A Novel Fourier Approach to Guitar String Separation

David Ramsay, Ted Burke, Dan Barry, and Eugene Coyle

Department of Electrical Engineering Systems

Dublin Institute of Technology

Dublin, Ireland

email : david.ramsay@case.edu, ted.burke@dit.ie

Abstract— **The ability to separate a single string's signal from the standard guitar output has several proven commercial applications, including MIDI control and effects processing. The current body of work in this area, based on expensive hardware modifications to the guitar, would greatly benefit from a generalised DSP-based solution. In this paper, we present a novel solution for the special case of separating the six open strings of a standard electric guitar, for particular use in a rehabilitation technology setting. By re-tuning the guitar strings to frequencies that are chosen at prime number multiples of the analysis window's frequency resolution, a rectangular window is able to capture over 97% of a sounding string's power, while minimising harmonic overlap. Based on this decomposition method, reconstruction of the string's harmonic series using sine wave tables is shown to closely recreate the original sound of the string. This technique for guitar signal resynthesis can robustly separate the signals from each of its six strings with multiple strings sounding, and faithfully reconstruct their sound at new fundamental frequencies in real-time. The combination of intelligent re-tuning and DSP algorithms as described in this paper could be extended to include fretted note detection with further applications in MIDI control or music transcription.**

Keywords – **String Separation, Rectangular Windowing, Hexaphonic Guitar.**

I INTRODUCTION

The output of a conventional guitar can be conceptualised as the summation of signals generated by each of the six strings. The separation of any individual string's signal from the composite output is a challenging signal processing problem, with several practical benefits. Hardware-based solutions have demonstrated the commercial potential in the areas of effects processing, MIDI interface control, and automatic re-tuning (examples include Line 6's Variax line of guitars and Roland's GK-3 interface) [1]. Automatic music transcription (i.e. tablature) relies on robust string separation [2], as do new techniques for remapping of the standard guitar control interface. In all of these cases, a DSP-based approach would eliminate the need for physical guitar modifications and reduce the expense associated with current hardware solutions, catalysing mainstream acceptance of these technologies.

A cursory look at the proposed problem will clarify why hardware solutions have become the standard. When sounding, an individual guitar string enters into highly periodic behaviour comprised of linearly spaced harmonic overtones above and including its fundamental oscillation frequency (F, 2*F, 3*F, etc.). Given the harmonic relationship between notes produced by a guitar that is played in any standardised tuning, the possibility of separating out the sound of an individual string (or harmonic overtone series) from the composite signal using Fourier techniques becomes difficult. There are two main reasons for this: 1) spectral overlap or near-overlap between series, and 2) spectral energy leakage across relevant frequency bins. Statistical approaches that might approximate the sound of an individual note in this context still lack a robust methodology to determine its originating string (as the same note may be fretted on different strings).

In this paper, we propose a DSP-based solution to robustly separate the signals of each of the unfretted, open strings of the guitar. This special case of the string separation problem is particularly applicable to an ongoing project for remapping the guitar's control interface. Specifically, this work aims to eliminate the need for the fretting hand by replacing it with a melody prediction algorithm, thus pitch shifting the sound of the individual strings in real-time. Our solution provides robust, real-time separation of all of the open strings, as well as a simple and effective means for reconstructing the actual sounding note at new frequencies. It effectively handles multiple strings sounding at once, and can shift them all to distinct new pitches in real-time while preserving their original acoustic sound. This interface substantially reduces the required dexterity to play any conventional guitar, and has particular use in rehabilitation technology.

Despite the focus on one application, this strategy for string separation has the potential to be generalised to include accurate recognition of any fretted note on any string. Practical use of this technique for MIDI control or automatic music transcription are well within reach.

