SRA: A Web-based Research Tool for Spectral and Roughness Analysis of Sound Signals*

Pantelis N. Vassilakis
DePaul University School of Music, 804 West Belden Ave. Chicago, IL 60614, USA, pantelis@acousticslab.org

***Abstract* — SRA is a web-based tool that performs Spectral and Roughness Analysis on user-submitted sound files (*.wav* and *.aif* formats). Spectral analysis incorporates an improved Short-Time Fourier Transform (STFT) algorithm [1-2] and automates spectral peak-picking using *Loris* open-source C++ class library components. Users can set three spectral analysis/peak-picking parameters: analysis bandwidth, spectral-amplitude normalization, and spectral-amplitude threshold. These are described in detail within the tool, including suggestions on settings appropriate to the submitted files and research questions of interest. The spectral values obtained from the analysis enter a roughness calculation model [3-4], outputting roughness values at user-specified points within a file or roughness profiles at user-specified time intervals. The tool offers research background on spectral analysis, auditory roughness, and the algorithms used, including links to relevant publications. Spectral and roughness analysis of sound signals finds applications in music cognition, musical analysis, speech processing, and music teaching research, as well as in medicine and other areas. Presentation of the spectral analysis technique, the roughness estimation model, and the online tool is followed by a discussion of research studies employing the tool and an outline of future possible applications.**

I. Introduction

A. Auditory Roughness: Definitions

The term *auditory roughness* was introduced in the acoustics and psychoacoustics literature by Helmholtz [5] to describe the buzzing, harsh, raspy sound quality of narrow harmonic intervals. Within the Western musical tradition, auditory roughness constitutes one of the perceptual correlates of the multidimensional concept of dissonance, concept that has historical, cultural, and cognitive bases, along with physical and physiological ones [3-4]. The dimension of dissonance correlating best with auditory roughness has been termed *sensory* or *tonal dissonance* [11] or *auditory dissonance* [14], to mark its dependence more on physical and physiological, rather than cognitive, historical, or cultural considerations.

A familiar example of a signal corresponding to a rough sound would be the signal of a harmonic minor second performed, for instance, on two flutes. Although a harmonic minor second will sound rough regardless of the sound sources involved, steady state sources such as singing voices, bowed strings, or winds (as opposed to impulse sources such as percussion, plucked strings, *etc.*) result in more salient roughness sensations [6-7]. At relatively low registers, wider intervals such as major seconds and minor thirds can also sound rough and, within the Western musical tradition, are usually avoided as dissonant. For example, the general practice in Western art music orchestration of spacing out harmonic intervals more at low registers than at high registers has its basis on roughness considerations.

More broadly, the term *auditory roughness* can be used to describe the buzzing sound quality of a variety of signals, beyond those of narrow harmonic intervals (*e.g.* signals corresponding to fast trills, fast vibrato, percussive rolls, rattles, *etc.*). Roughness is one of the perceptual manifestations of interference and, in the physical frame of reference it is usually described as a function of a signal's amplitude envelope (*i.e.* amplitude fluctuation rate and depth) and corresponding spectral distribution. As such, auditory roughness can also be considered a dimension of timbre.

The reason all complex signals, including the signals of chords, harmonic intervals, *etc.*, exhibit amplitude fluctuations is physical and is related to the phenomenon of interference. The reason why some of these signals correspond to rough sounds is physiological and has to do mainly with the properties of the inner ear (review in [3]).

B. Signal Amplitude Fluctuation, Critical Band, and Auditory Roughness

Amplitude fluctuations describe variations in the maximum value (amplitude) of sound signals relative to a reference point and are the result of wave interference. The interference principle states that the combined amplitude of two or more vibrations (waves) at any given time may be larger (constructive interference) or smaller (destructive interference) than the amplitude of the individual vibrations (waves), depending on their phase relationship. In the case of two or more waves with different frequencies, their periodically changing phase relationship results in periodic alterations between constructive and destructive interference, giving rise to the phenomenon of periodic amplitude fluctuations.

Amplitude fluctuations can be placed in three overlapping perceptual categories related to the rate of fluctuation. Slow amplitude fluctuations (~ ≤15 per second) are perceived as loudness fluctuations referred to as *beating*. As the rate of fluctuation is increased, the loudness appears to gradually become constant and the fluctuations are perceived as “fluttering,” “buzzing,” or *roughness*. As the amplitude fluctuation rate is increased further, the roughness reaches a maximum strength and then gradually diminishes until it almost disappears (~ ≥75-150 fluctuations per second, depending on the frequency of the interfering waves) [3, 6, 8-9].

