RESOLVING OVERLAPPING HARMONICS FOR MONAURAL MUSICAL SOUND SEPARATION USING PITCH AND COMMON AMPLITUDE MODULATION

John Woodruff and Yipeng Li
Dept. of Computer Science and Engineering
The Ohio State University
{woodrufj, liyip}@cse.ohio-state.edu

DeLiang Wang
Dept. of Computer Science and Engineering
& the Center for Cognitive Science
The Ohio State University
dwang@cse.ohio-state.edu

ABSTRACT

In mixtures of pitched sounds, the problem of overlapping harmonics poses a significant challenge to monaural musical sound separation systems. In this paper we present a new algorithm for sinusoidal parameter estimation of overlapping harmonics for pitched instruments. Our algorithm is based on the assumptions that harmonics of the same source have correlated amplitude envelopes and the phase change of harmonics can be accurately predicted from an instrument's pitch. We exploit these two assumptions in a least-squares estimation framework to resolve overlapping harmonics. This new algorithm is incorporated into a separation system and quantitative evaluation shows that the resulting system performs significantly better than an existing monaural music separation system for mixtures of harmonic instruments.

1 INTRODUCTION

Musical sound separation attempts to isolate the sound of individual instruments in a polyphonic mixture. In recent years this problem has attracted significant attention as the demand for automatic analysis, organization, and retrieval of a vast amount of online music data has exploded. A solution to this problem allows more efficient audio coding, more accurate content-based analysis, and more sophisticated manipulation of musical signals [1]. In this paper, we address the problem of monaural musical sound separation, where multiple harmonic instruments are recorded by a single microphone or mixed to a single channel.

A well known difficulty in music separation arises when the harmonics of two or more pitched instruments have frequencies that are the same or similar. Since Western music favors the twelve-tone equal temperament scale [2], common musical intervals have pitch relationships very close to simple integer ratios (≈3/2, 4/3, 5/3, 5/4, etc.). As a consequence, a large number of harmonics of a given source may be overlapped with another source in a mixture.

When harmonics overlap, the amplitude and phase of individual harmonics become unobservable. To recover an overlapped harmonic, it has been assumed that the amplitudes of instrument harmonics decay smoothly as a function of frequency [3]. Based on this assumption, the amplitude of an overlapped harmonic can be estimated from the amplitudes of neighboring non-overlapped harmonics of the same source. For example, Virtanen and Klapuri [4] estimated an overlapped harmonic through non-linear interpolation of neighboring harmonics. Every and Szymanski [5] used linear interpolation instead. Recently, Virtanen [1] proposed a system which directly imposes spectral smoothness by modeling the amplitudes of harmonics as a weighted sum of fixed basis functions having smooth spectral envelopes. However, for real instrument sounds, the spectral smoothness assumption is often violated (see Figure 1). Another method of dealing with overlapping harmonics is to use instrument models that contain the relative amplitudes of harmonics [6]. However, models of this nature have limited success due to the spectral diversity in recordings of different notes, different playing styles, and even different builds of the same instrument type.

Although in general, the absolute value of a harmonic's amplitude with respect to its neighboring harmonics is difficult to model, the amplitude envelopes of different harmonics of the same source exhibit similar temporal dynamics. This is known as common amplitude modulation (CAM) and it is an important organizational cue in human auditory perception [7] and has been used in computational auditory scene analysis [8]. Although CAM has been utilized for stereo music separation [9, 10], to our knowledge, this cue has not been applied in existing monaural systems. In this paper we demonstrate how CAM can be used to resolve overlapping harmonics in monaural music separation.

Many existing monaural music separation systems operate only in the amplitude/magnitude domain [5, 6, 11, 12]. However, the relative phase of overlapping harmonics plays a critical role in the observed amplitude of the mixture and must be considered in order to accurately recover the amplitudes of individual harmonics. We will show that the phase change of each harmonic can be accurately predicted from the signal's pitch. When this and the CAM observation are combined within a sinusoidal signal model, both the amplitude and phase parameters of overlapping harmonics can be accurately estimated.

