of 200 Hz, P and Q orders of 20 would be sufficient to model the AM and FM signal with a RMS error of < 1% of the signal's peak-to-peak value (see Fig. 2b and 2c). As the AM/FM frequencies reduce, the order of the L-STPSA could be reduced almost logarithmically while maintaining the same accuracy.

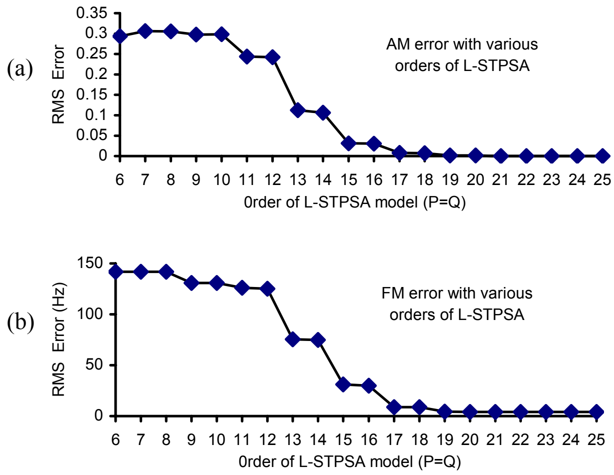

Figure 3. RMS errors using L-STPSA of different orders for the (a) amplitude envelopes and (b) instantaneous frequencies.

B. Noise sensitivity

The performance of the L-STPSA algorithm is compared with two well-known AM-FM demodulation algorithms, the smoothed energy operator separation algorithm (SEOSA) of Potamianos [5] and the Hilbert transform method [4]. White-Gaussian noise of varying variances were added to a 500-sampled discrete sinusoidal AM-FM signal with a carrier frequency w_c=2000 Hz, FM frequency w_{fm}=100 Hz and AM frequency w_{am}=200 Hz, similar to that shown in Fig. 6a. At various signal-to-noise ratios (SNR), the amplitude envelope and instantaneous frequency were estimated using the proposed L-STPSA model approach, the Hilbert transform method and the energy separation approach (i.e. SEOSA). The robustness of the L-STPSA approach in demodulating noisy AM-FM signals can be clearly observed in results shown in Fig. 4 and Fig. 5, which showed significantly lower RMS errors for both AM and FM estimates for a SNR between 0 dB to 90 dB.

Even at a SNR of 10 dB, the AM and FM estimates obtained using the L-STPSA models were reasonably accurate, as can be observed in Fig. 6b and 6c. Errors were observed mainly at the two ends of the analysis frame. This is to be expected since the signals are modeled using a power series. Small errors in the estimated coefficients will result in large errors as the value of n^k in (6) quickly increases as one moves away from the centre of the analysis frame, where $n = 0$. The results also show that the SEOSA approach outperforms the Hilbert transform method at high noise levels, which is consistent with that observed in [6] since the carrier frequency is reasonably high at 2 kHz.

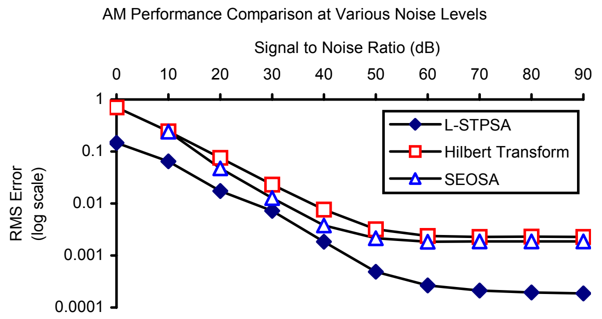

Figure 4. RMS errors (in log scale) of amplitude envelopes estimated at various SNR, using the L-STPSA models of order ($P=Q=20$), Hilbert transform and SEOSA methods.

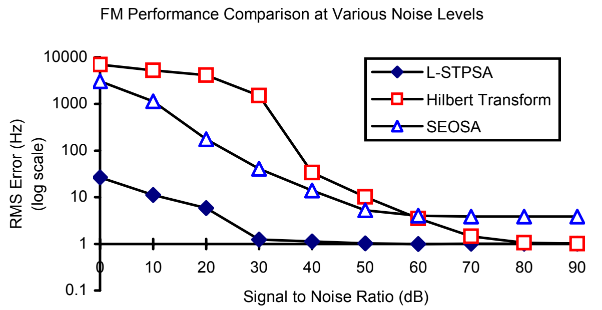

Figure 5. RMS errors (in log scale) of instantaneous frequencies estimated at various SNR, using the L-STPSA models of order ($P=Q=20$), Hilbert transform and SEOSA methods.

