PROCEEDINGS OF SCIENCE

Creating music from tweets with EGI and GÉANT: an artistic demo running on research and education networks

Domenico Vicinanza[1]

DANTE

126-130 Hills Road, CB2 1PQ, Cambridge, United Kingdom

E-mail: domenico.vicinanza@dante.net

Data audification is the representation of data by means of sound signals (waveforms or melodies typically). Although most data analysis techniques are exclusively visual in nature, data presentation and exploration systems could benefit greatly from the addition of sonification capabilities. Following a successful demonstration organised at the SC11 supercomputing conference in Seattle, music has been used to demonstrate how grid computing and research and education networking can cooperatively enable an artistic demonstration based on data audification at the EGI Community Forum in Munich. People attending the EGI conference were able to suggest texts, poetries and data to translate into sound. Grid computing and high-speed research networks (GÉANT and EU National Research and Education Networks) were used to run a grid-enabled sonification algorithm to turn tweets into music, on demand, in real time.

[1] Speaker

PROCEEDINGS OF SCIENCE

EGI Community Forum 2012 / EMI Second Technical Conference,
Munich, Germany
26-30 March, 2012

1. Introduction

Data sonification [1] (or audification) is the representation of data and information through sounds and melodies. Sonification can be considered as the acoustic counterpart of data graphic visualization, a mathematical mapping of information from data sets to sounds. In the past twenty years the word sonification has acquired a new meaning in the world of computer music, computer science, and auditory display application development.

Data sonification is currently used in several fields, for different purposes: science and engineering, education and training, in most of the cases to provide a quick and effective data analysis and interpretation tool. Although most data analysis techniques are exclusively visual in nature (i.e. are based on the possibility of looking at graphical representations), data presentation and exploration systems could benefit greatly from the addition of sonification capabilities. In addition, sonic representations are particularly useful when dealing with complex data or in data monitoring tasks where it is practically impossible to use the visual inspection. More interesting and intriguing aspects of data sonification concern the possibility of describing patterns or trends, through sound, which were hardly perceivable otherwise.

2.Social motivation

Data sonification is a powerful and extremely effective way to represent information (in particular sequential data, like series, trends or sequences [3]). Being so immediate and easy to communicate, sonification has gained a quite wide interest from both the scientific and artistic domains.

One of the most interesting social aspects related to the data sonification is the possibility offered to visually impaired or blind researchers to perform data analysis and data mining using the ears instead of the eyes. Listening to a melody coming from a sonification algorithm (as described in the next section) is actually listening to the original data or information. Interestingly, while visualisation tools are constantly developed and improved since the first attempt of the human being of representing information in a visual way, there wasn't a similar trend in auditory analysis until the last fifteen or twenty years, when the International Community for Auditory Display (ICAD) was born in 1992 [4].

Auditory display enables non visual representation of information. A rapid detection of acoustic signals and the omnidirectional feature of the sense of hearing can contribute to the effectiveness of an auditory display even when vision is available.

Alongside the scientific aspects of data sonification and its potential in enabling eyes-free data analysis, data audification techniques have been used by this research group in collaboration with visual artists [5], musicians [6] and dancers [7]. As an example, in a concert in Rome held this April, a violin, cello and clarinet trio played data coming from the ALICE experiment. The concert was the first public event based on scientific data coming from the particle physics world, used to create scores for a classic music concert for the general public [8].

3. Scientific motivation

In many cases human ears are used to discover slight changes in acoustic patterns [3]. One example is the medicine who routinely applies acoustic data analysis when uses the stethoscope to listen to breath noise and heart tones. Audification may give information about the inner structure of the represented data using the power of an abstract description.

From a scientific point of view, converting textual or numerical information into sound signals is the same as creating a graph, except that melody constituents are notes and tones instead of lines and points. In order to have a meaningful sonification, two conditions must be satisfied (in order to represent a set of data accurately):

- Uniqueness: a single item of data must be linked to one and only one point or sound
- Covariance: the graph or the melody must vary as quickly as the data. In music, this second condition is satisfied by frequency, timbre and volume.

The human ear is, to some extent, naturally trained to analyse series of data, detect anomalies and spot regularities. We can say that the ear constantly performs auditory data analysis, for example, every time we recognize a person on the phone.

The ear naturally recognizes patterns, structures and sequences. If we are searching for a particular value which, for some reason, is distinguished from a series of data, identifying it on a graph can be difficult, while finding it through a melody could result in a much easier task. Everybody can in fact spot a wrong note in a song, even without any special music training.

Sonification then may provide precious information about the inner structure of the represented data. Any kind of regularity in the original data set will be reflected to the aural signal generated by the sonification algorithm, if the uniqueness and covariance conditions are satisfied. One of the most important challenges is finding the proper balance between the amount of information that can be converted into an audio signal and the effective capability of that sound to communicate meaningful information to a listener.