This paper is organized as follows. Section 2 presents the sinusoidal model for mixtures of harmonic instruments. In Section 3 we justify the CAM and phase change prediction assumptions and propose an algorithm where these assumptions are used in a least-squares estimation framework for resolving overlapping harmonics. In Section 4 we present a monaural music separation system which incorporates the proposed algorithm. Section 5 shows quantitative evaluation results of our separation system and Section 6 provides a final discussion.

2 SINUSOIDAL MODELING

Modeling a harmonic sound source as the summation of individual sinusoidal components is a well established technique in musical instrument synthesis and audio signal processing [13, 14]. Within an analysis frame m where sinusoids are assumed constant, the sinusoidal model of a mixture consisting of harmonic sounds can be written as

$$x_m(t) = \sum_n \sum_{h_n=1}^{H_n} a_n^{h_n}(m)\cos(2\pi f_n^{h_n}(m)t + \phi_n^{h_n}(m)), \quad (1)$$

where $a_n^{h_n}(m)$, $f_n^{h_n}(m)$, and $\phi_n^{h_n}(m)$ are the amplitude, frequency, and phase of sinusoidal component h_n, respectively, of source n at time frame m. H_n denotes the number of harmonics in source n. The sinusoidal model of $x_m(t)$ can be transformed to the spectral domain by using the discrete Fourier transform (DFT). With an appropriately chosen time-domain analysis window (in terms of frequency resolution and sidelobe suppression), and assuming perfect harmonicity, the spectral value of $x_m(t)$ at frequency bin k can be written as

$$X(m,k) = \sum_n S_n^{h_n}(m)W(kf_b - h_nF_n(m)). \quad (2)$$

Here, W is the complex-valued DFT of the analysis window, f_b is the frequency resolution of the DFT, and $F_n(m)$ denotes the pitch of source n at time frame m. We call $S_n^{h_n}(m)$ the sinusoidal parameter of harmonic h_n of source n, where $S_n^{h_n}(m) = \frac{a_n^{h_n}(m)}{2}e^{i\phi_n^{h_n}(m)}$. As a proof of concept, we further assume that pitches of individual sources are known.

3 RESOLVING OVERLAPPING HARMONICS

Given ground truth pitches of each source, one can identify non-overlapped and overlapped harmonics. When a harmonic of a source is not overlapped, the estimation of the sinusoidal parameter, $S_n^{h_n}(m)$, from observed spectral values, $X(m,k)$, in corresponding frequency bins is straightforward (see Section 4). However, when harmonics from different sources overlap, finding $S_n^{h_n}(m)$ for each active harmonic is an ill-defined problem. To address this, we make use of non-overlapped harmonics of the same source as well as the phase change estimated from the pitch information.

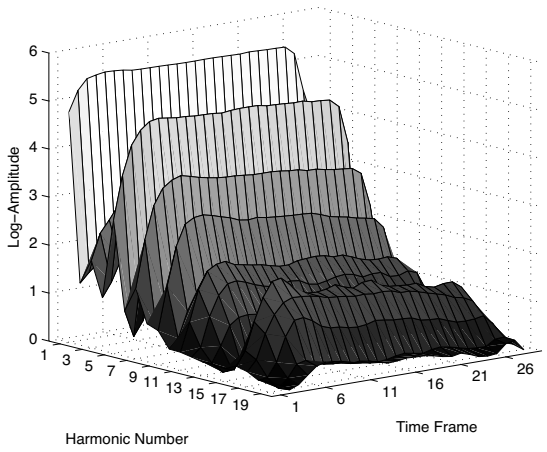

Figure 1. Logarithm of the amplitude envelopes for the first 20 harmonics of a clarinet playing a G#.

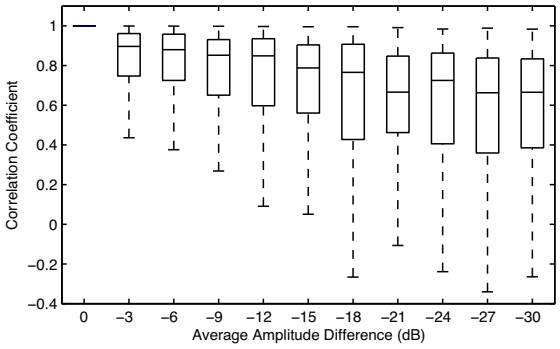

Figure 2. Box plots of correlation coefficient between different harmonics of the same source. Results are calculated using 40 five second instrument performances, with correlation calculated for each note of each performance.

