Evaluating apparatus for musification of *P. berghei* micro-organism motility with synchronous 3D holography

D. A. Williams[1] and L. G. Wilson[2],

[1] Interdisciplinary Centre for Computer Music Research (ICCMR), 310 Roland Levinsky Building, Plymouth University, Drake Circus, Plymouth, PL4 8AA UK

[2] Department of Physics, P/A007, University of York, Heslington, York, YO10 5DD, UK
duncan.williams@plymouth.ac.uk

Abstract. This paper describes the motivation behind the adoption of a musification approach to the auditory display of flagella movement (specifically, malaria in the form of *P. berghei*) as shown in high resolution 3D holography. The criteria applied to the data mapping process in order to achieve this musification are also described, with an approach to spatialization in order to represent three dimensional movement in the data in a soundfield. Details regarding the mapping are illustrated in relation to the different phases of the movement of the flagella. The results are then discussed, noting that adopting a musification approach might help to produce a more engaging audience experience whilst also working towards expressive solutions that might be useful when modeling other complex biomedical data or processes.

Keywords: Musification, biomedical sonification, pitch and rhythm mapping, multimodal musification, spatial localisation

1 Introduction

Analysing micro-organism motility with high-throughput fourier techniques like digital holographic microscopy allows for cell-level movement to be studied in two and three dimensions. Practical applications include real-world detection of flagellated pathogens, such as the malaria parasites of the genus *Plasmodium* in a blood sample, or *E. coli* in drinking water [1, 2].

Malaria is caused by a eukaryote parasite from the *Plasmodium* genus, of which *P. berghei* is a form which does not infect humans, but which is nevertheless useful in order to model malaria behaviour in laboratory conditions. The male, vector-associated reproductive forms of these microorganisms are composed of an isolated flagellum without an attached cell body. This relatively simple form exhibits a characteristic swimming behaviour in Newtonian fluids, providing a unique opportunity to model the behaviour of such flagella more generally.

Perceptually, clockwise and counter-clockwise motion are well-documented as difficult processes in terms of cognitive load, as shown in the so-called bistability visual illusion [3] which is perceptually analogous in some respects to Deutsch's tritone paradox [4]. Flagellar movement presents a related challenge in that their movement pattern is not readily resolved in two dimensions. Thus, 3-D holography offers the opportunity to capture the entirety of the flagellar movement, in a manner which is analogous to wavelet transforms in the auditory domain (representations include frequency, amplitude, and most importantly, phase, allowing for the full wave to be captured and re-created) [1]

System Criteria. We work under the premise that a multimodal representation of flagella movement might ultimately yield a practical application, (for example, auditory display in-field on a mobile device), and thus the audition should include all salient features from the 3D spectroscopic analysis – including frequency, amplitude, and phase. As well as the documented advantages to multimodal representations, an in-field tool for this type of auditory display could conceivably be extremely useful in that it might allow hands free listening in difficult environments. Additionally, this type of biomedical sonification presents a number of challenges which are unique to the discipline, outwith sonification for artistic or creative purposes: firstly, the principle of *primum non nocerere* ("firstly, do nothing noxious" or more colloquially in modern medicine, "do no harm") – in other words, any auditory display must not invalidate, or attempt to supersede, the existing data analysis. Secondly, with the goal of a system that might ultimately be useful in field, we consider it essential that the auditory display is perceptually pleasant to listen to for long periods of time, and that it requires a minor amount of computational processing overhead. In other words, a computationally expensive sonic alarm would be axiomatically inappropriate. The criteria for the auditory display documented here are therefore:

1. sonify all three wavelet features (frequency, amplitude, phase)
2. be able to provide clear auditory discrimination between (e.g.,) clockwise and counter-clockwise motion
3. be computationally inexpensive (portable enough to run on battery powered devices in-field)
4. be audibly 'pleasant' – a difficult aesthetic criteria, see musification
5. do no harm – be complimentary to any visual analysis without reducing the use of any visual analysis tools

Previous Work. *"Sonification conveys information by using non-speech sounds. To listen to data as sound and noise can be a surprising new experience with diverse applications ranging from novel interfaces for visually impaired people to data analysis problems in many scientific fields." [5]*

The use of sonification in biomedical applications is a growing and progressive field, with many existing applications. It is important to make the distinction between *sonification* and *musification*; both forms of auditory display which are alluded to above.

