

"PERCEPTIONS IN C": SUBMISSION FOR THE CONCERT OF SONIFICATIONS AT ICAD 2004

John Sanderson[†] and Tom Heuzenroeder[‡]

[†]CSIRO Division of Minerals, Box 312 Clayton South, Victoria, Australia 3169

[‡]Sonic Art Studios, Box 577 North Adelaide, S. A., Australia 5006

john.sanderson@csiro.au

tom@sonicart.com.au

ABSTRACT

"Perceptions in C" is a submission for the ICAD2004 Concert of Sonifications and was composed as a first experiment by us in the field of converting data to music. Our principle objective was to see if the ECG data (or part thereof) could be made at all "listenable" from a musical standpoint. Central to the data processing approach was the use of a colour note organ to directly interpret the plotted signals as sound. Discretizing the signal values to correspond to frequencies in the equal-tempered scale was our attempt to make the data sound musical. Some audio post-processing was used to improve discernment between different channels of data. We feel that in addition to being somewhat musical, the pitch of the resulting sounds effectively represents the general signal trends in the selected data.

1. INTRODUCTION

To be perfectly honest, the idea of listening to experimental data only occurred to us rather recently over coffee when we should have been discussing something else entirely. But curiosity got the better of us and we tried converting some pressure fluctuation data from experimental trials of a fluidized bed gasifier to a pcm wav file so we could "hear" what the data sounded like. Little did we know that a term was already coined for what we were doing - *sonification*.

We had already attempted various means of detecting changes in these data numerically [1] and then (recognizing the human brain as a great tool for use in pattern recognition), we became interested in graphical interpretations of data, such as [2]. However, listening to the data was a new option entirely, and interesting to us because of the brain's remarkable capacity for auditory discernment. It would seem to be an extraordinary coincidence that we then received information about the ICAD event a week or so later.

This submission, therefore, is the work of novices who are by their own admission poorly read in the field of sonification! Our approach was to try and get the data to sound at least to some extent musical and pleasing to the ear. To that end the data were discretized in order to correspond to the notes of the equal-tempered scale when converted to frequency via colour note organ (Coagula Ver. 1.666 [3]).

Only 7 selected channels of the full 36 channels were used in this composition. They are listed in Table 1. The majority of the numerical processing steps were carried out using standard features of Microsoft Excel and are described in Section 2 of this document. The processing of the graphical data via colour note organ is described in Section 3. Finally, as an attempt to enhance the output sound files and allow for some discernment

between channels, some audio post-processing was carried out, as described in Section 4.

2. NUMERICAL PROCESSING

The following seven channels were selected for sonification: Ch1, Ch17, Ch30, Ch31, Ch34, Ch35, Ch36, purely on the basis that this selection exhibited a wide variety of trends for the 300 second duration of the data set.

In the context of our processing method, the initial data sampling rate of 500 Hz may be considered very high. The first step of the process was to subsample the data to 1/20th of its original rate. This was carried out using a UNIX-based command line program that simply skips 19 out of every 20 points in the input file. Note that this is a *sub*-sampling process, not *re*-sampling involving averaging or interpolation.

Once the subsampling step was carried out, the data (now consisting of 7500 points per channel) were imported into Microsoft Excel for numerical processing.

For the selected datasets, a normalization step was carried out to establish a similar range of values for the signals via:

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

where Z is the instantaneous normalized value, X is the instantaneous value of the signal, μ is the mean of the entire signal and σ is the standard deviation of the entire signal. With the exception of the large negative spike in the Ch35 data, this mapped the signals to the range $(-5 < Z < 5)$.

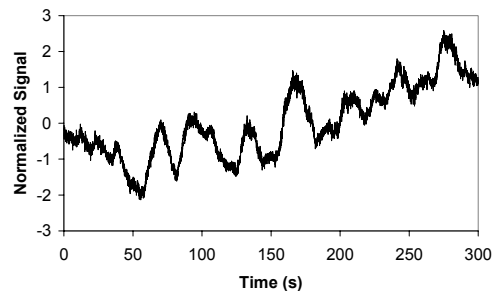


Figure 1. *Subsampled and normalized signal - Ch01-Fp1.*

Bearing in mind that the colour note organ works (by analogy) rather like a voltage to frequency converter, the next step was to re-scale the normalized data sets so that the dynamic range of the signal (Y-axis) spanned the range of audio frequencies normally associated with music (C₀ to D₈[#]; or 16.35 to 4978.03 Hz). This was done via:

$$P = (Z + C_1)m - C_2 \quad (2)$$

where P is the final value of the data point, Z is the input data point value, C_1 is an arbitrary constant, m is a multiplier and C_2 is a second constant for applying additional off-set. Values of the constants and multipliers used are given below.