II Decomposition Of The Signal

The method set forth in this section seeks to address the two difficulties of separating harmonically-related sounding strings: 1) the overlap of harmonic series, and 2) the spectral energy leakage resulting from windowing effects. In order to accomplish this, the guitar strings are re-tuned to bin frequencies of the FFT window used to analyse them, in such a way that minimises harmonic overlap. Any pitch whose fundamental falls exactly at a bin frequency will also have its constituent harmonic frequencies fall in bins; in other words, the individual sinusoidal frequency components of the re-tuned string will fit an integer multiple number of times into the sampling window, circumventing several negative windowing effects. Using a rectangular window, this allows us to perfectly capture the harmonic series of any string with no spectral leakage. Minimising overlap becomes an exercise in proper frequency bin selection for each fundamental, with favorable selections resulting from prime number ratios between frequency bins. The final constraint in selecting frequencies is to maintain proper string tension on the guitar, therefore new frequency selections must fall in the range of standard guitar tuning.

a) The Guitar Signal

The signal that results from striking a string is shown in Figure 1. Typically, the signal is divided into two main sections: 1) a < 100ms (40-60ms typical) attack, in which the amplitude envelope of the string rises from 10 to 90% of its RMS maximum value, and 2) a region of highly periodic behaviour that lasts until the signal decays or the string is physically muted [3].

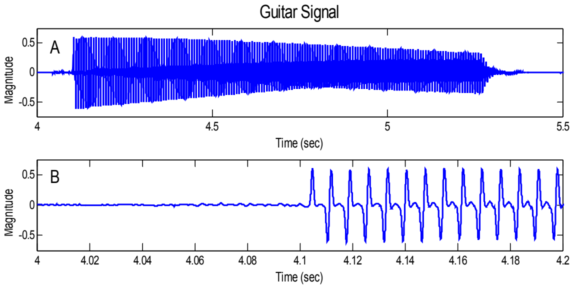

Figure 1: This is a time domain signal captured from an electric guitar. A) shows a full sounding note, while B) shows a close-up of the note onset.

Examination of several recordings from a clean electric guitar support the assumption of robust periodic behaviour. Additionally, we notice an immediate shift from the low amplitude broadband plectrum noise to periodicity, without a long or distinct attack. The lack of any transient, pitch-dependent information residing outside of the predicted overtone series is useful. With high reliability, the pitched information falls almost completely within the expected overtone series. This is crucial to the success of rectangular windowing as a means for separation of harmonic content.

The lack of pitch-dependent transients, as well as the predictable periodic behaviour, are assumptions based on the unaltered electric guitar signal. The complexities introduced by acoustic guitar resonances [4] or effect processors would impinge the efficacy of the following strategies. Thus, for the best results, these techniques should only be applied to an electric guitar signal before any effect pedals alter its character.

The guitar used in these and all future recordings is a Fender Standard Stratocaster HSS fitted with .009 GHS Boomer Strings.

b) Choosing Tuning Frequencies

Re-tuning the guitar strings to disparate frequencies will address the inherent difficulties of separating harmonically related strings. By choosing prime number bin frequencies from our FFT window for re-tuning, we can fully capture the sound of each string with minimal overlap.

Using a standard sampling rate of 44.1 kHz, the shortest power of two window length that is feasible with prime number bins is 4096 points. Prime number ratios with smaller FFT lengths fail to provide tuning possibilities that are realistic given the constraints of typical guitar strings. The 4096 point FFT gives us a frequency resolution of 10.77 Hz. By choosing our fundamental frequencies at prime number bins from 11 up to 29, we achieve a tuning range of 118.43 - 312.23 Hz, where each frequency is equal to the resolution multiplied by its

Guitar String	**Standard Tuning Frequency**	**Bin # of FFT**	**New Tuning Frequency**
E	329.63 Hz	29	312.23 Hz
B	246.94 Hz	23	247.63 Hz
G	196.00 Hz	19	204.57 Hz
D	146.83 Hz	17	183.03 Hz
A	110.00 Hz	13	139.97 Hz
E	82.41 Hz	11	118.43 Hz

Table 1. Re-tuned frequencies of the guitar compared to their standard tuning, as chosen by prime number bin ratios in the Fourier domain.

prime bin number (i.e. 11 * 10.77 Hz = 118.43 Hz). The standard guitar is tuned from 82.41-329.63 Hz, so these tunings are functional. Table 1 illustrates the final selections.