3.1 Common Amplitude Modulation (CAM)

CAM assumes that the amplitude envelopes of sinusoidal components from the same source are correlated. Figure 1 shows the envelopes of the first 20 harmonics of a clarinet tone. We can see that in this case the CAM assumption holds while the spectral smoothness assumption does not. As further support for the CAM assumption, we calculated the correlation coefficient between the strongest harmonic of an individual instrument tone with other harmonics in the same tone as a function of difference in amplitude. The amplitude envelope of each harmonic was calculated by predicting each harmonic's frequency from the ground truth pitch

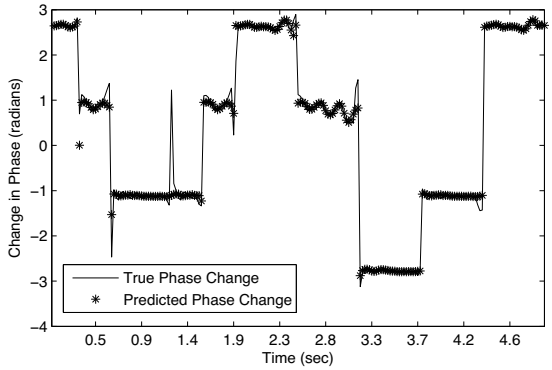

Figure 3. True and predicted phase change for harmonic 1 from a five second excerpt of a flute recording.

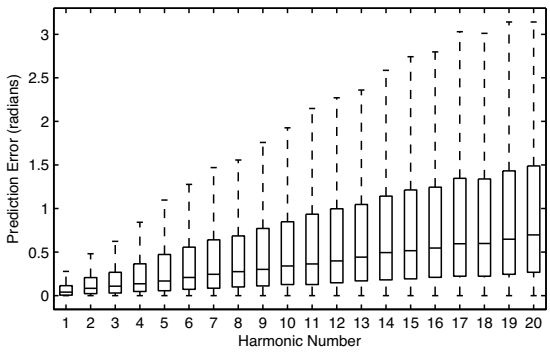

Figure 4. Box plots of phase change prediction error as a function of harmonic number. Results are calculated using 40 five second instrument performances, with error (in radians) calculated as absolute difference between true change in phase from frame m to $m+1$ and predicted change in phase for frame m.

and using the approach described in Section 4 for parameter estimation of non-overlapped harmonics (since every harmonic in a performance by a single instrument is non-overlapped).

Figure 2 shows box plots of the results obtained using 40 five second instrument performances with the correlation calculated for each note of each performance. The upper and lower edges of each box represent the upper and lower quartile ranges, the middle line shows the median value and the whiskers extend to the extent of the sample. For clarity, outliers are excluded from the plot. We can see that the correlation is high for harmonics with energy close to that of the strongest harmonic and tapers off as the energy in the harmonic decreases. This suggests that the amplitude envelope of an overlapped harmonic could be approximated from the amplitude envelopes of non-overlapped harmonics of the same source. Since the low-energy harmonics do not have a strong influence on the perception of a signal, the decreased correlation between the strongest harmonic and lower energy harmonics does not significantly degrade performance.

3.2 Predicting Phase Change using Pitch

According to the sinusoidal model in the time domain (see Equation (1)), the phase of a sinusoid at frame $m+1$ is related to the phase at frame m by

$$\phi_n^{h_n}(m+1) = 2\pi h_n F_n(m) T + \phi_n^{h_n}(m), \qquad (3)$$

where T denotes the frame shift in seconds. Equivalently we can write

$$\Delta\phi_n^{h_n}(m) = 2\pi h_n F_n(m) T. \qquad (4)$$

Therefore the phase change can be predicted from the pitch of a harmonic source. Figure 3 shows the phase change between successive time frames as measured from the first harmonic of a flute recording, and the predicted phase change using the true pitch of the signal. The predicted phase from Equation (4) is wrapped to $[-\pi, \pi]$. This example clearly shows that the phase change of a harmonic component can be accurately predicted from the pitch.

