MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-526

Can Sonification Become a Useful Tool for Medical Data Representation?

George I Mihalas[a,c,d], Minodora Andor[b], Anca Tudor[a], Sorin Paralescu[b]

[a] *Department of Medical Informatics, Victor Babes University of Medicine and Pharmacy, Timisoara, Romania*
[b] *Department of Medical Semiotics, Victor Babes University of Medicine and Pharmacy, Timisoara, Romania*
[c] *Commission of Medical Informatics and Data Protection, Romanian Academy of Medical Sciences, Bucharest, Romania*
[d] *Department of Biophysics, Vasile Goldis West University, Arad, Romania*

Abstract

This 'vision' paper refers to sonification – a novel method to represent data by sounds. A short theoretical background comprises the main features to attach sound to a set of data – how to map the correspondence between the sound parameters (pitch, duration) and the initial set of data. The classification of sonification methods is followed by a description of sound display tools – tempolenses and artifacts (saccadic display or loudness variations). The Results section comprises examples of sonification performed by our team: heart rate (HR), ECG signals, HR variations during exercise, including warning procedures. The procedure to evaluate the discriminant power of various sonification algorithms *is then described. As a 'vision' paper, the most important part is not represented by the results, but the potential future developments, presented in the Discussion section, which starts with a critical view of the present state and presents future potential applications of sonification in medicine.*

Keywords:

Sound; Heart Rate; Electrocardiography

Introduction

There are several intuitive means for data representation in science, developed for a better understanding of various phenomena, mostly based on performances of visual perception. Indeed, all visualization procedures have proved their capacity to give us a deeper insight of the processes and even reveal interesting features. But, the auditory system is also feeding our brains with different kinds of information, a way still insufficiently explored. Yet, some attempts have been made to develop such procedures, yielding a new method for data representation – sonification. It is defined as "the use of non-speech audio to transform data into an acoustic signal". Sonification is the core element of the "auditory display", which comprises "any technical solution for gathering, processing, and computing necessary to obtain sound in response to data" [1]. A comparison between visualization and sonification [2] can partially exculpate the much lower use of sonification for scientific purposes as compared to visualization. Nevertheless, the present technical advent and the permanent increase of expectations to extend the ways we interact with the environment open the doors for new developments in this direction.

This paper aims to present an overview of the potential approaches for various applications, to describe authors' experience and explore in more detail the potential future applications for medical data representation.

Methods

The good survey of the sonification domain, until 2011, was done by Hermann et al [1] to which one can add the news from the site of the International Community of Auditory Display [3].

There is a large variety of methods, but almost all start from the same basic principle: to associate the major physical parameter of a sound — frequency (pitch) — with a (major) property of the data to be represented. Other properties of the data set can be linked to other properties of the sound sequence – duration or intensity or even raising the complexity of representation by introducing timbre, rhythm, harmony (multiple sound channels), etc.

We will present here our approach [4], which has multiple similarities with the one in [1], starting with temporal sequences.

Formalization Levels

We will first refer to two major paramaters of sound – frequency (pitch) and duration.

Frequency

From a physical point of view, there are two distinct levels, corresponding to the sonic output: continuous or discrete frequency spectrum. However, for the purpose of potential applications we will prefer to define three levels:

1. *Acoustic level* – with a continuous spectrum,
2. *Sonic level* (S) – with a discrete spectrum, having values belonging to the musical scale, and
3. *Musical level* (M) – a complex level, with multiple channels, rhythm, and harmony.

As the original set of data can have values expressed by numbers in any region, they are usually normalized (y_i), to fit an interval – the simplest [0, 1] interval. A reference frequency (f_0) is also needed to yield results within the audible interval. Some practical values would be 440 Hz (A4 on musical scale) or 262 Hz (C4). Since the natural sound scale is exponential, for the acoustic level the sound will have the frequency f_i:

$$f_i = f_0 \times 2 \wedge (y_i) \quad (1)$$

For the S level, the fi value will be rounded to one of the values from the scale.

Duration

One can choose to display sound for data originating from time series with a duration equal to the real duration of the corresponding event. However, most of the times the data are discrete; even continuous processes are sampled. Hence, we need a conventional rule for the transition between two consecutive sounds. There are two major possible transitions:

1. Continuous transition (also called sublevel A), when for two consecutive points, at t_i and t_{i+1}, the frequency will vary continuously from f_i to f_{i+1}; and
2. Discrete representation, when frequency f_i will be displayed for the interval dt = (t_i, t_{i+1}), followed then by f_{i+1}. Usually the intervals dt are very short and this sub-level is called "quasicontinuous" (Q).

