WHAT ARE YOU REALLY THINKING?
A Sonification for ICAD2004

Gordon Monro

Sydney Conservatorium of Music
University of Sydney C41
NSW 2006 Australia
gmonro@mail.usyd.edu.au
www.gordonmonro.com

ABSTRACT

This note describes a response to the call for sonifications issued in connection with ICAD2004. The resulting piece is entitled "What are you really thinking?"

1. OVERVIEW

The call for sonifications issued in connection with ICAD2004 presented a challenge to the scientific and artistic community: to convert to sound 36 channels of data (mostly electro-encephalograms, but also heartbeat, respiration and other data) recorded while a subject was listening to a piece of music.

The sonification presented here, entitled "What are you really thinking?" attempts to use essentially all the data given on the ICAD2004 website [1].

For this sonification, a distinction was made between the **scalp signals**, the 26 data channels recorded from various locations on the scalp of the subject, and the **non-scalp signals**, including the recordings made near the eyes (the "eye signals"), the masseter muscle signal, the electro-dermal signal, breathing and heartbeat.

Each scalp signal is considered as a sound source and converted to a **stream of notes**, as explained below. Most of the non-scalp signals control aspects of the streams of notes, but do not generate sound directly. However, the heartbeat signal is also used as a sound source.

The streams of notes are spatialised according to the coordinates of the scalp electrode positions, but with modifications due to the eye signals.

The main aspect of the sonification is the representation of scalp EEG activity as pitches directly corresponding to the frequencies occurring in the scalp signals.

2. CONSTRUCTION OF A STREAM OF NOTES

2.1. Fourier Analysis

Each scalp signal was treated as follows. The slow drift visible on these files was removed, and a Fourier analysis was performed (150,000-point FFT). Frequencies below 0.02 Hz and above 149.8 Hz were removed (the upper limit was chosen to avoid the third harmonic of 50 Hz hum). In addition frequencies between 49.9 and 50.1 Hz were removed.

A low-frequency band covering the range 0.02 Hz to 0.5 Hz was treated separately. The range 0.5 Hz to 149.8 Hz was divided into 100 frequency bands, of linearly increasing width; each frequency band was used to generate a stream of notes of a specific pitch.

2.2. The Notes From One Frequency Band

For each frequency band between 0.5 Hz and 149.8 Hz, the signal in that band was reconstructed (inverse FFT). Its envelope was then found, and a threshold applied.

To construct the notes, the band-limited signal is scanned from the beginning. If the envelope of the signal exceeds the threshold for the band, a **note** is created. There is then a "refractory period" before another note can be started. The duration of a note is two periods of the centre frequency of the band; the pitch of the note is 30 times the centre frequency of the band. The timbre of the note is synthetic, made up of relatively sharp triangular spikes. The refractory period (the minimum time between the start of one note and the start of the next) depends on the centre frequency of the band and also on the value of the low-frequency band (0.02 – 0.5 Hz) at that point in the the signal.

The effect is that activity in a certain frequency band will produce repeated audible notes at a corresponding pitch. For example, activity in the frequency band 8.53 – 9.0 Hz will generate notes at a frequency of $30 \times 8.765 = 262.95$ Hz, which is close to that of Middle C.

The total number of notes is large: around 14,000 per scalp electrode over the 5 minute duration of the piece; over 360,000 in total. Most of the notes are quite high-pitched and very short.

As noted above, the pitches of the notes form the main aspect of the sonification: representation of scalp EEG activity as pitches corresponding to the frequencies occurring in the scalp signals.

2.3. The Volume-Control Effect

The volume of a note in the sonification depends on the value of the envelope of the signal of its frequency band. If this value is close to the threshold, the note is soft; if it is well above it, the note is loud.

For natural sounds and notes from musical instruments, loud sounds are not identical with soft sounds amplified, but have different timbres. With synthetic sounds, if the loud sounds are just the same as the soft sound amplified, the result sounds unnatural. This unnatural effect is called the **volume-control effect**.

The volume-control effect was avoided in "What are you really thinking?" by low-pass filtering the softer notes.

No attempt was made at naturalness in other aspects of the sonification: in particular linking reverberation to the respiration signal (see below) does not produce natural-sounding effects.

3. MODIFICATIONS OF THE NOTES

The non-scalp signals were used to modify the notes. Since the non-scalp signals were used as control signals, they were low-pass filtered (by Fourier filtering) to below 40 Hz. Most of them had a slow drift (trend line) removed, though this was not done for the electro-dermal (EDA) data.

The only signal not used at all was the Erbs point signal, as inspection showed it to be very similar to the heartbeat signal.

- If both the Orbicularis Oculi signal and the low-frequency band from the scalp signal are high, the positions of the spikes making up a note are jittered, resulting in noisy sounds.
- If the Masseter signal is high, the spikes making up a note become very narrow, resulting in a nasal quality.
- The Respiration signal was very noisy, but it was possible to extract a reasonably regular fluctuation with a period of around three seconds. This filtered Respiration signal controls **apparent distance**, in that when it is high the notes are louder and have little reverberation, and when it is low the notes are softer and have more reverberation.
- Large excursions in the Respiration signal additionally affect the pitch of the generated notes. This effect is only noticeable for the big excursion 95 seconds from the start of the data, where it can be heard as a rapidly rising sequence of pitches.

The use of the remaining non-scalp signals is described below.

