Sonification of NRL Dual-Task Data

Brian McClimens

Naval Research Laboratory
4555 Overlook Ave, Washington DC
brian.mcclimens@nrl.navy.mil

Derek Brock

Naval Research Laboratory
4555 Overlook Ave, Washington DC
derek.brock@nrl.navy.mil

1. INTRODUCTION

The addition of a head-tracker to the dual-task experiment reported in Brock et al., 2010 led to two distinct sonifications of the collected data [1]. Initial analysis of the head-tracking data was attempted in Microsoft Excel, and the size of the data sets made the visual comparison of multiple log files intractable. The head-tracker logged data approximately every 70 ms, and each condition was thirteen minutes long, so a graph of a single condition contained over eleven thousand points. This proved too large to be managed from within Excel. In exploring possible programs to view the data in we realized that our log files could easily be modified to conform to the .wav file format.

Initially, the realization that the head-tracking data was essentially a digitally sampled analogue signal was used only as a means transform the data into a form that could be visually examined from within sound editing software. Sound Forge™ allowed us to simultaneously view the data from multiple subjects. Although the data had been changed to a .wav format, it was not specifically intended as a sonification in the sense that no thought had been put into how the data would sound. Nevertheless, the format used happened to produce meaningful sound and with very little practice, several colleagues could easily differentiate between sonifications of conditions which had been shown to have statistically significant differences in the head movements of subjects.

This spurred the construction of a more detailed sonification of our data which included not only the data about the motion of subjects' heads, but also data collected about their tracking performance and a detailed event log from the radar task. The goal was to sonify as much of the collected data as possible, and then to see whether or not people could differentiate between data from each of the different conditions used in the experiment.

2. Data and Sonification Methods

The NRL Dual-Task consists of a tracking task and a radar task. The tracking task is a continuous task in which subjects are required to follow the movements of a target on the screen with a joystick. The radar task involves a series of classification events in which the subject responds to the behavior of blips on the screen by classifying them as either hostile or neutral. Each classification event requires two key presses, one to denote the determination of hostility, and one to identify the blip. Each condition had a duration of thirteen minutes, with a total of 65 classification events. A detailed description of the task can be found in Brock et al., 2010[1]. Two sonifications have been created from the data collected in this experiment. The first sonification only made use of the head-tracking data, but a second sonification used performance data from the tracking and radar tasks.

2.1. First Sonification

Subjects wore a pair of headphones fitted with a head-tracking device that logged the direction the subject was facing approximately every 70 milliseconds. This logging was slightly irregular, with some samples separated by as much as 100 milliseconds. In order to most effectively display this data as a .wav file, evenly spaced samples are needed. Because the magnitude of head movements between any two points in our dataset was small, we were able to use a simple linear interpolation between points to create a dataset with points spaced exactly 50 milliseconds apart. These data points were measured in radians and ranged from approximately -0.3 to .3 radians.

The initial sonification used the interpolated, evenly spaced data, and converted it to an audio format in the most direct possible way. Each data point was scaled and converted to a 16-bit integer between -32677 and 32678. These were then written as sample points in a 16-bit, 8kHz PCM format. Because each sample represented 50 milliseconds of time in our experiment, and 8kHz audio stream represented a playback speed four hundred times faster than real-time, and each thirteen minute condition was translated into a 1.95 second sound file.

As examples, eight auditory files are included. The first four are sonifications from a single subject of the four experimental conditions described in Brock et al., 2010 [1]. The conditions are referred to as the "no sound" condition (NS), the "non-augmented auditory reality" with three sounds (NAAR-3), the "augmented auditory reality" with three sounds (AAR-3), and the "augmented auditory reality" with one sound (AAR-1). For details about the difference of the conditions, please refer to the referenced paper. The key issue for the purpose of this example is that in the NS condition, subjects performed significantly worse than in others, and had a greater number of head turns.