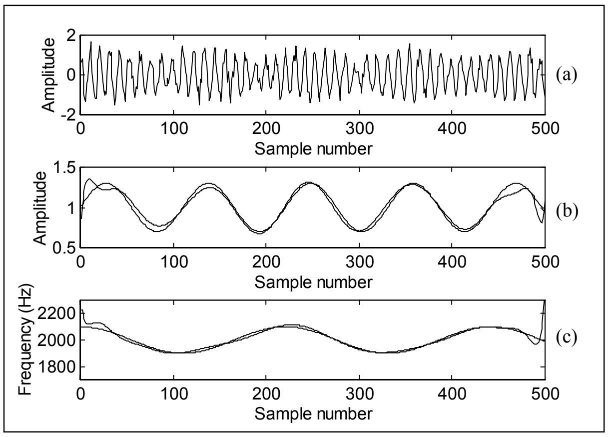

Figure 6. AM-FM estimates with L-STPSA model of orders ($P=Q=20$) in noisy signals. (a) Test signal with SNR=10dB. (b) The amplitude and (b) frequency. Original signals in solid lines and estimated signals in dash lines.

C. Computational Load

Fig. 7 shows that the computational load of the L-STPSA algorithm increases linearly with signal segment length. However, the increase in computational load as the L-STPSA model order increases in not linear since the number of elements in the coefficient matrix, whose inverse has to be computed grows by $O(n^2)$, where n is the order of the truncated power series model used.

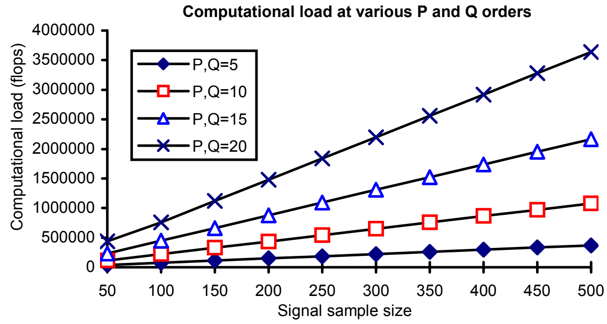

Figure 7. Computational load at various P-Q orders. Measured by the flops shown at the end of a Matlab routine execution.

D. Singing voice analysis

Fig. 8 shows a multicomponent AM-FM signal which is a short segment of a soprano's voice. The vibrato present in signal can be observed in spectrogram and it can be described by two parameters, namely the rate of vibrato and the depth of vibrato. Seashore [14] found that the mean vibrato frequency registered for 29 singers was 6.6 Hz, varying between extremes of 5.9 to 7.8 Hz. The average depth of vibrato was ±48 cents, varying between extremes of ±98 to ±31 cents (1 cent is the interval between two tones having the frequency ratio $1{:}2^{1/1200}$). The proposed technique was applied to the multicomponent signal which consist of the diphthong /oU/ (as in b<u>oa</u>t), sung at the key F5 (698 Hz).

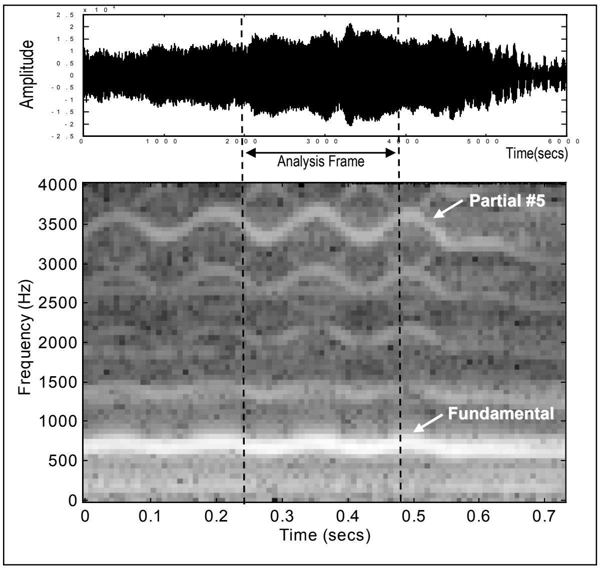

Figure 8. (a) Digitized waveform of a soprano's voice, which was sampled from [15] and (b) its corresponding spectrogram.

The extracted AM-FM sinusoid using polynomial models of orders P=30, Q=20 (see Fig. 9). Different partials of the multicomponent signal were extracted by starting the iterative estimation process using different estimated carrier frequency w_c in (1). The mean fundamental frequency was estimated to be about 700

Hz, which is close to the frequency of the key F5. The estimated vibrato frequency is 6.67 Hz and the vibrato depth is approximately ±84 cents. These values are within the limits observed in [14]. Fig. 9a and 9b show the vibrato parameters are relatively consistent between the higher partials and fundamental. Fig. 9c and 9d show the estimated amplitude envelope of the fundamental frequency and partial #5, which though greatly reduced in energy, still maintains a similar time-varying profile.