* Work supported by DePaul University and a Northwest Academic Computing Consortium grant to Dr. J. Middleton, Eastern Washington University.

Assuming the ear performs a frequency analysis on incoming signals [5-6, 8], the perceptual manifestations of amplitude fluctuation can be related directly to the bandwidth of the hypothetical analysis-filters, depending upon and defining what Zwicker [9] termed critical bandwidth. For example, in the simplest case of amplitude fluctuations resulting from the addition of two sine signals with frequencies f_1 and f_2, the fluctuation rate is equal to the frequency difference between the two sines $|f_1-f_2|$, and the following statements represent the general consensus:
(a) If the fluctuation rate is smaller than the critical bandwidth, then a single tone is perceived either with fluctuating loudness (beating) or with roughness. (b) If the fluctuation rate is larger than the critical bandwidth, then a complex tone is perceived, to which one or more pitches can be assigned but which, in general, exhibits little or no beating or roughness.

Psycho-physiologically, the roughness sensation can be linked to the inability of the auditory frequency-analysis mechanism to resolve inputs whose frequency difference is smaller than the critical bandwidth and to the resulting instability or periodic "tickling" [10] of the mechanical system (basilar membrane) that resonates in response to such inputs.

Along with amplitude fluctuation rate, the next most important signal parameter related to roughness is amplitude fluctuation degree [3, 7], that is, the level difference between peaks and valleys in signals with nonflat envelopes. The degree of amplitude fluctuation depends on the relative amplitudes of the components in the signal's spectrum, with interfering components of equal amplitudes resulting in the highest fluctuation degree and the highest roughness degree.

C. Auditory Roughness as Means of Musical Expression

The sensation of roughness has been explored more than any other perceptual manifestation of amplitude fluctuation and by numerous musical traditions, a practice that has only recently been documented and researched [3-4]. Manipulating the degree and rate of amplitude fluctuation helps create the buzzing sound of the Indian *tambura* drone and the rattling effect of Bosnian *ganga* singing, resulting in a sonic canvas that becomes the backdrop for further musical elaboration. It permits the creation of timbral variations (*e.g.* Middle Eastern *mijwiz* playing) and rhythmic contrasts (*e.g. ganga* singing) through gradual or abrupt changes among roughness degrees. Whether such variations are explicitly sought after, as in *ganga* singing and *mijwiz* playing, or are introduced more subtly and gradually, as may be the case in the typical chord progressions/modulations of Western music, they form an important part of a musical tradition's expressive vocabulary. Other examples include the *Quechua Haraui* songs of Peru, with their frequent use of narrow harmonic intervals, and the performance of the *taqara* flutes of the *Xingu* river in Brazil, where sonic effects similar to those produced with the *mizwij* are produced by two or more simultaneous performers.

II. Roughness Calculation Model

A. Background

Models that systematically quantify the roughness degree of a given sound permit the empirical testing of hypotheses that link roughness to musical variables and concepts. For example, a reliable roughness calculation model may be used to experimentally examine claims that link auditory roughness to (a) dissonance within the Western musical tradition, (b) patterns of tension and release in Near Eastern or North Indian musical pieces (as intended by performers and/or perceived by listeners), or (c) rhythmic/timbral effects in Balkan folk songs.

Numerous roughness calculation models have been proposed over the last ~100 years (*e.g.* [5, 11-15]). They have been employed in studies that attempt to link auditory roughness to auditory/sensory dissonance (*e.g.* [16-18]), demonstrating a relatively low degree of agreement between calculated and experimental data. Surprisingly for post-1960 models, the two principal studies [6-7] that have systematically examined the relationship between a signal's amplitude fluctuation degree and roughness have been overlooked. All the above models (a) overestimate the contribution of sound pressure level (*i.e.* absolute amplitude values of the interfering signals) to roughness, (b) underestimate the contribution of the degree of amplitude fluctuation (*i.e.* relative amplitudes values of the interfering signals) to roughness, and (c) often misrepresent the relationship between roughness and register (review in [3]).

SRA incorporates a new roughness calculation model, outlined below. Perceptual experiments testing the model indicate that it reliably and validly represents the perception of roughness, and performs better than previous roughness calculation models [3-4].