The first sound clip (Son1_1-1.wav), is a sonification of subject 1's head movement in the NS condition. When compared to the other three files (Son1_1-2.wav, Son1_1-3.wav, and Son1_1-4.wav), subtle differences are apparent. Although each condition sounds like low-frequency static, the 1-1 file is noticeably smoother. In the other conditions (NAAR-3, AAR-3, and AAR-1), subjects performed better, with fewer unnecessary head turns. In these sonifications, the relatively large gaps of time with no head movement producs short silences followed by a noise sounding like a pop or a click. With a small amount of practice, we found ourselves able to correctly identify sonfications of the NS condition. In order to test this for yourself, four additional auditory clips are provided (Son1_2-a.wav, Son1_2-b.wav, Son1_2-c.wav, Son1_2-d.wav). These four files are the Sonifications of another subject's head movement in the NS, NAAR-3, AAR-3, and AAR-1conditions. There are distinct differences between subjects, but you should be able to identify the sonification of the NS condition. The conditions represented by the four files are revealed in a text file (Son1_key.txt).

3. Second Sonification

A second sonification was created with the goal of including all data that had been collected during the experiment. In addition to the head-tracking data, this sonification would include data about the subjects' performance on both the tracking task and the radar task, as well as data describing the state of the radar task.

3.1. Head-tracker

The second sonification made use of the interpolated data points used in the first sonification described above. Instead of directly mapping those data points to samples, this sonification used simple properties of the data set to modulate a square wave. The value of the data point was used to position the square wave such that the left-right panning corresponded to the direction that the subjects' head was facing. In addition, the rate of head movement as measured by the magnitude of change between neighboring data points was used to modulate the frequency and amplitude of the square wave. In order to make movement seem continuous, all values are interpolated from the two nearest data points. The goal was to have a relatively subtle sound that would indicate the subject's current focus to a listener, and modify it so that head movements would be noticeable with larger head movements more salient than small ones.

3.2. Joystick

In the tracking task, a subject is required to keep a reticle over a target by controlling the reticle with a joystick. The log file for this task records the position of the target, position of the reticle, and the position of the physical joystick on both the x-axis and y-axis. Each of these values is recorded every 83 milliseconds (12 Hz). From this log, interpolated points are calculated every 50 milliseconds in a method equivalent to what was used for the head-tracking data. From those data points, we calculated the distance between the target and reticle at 50 millisecond intervals. White noise was then synthesized, with the distance between target and reticle used to modulate amplitude so that the noise is quiet for periods where the subject performs well, and increases in volume as a subject's performance decreases.

4. Radar data

In the radar task, subjects see blips of three different types traveling from the top of the screen to the bottom. They are required to classify these incoming blips as hostile or neutral based upon rules that differ for each blip type, and must enter their designation via the keyboard. Theses responses are not allowed until the blips are about halfway down the screen. Prior to this point the blips are grey, and when a response is allowed, they change color. In most conditions of the experiment, an alarm is played at the same time as the color change, but in one condition, no alarms are played. The primary metric used to evaluate subject performance on the radar task is the reaction time measured from when a blip changes color to when the subject completes a designation entry.

The log files for the radar task contain several types of information. Each time a blip appears on the radar screen, becomes active or is classified either incorrectly or correctly by the subject, an event is logged. In addition, every keystroke performed by the subject is recorded, and categorized. Each of the events is encoded and represented within the sonification as short sound clips. These are not synthesized or modulated in any way, but are simply a mapping of the possible event types used to trigger the playback of preset audio clips. Finally, a 440 Hz sine wave was used to indicate periods of time in which the task was waiting for a response from the subject. For all time periods in which a colored blip was present on the screen, a sine wave is generated. In this way, a sonification in which the sine wave was more prevalent reflected poorer performance by the subject.