We have used the subdivision into A and Q sublevels only for the acoustic level, using the S level long durations, to be perceived as separate sounds.

There are several cases when the structure of the input data can suggest various corresponding relations between data characteristics and sound characteristics (often called 'natural' correspondence). Some such examples will be given in the section on potential medical applications.

Sound display

The large diversity of mapping data over sound parameters is still increased when thinking about the possibilities to display the results.

Tempolenses

For the cases when the duration of the analysed phenomenon is either too long or very short, and the duration of the sonic display is different from the recording, we can use compression or dilation procedures, also called "tempolenses" [5]. The main parameter of a tempolens (TL) would be the "magnification", defined as the ratio between the sound display duration (t-repr) and the corresponding real process duration (t-real):

$$m = (t\text{-repr}) / (t\text{-real}) \qquad (2)$$

When m > 1 the TL dilates the signal, good for detecting details (especially for very short events); when m < 1, it compresses the signal – good for exploring long recordings in a short time.

An interesting case is represented by TLs with variable magnification (TL-v). Unlike the TL with fix magnification (f), for TL-v, the magnification is m > 1 in certain regions and m < 1 in other regions – a version recommended for detailed representation of fast process, without an overall sacrifice of the total displaying time; for instance, for the ECG sonification, the QRS complex is a fast event and can be dilated, while the T–P interval does rarely contain relevant events and can be compressed. A tempolens with a variable magnification is like a "distorsion" tool, with potential applications.

Sound artifacts

Depending on the potential use, the conventionality degree can be increased by introducing various artifacts, with specific significance. This would be especially useful in monitoring systems associated with warnings. Two such artifacts are described here.

- *Saccadic display*

The auditory system contains phasic receptors, hence it has an efficient adaptation (decrease of the response at a constant value of stimulus intensity), hence an interrupted sound would be better perceived than a continuous sound. Thus, instead of a single sound, we can introduce a "saccad" of two or three sounds as a warning. Moreover, some patterns can be used for different situations; these patterns can easily be learned (when not too many of them).

- *Intensity (loudness)*

For different levels of warning, the intensity of the sound can also be varied, especially when reaching "alert" regions of the followed parameter (for example, heart rate during exerceise).

Differential display

One of the potential future developments might rely on the capacity of the ear to distinguish close frequencies. We can compute an average normalized period for periodic signals, then display it repeatedly on one channel while the real signal is displayed on a second channel so that even small deviations will be perceived very easily.

Non temporal sequences

A relation similar to (1) for frequencies can be used for non temporal sequences such as the primary structure of nucleic acids or proteins, but the (y_i) array is not an orderred array. One can introduce an "order relation" into the original data starting from some properties which might be relevant such as molecular weight, hydrophobicity index, or percent abundance in a certain structure. The correspondence map does not have to follow the ordered musical scale, it might also be linked by some other properties such as probabilities to find a certain note in a musical composition. One can imagine the tremendous variety of mapping systems and it is difficult to qualify/score them.

Transforming digital images into sounds

A serial transmission of an image is actually a signal which can be sonified by any of the procedures described above, and therefore, indeed, such transformations might be useful.

Results

As this paper is submitted as a "vision" paper, this section does not present only our own results [6], but also other applications, all are regarded as examples to be discussed. All examples refer only to applications in medicine and life sciences.

Sonification of cardiac signals

Cardiac signals, especially the electrocardiogram (ECG) are periodic signals, often used for diagnosis in cardiovascular diseases.

Heart rate analysis

Cardiac heart rate (HR) is one of the simplest signals to record, either from pulse oxymeter devicees or from ECG. It was one of the first attempts to sonify a physiological signal done by Ballora [7] in which he tried to raise the complexity up to the musical level. Instead, we tried to keep closer to the original signal [4], using a couple of different sets of sonification parameters and compared the results. Thus, we realized the sound produced by the acoustic level A would not be attractive (sounds like whistling), the same for Q sublevel with very short durations. Sublevel Q with larger durations or level S seem appropriate for detecting deviations from normal (sinus) rhythm. For recorded signals (we have used RRs and ECGs from Physiobank [8]), a compressed sonic display (m=1:6 to 1:4) seemed to be most appropriate.

ECG analysis

We had bigger expectations to get better results from ECG analysis. Indeed, interpreting the classical recording was not an easy task, some modifications were very small, it was important to analyse both amplitudes and intervals, one had to look over the recordings in several leads. What would be the most convenient sonification procedure to reveal all these aspects?

- *Integrated display*

We tried to apply several sonification parameter sets for the same signals [6, 9], including various types of tempo-lenses [5], and compare them (Figure1) using an integrated display – both visual and sonic; a library of various signals and their sonification was created [10].

