SoundExplorer:
A Workbench for Investigating Source Separation

David K. Mellinger and Bernard M. Mont-Reynaud
Center for Computer Research in Music and Acoustics
Stanford University
Stanford, CA 94305-8180
davem@ccrma.stanford.EDU, bmr@ccrma.stanford.EDU

Abstract

SoundExplorer is a system for computing and displaying information for audio source separation. It provides an interactive workbench with which you can listen to a sound or view any of several representations of it, process the sound with filters for various source-separation tasks, and view the outputs of this filtering. SoundExplorer displays the variables that control the operation of these filters, allowing you to read and change them interactively to see what the effects are. By using a standard for storing computed values and parameters, parts of SoundExplorer are portable to different types of computers.

Introduction

Source separation in human hearing is the process by which a mixture of sounds arriving at the ear is separated into its constituent components [Bregman90]. Simultaneous, as contrasted with sequential, separation is the part of this process that operates at very short time intervals. Simultaneous separation works by extracting from the sound mixture several different types of sound cues that provide evidence about which parts of the sound spectrum come from one source and which from another.

SoundExplorer, working from input of a sampled, recorded sound, computes (using computational models of auditory processes) and displays several of the cues used for simultaneous source separation: harmonicity, common frequency change, common onset and offset, common amplitude modulation, and so on. It shows visual images of the computed results; it does not currently have a method for re-synthesizing sound from the separation images.

Feature Maps

Data in SoundExplorer is organized in feature maps. A feature map is a rectangular array of up to (currently) three dimensions. The elements of this array are real values (floating-point numbers) that represent intensity of one sort or another. These values are produced by the various computational filters in SoundExplorer. For example, one feature map for a snippet of piano music is displayed in fig. 1. It is a two-dimensional intensity map with logarithm of frequency on the vertical axis and time on the horizontal. SoundExplorer is a system for computing, displaying, and storing feature maps.
The Computation Engine

The backbone of SoundExplorer is a set of programs for computing various types of feature maps for studying source separation. To study a sound in SoundExplorer, you must first obtain a sampled, recorded version of the sound in Next Soundfile format. SoundExplorer uses the recorded sound as input to a computational model of the cochlea [Lyon82, Slaney88], which produces the first two feature maps: a two-dimensional image (time x frequency) of cochlear neuron firing rates, and a three-dimensional image (lag x frequency x time) of the per-frequency-channel autocorrelation of the neural firings.

From these images, you can use SoundExplorer to do any of a variety of things to process the above images to useful extract information. Each further processing step takes as input one or more of the existing feature maps and produces as output one or more feature maps. You run each filter by typing its name, any necessary parameters, and name of the sound. The program reads the sound file and any feature maps it needs, computes its output feature map(s), and writes the results to a new files.

Each type of filters in SoundExplorer extracts a different type of information useful for sound source separation. Details of these filters is beyond the scope of this paper. It is fairly easy to add new filters to SoundExplorer to extend its capabilities.

Running all of the filters in SoundExplorer takes up to 16 hours for one second of sound in the worst case. You can speed up this process considerably if other Unix computers are available on a local network. The slower filters are capable of spreading out their workload over locally networked machines. Since SoundExplorer's computation engine is portable to other kinds of computers, the others need only be connected by a network.

With all of the above options available, you easily forget which filters need which input parameters, which filters must be run before which others, and even the names of the filters.
SoundExplorer has several solutions to this problem. First, each filter program knows the standard values for its parameters, so you need only provide the name of the sound to make the filter do the standard filtering. Second, a program named crunch runs any or all of the filters in order. For example, you can say "crunch on ear miles" to run the autocorrelation and onset filters in that order, as necessary, or "crunch all miles" to run every existing filter. Finally, you can avoid all of these mnemonic commands entirely by using SoundExplorer’s interactive browser.

The Interactive Browser

Sitting down to use SoundExplorer, you see the interactive browser. This part of the system displays feature maps and allows you to examine and change them in various ways.

Feature maps come in one, two, and three dimensions. The browser displays one-dimensional maps in the ways shown in fig. 2, where the height of the lines represents intensity. Two-dimensional maps are shown as grey-level images, as seen in fig. 1 above. In these images, the grey level shows intensity, with darker areas for higher (more intense) values. You can change the transfer function of intensity values to grey levels, enhancing low-contrast areas of the image.

Three-dimensional maps are displayed as two-dimensional animated images. Since all types of three-dimensional feature maps currently used have a time axis, it works quite naturally to display the time dimension of the feature map as the time dimension in animation.

The remaining two dimensions map to the screen.

Each map displayed has a set of associated parameters. The set varies depending on the type of filter used to compute the map. The SoundExplorer browser, when displaying a map, allows you to see the type of map and its associated parameter values. These values are editable text, so you can simply point the mouse at one of them, type in a new value. Some types of maps also have additional graphical information that SoundExplorer displays with the map. For example, fig. 3 shows the result of a convolution operator applied the cochlear model map for a sound, together with the kernel of this two-dimensional operator.

Each feature map has its own window on the Next machine screen. This lets you hide the maps you no longer need or re-display the ones you need again, and makes it easy to move them around independently. The latter feature has proves especially useful, since you can line up different images next to each other on the screen for comparison.

Another useful feature is the mouse tracker. Any time you click the mouse button on a feature map image, the intensity of the map at that point is recovered and displayed, along with the mouse's position with the map. Fig. 4 shows the mouse position in frequency and time and the map value at that position. It also shows the horizontal and vertical slices through this time/frequency map – a frequency slice and spectrum, respectively.

ICMC 92
Figure 3. FM map with convolution kernel.  
Figure 4. Mouse position data.

You can also perform filtering with SoundExplorer's computational engine via the interactive browser. After changing parameter values for a filter, you can press a "Compute" button, which makes the browser take all the relevant parameter values and start up a new process that performs the filtering you asked for. When the computation is finished, the resulting feature map is displayed for further perusal.

Map Storage Standard

One goal of SoundExplorer is to have the computational engine be portable to different types of computers, so that it's not restricted to just Next machines. For this to be possible, the files that SoundExplorer stores its feature maps in must be standardized. SoundExplorer's standard has three parts: a standard way to store feature maps, a standard way to store the parameters associated with a map, and a standard way of placing map files in a directory structure to keep track of them all. Details are available on request.

This work was supported in part by National Science Foundation grant IRS-8613574.

References