Summarising, we can say that sonification can be crucial in developing strategies to translate data into audible information to help researchers carry out comprehensive data analyses exclusively from sounds.

Compared to other sonification methods, the MIDI melodisation has the great advantage to code information into one of the most widely used musical format. Almost all the applications for music analysis/sequencing are able to operate and play MIDI files in extremely customizable manners. Moreover many programming languages, as Java for example, have the possibility to load and use special libraries to efficiently manage them (see for example jMusic [9]).

Finally several tools have been developed to analyse and process MIDI files, to discover self-similarities, repetition of microstructures, statistical diagrams (one for all, the well known ”MIDI toolbox” written by Petri Toiviainen [10]).

4.Sonification in linguistics

The audification technique used for the demo described in this paper is based on text sonification. Text sonification is a particularly relevant sonification process dating back from the Middle Age. During the first decades of the XI Century, Guido Monaco, also known as Guido d'Arezzo (990-1050 ca), a Benedictine Monk who invented the system of staff-notation still in use, proposed (and used in several musical pieces) a method to associate a note on the pentagram to each vowel of text to sing. The duration of the notes was depending upon the number of consonants between each vowel.

In the text sonification, the most natural choice is to associate to the same set of characters exactly the same musical phrase (i.e. the same sequence of notes). Sonification would then allow to (acoustically) recognising words which start with the same root, or end with the same letters, as rhymes, for examples. In some sense one can listen to the metric properties of a poetry just using melodies coming from the sonification process.

5.Sonification on the EGI infrastructure: the live demo at Supercomputing 2011 and EGI Community Forum 2012

The sonification described in this work was initially shown at the Supercomputing 2011 (SC11) conference in Seattle, and presented as a live demo at the EGI Community Forum Conference (http://cf2012.egi.eu/) held in Munich at the Leibniz Supercomputing Centre (LRZ), from 26 to 30 March 2012. The demo used a grid-enabled version of the original Guido's method running on the EGI infrastructure running on GÉANT (http://www.geant.net).

GÉANT [11] is the pan-European GÉANT backbone for research and education and it created the seamless fabric layer enabling linking together multiple computers in different locations via high speed networks, combining their processing power to deliver faster results when analysing enormous volumes of data.

The pitch of the note was chosen according to the position of the letter into the English alphabet, so a "S" will sound a scale step higher with respect to a "R", a "T" two tones higher, and so on. The duration of the notes was:

- 1/16 (semiquaver) for any consonant
- 1/8 (quaver) for any vowel.

Visitors who attended the conference were able to enter their own data to sonify using a special web portal, http://gw.ct.infn.it/, who allowed users to directly submit, inspect and retrieve the result of sonification grid jobs. The portal gw.ct.infn.it is the Science Gateway to the Italian Grid Infrastructure (IGI) and it uses GILDA [7], the EGI training infrastructure. From this Science Gateway it was possible to access in an easy way the Italian Grid Infrastructure and run on it a set of scientific applications belonging to different domains.

The portal worked through servlets created by the INFN Catania team to run the data sonification process on the grid distributed cluster for massive computation. Thanks to the web interface the users at the demo venue were able, in a transparent and easy way, to specify the input settings, start the sonification process and monitor the grid process. Once the sonification was completed, the audio produced during the computation was made available on the portal, ready to be downloaded by the users.

The actual sonification package was written in Java using jMusic [10] and supported an upload functionality, allowing users to upload an existing ASCII file to be sonified.