---

[1] This accounts for the name 'holography', which compounds the Greek words for 'whole' and 'drawing'. Holography has been around for some time (Dennis Gabor won the Nobel prize in 1971 for inventing it), and there are many different variations, but the data from which these musifications are derived uses a fairly simple implementation in order to derive the three wavelet parameters.

In a *musification* the data is not just auralized as in a sonification, but instead, various constraints are created and applied in order to create a musical performance of the sonic data (see System Overview for full details of this process). This is in keeping with the system criteria above, in particular, the requirement for audible ‘pleasantness’. Imagine being an in-field analyst who was subjected to audible noise or synthesised test tones for 8-10 hours in-field; even anecdotally it is not difficult to see how this would become tiresome for the listener, and thus more of a hinderance than a help – again, this is related to the *primum non nocere* concept which is embedded in the system criteria as a requirement.

Previous work has suggested that musification of complex biomedical data is a useful way of presenting extremely complicated data, particularly data in 4 dimensions (time, frequency, amplitude and phase in the unfolding and fibrilisation processes in amino acid wrapping), as musification allowed for these processes to be analysed quite readily by a casual listener [6]. The process was broadly as follows: the data was first sonified, and then subjected to a set of musical constraints (pitches, modes, and rhythmic patters)., such that a musical representation was audialized.

Musification allows the listener to engage with complex data by allowing them to facilitating the listening process that listeners do automatically and intuitively as part of their everyday listening process in the real world. This philosophy is common to many auditory display projects making use of multimodal techniques in the biomedical arena.

“The idea behind sonification is that synthetic non-verbal sounds can represent numerical data and provide support for information processing activities of many different kinds.” [7]

Combining a musification with a visual representation of the datastream further targets one of these functions; that of increased audience engagement through a multimodal experience. Multimodality is a complimentary human perceptual process which has also been previously well-exploited by the bio-medical world, e.g., [5, 7, 8].

"The human auditory system is very good at detecting minute changes in audio signals and can also monitor several parallel audio streams at one time. This means hearing offers an exciting opportunity for parallel information interpretation tasks to complement existing visual representational techniques.” [9].

2 System Overview

Our system proposes a musification whereby musical parameters are derived from the wavelet properties of the source data; frequency (pitch), amplitude (volume), and phase (spatial location in a surround sound field). Furthermore, note durations and rhythmic properties are also derived from the four-value array of data such that musical passages might be generated from the data in direct correlation to the visualisation of flagella movement. Changes in the direction of movement which might otherwise be obscured by the visual illusion of clockwise/counter-clockwise movement, are represented by musical modulations, therefore pitch values are also quantised to scales and modes.

Fig. 1 shows an overview of the system and the signal flow involved.

