9th International Conference on Music Perception and Cognition

Alma Mater Studiorum University of Bologna, August 22-26 2006

Recognition of similarity relationships between time-stretched spectral structures

Renato Messina
Istituto Superiore di Studi Musicali "Vincenzo Bellini"
Catania, Italy
renatomessina@imuter.it

ABSTRACT

The purpose of this study is to present a test-based approach for recognizing from the spectral structure of unpitched short sounds (in the 50-100 milliseconds range) whether several time-stretching response functions are perceptually similar. For example, in a spectral composition the problem may be to determine whether the harmonicity/inharmonicity percentage defined in a short temporal scale by mathematical ratios, apply to each different time-stretching factor, preserving the object identity. This investigation proposes to evaluate the perceptual information obtained from the audible similarity comparison between the time-stretching models and their reference patterns. Let ($a_1, \ldots, a_m$) *be p-dimensional parameters associated with m response models of the same type. This study is concerned with the morphological comparison of* $a_1, \ldots, a_m$ *to individuate the timbral properties that are salient for similarity or identity judgments.*

Keywords

Time-stretching, spectral processes, similarity judgments.

MORPHOLOGICAL STRUCTURE

Certain morphological characteristics are selected to test subjective salient similarities between original spectral patterns and their time-stretching derivatives.

In: M. Baroni, A. R. Addessi, R. Caterina, M. Costa (2006) Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9), Bologna/Italy, August 22-26 2006.

Frequency to time domain conversion: formal attributes

The parameters selection is based upon the idea of a possible derivation of formal time-invariant relationships from audio data structured in temporal matrices perceptually identifiable on a short time scale. Several examples of such functional organization of music processing, like models of organic distribution or thematic developments of cellular structures, are observable in methods of spectral exploration developed by composers like Gerard Grisey (e.g. section I of *Partiels*, 1975) or Tristan Murail (e.g. section III of *Désintegrations*, 1983). The parameters, here described in a synthetic representation to verify the perceptual coherence of these procedures, are: 1) Distribution level of both frequencies and formant regions in different critical bands with harmonic relations; 2) Differentiation level of extension of the frequencies bandwidth; 3) Differentiation level of extension of the amplitude bandwidth; 4) Level of acoustical effects produced by spatial panning and reverberation processing; 5) Level of synchronisation between the piecewise-linear segments determining the phases of modifications in the partials temporal envelopes; 6) Percentage of energetic distribution in ADSR amplitude envelopes. These parameters are classifiable in both frequency and short time domain (see Table 1).

Table 1. Parameters selected for each analyzed domain.

	Frequency domain saliency parameters
1	Harmonicity and critical bands differentiation
2	Frequency bandwidth
3	Amplitude bandwidth
4	Acoustical effects
	Time domain saliency parameters

5	Synchronization of the partials envelopes
6	Energetic distribution in ADSR

Stretching models

Two principal time-stretching models are achieved by temporal extension of 2000% and of 60000% of the references short time patterns, variables from 50 to 100 milliseconds. These imply two responsive structures variables from 1000 to 2000 and from 30000 to 60000 milliseconds.

METHODS

The parametric relevance of the inspected formal attributes is here verified by assuming an increment of similarity derived from a parallel increment of the attribute prominence in the spectral patterns.

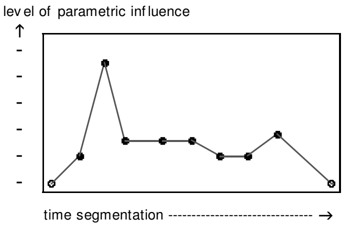

Figure 1. Temporal segmentation of a single function.