In Figure 4 we show box plots of the error between the true phase change of a harmonic component and the predicted phase change of a harmonic component as a function of harmonic number. The results are taken over the same performances as in Figure 2. As can be seen, for lower-numbered harmonics, the predicted phase change matches well to the true changes.

3.3 Estimating Amplitudes and Phases of Overlapped Harmonics

Using the amplitude envelope and phase change information, we can express the sinusoidal parameter of harmonic h_n, $S_n^{h_n}(m)$, in terms of a reference time frame m_0 as follows:

$$S_n^{h_n}(m) = S_n^{h_n}(m_0) r_{m_0 \to m}^{h_n} e^{i \sum_{l=m_0}^{m} \Delta\phi_n^{h_n}(l)}. \qquad (5)$$

Here, $r_{m_0 \to m}^{h_n} = \frac{a_n^{h_n}(m)}{a_n^{h_n}(m_0)}$ is the amplitude scaling factor between frames m_0 and m for harmonic h_n. The discussion in Section 3.1 suggests that the scaling factor of harmonic h_n can be approximated from the scaling factor of another harmonic of source n, i.e., $r_{m_0 \to m}^{h_n} \approx r_{m_0 \to m}^{h_n^*} = \frac{a_n^{h_n^*}(m)}{a_n^{h_n^*}(m_0)}$, where h_n^* is a non-overlapped harmonic with strong energy. As discussed in Section 3.2, the phase change of harmonic h_n can be predicted using the pitch $F_n(m)$. If we write

$$R_n(m,k) = r_{m_0 \to m}^{h_n^*} e^{i \sum_{l=m_0}^{m} \Delta\phi_n^{h_n}(l)} W(k f_b - h_n F_n(m)), \qquad (6)$$

then Equation (2) becomes

$$X(m,k) = \sum_n R_n(m,k) S_n^{h_n}(m_0). \qquad (7)$$

If harmonics from different sources overlap in a time-frequency (T-F) region with time frames from m_0 to m_1 and frequency bins from k_0 and k_1, we can write Equation (7) for each T-F unit in the region and the set of equations can be represented as

$$\mathbf{X} = \mathbf{RS}, \tag{8}$$

where,

$$\mathbf{X} = \begin{pmatrix} X(m_0, k_0) \\ \vdots \\ X(m_0, k_1) \\ \vdots \\ X(m_1, k_1) \end{pmatrix}, \tag{9}$$

$$\mathbf{R} = \begin{pmatrix} R_1(m_0, k_0) & \dots & R_N(m_0, k_0) \\ \vdots & & \vdots \\ R_1(m_0, k_1) & \dots & R_N(m_0, k_1) \\ \vdots & & \vdots \\ R_1(m_1, k_1) & \dots & R_N(m_1, k_1) \end{pmatrix}, \text{and} \tag{10}$$

$$\mathbf{S} = \begin{pmatrix} S_1^{h_1}(m_0) \\ \vdots \\ S_N^{h_N}(m_0) \end{pmatrix}. \tag{11}$$

The coefficient matrix $\mathbf{R}$ is constructed according to Equation (6) for each T-F unit. $\mathbf{X}$ is a vector of the observed spectral values of the mixture in the overlapping region. We seek a solution for $\mathbf{S}$ to minimize the sum of squared error

$$J = (\mathbf{X} - \mathbf{RS})^{\mathrm{H}}(\mathbf{X} - \mathbf{RS}). \tag{12}$$

The least-squares solution is given by

$$\mathbf{S} = (\mathbf{R}^{\mathrm{H}}\mathbf{R})^{-1}\mathbf{R}^{\mathrm{H}}\mathbf{X}, \tag{13}$$

where H denotes conjugate transpose. After $S_n^{h_n}(m_0)$ is estimated for each of the sources active in the overlapping region, we use Equation (5) to calculate $S_n^{h_n}(m)$ for all $m \in [m_0, m_1]$.