Since the fundamental note frequencies are chosen to fit an integer multiple of times into the rectangular window, the harmonics will also fall exactly into bins. For example, the high E string, now tuned to the 29th bin, will have harmonics falling in the 58th, 87th, etc bins. Since this applies to all of the open strings and their overtones, rectangular windowing of the output will provide no spectral leakage of harmonic content. Since we have chosen prime ratio bins, the first overlap of overtones will be at the 143rd bin, or 1540.11 Hz. This is the 11th harmonic of the fundamental starting on the 13th bin, colliding with the 13th overtone of the lowest string (whose fundamental lies on the 11th bin). For all strings, the lowest-order harmonic possibly compromised is the 11th. Given these conditions, we have created a framework in which perfect separation and reconstruction of pitched material up to the 11th overtone, even with multiple strings sounding, is easily achievable.

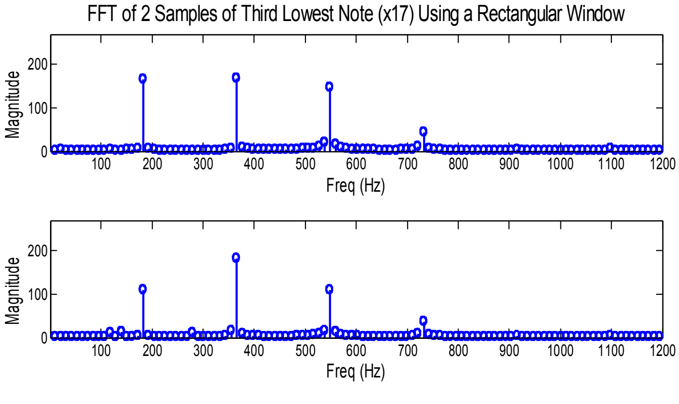

Figure 2: Frequency spectra of a recorded guitar signal resulting from the use of a rectangular window.

c) Verification of Model

In order to verify that this theoretical framework holds true in real-world situations, an electric guitar was recorded after being tuned to the frequencies listed in Table 1. Tuning was done by ear, comparing the string's pitch to a synthesised reference, and FFTs were performed on several samples of each note. The magnitude frequency spectrum of two such samples, taken of the third lowest string, is shown in Figure 2.

The two frequency spectra in Figure 2 represent different tone settings on the guitar, hence the variation in spectral quality. These spectra show that for a well-tuned string, the previous assumptions about rectangular windowing withstand real-world conditions. In the periodic regime of these two test cases, the relative power [5] of the signal falling within the expected harmonic bins is 97.89% and 97.47% of the total signal power. The remaining 2.1-2.5% of signal power is generally found in neighboring bins due to the effect of windowing an imperfectly tuned string (i.e. one whose harmonics do not coincide perfectly with DFT bins), though small transients of non-harmonic content may also play a small role.

For strings that are not as well-tuned, larger amounts of leakage appear. Using psychoacoustic principles, the expected range of tuning error by ear can be estimated for these cases. Research presented in [6] demonstrates that human hearing has a resolution of about 3.6 Hz below 500 Hz, and (0.0007 * frequency) above it. This work also concludes that musicians rely on upper harmonics while tuning. Using this information, we approximated the expected error range of a string's fundamental by looking at the sounding harmonic with the highest resolution (e.g. the 3rd or 4th harmonic in the signals shown in Figure 2). The results show expected errors between ±0.5 Hz and ±1.1 Hz, growing linearly with string frequency. In test cases combining the largest expected human error with an unfavorable input signal, 78% to 96% of the signal power still remained in the proper harmonic bins (with the lowest four strings all $\geq$ 90%). Proper tuning is critical, and this calculation demonstrates that the resolution of the human ear is sufficient for a well functioning implementation, though not an ideal one. Using an electronic tuning aid, it would be possible to ensure very robust (i.e. $\geq$ 97%) separation for each of the string signals.

III RECONSTRUCTION

After re-tuning the guitar so that the energy of each string falls into non-overlapping frequency bins, we can successfully reconstruct the harmonic behaviour of any individual string at a new fundamental frequency using sine wave tables. Standard additive resynthesis [7] using the first five magnitude values from string in question is an effective and computationally efficient strategy. In this section, we will show that this method for reconstruction is practical in a real world context.

a) Methods

Empirical observation of the frequency spectra of multiple strings results in the conclusion that harmonics above the first five contain insufficient energy to impact signal reconstruction under normal conditions. In the two examples shown in Figure 2, 97.88% and 97.46% of the respective signal power were found in the first five harmonic bins (compared to 97.89% and 97.47% falling in *all* of the harmonic bins). In these cases, the .01% of extra power found in harmonic bins above the first five is split between 115 bins (with each carrying about 0.00009% of the original signal's power). This behaviour is typical of all observed signals. Tests with brand new strings revealed minor but audible content from the sixth to tenth harmonics, which diminished quickly as the strings were worn. Thus, this assumption about harmonic content is more robust after new strings have been sufficiently broken in.