The patterns spectral structures, implemented in Max/MSP environment, are based on additive synthesis of sine wave partials and share the same frequency band selected by the subdivision of the audible frequency range into the twenty-four critical bands (Zwicker et Feldtkeller, 1981). Such structures are arranged both in parallel and in sequential levels. The parallel level controls additive synthesis components in the frequency domain; the sequential level implements piecewise-linear envelopes according to three stages in the time domain: 1) attack transients; 2) stationary state with a specific formant region; 3) decay transients (see Figure 1). The formant regions are realized using a quasi-synchronous granular synthesis technique (QSGS) with a grain size variable from 1 to 15 ms; the grain separation time is calculated to determine the formant peaks by the ratio: grain separation (ms) = sample_rate/pitch (Hz). To make the static synthetic sound more lively a frequency shift of a few Hz is employed, using a single-side band (SSB) modulation with two allpass filter to generate 90-degree phase difference (Bode and Moog, 1972). Using two spectrum shifter in each band, one shifting in the reverse direction of the other and mixing those to left and right channels an ITD spatialization cues is also implemented. The time-stretching models are obtained by multiplying all the weighted segments of the sequential level functions by a common coefficient (approximately from 2000% to 60000% of the original tempo).

TEST PACKAGE

The test package is based on a linear increment in the structure envelopes complexity: s1=p1; s2=p1*2; etc., where *s* is the structure and *p* is the number of parameters of both sequential and parallel functions (see Figure 2). If the number of correct listener responses to each parameter increases growing the envelopes complexity then an attribution of relevance is determined for the tested attribute.

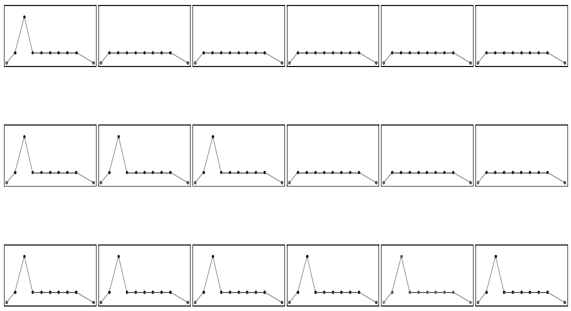

Figure 2. Sequential functions. Example of parametric density distribution - low, middle, high - for three different trial tests.

RESULTS

The test results are interpretable: (1) in relation to the morphological characteristics of the structure of how the subjects can detect the transformations by each stretching models proposed; (2) in relation to perceptual performances comparison between different levels of complexity. The results may also suggest to investigate the perception thresholds in an inverse mapping of large-scale temporal events converted into compacted data structures.

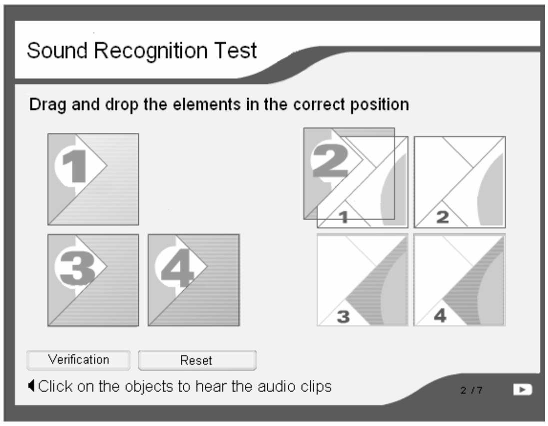

Figure 1. Graphic interface.

ACKNOWLEDGMENTS

Thanks to composer Emanuele Casale (Department of Computer Music, "V. Bellini" Institute of Music, Catania, Italy).

REFERENCES

Xenakis, I., (1971). *Formalized Music. Thought and Mathematics in Composition*, Bloomington, Indiana: Indiana Univ. Press.

AA.VV. (2000). *Gérard Grisey,* Quaderni della Civica Scuola di Musica di Milano, Anno, 15, N.27, Milano.

Zwicker, E., Feldtkeller, R., (1981). *Psychoacoustique. L'oreille récepteur d'information*, Paris: Masson.

De Poli, G., Piccialli, A., Roads, C. (1991). *Representations of Musical Signals.* Cambridge, Massachussets: MIT Press.

Roads, C. (1996). *The computer music tutorial*, Cambridge, Massachussets: MIT Press.

Roads, C. (2001). *Microsound,* Cambridge, Massachussets: MIT Press.

Bode, H., Moog, R. A. (1972). A High-Accuracy Frequency Shifter for Professional Audio Applications, *J. Audio Eng. Soc*., vol. 20, (1972 July/ Aug.), pp. 453-458.

ISBN 88-7395-155-4