B. Proposed Roughness Calculation Model Outline

The roughness **R** of a signal whose spectrum has two sinusoidal components with frequencies f_1, f_2 and amplitudes A_1, A_2, where $f_{\min} = \min(f_1, f_2)$, $f_{\max} = \max(f_1, f_2)$, $A_{\min} = \min(A_1, A_2)$, and $A_{\max} = \max(A_1, A_2)$, is [3]:

$$\mathbf{R} = \mathbf{X}^{0.1} * 0.5(\mathbf{Y}^{3.11}) * \mathbf{Z} \tag{1}$$

where:

$$\mathbf{X} = A_{\min} * A_{\max} \tag{1a}$$

The term $\mathbf{X}^{0.1}$ in (1) represents the dependence of roughness on intensity (related to the amplitude of the added sines). It is based on [7], adjusted [3, 19] to account for the quantitative difference between modulation depth, used in [7], and amplitude fluctuation degree, the signal parameter influencing roughness.

$$\mathbf{Y} = 2A_{\min} / (A_{\min}+A_{\max}) \tag{1b}$$

The term $\mathbf{Y}^{3.11}$ in (1) represents the dependence of roughness on amplitude fluctuation degree (related to the amplitude difference of the added sines). It, too, is based on [7], adjusted to account for the quantitative difference between modulation depth and amplitude fluctuation degree [3, 19].

$$\mathbf{Z} = e^{-b1s(f\max - f\min)} - e^{-b2s(f\max - f\min)} \tag{1c}$$

[where b1 = 3.5; b2 = 5.75; s = 0.24/($s_1 f_{\min}$ + s_2); s_1 = 0.0207; s_2 = 18.96]

The term **Z** in (1) represents the dependence of roughness on amplitude fluctuation rate (frequency difference of the added sines) and register (frequency of the lower sine). It is based on Sethares's [15] modeling of the roughness curves in Fig. 1, curves that have been derived from multiple perceptual experiments examining the roughness of pairs of sines [11-15].

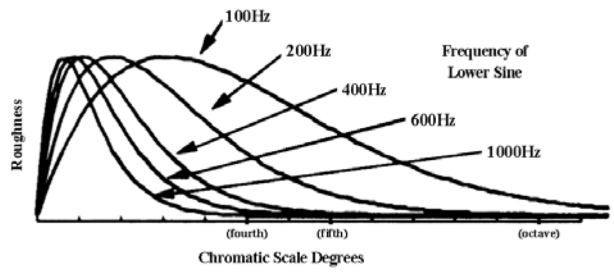

Fig. 1. Roughness curves plotting observed roughness (arbitrary measure, ***y*** axis) of a pair of equal-amplitude sines, as a function of frequency separation (*x* axis) and frequency of the lower sine (in [3] after [15]).

The roughness of signals corresponding to spectra with more than two sine components is calculated by summing the roughness of all sine-pairs in the spectrum. Although it has been argued that, depending on the relative phase of the respective amplitude fluctuations, the total roughness can be less than the sum of the roughness values for individual sine-pairs [6], several studies [7, 20] and pilot experiments [3] indicate otherwise. More specifically, [20] concluded that the total roughness is summed over all auditory filters. In addition, since roughness modeling is meaningful to roughness comparisons among multiple signals, rather than to roughness calculations of isolated signals, any potential signal-envelope phase effects are more likely to be diffused across the signals of interest, the more complex the signals.

The phase of a signal's spectral components is not included as a parameter in the roughness calculation. According to [21], the relative phase of the components of a three-component spectrum influences the complex signal's overall envelope shape and/or amplitude fluctuation degree, consequently influencing the signal's roughness, especially when three or more sine components fall within the same critical band. In spite of this observation, the absence of the phase parameter from the model does not significantly distort the model’s calculations. For the types of signals submitted to the calculator (synthetic signals, where the phase relationship of the components can be controlled and remain the same for all, or natural signals from polyphonic passages, where the phase relationships are more likely to be random than systematic), differences in the roughness phase effects among the signals to be examined are either controllable or defused. This supports valid comparisons of the resulting relative roughness values.

As is the case with all roughness calculation models, the absolute roughness values calculated by the model are arbitrary and are only useful for roughness comparisons among signals that have been analyzed using consistent analysis parameters. The roughness calculations of the above model correlate very well (r = 0.98) with roughness ratings obtained in a set of perceptual experiments [3-4], better than predictions by [5] (r = 0.73) and [14] (r = 0.87).

III. Spectral Analysis Method

The roughness model calculates the roughness of sound signals using spectral information (frequency and amplitude values of a signal’s spectral components). Spectral analysis in SRA uses an improved Short-Time Fourier Transform (STFT) algorithm, which is based on reassigned bandwidth-enhanced modeling [1-2; 22-24], and incorporates an automatic spectral peak-picking process to determine which frequency analysis bands correspond to spectral components of the analyzed signal.