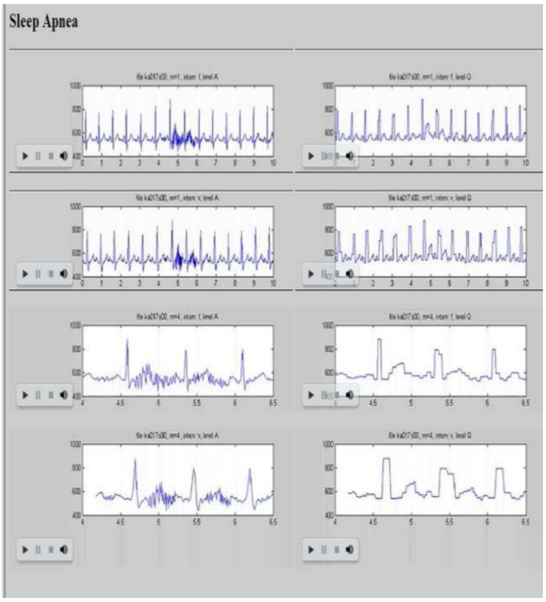

Figure 1 - Example of a screen for ECG signal sonification

- *Discriminant power*

An important issue in our project was to test the discriminant power of various sonic representations i.e. the capacity of listeners to distinguish details and recognize various types when the same signal is represented in different ways [11]. An example is presented in Figure 1: a 10 sec ECG signal was displayed in eight different forms (A or Q mode, f or v display, TL 1× or 4×). For Q mode we had dt = 0.4 sec. The sounds were presented in two groups, one of physicians (residents) and on of music students. We present here some relevant conclusions from this study:

- Representation in A mode was less preferred by both groups (it sounds like a whistle); however it had a higher discriminant power in sleep apnea, during the obstructive episodes;
- Tempolenses with variable magnification did not bring the expected increase in resolution of the QRS complex;
- Short durations (less than 0.2 sec) in Q mode sound like A;
- The distinction between Q and S modes was much clearer for the group of musicians; small differences for other cases.

Our studies showed that there were clear distinctions between the sonic display of different signals (normal sinus rhythm, arrhythmia, and atrial fibrilation), such a simple “additional sound” did not prove to be very attractive to physicians, since their classical procedures did fully satisfy their needs.

Cardiac parameters variation during exercise

We found an interesting application, with a higher user acceptance for monitoring cardiac parameters rate during exercise. Professional equipment can trace various parameters important for the exercise tests on cardiac patients, like the heart rate HR or depression of the ST segment in ECG recordings, with warnings when some parameters exceed thresholds. But this information is mostly visually displayed and the patient is usually kept passive.

We have developed an application to add sounds for various thresholds of HR [12] or ST segment [13]. The threshold warnings for HR have been established using Kevonen relations [14], separating four exercise levels: quasi-rest/[start threshold]/aerobic mild exercise/[attention threshold]/hard exercise anaerobic zone/[alert threshold]/risk zone with HR > HR_{max}. Each level (zone) has a specific intensity, from very low to loud, and saccadic display (no saccades/1/2/3 saccades). The original record is trimmed, keeping the regions comprising transitions between zones.

An example is given in Figure 2 and the sonic displays can be downloaded from [10]. Three types of sonic displays have been tested: with various patterns of pitch/saccades/intensity. The preferences were oriented towards a warning display using four distinct patterns: no sound or low intensity continuous sound for normal (safe) domain, and increasing the number of saccades, the frequency and the intensity for each higher level.

The patients as well as the physicians monitoring the exercise tests appreciated the warning system.

An extension of the application for the self monitoring of the HR during daily individual exercise (jogging) has also been tested.

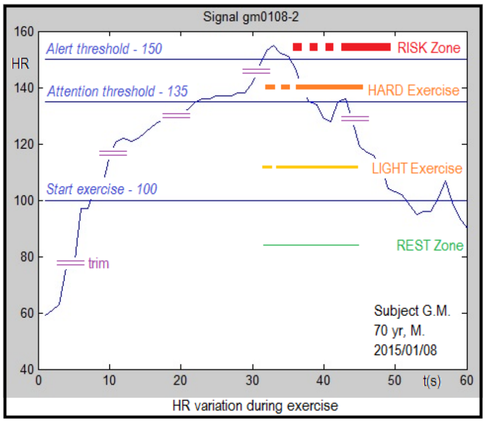

Figure 2 - HR zones during exercise and sound scheme

Molecular sequences

One of the most prolific areas of sonification was inspired by the similarity between musical sequences and biomolecular structures, especially DNA or proteins. Among early attempts we can cite Susumu Ohno’s mapping of DNA [15], associating the nucleotides with notes of the musical scale (C cytosine to ‘do/C’, A adenine to ‘re/D or mi/E’, G guanine to ‘fa/F or sol/G’ and T thymine to ‘la/A or si/B’) and applying some rules for durations linked to the repetitions in sequences. Several other algorithms have been developed: ‘gene-2-music’ project at UCLA (R Takahashi, JH Miller and F Pettit) [16], ‘Algorithmic Arts’ for protein sequences, by John Dunn [17], ‘genetic-musicproject’ by Greg Lukianoff [18].