Since the audible sound is typically made of only five harmonics, and their exact magnitude information is available using the rectangular windowing method for deconstruction, it is simple to accurately reconstruct the waveform using additive resynthesis. A bank of five sine oscillators is used to reconstruct the signal, with linearly spaced oscillation frequencies starting at the fundamental frequency of the desired new pitch. The magnitudes of each of these frequencies is set by the corresponding magnitude of the actual vibrating string. In music, phase distortion is only audible in rare cases [8], thus the outlined approach should supply an indistinguishable resynthesis of the original signal without including phase, provided the string is in tune. When reconstruction occurs over multiple windows, changes in magnitude between them will result in small discontinuities if they are not handled properly. Solutions to this include 1) low-pass filtering of the magnitude multipliers so they change continuously and gradually, or 2) waiting for zero-crossings in the reconstructed signal before shifting coefficients. In our implementation, we opted for the latter technique.

b) Verification

In order to verify that this resynthesis technique was working properly, string signals were reconstructed at their original frequency and compared visually and audibly to the guitar input. Figure 3 shows an example of a real sounding string in black, and its reconstruction using five sine waves in green. The reconstructed wave closely mirrors the original wave, successfully capturing all of its main features with slight visual error. In these examples phase information was used, so that the physical characteristics of the wave could be usefully compared, though this is not the case during normal resynthesis.

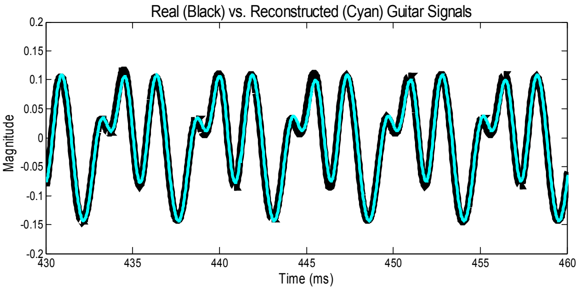

Figure 3: Additive resynthesis of the original guitar signal (black) using just five sine oscillators (green) accurately captures the features of the waveform.

More significantly, the original waveform and the reconstructed waveform were each looped and then stitched together in a .wav file to observe audible difference between them. For the example in Figure 3, the shift from the actual to the resynthesised regime is unnoticeable. In cases where the string was not as well-tuned (i.e. around 90% of the power preserved), a slight change in timbre was detectable to the highly-critical listener. This minute change under poor test conditions substantiates this method as a robust and practical resynthesis system.

IV Real-Time Implementation

Ensuring that the signal resynthesis can be applied in real-time is crucial to this application. Considering the prime number tuning implemented, the extreme dissonance of the open strings must be completely masked by their amplified, pitch-shifted versions to be functional. In its current form, the 4096 point FFT results in a 92.9ms delay at 44.1 kHz, which is very noticeable.

In order to adapt this technique for real-time use, it is possible to shorten our FFT to 2048 points and halve the delay, reducing the latency to within the threshold of real-time music generation [9]. This requires the concession of prime number ratio bins as tuning frequencies, but retains the separability of the first five harmonics of every string with one exception.

With a resolution of 21.53 Hz, the 2048 point FFT can be tuned to bin frequencies at the 4th, 5th, 7th, 9th, 11th, and 13th bins. This gives fundamental frequencies ranging from 86.12 to 279.89 Hz, with only one overlapping harmonic of importance at 430.6 Hz. This is the frequency at which the 5th harmonic of the lowest string collides with the 4th harmonic of the second-lowest. A weighted distribution of the energy in this bin based on the energy in the overlapping strings' other harmonics should provide a reasonable approximation for this bin. This small trade-off is well-worth the 46.5ms improvement in latency.