Frequency reassignment [25] works differently from traditional Fast Fourier Transform (FFT) and has more in common with phase vocoder methods. For example, as in traditional FFT, frequency resolution of 10Hz will not be able to resolve frequency components laying less than 10Hz apart. But, unlike traditional FFT, the precision of the frequency values returned will not be limited by this 10Hz “bandwidth,” since the frequency band boundaries are floating rather than fixed. This (a) fine-tunes the frequencies reported and (b) practically eliminates spectral smearing, since the method ensures that the standard assumption of all energy being located at the high-frequency end of an analysis band can be fulfilled.

Similarly, as in traditional FFT, a given analysis window length determines the length of the shortest signals that can be reliably analyzed. But, unlike traditional FFT, the temporal resolution of a signal's spectral (and therefore roughness) time-profiles is not limited by this “window length,” since the frequency and amplitude estimates are not time-window averages but instantaneous at the time-window's center. This (a) pin-points time with much higher precision than implied by the window length and (b) practically eliminates temporal smearing, since the spectra estimated through time-window overlaps do not involve averaging over the entire analysis windows [1-2, 22-24].

In practical terms, spectral analysis results are fine-tuned through the incorporation of a dual STFT process. Frequency values reported correspond to the time derivative of the argument (phase) of the complex analytic signal representing a given frequency bin. Similarly, time values reported correspond to the frequency derivative of the STFT phase, defining the local group delay and applying a time correction that pinpoints the precise excitation time.

IV. SRA Application Outline (Fig. 2)

A. 1: *File Input*

SRA can process mono, uncompressed sound files, saved in the *.wav* or *.aif* formats. Only the left channel of stereo files submitted is processed. Sound files must be at least 316ms-long for 10Hz frequency resolution (158ms-long for 20Hz frequency resolution). The maximum size of files that can be submitted to the server is 12Mb, corresponding to ~ 2minutes of mono files at 48Ksmp/sec, 16bit.

The application will calculate roughness (or roughness profiles) based on the instantaneous spectrum at the user-specified point(s) in time (see Section III). Files submitted for analysis are transferred to the application server via the network. After the analysis has been completed, the submitted files are automatically deleted. Multiple analyses of the same sound file require resubmission of the file.

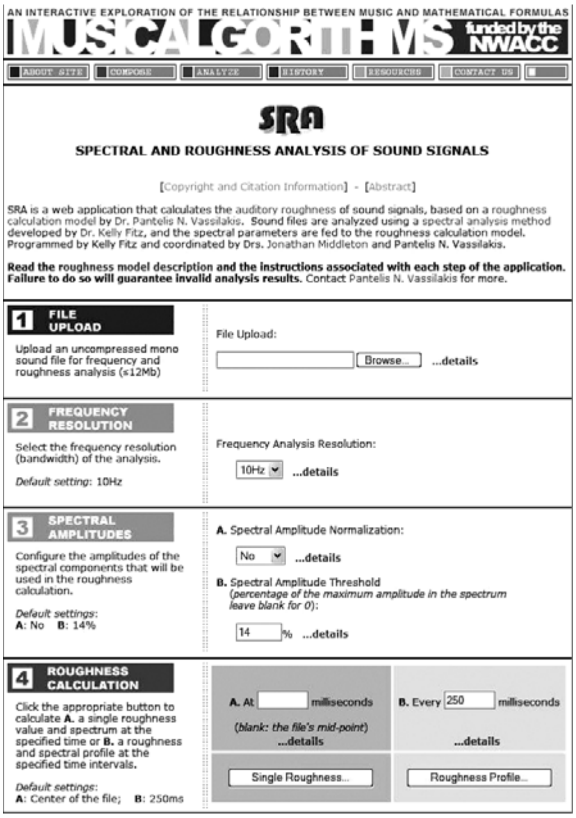

Fig. 2. Screen capture of SRA's main page. Application available at http://musicalgorithms.ewu.edu/algorithms/roughness.html and http://www.acousticslab.org/roughness [26].

B. 2: *Frequency Analysis Resolution*

The frequency resolution (bandwidth) of the analysis in SRA can be set to either 10Hz or 20Hz, with the shortest files that can be analyzed being 316ms- and 158ms-long respectively. It is important to note that this length limit does not apply to the time interval for roughness and spectral profile calculation, which can be as short as 1ms, although such short time intervals may be both impractical (in terms of the amount of data reported) and meaningless (in terms or roughness changes). Rather, the limit applies to the shortest file that can be submitted for analysis and to the earliest and latest points within a file for which roughness values and spectra can be calculated.