5. Auditory Examples

Two 16-bit, 44.1kHz .wav file examples of this sonification are included (1-1_20-mix.wav, 1-2_20-mix.wav). The algorithm used to sonify the NRL Dual Task data has a parameter for temporal compression. If the temporal compression is set to 1, then the resulting sonification will be equal in length to the data provided. A temporal compression of 400 would result in a file with the same duration as the first sonification. The examples provided are presented with a temporal compression rate of 20.

Though we hoped to create sonifications that would comfortably include all of the different types of data from the experiment, I find that the sonifications are easier to understand when they are split into data from the head-tracker, joystick data and radar data.

5.1.1. Head-tracker audio

The first data file (HT_1.wav) is a sonification in which the data is being presented in real time. This sonification is a short segment of data from subject 1's head motion during the NS condition. When presented with no temporal compression in the sonification algorithm, it is relatively easy to keep track of the head's position, and to get a sense of the speed with which the head was moved for any individual head turn. Unfortunately, it is very difficult to compare two of these sonifications and notice differences in a subject's performance or behavior based upon them. Increasing the compression improves the ability to compare between sonifications, while giving a less clear picture of head positioning. As an example, two sound clips of sonifications with a temporal compression rate of 20 are included (HT_20_1-1.wav, HT_20_1-2.wav), along with two examples at a compression rate of 400 (HT_400_1-1.wav, HT_400_1-2.wav). These files represent the head movements of subject 1 in the NS and AAR-3 conditions at two different speeds. The files with a compression rate of 400 contain the entire data set for their conditions. At a compression rate of 20, individual head turns are difficult, but possible to perceive. However, it is also difficult to distinguish between the two conditions. The files with a compression rate of 400 make it impossible to identify individual motions, but a different property emerges which makes it easy to distinguish between conditions. At the higher compression rate, the average position of a subject's head can be heard more clearly, with the subject's head position falling far right in conditions other than NS due to the smaller number of head turns.

5.1.2. Joystick audio

Of all the portions of the sonification, the joystick data changed least with different temporal compression rates. The joystick sonification is a direct measure of a subject's performance on the joystick task but in most cases, the differences between conditions in the joystick files are not apparent. Included are the four sonifications of subject 9's performance in each condition (JS_9-1.wav, JS_9-2.wav, JS_9-3.wav, JS_9-4.wav). These files have a compression rate of 400, and represent the entire data set for subject 9.

5.1.3. Radar audio

The main quality that stands out as a difference between conditions in the sonification of the radar data is the presence or absence of the sine wave indicating a colored blip on the screen. This closely related to the subject's reaction time since the completion of a response (whether correct or incorrect) will remove a blip from the screen and results in less time with a colored blip onscreen. Again, the differences between conditions are more pronounced when the compression rate is higher. Another indication of condition type is the presence of the (Miss.wav) sound. Most misses occurred in the NS condition. The number of occurrences is small though, and some subjects were able to complete the NS condition without missing any blips, while others made a few errors of omission in other conditions as well. Two radar files at compression rates of 400 are included for comparison (R_400_1-1.wav, R_400_1-2.wav)

6. Discussion

In the process of sonifying the NRL Dual-Task data, we found it possible to highlight all of the statistically significant features reported on in Brock et al., 2010. We had hoped that these sonifications would provide some insight that might compliment our data analysis. Unfortunately, no distinguishing qualities between conditions have yet been noted in the sonifications that was not present in our statistical analysis. One difficulty with exploring the data was that parameters' settings had to be chosen in advance, and each time we wanted to change a parameter, the program needed to be recompiled, and the sonifications regenerated. This process was quite time consuming, especially for the sonifications with lower temporal compression rates. In order to explore this data more effectively, a dynamic interaction between the listener and the sonification may help. A sonification that could be generated in real-time, with controls to change parameters might allow for a more thorough exploration, and may enable a listener to discover features of these data sets that have not yet been identified.

7. REFERENCES

[1] Brock, D., McClimens, B., and McCurry, M. “Virtual auditory cueing revisited,” In *Proceedings of the 16th International Conference on Auditory Display*. Washington, DC, June 9-15, 2010.