We have also tried an algorithm based on the relations between physico-chemical properties of amino acids (hydrophobicity, polarity, size, and abundance) and the musical scale [10].

Ohno and Midori [15] have also tried the reverse correspondence – to transpose a musical piece (Chopin’s Nocturne op.55) into a DNA sequence and compare the simulated molecular structure with real molecules, obtaining a certain resemblance with the structure of an enzyme (mouse DNA polymerase II).

There are now several sites (such as [17, 18]) allowing you to upload a sequence (real or imaginary), which generates a musical sequence. The results are attractive, but, despite the attempt to build algorithms based on ‘natural’ maps, often using statistical data processing or even more sophisticated procedures, the impact on scientific community remains modest. This might be, at least partially, attributed to the low degree of usability.

Sonification of other types of data

Our experience include also other applications (not detailed here), like pulse wave signals in experiments on mice, cellular kinetics (especially protein-protein interaction for p53-mdm2 system), and molecular sequences (work in progress now).

We should mention here that there are several teams working in this field and we can cite several papers from the literature. A comprehensive site is the one of ICAD [3].

There are two other interesting applications in the biomedical domain:

1. Analysis of fetal heart rate, Paracelsus Clinic, Austria [19], and
2. Bio-rhythm analysis [20].

Discussion

The results presented above as well as a full list of references would rather give the impression, at a first glance, of a well set domain, with well defined tracks to be developed in the future. However, a thorough and fair analysis would reveal several weaknesses and inconsistencies that actually brought such a limited applicability in the bio-medical domain.

Analysis of the present state

We can point to the following as among the potential reasons to account for the reduced use of sonification:

- The technological support was not fully available for handling sounds. This point is no longer an issue.
- Most authors had tendency/ambition to reach the musical level. It seems attractive, indeed, but it introduces a high degree of arbitrary/conventional rules, moving us farther from the real signal. Not all real data carry harmony! Moreover, quite often the algorithms generated 'atonal' values that have been rounded to make them musical.
- A new tool would be accepted only if it brings something new, which is not easily achieved by present methods. Currently, visualization, complemented by the well-developed technical support, meets expert expectations in bio-medical practice.
- Even when a new method brings more information, the ones in use have several objective and subjective barriers.
- The limited usability is, perhaps, the most relevant factor; most of the work done in the sonification area started as an 'interesting' approach, almost as a curiosity, and not as a real need to solve a practical problem in a better (or cheaper or more convenient) way than by other methods.

Some such potential applications will be presented bellow.

Let's learn from visualization

A comparison between the two methods might help us in finding better ways to strengthen the usefulness and use of sonification as a method for data representation.

Table 1– Comparison between Visualization and Sonification

Property	Visualization	Sonification
Best representation	spatial	temporal
Prevalent elements	static	dynamic
Type of attention	focused	distributed
Recommended for	structures	processes
Detection of small difference	poor	good
Specific elements	brightness texture color shape	loudness harmony timber rhythm

A deeper insight into this table will let us propose a set of properties of a sonic display, similar to that defined by Tufte for graphical/visual displays [21]; some requirements are valid for both types of displays (marked with *).

Thus, we can consider that a 'sonic display' should:

- (*) represent the data
- (*) avoid distorting what they do represent
- (*) make large data sets coherent
- (*) serve a clear purpose: description, exploration
- Preserve the initial sequence
- Bring out information that is inaccessible (regardless the reason) or unavailable by other means
- Be integrated with other informative displays, when possible.

The list above can (should) be refined for a more comprehensive and clearer meaning.

Potential bio-medical applications of sonification

The experience accumulated during the research—both positive and negative—would let us find the directions to be followed for future successful applications in the bio-medical domain.

One category of applications refers to the cases when the visual system is engaged in other activities, such as providing information during driving or jogging.

A second category would refer to all kinds of warnings, based on the property that sonic/acoustic warnings do not need a focused attention. It might become a useful tool in the context of the rapid development of wearable devices for monitoring various physiological parameters. We think this direction might turn into a prolific domain.