For example, given frequency resolution of 10Hz, the earliest (latest) point within a file for which a roughness value can be calculated will be 158ms from the beginning (end) of the file, corresponding to the mid-point of the shortest file that can be analyzed. More importantly, although 10Hz frequency resolution means that frequency components less than 10Hz apart will remain unresolved, the frequency values reported will not necessarily be in multiples of 10 but will have <1Hz precision (see Section III).

The default value for this parameter is "10Hz." The value "20Hz" should be selected only if it is certain that the sound files submitted for analysis contain no components separated by less than 20Hz in frequency. For example, 20Hz represent ~ a minor 2nd harmonic interval for fundamental frequencies around F4. With the analysis bandwidth set at 20Hz, minor 2nd harmonic intervals below F4 will correspond to fundamental frequencies that will be interpreted as unisons by the analysis, distorting the roughness calculation results. The lower the fundamental frequency of the tones making up the sound files analyzed, the more severe the distortion of the roughness calculation results will be. In fact, for fundamental frequencies around F2, 20Hz bandwidth represents ~ a major 3rd, and would wrongly interpret this and all narrower intervals as unisons. The only benefits of selecting 20Hz over 10Hz resolution are (a) the ability to analyze shorter files and (b) faster calculations.

C. 3A: *Spectral Amplitude Normalizarion*

For the vast majority of cases, the spectral amplitude normalization parameter should be set to "No." Selecting "No" will calculate the amplitudes of the submitted signal's spectral components based on the amplitude envelope of the signal-portion analyzed, assuming that the maximum possible signal amplitude is 1. Complex signals with maximum amplitude ≤1 have spectral components whose amplitudes are necessarily <1. Selecting "No" is appropriate if preserving the intensity relations among the files/file-portions analyzed is significant to the research question of interest. It is also appropriate to most analysis contexts, whether the application will be calculating a single roughness value at a user-defined point in time or a series of roughness values (*i.e.* roughness profile) at user-defined, regular time intervals. Selecting "No" is strongly recommended when calculating roughness values for signals that have been used as stimuli in a perceptual experiment.

Selecting "Yes" will scale up the amplitudes of the signal's spectral components so that the amplitude of the strongest component will be equal to 1. This selection is appropriate if preserving the intensity relations among the files/file-portions analyzed is not significant to the research question of interest. For example, selecting "Yes" would be appropriate for theoretical roughness comparisons among isolated vertical sonorities (harmonic intervals, chords, *etc.*), different isolated orchestrations of the same vertical sonorities, *etc.* "Yes" essentially performs intensity (not loudness) equalization over time on the submitted signal's spectral time-profile and is most useful when addressing theoretical questions that can benefit from such equalization. For example, calculating two roughness profiles for the same signal, with the spectral amplitude normalization parameter set to "No" and "Yes" respectively, will permit examination of the contribution of spectral intensity variations to the roughness profile of the submitted signal.

Note that spectral amplitude normalization, discussed here, is not equivalent to signal normalization. Regardless of the setting for the spectral normalization parameter, it is advised to adjust signal levels so that their maximum level falls within the range -3 to -10dB before submitting them for roughness profile calculation. This will ensure that the dynamic range of the analysis will be utilized most efficiently (see also below).

D. 3B: *Spectral Amplitude Threshold*

The value (x) entered in this field will remove from the spectral and roughness analysis results of each signal portion all spectral components with amplitudes below that portion's maximum spectral amplitude, multiplied by x. The default value (14%) has been selected based on analyses of over 100 sound signals spanning a wide variety of spectral complexity. It was determined to be the most reliable and valid value, in terms of roughness calculation, for signals with maximum levels within the range -3 to -10dB. The default value may be changed for analyses of:

(a) Synthesized complex signals, containing only discrete sinusoidal components. In such cases, the value may be set to 0.5-5%.
(b) Signals with maximum levels above -3dB (or below -10dB). In such cases the value should be progressively decreased (or increased). Alternatively, the overall signal level may be adjusted prior to analysis so that the maximum level falls within the range -3 to -10dB.
(c) Signals for which access to all spectral components returned from the analysis (up to 50) is desired. In such cases, the value may be set to 0%.