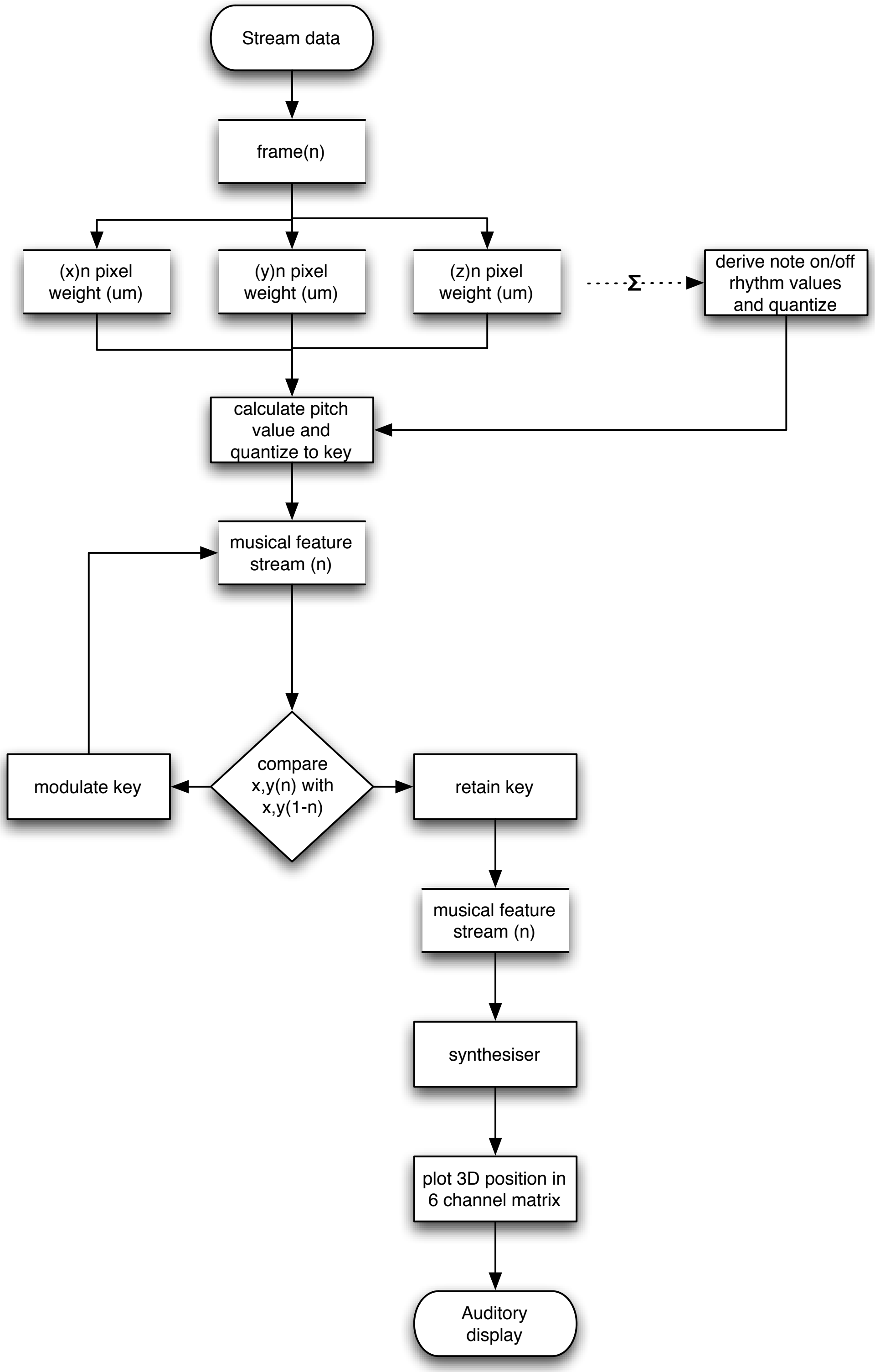
Stream data
frame(n)
(x)n pixel weight (um)
(y)n pixel weight (um)
(z)n pixel weight (um)
Σ
derive note on/off rhythm values and quantize
calculate pitch value and quantize to key
musical feature stream (n)
compare x,y(n) with x,y(1-n)
modulate key
retain key
musical feature stream (n)
synthesiser
plot 3D position in 6 channel matrix
Auditory display

Fig. 1. Signal flow of system. Musical features are derived on a frame-by-frame (n) basis. A Δ change is derived and used as the basis for mapping a number of musical features such that significant changes in direction on a frame-by-frame basis result in different melodic and rhythmic features.

Example musical figures are shown in Fig. 2-5., showing the use of pitch and rhythm quantisation to generate a variety of different melodic and rhythmic effects in the musification.

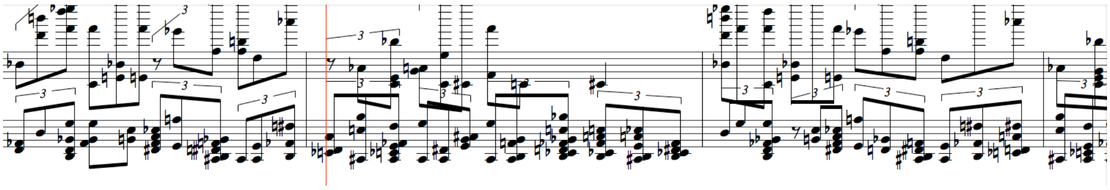

Fig. 2. Raw musical transcription of weighting and 3D co-ordinates.