V DISCUSSION

In this paper, we have put forth a framework for re-tuning any electric guitar so that its open strings become separable. This technique allows accurate resynthesis of multiple sounding strings at arbitrary pitches, by combining rectangular windowing with a sine oscillator bank. We have shown that in practical situations, re-tuning of the guitar by ear introduces only small amounts of error into this model (with four strings still at $\geq 90\%$ power in their expected harmonic bins in the worst case). Using this method of decomposition and resynthesis, the threshold for very slight but audible changes in timbre between the original string sound and the reconstructed occurs around the 90% power mark. With a digital tuning aid, it should be possible to ensure that $\geq 97\%$ power of each string falls into the expected bins, resulting in no perceptible change in the sound of the reconstructed string.

Based on empirical data about the clean electric guitar signal (acoustic guitar resonances break down these assumptions), we found that the first five harmonics contain the necessary information for satisfactory signal reconstruction during quasi-periodic, sounding behaviour. Using this knowledge, we could relax our constraints on harmonic overlap, and thus reduce our FFT length from 4096 to 2048 points. The 46.5ms in latency this affords is crucial for achieving results that are as perceptibly 'real-time' as possible and masking the dissonances of the open strings.

This research has been carried out in the context of developing a guitar interface that does not require the use of a fretting hand, for rehabilitation purposes. Work has been done to combine these ideas with melodic prediction algorithms (which require robust transient-attack detection [10] to track a melody properly) and transient-attack resynthesis. Currently, the authors are developing a real-time implementation using an Analog Devices ADSP-21262 SHARC processor EZ-Kit.

This approach to the string separability problem works well for the open string case. Extending this idea to minimise frequency overlap of all of the fretted notes is the next logical step for this work, though there are challenges in real-time recognition of such high resolution frequency steps. The most feasible approach would combine fret and string recognition with magnitude and duration tracking. An implementation of this nature would have tangible benefits in applications in which fingering, and not accurate note resynthesis, is the focus. Potential examples include MIDI or video game control and tablature transcription.

ACKNOWLEDGEMENTS

We would like to thank Richard Hayes for his insight and guidance throughout the project. In addition, David would like to acknowledge the Fulbright Commission and the American Institute of International Education for their funding and generous support.

REFERENCES

[1] E. Bates, D. Furlong, and D. Dennehy. “Adapting polyphonic pickup technology for spatial music performance.” In *Proceedings of the 2008 International Computer Music Conference*, S.A.R.C. Belfast, Aug. 2008.

[2] P. D. O’Grady and S. T. Rickard. “Automatic hexaphonic guitar transcription using non-negative constraints.” In *Proceedings of the Irish Signal and Systems Conference*, 2009.

[3] R. M. French. “Elements of Sound Quality: Time Domain Descriptions,” in *Engineering the Guitar: Theory and Practice*. New York, Springer, 2009, pp. 182-186.

[4] H. Penttinen and V. Välimäki. “Time-Domain Approach to Estimating the Plucking Point of Guitar Tones Obtained with an Under-Saddle Pickup.” *Applied Acoustics,* vol. 65, pp. 1207–1220, 2004.

[5] J. D. Bronzino. “Frequency Analysis of the EEG,” in *The Biomedical Engineering Handbook: Volume 1,* 2nd ed. Florida, CRC Press LCC & Springer-Verlag, 2000, pp. 15.7-15.8.

[6] H. Fastl and E. Zwicker. “Just Noticeable Changes in Frequency,” in *Psychoacoustics, Facts and Models,* 3rd ed. Berlin, Springer-Verlag, 2007, pp. 182-185.

[7] C. Lippe. “Music for Piano and Computer: A Description.” *Information Processing Society of Japan SIG Notes,* vol. 97(122), pp. 33-38, 1997.

[8] S. P. Lipshitz, M. Pocock, and J. Vanderkooy. “On the Audibility of Midrange Phase Distortion in Audio Systems.” *Journal of the Audio Engineering Society*, vol. 30, pp. 580–595, Sept. 1982.

[9] C. Chafe, M. Gurevich, G. Leslie, and S. Tyan. "Effects of Time Delay on Ensemble Accuracy." In *Proceedings of the International Symposium on Musical Acoustics (ISMA2004)*, Nara, Japan, Apr. 2004.

[10] D. Barry, D. Dorran, and E. Coyle. "Time and Pitch Scale Modification: A Real-Time Framework and Tutorial." In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, Espoo, Finland, Sept. 2008.