Data File	C_1	C_2	m
Ch01-Fp1	5	1220	500
Ch17-CPz	5	2257	1000
Ch30-HNHR	5	50	50
Ch31-Erbs	5	4000	1000
Ch34-EDA	5	3500	1000
Ch35-Resp	5	-50	10
Ch36-ECG	5	4000	1000

Table 1: Data sets used in the sonification with associated re-scaling constants and multipliers (Eqn 2).

A trial and error approach was used for the selection of the values of C_1 , C_2 and m , with the intention being to create data that varied in an interesting way throughout the abovementioned audio frequency range.

To remove the influence of high frequency fluctuations in the signal (these would translate to lots of very rapidly changing notes in the final composition), a moving average (M.A.) was applied to the re-scaled data, with the exception of Ch17 and Ch35 where no averaging was used. Again, the length of the averaging was selected on a trial-and-error basis. Table 2 provides the details:

Data File	M.A. length
Ch01-Fp1	20 points
Ch17-CPz	N/A
Ch30-HNHR	20 points
Ch31-Erbs	100 points
Ch34-EDA	100 points
Ch35-Resp	N/A
Ch36-ECG	100 points

Table 2: The length of moving averages (M.A.) applied to the re-scaled data.

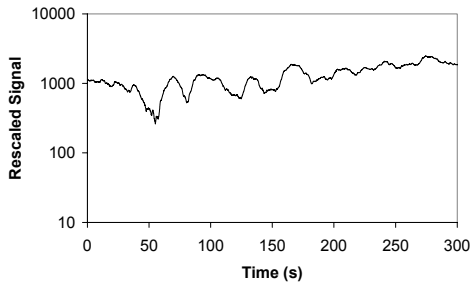


Figure 2. Signal after re-scaling to audio frequency range and applying 20-point moving average (plotted on log axis) - Ch01-Fp1.

The data were then discretized to the frequency values corresponding to notes on the equal-tempered scale (see Appendix 1). To keep things simple we used the C major scale, C minor scale and chords based on the C major or C minor triad (eg $E_3 C_3 G_3$ for the former). The 1 minute in minor key was a deliberate sub-experiment in order to evaluate how clear this small alteration to the discretization function would be to the ear. The data were discretized by use of the Excel VLOOKUP worksheet function whereby for a given input value from the

data set, the function returned the nearest value from the specified table of frequencies (eg the list of frequencies corresponding to the C major scale). The discretization was allocated to the channels and with respect to time as shown in Table 3.

Data File	Discretization		
	0-59 sec	60-119 sec	120-300 sec
Ch01-Fp1	C Maj. (S)	C Min. (S)	C Maj. (S)
Ch17-CPz	C Maj. (T)	C Min. (T)	C Maj. (T)
Ch30-HNHR	C Maj. (S)	C Min. (S)	C Maj. (S)
Ch31-Erbs	C Maj. (T)	C Min. (T)	C Maj. (T)
Ch34-EDA	C Maj. (T)	C Min. (T)	C Maj. (T)
Ch35-Resp	C Maj. (T)	C Min. (T)	C Maj. (T)
Ch36-ECG	C Maj. (T)	C Min. (T)	C Maj. (T)

Table 3: Summary of discretizations applied to the data. S = scale, T = triad.

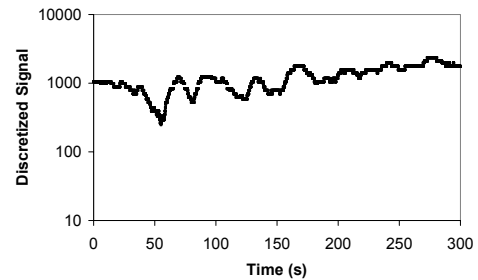


Figure 3. Signal after discretization to C Major scale (time ranges 0-59 seconds and 120 - 300 seconds); and C Minor Scale, (60 - 119 seconds) - Ch01-Fp1.