The need for non-zero spectral amplitude threshold is due to the fact that, for natural signals, several of the up to 50 spectral components that may be returned from the analysis will represent signal and/or analysis noise, and will have very low and similar amplitudes. Since sine-pairs of equal or almost equal amplitudes result in maximal calculated roughness (1b), the roughness contribution of 'noise' components will overestimate the total calculated roughness. As a result, for example, the roughness calculated for a quiet, solo violin passage may end up higher than the roughness calculated for a strong dissonant chord performed by an entire string orchestra. To avoid such invalid results, the amplitude threshold value is designed to remove the 'noise' components (components with very low and almost equal amplitudes) from the roughness calculations, while retaining the components that contribute to a given signal's perceived roughness.

E. 4A: *Single Roughness Value Calculation*

Clicking on the 'Single Roughness' button will send the selected file to the application server for analysis and present the 'Results' page. The results will include the roughness value and spectral distribution (frequency and amplitude values, by descending amplitude) of the submitted file at the user-defined point in time, and a list of the roughness contributions pair sine-pair within the signal's spectrum. The analysis window length will be determined by the frequency resolution setting (316ms for 10Hz and ~158ms for 20Hz) and be centered at the user-defined point in time. If the 'milliseconds' field is left blank, the analysis window will be centered at the file's mid-point. The earliest (latest) point within a file for which a roughness value can be calculated is 158ms (for 10Hz frequency resolution; 79ms for 20Hz frequency resolution) from the beginning (end) of the submitted file.

The degree of precision of the spectral analysis (in terms of time, frequency, and amplitude estimates) is much finer than the one required for roughness calculation (see Section III). It captures fast, fine spectral time-variance that does not necessarily correspond to perceivable roughness variations. To reduce this effect, the results reported from the analysis at a requested point in time within a file reflect:

(a) the median of five roughness values calculated from five equally-spaced spectra within an 100ms window (*i.e.* at -50ms, -25ms, 0ms, +25ms, and +50ms), centered at the time requested,
(b) the time corresponding to the reported roughness value [possibly different from the requested time – see (a)], and
(c) the spectrum used to calculate the reported roughness value.

F. 4B: *Roughness Profile Calculation*

Clicking on the 'Roughness Profile' button will send the selected file to the application server for analysis and present the 'Results' page. The results will include a set of time/roughness pairs (by ascending time) at the user-defined time interval (Fig. 3), followed by a list of the corresponding spectra (frequency and amplitude values, by descending amplitude). The first time/roughness pair and spectrum will be estimated on or after the time corresponding to half the length of the analysis window. As is the case with single roughness calculation, the analysis window length will be determined by the frequency resolution setting (316ms for 10Hz and ~158ms for 20Hz) and be centered at times determined by the 'time interval' setting. The 'time interval' setting (default: 250ms) is not limited by the analysis window length and can be as short as 1ms, although such short time intervals may be both impractical (in terms of the amount of data reported) and meaningless (in terms of changes in perceived roughness). If the time-interval field is left blank, the analysis will return an error. The earliest (latest) point within a file for which a roughness value can be calculated is 158ms (for 10Hz frequency resolution; 79ms for 20Hz frequency resolution) from the beginning (end) of the submitted file.

As is the case with single roughness calculation, the degree of precision of the spectral analysis (in terms of time, frequency, and amplitude estimates) is much finer than the one required for roughness calculation. It captures fast, fine spectral time-variance that does not necessarily correspond to perceivable roughness variations. To reduce this effect, the roughness value reported at any given time corresponds to the median of five roughness values calculated from five equally-spaced spectra within an 100ms window (*i.e.* at -50ms, -25ms, 0ms, +25ms, and +50ms), centered at the time reported. The precise time for each spectral distribution used to calculate the reported roughness values is recovered in the spectral distribution reports.

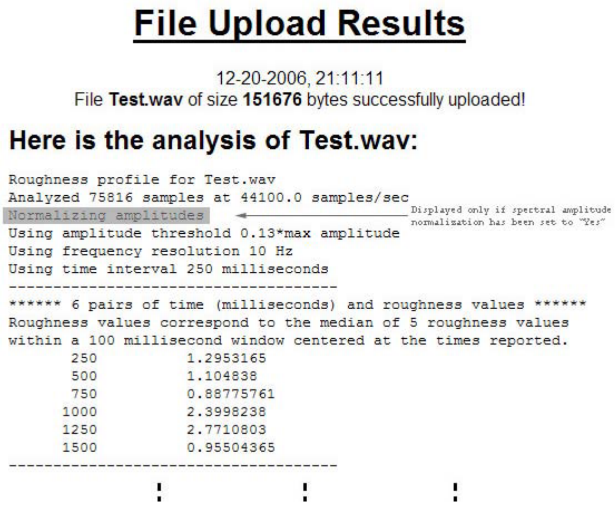

Fig. 3. Screen capture from a portion the 'Roughness Profile' results page returned by SRA. Roughness profiles are presented as time-roughness arrays and can be easily converted into graphical form (see Fig. 4).