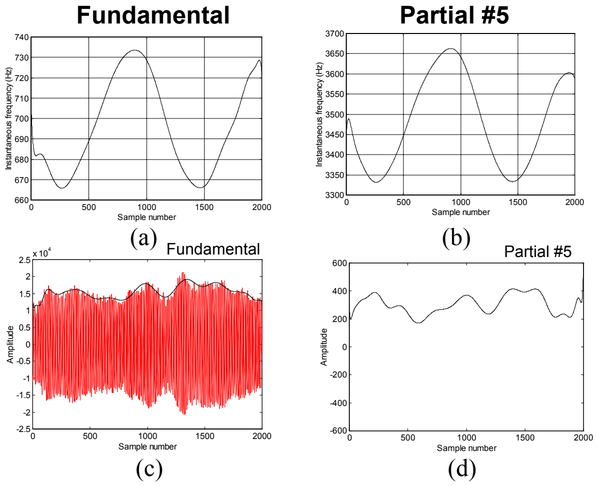

Figure 9. Time-varying AM-FM plots extracted from a singing voice. (a) Fundamental frequency and (b) partial #5, after 20 iterations with w_c = 684 Hz and w_c = 3344 Hz, respectively. Amplitude envelope of (c) the fundamental component, shown overlaid over the original waveform (red) and (d) partial #5.

IV. CONCLUSIONS

This paper presents a parametric approach for demodulating multicomponent AM-FM sinusoidal signals. The quadrature amplitude components of an AM-FM sinusoidal signal are modeled using a pair of truncated power series. Based on the instantaneous phase derived from the modeled quadrature components, an iterative framework was proposed in which the time-varying instantaneous frequency of the signal is progressively estimated through repeated computation and substitution of amplitude and phase parameters. The truncated power series coefficients are estimated using linear least-squares minimization, resulting in an algorithm that is more robust to additive white Gaussian noise compared to demodulation algorithms such as SEOSA and Hilbert transform.

The proposed technique is also able to extract AM-FM parameters of sinusoids within multicomponent signals as was seen in the vibrato analysis performed on a segment of singing voice. The inherent problems of non-parametric time-frequency distributions such as time-frequency resolution has been alleviated using the proposed parametric modeling technique, as the vibrato depths of both the low frequency fundamental and its high frequency partial (partial #5) were estimated with high precision. This was also done without the need for distortion inducing bandpass filtering pre-processing.

Unfortunately, the truncated power series model constrains the proposed technique to sinusoidal signals that are continuously differentiable in both its amplitude and frequency variations. Signals with singularities in the form of sharp transients are problematic. In such cases, it is possible to view these signals as piecewise-analytic and the proposed analysis can still be applied to segmented regions of the signal which do not straddle singular points. Some means of onset detection could be used in the pre-processing stage. One example is the onset map proposed by Brown and Cooke [16], which is derived using physiologically-motivated filters.

REFERENCES

[1] B. Friedlander and J. Francos, "Estimation of amplitude & phase parameters of multicomponent signals", *IEEE Trans. Signal Processing*, vol. 43, pp. 917-925, 1995.

[2] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 899-911, Aug. 1983.

[3] S. Mukhopadhyay and P. Sircar, "Parametric modelling of non-stationary signals: a unified approach", *Signal Processing*, vol. 60, pp. 135-152, 1997.

[4] S. L. Hahn, *Hilbert transforms in signal processing*. Boston, MA: Artech House, 1996.

[5] A. Potamianos and P. Maragos "A Comparison of the Energy Operator and the Hilbert Transform approach to Signal and Speech Demodulation", Signal Processing, vol. 37(1), pp. 95-120, 1994.

[6] T. F. Quatieri, T. E. Hanna and G. C. O'Leary, "AM-FM separation using auditory-motivated filters", *IEEE Trans. Speech Audio Processing*, vol. 5, pp. 465-480, Sep. 1997.

[7] P. Maragos, J. F. Kaiser and T. F. Quatieri, "On amplitude and frequency demodulation using the energy operators", *IEEE Trans. Signal Processing*, vol. 41, pp. 1532-1550, 1993.

[8] P. Maragos, J.F. Kaiser and T.F. Quatieri, "Energy separation in signal modulations with application to speech analysis", *IEEE Trans. Signal Processing*, vol. 41, pp. 3024-3051, 1993.

[9] H.M. Hanson, P. Maragos, and A. Potamianos, "A system for finding speech formants and modulations via energy separation", *IEEE Trans. Speech Audio Processing*, vol. 2, pp. 436-443, 1994.

[10] R. B. Sussman and M. Kahrs, "Analysis and resynthesis of musical instrument sounds using energy separation", *Proc. IEEE-ICASSP'96*, pp. 997-1000, 1996.

[11] J. S. Marques and L. B. Almeida, "Frequency-varying sinusoidal modeling of speech", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 763-765, 1989.

[12] J. D. Paliouras, *Complex variables for scientists and engineers*. New York: Macmillan Publishing Co., 1975.

[13] W.B.Goh and K.Y.Chan, "Amplitude Modulated Sinusoidal Modeling using Least-square Infinite Series Approximation with Applications to Timbre Analysis," *Proc. IEEE-ICASSP'98*, vol. 6, pp. 3561-3564, 1998.

[14] C.E. Seashore, *Psychology of music*, McGraw-Hill, 1967.

[15] K. Battle, *Grace*, CD Album produced by D. Frost and S. Paul, Sony Music Entertainment Inc., 1997

[16] G. J. Brown and M. Cooke, "Physiologically-motivated signal representation for computational auditory scene analysis", in *Visual representation of speech signals*, Chichester, UK: John Wiley & Sons, 1993.