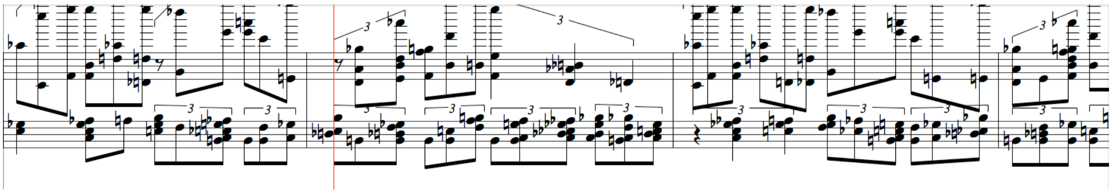

Fig. 3. Musical transcription of weighting and 3D co-ordinates with pitch quantisation applied – notes are now forced to Bb major.

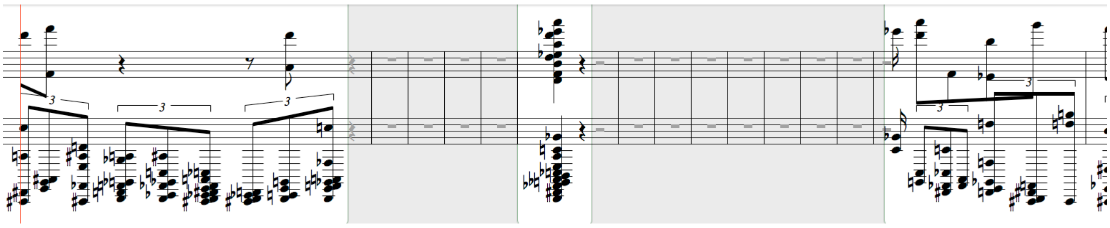

Fig. 4. Rhythm duration quantisation. Some bars are now rests (highlighted in shaded portions of the stave), with 'noise' chords generated by combining all notes from the melodic passages at the same time (middle section). Other notes have the same triplet groupings which are directly from the data with no rhythm quantisation.
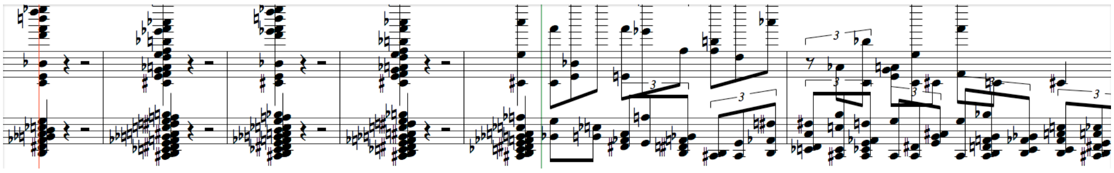
Fig. 5. Enlargement of same quantisation pattern shown in Figure 4, whereby noise chords are extended into triplet-derived melodies (typically compound time).

Spatialisation. Some studies have shown that listeners can localize sounds, especially if they are positioned in the panorama dynamically, with great efficiency if they are presented via individual point sources (i.e., panned to single loudspeakers at any point in the composition).

The need for a matrix arose in this musification in order to diffuse sound with more control over the intended localization, such that the auditory display would begin with an enveloping sound field, and that the degree of localization would then reflect the movement of the flagella. In the initial stages of installation, it was the intention that the audience should not simply be able to look to a speaker when attempting to localize sound sources. This is an issue for many composers working with acousmatic reproductions of spatialised pieces (see [10] p136.). Therefore, this system proposed that in a six channel surround sound speaker array, initially a minimum of three speakers should be used at any one time (stereophonic, binaural represenations of this process are also possible by combining various head-related transfer functions).

A matrix for passing audio to 6 speakers in combinations of triplets was implemented, so that listeners could not initially easily localize source sounds to single point sources at the beginning of the musification. This matrix is illustrated in Table 1 below. Each discrete speaker combination is presented to the audience via a synchronous consolidated audio and video presentation in the installation.

Table 1. Spatialisation array across 6 loudspeaker channels.