Note that, depending on the roughness profile time-interval selected and on the research question of interest, the roughness profile obtained may need to be modified by using 5- or 7-point running-average values in order to better reveal data trends.

V. SUMMARY AND RESEARCH APPLICATIONS OF SINGLE ROUGHNESS AND ROUGHNESS PROFILE CALCULATIONS

Spectral analysis of sound signals and calculation of auditory roughness date back to the 19th century. SRA is a web-based research application that performs spectral analysis on submitted sound signals and calculates their auditory roughness, either as a single roughness value or in the form of a roughness time-profile.

The spectral analysis portion of the application uses an improved STFT algorithm, based on reassigned bandwidth-enhanced modeling. It can pinpoint the instantaneous frequency and amplitude of a signal's spectral components with minimization of spectral and temporal smearing.

The roughness calculation portion incorporates a new roughness calculation model, which represents perceptual roughness more reliably and validly than previous auditory roughness models.

The implementation of spectral and roughness analysis in SRA has already found application in studies addressing a wide variety of research topics, ranging from timbre analysis [28-29] and sound synthesis [30] to dissonance [4] and musical tension [27] (see, for example, Fig. 4). The reliability and validity of the spectral and roughness time-profiles returned by SRA can support work in acoustics, music and speech perception, voice pathology, or cross-cultural music research, and address questions that could not previously be tackled in a systematic manner.

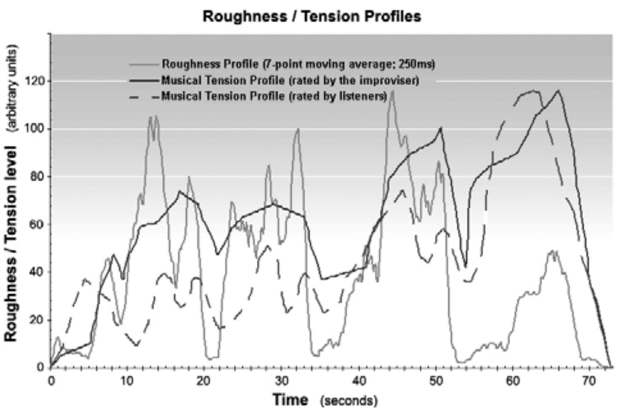

Fig. 4. Roughness profiles obtained by SRA in numeric form (see Fig. 3) can be easily converted into a graphical form. The above results are adapted from a study examining the relationship among the roughness profile (calculated by SRA) and musical tension profiles (indicated by participants in a perceptual experiment) of an improvisation on the Middle Eastern *mijwiz* (adapted from [27]).

ACKNOWLEDGMENT

The author thanks Dr. Kelly Fitz for programming SRA and for incorporating the Frequency Reassignment algorithm to SRA's spectral analysis portion, and Dr. Jonathan Middleton for sponsoring SRA and including it in *Musicalgorithms™*, an online interactive exploration of the relationship between music and mathematical formulas [http://musicalgorithms.ewu.edu].

REFERENCES

[1] S. A. Fulop and K. Fitz, "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *J. Acoust. Soc. Am.*, vol. 119(1), pp. 360-371, 2006.

[2] S. A. Fulop and K. Fitz, "Separation of components from impulses in reassigned spectrograms," *J. Acoust. Soc. Am.*, vol. 121(3), pp. 1510-1518, 2007.

[3] P. N. Vassilakis, *Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance*. Doctoral Dissertation. Los Angeles: University of California, Los Angeles, 2001.

[4] P. N. Vassilakis, "Auditory roughness as a means of musical expression," *Selected Reports in Ethnomusicology*, vol. 12 (Perspectives in Systematic Musicology), pp. 119-144, 2005.

[5] H. L. F. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, 2nd ed., A.J. Ellis, Trans. New York: Dover Publications, Inc., 1885 (1954).

[6] G. von Békésy, *Experiments in Hearing*. New York: Acoustical Society of America Press, 1960 (1989).

[7] E. Terhardt, "On the perception of periodic sound fluctuations (roughness)," *Acustica*, vol. 30(4), pp. 201-213, 1974.

[8] R. Plomp, "The ear as a frequency analyzer," *J. Acoust. Soc. Am.*, vol. 36(9), pp. 1628-1636, 1964.