Figure 5 shows the effectiveness of the proposed algorithm in recovering two overlapping harmonics for two instruments. In this case, the third harmonic of the first source overlaps with the fourth harmonic of the second source. Figure 5(c) shows the magnitude spectrum of the mixture in the overlapping region. Note that the amplitude modulation results from the relative phase of the two harmonics. The estimated magnitude spectra of the two harmonics are shown in Figure 5(d) and (e). For comparison, the magnitude spectra of the two sources obtained from pre-mixed signals are shown in Figure 5(a) and (b). It is clear that the estimated magnitude spectra are very close to the true magnitude spectra.

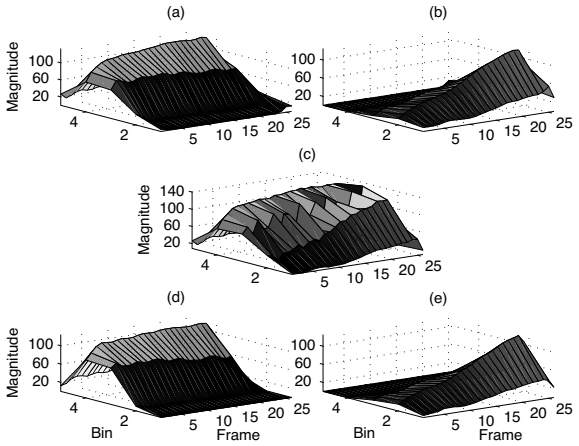

Figure 5. LS estimation of overlapped harmonics. (a) The magnitude spectrum of a harmonic of the first source in the overlapping T-F region. (b) The magnitude spectrum of a harmonic of the second source in the same T-F region. (c) The magnitude spectrum of the mixture at the same T-F region. (d) The estimated magnitude spectrum of the harmonic from the first source. (e) The estimated magnitude spectrum of the harmonic from the second source.

4 A MONAURAL MUSIC SEPARATION SYSTEM

We incorporate the proposed algorithm into a monaural music separation system to evaluate its effectiveness. The diagram of the system is shown in Figure 6. The input to the system is a polyphonic mixture and pitch contours of individual sources. As mentioned previously, we use ground truth pitch estimated from the clean signals for each source. In the harmonic labeling stage, the pitches are used to identify overlapping and non-overlapping harmonics.

To formalize the notion of overlapping harmonics, we say that harmonics h_{n_1} and h_{n_2} for sources n_1 and n_2, respectively, overlap when their frequencies are sufficiently close, $|f_{n_1}^{h_{n_1}}(m) - f_{n_2}^{h_{n_2}}(m)| < \theta_f$. If one assumes the signals strictly adhere to the sinusoidal model, the bandwidth of W determines how many frequency bins will contain energy from a sinusoidal component and one can set an amplitude threshold to determine θ_f.

For non-overlapped harmonics, sinusoidal parameters are estimated by minimizing the sum of squared error between the mixture and the predicted source energy,

$$J = \sum_{k \in K_n^{h_n}(m)} |X(m,k) - W(kf_b - h_n F_n(m)) S_n^{h_n}(m)|^2, \tag{14}$$

where $K_n^{h_n}(m)$ is the set of frequency bins associated with harmonic h_n in frame m. The solution is given by:

$$S_n^{h_n}(m) = \frac{\sum_{k \in K_n^{h_n}(m)} X(m,k) W(hF_n(m) - kf_b)}{\sum_{k \in K_n^{h_n}(m)} |W(hF_n(m) - kf_b)|^2}. \tag{15}$$

As described in Section 3.3, we utilize the amplitude envelope of non-overlapped harmonics to resolve overlapping harmonics. Since the envelope information is sequential,