6.Screenshots from the sonification demo at EGI

Figure 1: The sonification portal (http://gw.ct.infn.it), running on the IGI

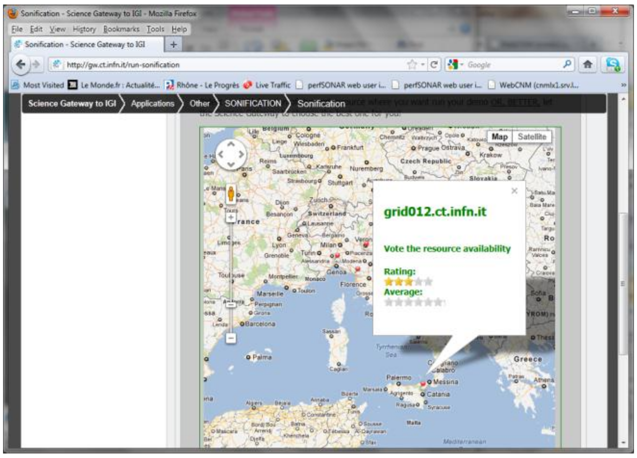

Figure 2: Selecting the grid resource on the sonification portal

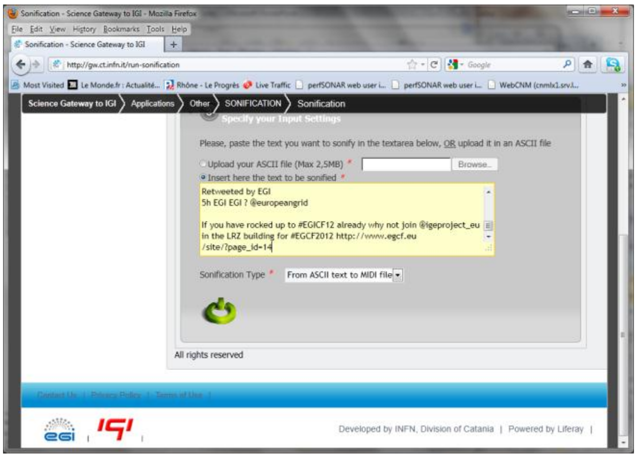

Figure 3: Entering the text to sonify

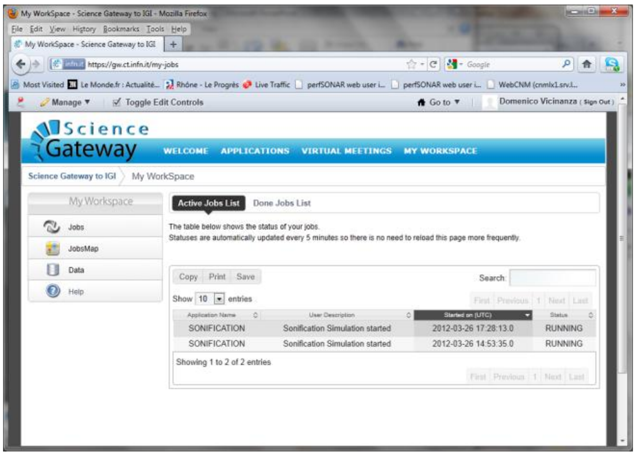

Figure 3: The successful job submission of the sonification algorithm

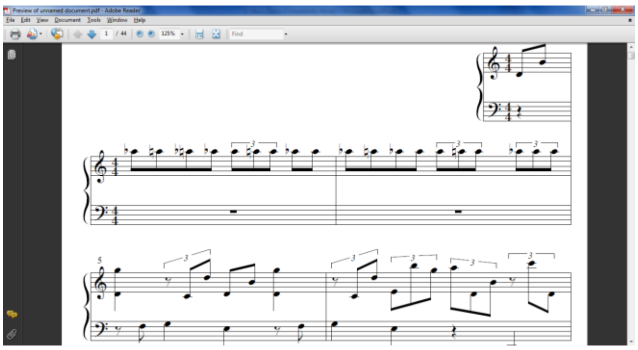

Figure 4: The final score

7.Acknowledgments

The author is particularly grateful to Giuseppe La Rocca from INFN Catania and Mariapaola Sorrentino from LHC Open Symphony project, Cambridge for their contribution to set up the sonification algorithm and make it available on the IGI Scientific Gateway.

8.Conclusions

Sonification, the data representation through sound signals, is now being recognised as a useful tool in many research areas and the range of data which can be analysed is unlimited. Its use for and by the blind and visually impaired is of particular relevance, along with its application to studies of other disciplines.

The demo described relied on several elements that underpin research every day: high-speed research networks including GÉANT, EGI grid computing and modern portals to access transparently grid infrastructures.

GÉANT and the European NRENs are a vital part of the EGI grid infrastructure as they provides the real time, high bandwidth transport layer to link distributed processing power, removing latency and avoiding data bottlenecks.

An example of the audio file created during the demo can be downloaded here: http://www.bubblechambermusic.org/audio/supercomputing2011.mp3

References

[1] Data sonification, definition and use cases: *http://en.wikipedia.org/wiki/Sonification*

[2] Data sonification, definition and use cases: *http://en.wikipedia.org/wiki/Audification*

[3] Sonification and auditory display examples: *http://sonification.de/*

[4] International Community for Auditory Display (ICAD) official website: *http://www.icad.org/*

[5] Observatory, by Paolo Parisi and Domenico Vicinanza: *http://ita.fondazionecasa.com/paolo-parisi/*

[6] Higgs boson sonification: *http://news.discovery.com/space/listen-to-the-higgs-boson-120710.html* and *http://lhcopensymphony.wordpress.com/*

[7] GILDA the GRID INFN Laboratory for Dissemination activities, *http://gilda.ct.infn.it/*

[8] La musica della natura, sonification concert held in Rome, translating scientific data into music *http://www.auditorium.com/eventi/5187715*

[9] Midi toolsbox website: *http://www.jyu.fi/musica/miditoolbox/*

[10] jMusic (Java API for MIDI notation) website: *http://explodingart.com/jmusic/*

[11] GÉANT, the pan-European backbone for research and education: *http://www.geant.net*