

SOFTWARE FOR THE ANALYSIS OF SPECIES-SPECIFIC VOCALIZATIONS

Daniel Šuta^{1,2}, Martin Komárek³, Milan Jilek², Josef Syka²

Charles University in Prague, 3rd Faculty of Medicine, Czech Republic: Institute of Medical Biophysics and Informatics¹; Academy of Sciences of the Czech Republic, Institute of Experimental Medicine, Czech Republic: Department of Auditory Neuroscience²; Czech Technical University, Faculty of Electrical Engineering, Czech Republic: Department of Cybernetics³

Summary: Vocalization calls are behaviorally relevant complex sounds that typically contain several harmonics and show frequency and amplitude modulation. In this paper, an introduction to a software tool for the analysis of species-specific vocalizations is presented. The algorithm automatically or under user supervision detects time-varying amplitude and frequency parameters, which can serve for the statistical analysis of calls or as the substrate for the manipulation and synthesis of artificial calls. The described program and its results will be used in studying the representation of complex sounds in the central nervous system.

Key words: Acoustical signal; Communication sound; Neural system; Software; Species-specific vocalization

Introduction

Natural environmental sounds, such as species-specific vocalizations, are behaviorally important sounds with complex acoustical patterns. When studying the neural representation of such spectrotemporally complex stimuli, a fundamental question is to understand which aspect or parameter of the stimulus neurons are sensitive to. This requires the ability to systematically change the stimulus in many aspects. Traditional methods of call modification are based on various filtering strategies or global operations such as signal time reversal, time compression, or time expansion, which are typically performed on examples of representative calls. There are at least two limitations to this approach. First, when an example of the call is used, it may not be statistically accurate in representing a particular type of species-specific vocalizations. Second, there are many aspects of a complex call that cannot be changed by filtering or global operations. In order to address these limitations, a program tool has been developed for analyzing the spectrotemporal pattern of species-specific vocalizations. The design of the program was adjusted for its application in the field of auditory neuroscience.

Materials and methods

The software was programmed in Matlab version 5.3. (and also tested in version 6.5) with the Signal Processing Toolbox as part of a complex software solution for processing communication sounds, containing routines for record-

ing, analysis, modification and synthesis. The user interface employs a graphic environment for the interactive control of sound processing. The file input subroutine can read sound files in the standard wave format (*.wav) or in Matlab binary format.

Before the analysis starts, the segmentation subroutine can be used for separating individual calls or call phrases in the sound record. The segmentation is performed in the time domain, and individual sound segments can be saved into separate files in wave format (wav) or Matlab binary format.

The analysis algorithm is essentially a peak-picking contour extractor. The position of the fundamental frequency as well as the positions of higher harmonics are detected in the short-time sound spectra obtained by windowed fast Fourier transform (FFT). The user can control almost all algorithm parameters, such as the number of points and the overlap of the FFT or the number of harmonics. Individual frequencies are detected as positions of local energy maximums in the sound spectrum. A higher accuracy is obtained by tracking the energy peaks over the entire duration of the sound with a correction performed under the assumption of a continuous time course for all frequencies and the harmonic structure of the spectrum. Any part of the signal whose total energy falls below a certain threshold is assumed to be a silent period and does not affect the analysis algorithm.

Results

The application of the program was tested using calls of adult guinea pigs and very young rats (P10). Figure 1 shows

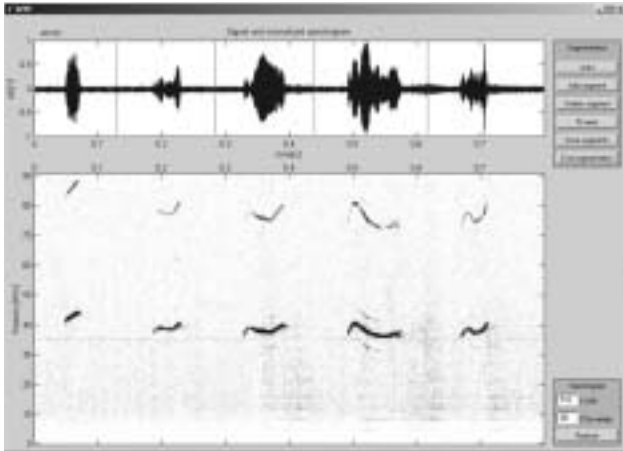


Fig. 1: Copy of the computer screen showing segmentation performed on a ~ 0.8 s long part of a rat vocalization. The upper panel shows the waveform of the signal; the lower panel shows the spectrogram of the signal. Vertical lines in the waveform represent the limits of each time segment as set by the user.