There are still less explored directions for applications to categories of people with attention disorders, like elderly people or children with ADHD.

Yet, one can also consider the potential exploration of large databases to find small differences between patterns or to detect mutations in molecular structures.

Last, but not least, we can also include several applications for visually impaired people – another direction which has not been explored enough.

This list is, of course, incomplete, and will have hopefully new items in the future.

Conclusion

Sonification is a method with several potential applications in medicine, yet it is not explored enough. The large palette of sonification methods, without well-defined procedures or standards and the quasi-random examples have unwittingly limited its application, generating a low user acceptance.

The discussion section, based on the present data and the previous work of the authors, tried to analyze the present state of the domain and anticipate future directions of work.

References

[1] T. Hermann, A Hunt, JG Neuhoff (eds.), *The Sonification Handbook.* Berlin: Logos-Verlag, 2011.

[2] S. Pauletto and A. Hunt, A comparison of audio and visual analysis of complex time series data sets, *ICAD05* (2005), 175-181.

[3] International Community for Auditory Display, Available at: www.icad.org (as on 04/12/2017).

[4] G.I. Mihalas, S. Paralescu, N. Mirica, D. Muntean, M. Hancu, A. Tudor, and M. Andor, Sonic Representation of Information: Application for Heart Rate Analysis, *Quality of Life through Quality of Information. Proceedings of MIE2012* ID_403, Amsterdam: IOS Press, 2012.

[5] G.I. Mihalas, S. Paralescu, M. Andor, and et al, Tempolenses with Variable Magnification for Sonic Representation of Medical Data. Application for Cardiac Signals, *Stud Health Technol Inform* **186** (2013),78-82.

[6] G.I. Mihalas, M. Andor, S. Paralescu, and et al, Adding Sound to Medical Data Representation, *Proc. 21st Intl Conf on Auditory Display ICAD 2015*: 325-326, 2015.

[7] M. Ballora, B. Pennycook, P.C.H. Ivanov, and et al, Heart Rate Sonification: a New Approach to Medical Diagnosis, *Leonardo* **37**(1) (2004), 41-46.

[8] A.L. Goldberger, L.A.N. Amaral, L. Glass, and et al, PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, *Circulation* **101**(23) (2000), e215-e220.

[9] G.I. Mihalas, S. Paralescu, A. Tudor, and M. Andor. Adding Sound to ECG, *Stud Health Tehnol Inform* **216** (2015), 1044.

[10] Victor Babes University of Medicine and Pharmacy, Available at: http://www.umft.ro/dim/sonification (as on 04/12/2017).

[11] M. Andor and et al, Testing Discriminant Power of Sonic Representation of ECG Signals, *Cercet. Exp. Medico-Chirurg* Available at: http://jmed.ro/articole_en.php?an=toti&id=365, 2013.

[12] M. Andor, A. Tudor, S. Paralescu, and G.I. Mihalas, Methods for Sonic Representation of Heart Rate During Exercise, *Stud Health Tehn Inform* **210** (2015), 60-64.

[13] M. Andor, A. Tudor, S. Paralescu, and G.I. Mihalas, Methods for Sonic Representation of ST Depression During Exercise, *Stud Health Tehnol Inform* **216** (2015), 1041.

[14] American Council of Exercise, Available at: www.acefitness.org/blog/1179/measuring-intensity, (as on 04/12/2017).

[15] S. Ohno and M. Ohno, The all pervasive principle of repeti-tious recurrence. *Immunogenetics* **24** (1986), 71-78.

[16] UCLA Microbiology, Immunology, & Molecular Genetics, Available at: http://www.mimg.ucla.edu/faculty/miller_jh/gene2music/projectdevelopment.htm (as on 04/12/2017).

[17] Genetic Music from DNA and Protein, Available at: http://www.algoart.com/music.htm (as on 04/12/2017).

[18] Genetics Project Music, Available at: http://www.geneticmusicproject.com (as on 04/12/2017).

[19] Heartmusic: Graphic & Audio Display of Maternal & Fetal HRV, Available at: www.michaelfalkner.de/herzmusik/heartmusicinfo.html

[20] R. McGee, Y.Y. Fan, and R. Ali, BioRhythm: a Biologically-inspired Audio-Visual Installation, Available at: www.mat.ucsb.edu/Publications/McGee_NIME_2011.pdf (as on 04/12/2017).

[21] E. Tufte, *The Visual Display of Quantitative Information, Cheshire* Connecticut: Graphics Press, 1983.

Address for correspondence

Corresponding author: George Mihalas, email at: mihalas@gmail.com.