Finally, the data were plotted using the Excel charting function. The formatting for the charts is of particular importance since this dictates the way the colour organ will interpret the data. For each channel of discretized data, the chart was plotted:

- for the time range 0 to 300 seconds
- with logarithmic y-axis scaling and range from 10 to 10,000.
- with points represented using black 2-point size square blocks and no interconnecting lines

The Excel chart for each channel was then copied to the Microsoft Windows Paint accessory program, where the axes were cropped off, leaving only the chart area with the data points to be saved as a *.bmp file.

3. COLOUR NOTE ORGAN

The Coagula colour note organ was downloaded from the Internet for the task of converting the graphical data to sound. The colour note organ is an image synthesizer that can be used to create and manipulate images and then generate sound from those images. Generated sound is based on the addition of sinewaves. For a given pixel within the image, the vertical position determines its frequency, the horizontal position corresponds to time and the brightness controls the volume.

(Note that colour can be used to control stereo placement and add noise effects - in our case the colour features were not used.)

In the current application, the conversion process was a simple one, carried out as follows:

- The bitmap file chart of the 300 seconds of data for a certain channel was opened in Coagula.
- The Invert Colour function was applied (this turns the original chart into white dots on a black background rather than the reverse).
- The image was rendered to sound with the following settings:
 - Time = 300 sec
 - Amplification Factor = 5.00
 - Noise Bandwidth = minimum
 - Soft Envelope Sweeps = disabled
 - Pitch Range...10,000Hz to 10Hz

Note that because Coagula uses a logarithmic pitch scale (on the Y-axis), with the above settings the charts plotted using the original logarithmic scale in Excel will map faithfully to the respective notes assigned to each horizontal level of the Coagula image.

Once the image for a given channel of data had been rendered to sound, it was exported from Coagula as a *.wav file for audio post-processing.

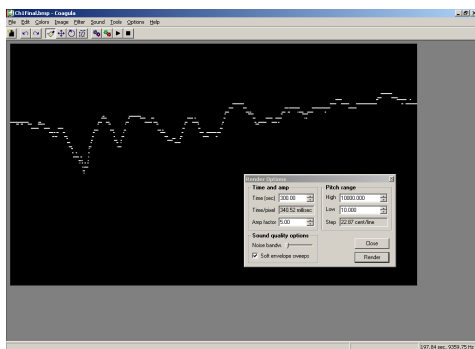


Figure 4. Screenshot from Coagula of signal graph after colour inversion - Ch01-Fp1.

4. AUDIO POST-PROCESSING

The motivation behind the audio post-processing steps was to provide some perceptible musical differences to the 7 sequences of tones produced from Coagula. This was in order to assist the listener in distinguishing one line of data from another as one would distinguish different types of musical instrument in an ensemble.

On their own the sinewave-based Coagula generations are very pure, and unlike the acoustically-generated sounds that we are used to, each tone within them has no real harmonic structure. Harmonic structures differ depending on what we listen to, and help us in the physical world to determine what instrument we are hearing. Even synthesizers are programmed mostly to generate harmonically rich sounds, so in these audio post-processing steps we are attempting to adhere to a more musical discipline.

The approach taken was to translate the musical tones of some of the lines of Coagula output (namely Ch1, Ch17, Ch30, Ch31) directly to MIDI messages so that synthesizer instruments could be "played", whilst distortion or compression effects were added to the other lines (Ch34, Ch35 and Ch36).

Below, the specific settings for all the audio post-processing steps are given for each of the 7 lines of Coagula pcm wav output files.

Ch1 - Fp1

This file was imported into a demo version of WIDI Ver. 3.0 music recognition system - a program that translates tonal polyphony in the form of pcm wav files, into MIDI messages. The MIDI message was then sent to a Roland JV1010 synthesizer module, through Patch 23 of the Session 8 bank of preset sounds (namely Synth Choir).