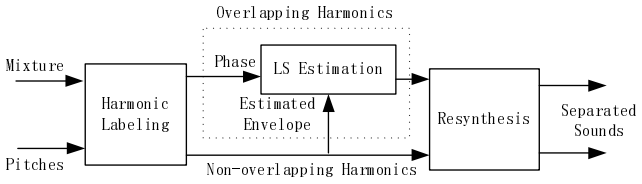

Figure 6. System diagram

we resolve overlapped h_n for time frames $[m_0, m_1]$ using a non-overlapped harmonic h_n^*. To determine appropriate time frames for this processing, we first identify sequences of time frames for which a harmonic h_n is overlapped with one or more other harmonics. If the pitch of any of the sources contributing to the overlapping region changes, we break the sequence of frames into subsequences. Given a sequence of frames, we choose the strongest harmonic for each source that is unobstructed in the entire sequence as h_n^*. We use θ_f to determine the bin indices, $[k_0, k_1]$, of the overlapping region.

For each overlapping region, we perform least-squares estimation to recover the sinusoidal parameters for each instrument's harmonics. To utilize the mixture signal as much as possible, we estimate the source spectra differently for the overlapped and non-overlapped harmonics. For all non-overlapped harmonics, we directly distribute the mixture energy to the source estimate,

$$\hat{Y}_n^{no}(m,k) = X(m,k) \quad \forall k \in K_n^{h_n}(m). \tag{16}$$

For the overlapped harmonics, we utilize the sinusoidal model and calculate the spectrogram using

$$\hat{Y}_n^{o}(m,k) = S_n^{h_n}(m) W(kf_b - f_n^{h_n}(m)). \tag{17}$$

Finally, the overall source spectrogram is $\hat{Y}_n = \hat{Y}_n^{no} + \hat{Y}_n^{o}$ and we use the overlap-add technique to obtain the time-domain estimate, $\hat{y}_n(t)$, for each source.

5 EVALUATION

5.1 Database

To evaluate the proposed system, we constructed a database of 20 quartet pieces by J. S. Bach. Since it is difficult to obtain multi-track recordings, we synthesize audio signals from MIDI files using samples of individual notes from the RWC music instrument database [15]. For each line selected from the MIDI file, we randomly assign one of four instruments: clarinet, flute, violin or trumpet. For each note in the line, a sample with the closest average pitch is selected from the database for the chosen instrument. We create two source mixtures (using the alto and tenor lines from the MIDI file) and three source mixtures (soprano, alto and tenor), and select the first 5-seconds of each piece for evaluation. All lines are mixed to have equal level, thus lines in the two instrument mixtures have 0 dB SNR and those in the three instrument mixtures have roughly -3 dB SNR. Details about the synthesis procedure can be found in [16]. Admittedly, audio signals generated in this way are a rough approximation of real recordings, but they show realistic spectral and temporal variations.

	SNR improvement
Virtanen (2 sources, 2006)	11.1 dB
Proposed System (2 sources)	14.5 dB
Proposed System (3 sources)	14.7 dB

Table 1. SNR improvement

5.2 Results

For evaluation we use the signal-to-noise ratio (SNR),

$$\text{SNR} = 10 \log_{10} \frac{\sum_t y^2(t)}{\sum_t (\hat{y}(t) - y(t))^2}, \tag{18}$$

where $y(t)$ and $\hat{y}(t)$ are the clean and the estimated instrument signals, respectively. We calculate the SNR gain after separation to show the effectiveness of the proposed algorithm. In our implementation, we use a frame length of 4096 samples with sampling frequency 44.1 kHz. No zero-padding is used in the DFT. The frame shift is 1024 samples. We choose $\theta_f = 1.5 f_b$, one and half times the frequency resolution of the DFT. The number of harmonics for each source, H_n, is chosen such that $f_n^{H_n}(m) < \frac{f_s}{2}$ for all time frames m, where f_s denotes the sampling frequency.