Flagella movement type	Spatialisation triplet
A	1+6+2
B	2+1+3
C	3+2+4
D	4+3+5
E	5+4+6
F	6+5+1

An additional requirement in the spatialisation was that the new matrix should not force unintended amplitude changes on the musification – therefore a normalisation factor is required. A simple normalisation was employed, as shown in Equation (1), whereby the signal sent to any given loudspeaker is divided by the total number of loudspeakers receiving a signal.

$$output_s = \sum \frac{a_s}{a_n} \quad . \tag{1}$$

Equation 1. Where s = speaker number, and n is the number of speakers receiving a signal from the matrix

This normalised output, combined with the triple speaker diffusion matrix, allows the musification to be generate a convincing phantom sound image in a diffuse soundfield using only amplitude based panning.

Mapping Melodic and Rhythmic Features. A custom Max[2] patch was designed to translate the numeric data taken from the matrix of flagella movement into respective MIDI values, which are then used to generate the musical score and provide an input source for control of sound synthesis.

Data mapping to musical features is organized in a layered framework. In the first layer, direct physical parameters on a frame-by-frame basis are used to assign basic musical features – pitch from the weighting of each frame, and duration/velocity from the correlates of each frame in the x,y,z(n) array where n is the frame and x,y,z, are individual values for each pixel in a given frame as a weighted 3D representation. The main melodic material is defined by a non-linear absolute mapping as shown in Table 2, whereby weighting values for each frame (from a range of 1-255) are associated to each note of an ascending chromatic scale where a value of 2 is equal to one semitone (giving the full range of 128 MIDI pitch values, or 11 octaves with a fundamental frequency range of approximately 8-12.5kHz). At this stage, the data is already therefore somewhat musified in that each value has a resulting pitch index and absolute frequency value in the MIDI format.

Table 2. Pitch mapping for a sample of frames. Weightings are adapted to MIDI values with a ratio of (2:1), and subsequently ascribed a pitch index, octave number, and absolute frequency value for control of sound synthesis.

Frame weighting	MIDI value	Pitch Index / Octave	Absolute Freq
8	4	E / -5	10.3
34	17	F / -4	21.8
50	25	C# / -3	34.6
14	7	G / -5	12.2
59	27	D# / -3	38.7
121	60	C / 0	261.6

The average shift in each frame is used to calculate duration features, such that larger shifts will be assigned longer note values. Subsequently, a set of five discreet

---

[2] http://cycling74.com/products/max/

durations was defined in order to obtain more consistent rhythmic patterns. The values for each frame are calculated and then mapped to the five durations in milliseconds, which correspond to the following note values at 120 beats per minute: sixteenth note, eighth note, quarter note, half note and whole note, as shown in Table 3.

Table 3. Duration extraction for a sample of frames. Average shift across a frame is adapted to musical values via an absolute duration in ms at 120bpm.

Average shift across frame	Duration (ms)	Note value @ 120 bpm
> 1 octave	125	semiquaver
> 2 octaves	250	Quaver
> 3 octaves	500	Crotchet
> 4 octaves	1000	Minim
> 5 octaves	2000	Semibreve

The note MIDI velocity (i.e. the intensity and resultant volume of each note) is determined by convolving the weighting factor for each frame with the z value (the final value in the 3D array of pixel values). The velocity corresponding to each frame is thus determined by the movement and weighting of each kth frame as $v = (z*n)k$.

Each one of the frames in the sequence is played as a note and the resulting melody represents a musical translation of the flagellum’s movement in 3-D space. The melody is then fed through a ‘tonalizer’ algorithm, which maps the incoming notes to various modal scales. If any alteration in clockwise/counter-clockwise motion is detected in the swimming motion of the flagella, a pitch modulation can be engaged (see Fig. 2., and Fig. 3.). To render this event musically, the note corresponding to the modulating frame is considered the root of the scale in which the whole melody is played. The pitch of subsequent events are then subject to a modifier (+/-1) which causes the melody to change key. As the modulation affects the subsequent data stream as a whole, this tonal modulation has an impact on the entire melody and is thus clearly delineated to the audience. The rhythmic features are also subject to note quantisation, such that in sequences were there is little or no movement in a frame, as determined by the weighted average, note streams are subjected to an increasingly large amount of rhythmic quantisation. These long durations are used to create ‘noise’ chords, whereby all of the notes from a subsequent stream are summed and played simultaneously, before the amount of quantisation is gradually reduced when more movement is present in the frames.