[9] E. Zwicker, "Subdivision of the audible frequency into critical bands," *J. Acoust. Soc. Am.*, vol. 33(2), pp. 248-249, 1961.

[10] M. Campbell and C. Greated, *The Musician's Guide to Acoustics*, 2nd ed. New York: Oxford University Press, 1994.

[11] R. Plomp and W. J. Levelt, "Tonal consonance and critical bandwidth," *J. Acoust. Soc. Am.*, vol. 38(4), pp. 548-560, 1965.

[12] A. Kameoka and M. Kuriyagawa, "Consonance theory, part I: Consonance of dyads." *J. Acoust. Soc. Am.*, vol. 45(6), pp. 1451-1459, 1969.

[13] A. Kameoka and M. Kuriyagawa, "Consonance theory, part II: Consonance of complex tones and its calculation method," *J. Acoust. Soc. Am.*, vol. 45(6), pp. 1460-1469, 1969.

[14] W. Hutchinson and L. Knopoff, "The acoustic component of Western consonance," *Interface*, vol. 7, pp. 1-29, 1978.

[15] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*. London: Springer-Verlag, 1998.

[16] E. Bigand, R. Parncutt, and F. Lerdahl, "Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training," *Perception and Psychophysics*, vol. 58, pp. 125-141, 1996.

[17] J. Vos, "Purity ratings of tempered fifths and major thirds," *Music Perception*, vol. 3(3), pp. 221-258, 1986.

[18] N. Dibben, "The perception of structural stability in atonal music: The influence of salience, stability, horizontal motion, pitch commonality, and dissonance," *Music Perception*, vol. 16(3), pp. 265-294, 1999.

[19] P. N. Vassilakis, "Amplitude modulation depth versus degree of amplitude fluctuation: implementation error, adjustment and implications," *J. Acoust. Soc. Am.*, vol. 108(5/2), p. 2597, 2000

[20] J. Y. Lin and W. M. Hartmann, "Roughness and the critical bandwidth at low frequency," *J. Acoust. Soc. Am.*, vol. 97(5/2), p. 3274, 1995.

[21] D. Pressnitzer and S. McAdams, "Two phase effects in roughness perception," *J. Acoust. Soc. Am.*, vol. 105(5), pp. 2773-2782, 1999.

[22] K. Fitz and L. Haken, "On the use of time-frequency reassignment in additive sound modeling," *Journal of the Audio Engineering Society*, vol. 50(11), pp. 879-893, 2002.

[23] K. Fitz, L. Haken, S. Lefvert, C. Champion, and M. O'Donnell, "Cell-utes and flutter-tongued cats: Sound morphing using Loris and the Reassigned Bandwidth-Enhanced Model," *Computer Music Journal*, vol. 27(4), pp. 44-65, 2003.

[24] S. A. Fulop and K. Fitz, "A spectrogram for the twenty-first century," *Acoustics Today*, vol. 2(3), pp. 26-33, 2006.

[25] F. Auger and P. Flandrin, "Improving the readability of time frequency and time scale representations by the reassignment method," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1068-1089, 1995.

[26] P. N. Vassilakis and K. Fitz, "SRA: A Web-based Research Tool for Spectral and Roughness Analysis of Sound Signals." Supported by a Northwest Academic Computing Consortium grant to J. Middleton, Eastern Washington University, 2007. [http://musicalgorithms.ewu.edu/algorithms/roughness.html] or [http://www.acousticslab.org/roughness]

[27] P. N. Vassilakis, "An improvisation on the Middle-Eastern mijwiz; auditory roughness profiles and tension/release patterns," *J. Acoust. Soc. Am.*, vol. 117(4/2), p. 2476, 2005. Available online through the *Acoustical Institute of Physics Press Room* [http://www.aip.org/149th/Vassilakis.html]

[28] D. Bolger and N. Griffith, "Multidimensional timbre analysis of shakuhachi *hohkyoku*," *Proceedings of the Conference on International Musicology*, Montréal, Canada, 2005. [http://www.oicm.umontreal.ca/cim05/cim05_articles/BOLGER_D_CIM05.pdf]

[29] W. Hsu, "Managing gesture and timbre for analysis and instrument control in an interactive environment," *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression*, pp. 376-379, Paris, France, 2006.

[30] G. D'Inca and L. Mion, "Expressive audio synthesis: From performances to sounds," *Proceedings of the 12th International Conference on Auditory Display*, pp. 182-186, London, UK, 2006.