Performance results are shown in Table 1. The first row of the table is the SNR gain for the two source mixtures achieved by the Virtanen system [1], which is also based on sinusoidal modeling. At each frame, this approach uses pitch information and the least-squares objective to simultaneously estimate the amplitudes and phases of the harmonics of all instruments. A so-called adaptive frequency-band model is used to estimate the parameters of overlapped harmonics. To avoid inaccurate implementation of this system, we asked the author to provide separated signals for our set of test mixtures. The second row in Table 1 shows the SNR gain achieved by our system. On average, our approach achieved a 14.5 dB SNR improvement, 3.4 dB higher than the Virtanen system. The third row shows the SNR gain of our system on the three source mixtures. Note that all results were obtained using ground truth pitches. Sound demos of the our separation system can be found at: www.cse.ohio-state.edu/∼woodrufj/mmss.html

6 DISCUSSION AND CONCLUSION

In this paper we have proposed an algorithm for resolving overlapping harmonics based on CAM and phase change estimation from pitches. We incorporate the algorithm in a separation system and quantitative results show significant improvement in terms of SNR gain relative to an existing

monaural music separation system. In addition to large increases in SNR, the perceptual quality of the separated signals is quite good in most cases. Because reconstruction of overlapped harmonics is accurate and we utilize the mixture for non-overlapped harmonics, the proposed system does not alter instrument timbre in the way that synthesis with a bank of sinusoids can. A weakness of the proposed approach is the introduction of so-called *musical noise* as performance degrades. One aspect of future work will be to address this issue and create higher quality output signals.

In this study we assume that the pitches of sources are known. However, for practical applications, the true pitches of sources in a mixture are not available and must be estimated. Since our model uses pitch to identify overlapped and non-overlapped harmonics and pitch inaccuracy affects both the least-squares estimation and phase change prediction, good performance is reliant on accurate pitch estimation. We are currently investigating methods that relax the need for accurate prior knowledge of pitch information. Preliminary results suggest that performance similar to the Virtanen system using ground truth pitch can still be achieved by our approach even with prior knowledge of only the number of sources (when combining our system with multi-pitch detection) or rough pitch information (as provided by MIDI data).

Acknowledgment

The authors would like to thank T. Virtanen for his assistance in sound separation and comparison. This research was supported in part by an AFOSR grant (F49620-04-1-0027) and an NSF grant (IIS-0534707).

7 REFERENCES

[1] T. Virtanen, "Sound source separation in monaural music signals," Ph.D. dissertation, Tampere University of Technology, 2006.

[2] E. M. Burns, "Intervals, scales, and tuning," in *The Psychology of Music*, D. Deutsch, Ed. San Diego: Academic Press, 1999.

[3] A. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 804–816, 2003.

[4] T. Virtanen and A. Klapuri, "Separation of harmonic sounds using multipitch analysis and iterative parameter estimation," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2001, pp. 83–86.

[5] M. R. Every and J. E. Szymanski, "Separation of synchronous pitched notes by spectral filtering of harmonics," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, pp. 1845–1856, 2006.

[6] M. Bay and J. W. Beauchamp, "Harmonic source separation using prestored spectra," in *Independent Component Analysis and Blind Signal Separation*, 2006, pp. 561–568.

[7] A. S. Bregman, *Auditory Scene Analysis*. Cambridge, MA: MIT Press, 1990.

[8] D. L. Wang and G. J. Brown, Eds., *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. Hoboken, NJ: Wiley/IEEE Press, 2006.

[9] H. Viste and G. Evangelista, "Separation of harmonic instruments with overlapping partials in multi-channel mixtures," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2003, pp. 25–28.

[10] J. Woodruff and B. Pardo, "Using pitch, amplitude modulation and spatial cues for separation of harmonic instruments from stereo music recordings," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, 2007.

[11] S. A. Abdallah and M. D. Plumbley, "Unsupervised analysis of polyphonic music by sparse coding," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 179–196, 2006.

[12] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 1066–1074, 2007.

[13] R. McAulay and T. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 4, pp. 744–754, 1986.

[14] X. Serra, "Musical sound modeling with sinusoids plus noise," in *Musical Signal Processing*, C. Roads, S. Pope, A. Picialli, and G. Poli, Eds. Lisse, The Netherlands: Swets & Zeitlinger, 1997.

[15] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC music database: Music genre database and musical instrument sound database," in *International Conference on Music Information Retrieval*, 2003.

[16] Y. Li and D. L. Wang, "Pitch detection in polyphonic music using instrument tone models," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2007, pp. II.481–484.