3 Conclusions and Further Work

The current pilot system is capable of generating novel musical score from the existing data illustrating flagellar motion. However, the system has yet to be evaluated formally. One approach would be to compare video and combined audio/video stimuli in order to evaluate the effectiveness of the auditory display. At present, the parameters are mapped to relatively simple synthesis control (melodic, frequency, and rhythm parameters, as well as basic spatialisation in a surround sound speaker array). A third layer of mapping between the input parameters and control of wavelet synthesis would be a useful avenue to explore in further work, for example, in order to give specific control of timbral features in the resulting auditory display. There is also the potential to develop a system that processes pre-recorded audio streams in real time – e.g., by applying selective filtering or phasing to a playback stream when flagella movement of the correct type is identified. We consider this an area for further work, particularly relevant to the criteria of in-field analysis, and a system which might musify the data whilst still satisfying the criteria for ‘pleasant’ listening experience (the users own selection of music, for example).

We present a musification of 3-Dimensional flagellar movement such that a complex series of biomedical data can be auditioned by means of a stream of musical data. The resulting musification employed modal scales, discreet durations and a rhythmic grid to create a musical mapping that attempts to represent the complexity of the source data. The sequences of the piece would be relatively difficult to perform for a human musician, even though they sound ‘easy’ and listenable to a non-musical audience.

A diffusion matrix was proposed with a small number of specific goals.

1. To allow the system to project video synchronous imagery such that the ‘subject-position’ of audience members might be informed or misdirected by video synchronous projection (synchresis).
2. To negate the tendency of listeners to localize sounds to a single loudspeaker when presented with multichannel diffusion, particularly in the initial stages of the musification.
3. To give an approach to spatialisation which allows the overall spatial impression to be gradually improved (and to become more localized) in keeping with the focus of the 3D hologram.

Adopting a musification approach not only allowed for a more accessible musical result but also provided expressive solutions for modeling complex events critical to understanding the movement of the flagellum, such as the change of rotation which was represented by a tonal modulation.

References

1. Wilson, L.G., Carter, L.M., Reece, S.E.: High-speed holographic microscopy of malaria parasites reveals ambidextrous flagellar waveforms. Proc. Natl. Acad. Sci. 110, 18769–18774 (2013).
2. Wilson, L.G., Martinez, V.A., Schwarz-Linek, J., Tailleur, J., Bryant, G., Pusey, P.N., Poon, W.C.: Differential dynamic microscopy of bacterial motility. Phys. Rev. Lett. 106, 018101 (2011).
3. Parker, A.J., Krug, K.: Neuronal mechanisms for the perception of ambiguous stimuli. Curr. Opin. Neurobiol. 13, 433–439 (2003).
4. Deutsch, D.: A musical paradox. Music Percept. 275–280 (1986).
5. Toharia, P., Morales, J., Juan, O., Fernaud, I., Rodríguez, A., DeFelipe, J.: Musical Representation of Dendritic Spine Distribution: A New Exploratory Tool. Neuroinformatics. 1–13 (2014).

6. Visi, F., Dothel, G., Williams, D., Miranda, E.: UNFOLDING| CLUSTERS: A MUSIC AND VISUAL MEDIA MODEL OF ALS PATHOPHYSIOLOGY. Proceedings of the SoniHED Conference on Sonification of Health and Environmental Data. University of York, York (2014).
7. Mihalas, G.I., Paralescu, S., Mirica, N., Muntean, D., Hancu, M., TUDOR, A., ANDOR, M.: Sonic Representation of Information: Application for Heart Rate Analysis. Proceedings MIE (2012).
8. Jovanov, E., Starcevic, D., Marsh, A., Obrenovic, Z., Radivojevic, V., Samardzic, A.: Multi modal viewer for telemedical applications. Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE. pp. 1254–1257. IEEE (1998).
9. Vickers, P.: Sonification for process monitoring. (2011).
10. Wishart, T.: Sound composition. Orpheus the Pantomime, York (2012).