WIDI used its "Sensors" Algorithm, developed in an earlier version of the program to detect the tones present within the pcm wav file. A time dicing value or "tick-time" of 50ms was used to calculate re-triggering frequency and minimum MIDI note length. The output of the synthesizer module was then recorded back into another pcm wave file, such that it could be easily blended and delivered with the rest of the material.

Ch17 - CPz

The same approach was employed here as for Ch1 (above), except that this time the MIDI data were sent through Patch 16 (Strings) in the Session 8 Bank of preset sounds in the JV1010 synthesizer.

Ch30 - HNHR

The same approach was employed here as for Ch1, except this time a mellow synth brass sound was used (Patch 42 of the Session 8 Bank of presets in the JV1010 synthesizer).

Ch31 - Erbs

Again, the same approach was used here as in Ch1Final.wav, except this time a woodwind/brass sound was used (patch 40 of the Session 8 Bank of presets in the JV1010 synthesizer).

Note that for each of the MIDI triggered lines, the internal JV1010 reverb was set to maximum.

Ch34 - EDA

Using Cool Edit Pro Ver. 2.1 distortion algorithms a special ICAD distortion preset made with common settings to all files where distortion was used:

- Lite Flange (Using Cool Edit's flange algorithms)
- Original = 63%
- Delayed = 37%
- Initial Mix delay = 4ms
- Final Mix delay = 6.24ms
- Feedback = 0
- Mode = Only Sinusoidal CHECKED
- Rate Frequency = 0.1Hz
- Period = 10000ms Total cycles = 30

Additionally, a phaser effect was added to the Ch34 tones using Cool Edit's phaser algorithms viz:

- Sweep gain = -17.5dB
- Center Frequency = 1000Hz
- Depth = 1.5/Q
- Resonance = 4.02Q
- Sweeping rate = 2.5Hz
- Period = 400ms
- Tempo = 150bpm
- Stereo Difference = N/A
- Sweep modes = sinusoidal; Log Freq Sweep
- Filter type = Band Pass
- Master Gain = 0

Ch35 - Resp

The sound here was compressed using Digilogue's DirectX compressor plug-in Ver. 1.02.6 with the following settings:

- Drive gain = +1.2dB
- Threshold = -29.0dB

- Ratio = 1:16
- Knee Softness = 100%
- Attack = 0.1ms
- Release = 0.5ms
- Makeup Gain = 5.5db

Ch36 - ECG

This was processed using the ICAD distortion preset described in Ch34.

5. INTERPRETATIONS FROM THE MUSIC

We feel that in addition to being somewhat musical, the pitch of the resulting sounds effectively represents the general signal trends in the selected data as follows:

Ch1 synth choir pitch indicates the general long-term fluctuating nature of the left forehead signal.

Ch17 strings pitch indicates the general and gradual increase in the central parietal midline signal with a sudden increase around the 230-second mark.

Ch30 synth brass pitch indicates the fluctuating and increasing trend of the horizontal right (outside of right eye) signal with an increasing level of activity (more cadences) in the last 2 minutes of the piece.

Ch34 woodwind/brass sound pitch highlights the increase/decrease in the overall signal value for the first part of the piece and then tracks the gradual increase in signal from 100 seconds until the end.

Ch35 low frequency tones indicate the generally constant respiration signal with rumble and pitch change denoting sudden spikes and a signal increase at the 100-second mark.

Ch36 tones start in the low range and follow long-term fluctuations in the ECG data before pitching up around the 90-second mark due to a step change in the signal at that point.

6. LAKE HURON COORDINATES

PCM File	Coordinates
Ch01-Fp1.wav	(1.0, 135, 30)
Ch17-CPz.wav	(1.0, 45, 45)
Ch30-HNHR.wav	(1.0, 225, 30)
Ch31-Erbs.wav	(1.0, 315, 45)
Ch34-EDA.wav	(0.7, 180, 45)
Ch35-Resp.wav	(1.0, 0, 0)
Ch36-ECG.wav	(1.0, 0, 90)