4. SPATIALISATION

In general each stream of notes corresponding to a scalp signal is placed at the position given in spherical polar coordinates on the ICAD2004 website for the corresponding electrode.

For electrodes forward of the line joining the ears (temporal left to temporal right), the position of the stream of notes is moved dynamically according to the eye signals.

The two control signals derived from VPVA and VNVB were highly (inversely) correlated (correlation -0.879), so a control signal was formed by the difference VNVB minus VPVA. If this difference signal is high, the streams of notes for the forward electrodes are pulled upwards (increasing elevation: towards the zenith); if it was low the elevations are decreased.

Similarly the control signals derived from HPHL and HNHR were found to be highly correlated, and a difference signal was formed. This difference signal pulls the streams of notes towards the right or the left, depending on whether it is high or low.

5. THE HEARTBEAT

The heartbeat channel was used directly as a sound source. The volume, and amount of reverb added, were controlled by the EDA signal. The heartbeat sound is not given a particular spatial location, but is present in all the ear-level (lower ring) speakers.

6. SUMMARY OF THE SONIFICATION

- The activity at each scalp electrode is represented by a stream of notes. The pitch (frequency) of a note indicates the frequency band in which activity is occurring: the frequency of the note is 30 times the centre frequency of the corresponding frequency band.
- The direction from which a note is coming corresponds to the location of the scalp electrode (in polar coordinates), except that eye movement signals cause notes corresponding to electrodes near the front of the head to be pulled up or down, and to the left or to the right.
- Noisy sounds indicate a high value of the Orbicularis Oculi signal.
- Nasal sounds indicate a high value of the Masseter signal.
- Fluctuation between quieter and more reverberant sounds and louder and "drier" sounds is controlled by the Respiration signal.
- The heartbeat is more prominent for high values of the EDA signal.

7. THE START AND END OF THE PIECE

All the above transformations of the data to sound were carried out completely mechanically (and therefore reproducibly), by a sequence of computer programs. The one piece of information not used in this process is the fact that the subject was listening to a piece of music.

It is not to be expected that a sonification produced in a deterministic manner from the data will have any of the normal characteristics of a piece of music. In particular, though there may be short-term order, the sort of larger-scale structure expected in a piece of music will be absent. Specifically, a piece of music has a start and an end, while the EEG data is essentially an arbitrary slice from an unending stream (despite the fact that in this case it corresponds in time to a particular listening experience).

The only "musical" characteristics I have added to the sonification are a start marker and an end marker. The start marker just consists of a quick fade-in, with the heartbeat sound being faded up before the note streams are.

The end marker consists of the fading down of the sonification, and the introduction of some sounds from the outside world. Apart from marking the end of the piece, the introduction of these sounds is intended to remind us that despite the large amount of data available for the sonification, we still have no access to the subject's thoughts, feelings or perceptions; in other words to the contents of the subject's consciousness. The title "What are you really thinking?" refers to this fact.

8. SOFTWARE USED

The work was carried out on a PC running Linux. Almost all the processing was carried out by programs in C++ written by the author. Processing was generally carried out using double precision floating-point arithmetic, at a sample rate of 48kHz. This sample rate was used because it is an integral multiple of the data sampling rate.

- The Fourier transforms were carried out using the FFTW3 library [2], version 3.01. This library handles transforms of any length, not just powers of 2.
- The spatialisation across the hemisphere of speakers was done by dividing the hemisphere up into triangles with the speakers at the vertices (obtaining 18 triangles) and then using the vector base amplitude panning method of Pulkki [3].
- The gnuplot program [4], version 3.7, was used to view data.
- The Audacity program [5], version 1.2.0-pre1, was used to view and play sound files. See also [6].
- The final conversion from 48kHz to 44.1 kHz was carried out by CoolEdit Pro LE (under Windows). This software is apparently no longer available.

9. LIST OF FILES

The 16 files provided correspond to the 15 speakers, plus a subwoofer channel. The files are:

speaker01.wav (ear level front centre)
speaker02.wav (ear level front left)
speaker03.wav (ear level front left wide)
speaker04.wav (ear level rear left wide)
speaker05.wav (ear level rear left)
speaker06.wav (ear level rear centre)
speaker07.wav (ear level rear right)
speaker08.wav (ear level rear right wide)
speaker09.wav (ear level front right wide)
speaker10.wav (ear level front right)
speaker11.wav (upper level front)
speaker12.wav (upper level left)
speaker13.wav (upper level rear)
speaker14.wav (upper level right)
speaker15.wav (zenith)
speaker16.wav (subwoofer)

10. REFERENCES

[1] "ICAD 2004 Concert: call for sonifications" [website], http://www.icad.org/websiteV2.0/Conferences/ICAD2004/concert.htm (accessed 29 March 2004).

[2] M. Frigo and S. G. Johnson, FFTW ("The Fastest Fourier Transform in the West") [website], http://www.fftw.org/ (accessed 29 March 2004).

[3] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning", *J. Audio Eng. Soc.,* vol. 45, no. 6, pp. 456–466, June 1997.

[4] "Gnuplot Central" [website], http://www.gnuplot.info/ (accessed 29 March 2004).

[5] "Audacity" [website], http://audacity.sourceforge.net/ (accessed 29 March 2004).

[6] D. Mazzoni and R. Dannenberg, " A Fast Data Structure for Disk-Based Audio Editing", *Computer Music J.,* vol. 26, no. 2, pp. 62–76, Summer 2002.