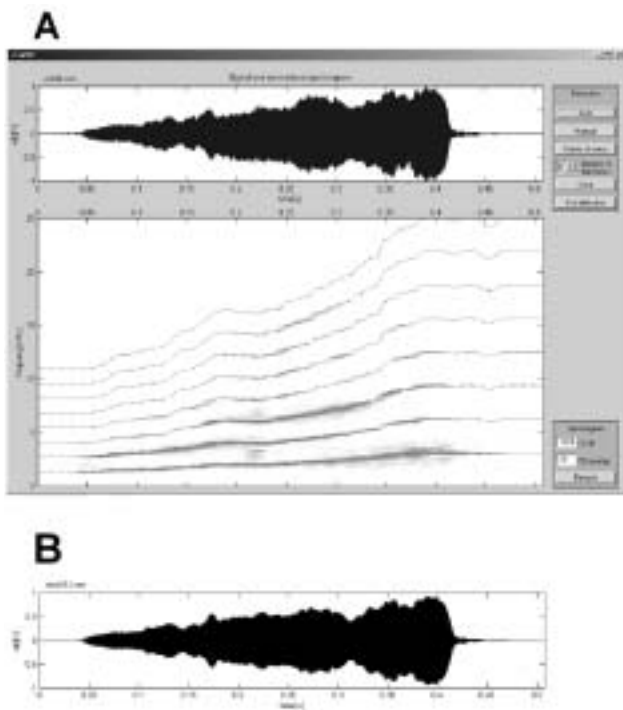


Fig. 2: A. Copy of the computer screen showing the result of signal detection performed on the guinea pig vocalization call 'whistle'. The upper panel shows the waveform of the signal; the lower panel displays the spectrogram of the signal with superimposed lines representing the time courses of individual harmonics as obtained by automatic detection.

B. The waveform of the artificial 'whistle' synthesized from the parameters obtained by the detection algorithm.

the segmentation of a rat vocalization signal sampled at 181 kHz. The signal is graphically represented by its waveform and spectrogram. The signal can be divided into several phrases by automatic operation or manually by introducing marks (shown as vertical lines in the waveform) with a computer mouse. There are two possible reasons to separate individual phrases of calls consisting of many phrases. First, it makes it possible to separate a time segment of interesting signals from uninteresting ones (e.g., a call from noise or another call). Second, if a vocalization consisting of many phrases is divided into individual phrases, the order of phrases can later be easily changed or randomized when altered call is synthesized.

The result of the detection algorithm performed on the guinea pig vocalization call 'whistle' (sampled at 50 kHz) is shown in Fig. 2A. Lines in the spectrogram indicate individual harmonics as found by the automatic detection algorithm. The accuracy of the detected sound parameters is illustrated on the artificial 'whistle' (Fig. 2B), which was synthesized [4] from the detected parameters of the natural 'whistle'. The waveform of the artificial 'whistle' (Fig. 2B) closely resembles the waveform of the original (natural) 'whistle' (Fig. 2A).

The output of the parameter identification module serves for the subsequent statistical analysis of individual types of species-specific vocalizations. The significance of this approach is that it accurately captures the statistical properties of species-specific vocalizations and allows the generation of both representative calls and their artificial variants for studying the representation of this class of complex sounds at various stages of the central auditory system.

Discussion

The algorithm represents a general-purpose solution for the description of species-specific vocalizations, e.g. of bat (3), guinea pig (5,6), rat (4), dolphin (2) or primates (7), but there are also some limitations in its application. First, the procedure assumes an appropriate quality of the recorded sounds. This means a reasonable signal-to-noise ratio and also the isolation of individual calls in the record because an overlap of two or more sounds in the record can lead to improper results (1). The negative influence of these factors can be significantly reduced by user supervision, thus correcting errors in the results obtained by fully automatic detection.

The second limitation is due to the fact that not all communication call types are suitable for 'tone-based' modeling because of their essentially noisy character. Such call types require a different model for analytical description (3). In general, the principle of analytical description can be also applied in the case of 'noisy' calls, but with a more sophisticated algorithm.

References

1. Komárek M. Program tools for processing of the communication sounds of laboratory animals (Diploma thesis). Prague, Czech republic: Czech Technical University, 2003:53pp.

2. Morgenbesser HB, Tyack PL. Synthesis and modification of the whistles of the bottlenose dolphin, *Tursiops truncatus*. *J Acoust Soc Am* 2000;108:407-16.
3. Ohlemiller KK, Kanwal JA, Butman JA, Suga N. Stimulus design for auditory neuroethology: Synthesis and manipulation of complex communication sounds. *Audit Neurosci* 1994;1:19-37
4. Šuta D, Komárek M, Jilek M, Syka J. Software tools for recording and modifying species-specific communication sounds (Abstract). *Plasticity of the Central Auditory System and Processing of Complex Acoustic Signals*, Prague 2003;p56.
5. Šuta D, Kvašňák E, Popelář J, Syka J. Representation of species-specific vocalizations in the inferior colliculus of the guinea pig. *J Neurophysiol* 2003;90: 3794-808.
6. Syka J, Popelář J, Kvašňák E, Šuta D, Jilek M. Processing of species-specific vocalizations in the inferior colliculus and medial geniculate body of the guinea pig. In: Syka J, ed. *Acoustical Signal Processing in the Central Auditory System*. New York: Plenum Press, 1997:431-41.
7. Wang X. On cortical coding of vocal communication sounds in primates. *Proc Natl Acad Sci USA* 2000;97:11843-9.

***Dr. Ing. Daniel Šuta,
Charles University in Prague,
3rd Faculty of Medicine,
Institute of Medical Biophysics and Informatics,
Ruská 87, 100 00 Prague 10,
Czech Republic.
e-mail: daniel.suta@lf3.cuni.cz***