7. REFERENCES

- [1] P. J. Sanderson, G. D. McBain and K. Seng Lim "Detecting Localised Agglomeration Trouble in Bubbling Fluidized Beds" *Proceedings, 31st Australasian Chemical Engineering Conference*, Paper 114, 2003.
- [2] Eckmann, J.-P., S. Oliffson Kamphorst and D Ruelle "Recurrence plots for dynamical systems" *Europhysics Letters* 4(9) 973 - 977, 1987.
- [3] R. Ekman "Coagula - Industrial Strength Colour Note Organ - Light Version 1.666", May 22, 2003 available for download at: <http://hem.passagen.se/rasmuse/Coagula.htm>

8. APPENDIX A

Frequencies for Equal Tempered Scale

Note	Frequency (Hz)	Note	Frequency (Hz)
C ₀	16.35	D ₄	293.66
C [#] ₀ /D ^b ₀	17.32	D [#] ₄ /E ^b ₄	311.13
D ₀	18.35	E ₄	329.63
D [#] ₀ /E ^b ₀	19.45	F ₄	349.23
E ₀	20.60	F [#] ₄ /G ^b ₄	369.99
F ₀	21.83	G ₄	392.00
F [#] ₀ /G ^b ₀	23.12	G [#] ₄ /A ^b ₄	415.30
G ₀	24.50	A ₄	440.00
G [#] ₀ /A ^b ₀	25.96	A [#] ₄ /B ^b ₄	466.16
A ₀	27.50	B ₄	493.88
A [#] ₀ /B ^b ₀	29.14	C ₅	523.25
B ₀	30.87	C [#] ₅ /D ^b ₅	554.37
C ₁	32.70	D ₅	587.33
C [#] ₁ /D ^b ₁	34.65	D [#] ₅ /E ^b ₅	622.25
D ₁	36.71	E ₅	659.26
D [#] ₁ /E ^b ₁	38.89	F ₅	698.46
E ₁	41.20	F [#] ₅ /G ^b ₅	739.99
F ₁	43.65	G ₅	783.99
F [#] ₁ /G ^b ₁	46.25	G [#] ₅ /A ^b ₅	830.61
G ₁	49.00	A ₅	880.00
G [#] ₁ /A ^b ₁	51.91	A [#] ₅ /B ^b ₅	932.33
A ₁	55.00	B ₅	987.77
A [#] ₁ /B ^b ₁	58.27	C ₆	1046.50
B ₁	61.74	C [#] ₆ /D ^b ₆	1108.73
C ₂	65.41	D ₆	1174.66
C [#] ₂ /D ^b ₂	69.30	D [#] ₆ /E ^b ₆	1244.51
D ₂	73.42	E ₆	1318.51
D [#] ₂ /E ^b ₂	77.78	F ₆	1396.91
E ₂	82.41	F [#] ₆ /G ^b ₆	1479.98
F ₂	87.31	G ₆	1567.98
F [#] ₂ /G ^b ₂	92.50	G [#] ₆ /A ^b ₆	1661.22
G ₂	98.00	A ₆	1760.00
G [#] ₂ /A ^b ₂	103.83	A [#] ₆ /B ^b ₆	1864.66
A ₂	110.00	B ₆	1975.53
A [#] ₂ /B ^b ₂	116.54	C ₇	2093.00
B ₂	123.47	C [#] ₇ /D ^b ₇	2217.46
C ₃	130.81	D ₇	2349.32
C [#] ₃ /D ^b ₃	138.59	D [#] ₇ /E ^b ₇	2489.02
D ₃	146.83	E ₇	2637.02
D [#] ₃ /E ^b ₃	155.56	F ₇	2793.83
E ₃	164.81	F [#] ₇ /G ^b ₇	2959.96
F ₃	174.61	G ₇	3135.96
F [#] ₃ /G ^b ₃	185.00	G [#] ₇ /A ^b ₇	3322.44
G ₃	196.00	A ₇	3520.00
G [#] ₃ /A ^b ₃	207.65	A [#] ₇ /B ^b ₇	3729.31
A ₃	220.00	B ₇	3951.07
A [#] ₃ /B ^b ₃	233.08	C ₈	4186.01
B ₃	246.94	C [#] ₈ /D ^b ₈	4434.92
C ₄	261.63	D ₈	4698.64
C [#] ₄ /D ^b ₄	277.18	D [#] ₈ /E ^b ₈	4978.03