Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 1 (2012) 397–406

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2010

Data Sonification of Volcano Seismograms and Sound/Timbre Reconstruction of Ancient Musical Instruments with Grid Infrastructures

Salvatore Avanzo[a], Roberto Barbera[b,c], Francesco De Mattia[d], Giuseppe La Rocca[b], Mariapaola Sorrentino[e], Domenico Vicinanza[e,]*

[a]*Dept. of Mathematics and Computer Science of the University of Salerno and SoundLab, Via Ponte don Melillo, Fisciano (SA) I-84084, Italy*
[b]*Italian National Institute of Nuclear Physics, Via S. Sofia 64, Catania I-950123, Italy*
[c]*Department of Physics and Astronomy of the University of Catania, Viale A. Doria 6, Catania I-95125, Italy*
[d]*Conservatory of Music of Parma, Via del Conservatorio 27/A, Parma I-43100, Italy*
[e]*DANTE, City House 126-130 Hills Road, Cambridge CB2 1PQ, UK*

Abstract

Recently, the scenario of international collaboration in scientific research has swiftly evolved with the gradual but impressive deployment of high bandwidth networks and Grid infrastructures towards what is currently called e-Science. So far, several scientific domains, such as Life Sciences, High Energy Physics, Computational Chemistry, Earth Science, etc. are benefiting of e-Infrastructures to tackle new global challenges, particularly those that have high societal and economic impact, with a truly multidisciplinary approach. Much less has been done, however, in the field of Humanities. In this paper we present some use cases of how the EU funded e-Infrastructures have been used to support both the Cultural Heritage community in reconstructing the sound of ancient musical instruments and the scientists belonging to the Earth Science domain to better understand the behavior of volcanoes close to eruptions translating the patterns of volcanic seismograms into a set of hearable sound waves.

Keywords: Grids: Data audification; Recover the sound from our historical memory; Music from volcanoes; the ASTRA project.

1. Modelling the past with the ASTRA Project

The ASTRA (Ancient instruments Sound/Timbre Reconstruction Application) project (www.astraproject.org) is coordinated by the Conservatory of Music of Parma (www.conservatorio.pr.it) in close synergy with the Conservatory of Music of Salerno (www.consalerno.com), the Division of Catania of the Italian National Institute of Nuclear Physics (www.ct.infn.it) and other international partners. The project, running since 2006, aims at providing the Arts and Cultural Heritage community with an application to reconstruct the sound/timbre of ancient musical instruments. By applying the physical modelling synthesis, a complex digital audio rendering technique which

* Corresponding author. Tel.: +44 (0)1223 371 300; fax: +44 (0)1223 371 371
E-mail address: Domenico.Vicinanza@dante.net

1877-0509 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.procs.2010.04.043

allows modelling the time-domain physics of an instrument, the experts who are carrying out the project can recreate models of some musical instruments that were lost for ages (since hundreds and hundreds of years) and reproduce their sounds by simulating their behaviours as mechanical systems. The application is very computer intensive and this is the main reason why it runs over research and academic networks such as GÉANT (www.geant.net) and EUMEDCONNECT2 (www.eumedconnect2.net) operated by DANTE (www.dante.net). The work performed by the parties involved in this project to address this goal is sketched in fig.1.

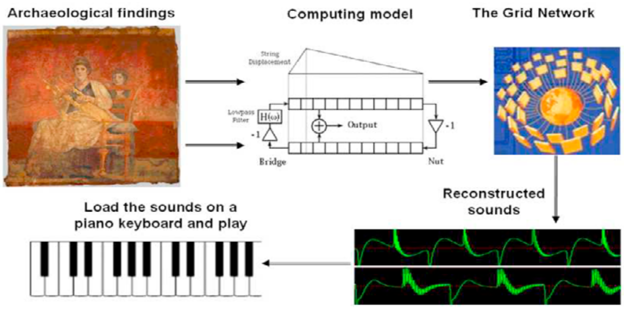

Fig. 1. Modelling and computation of ancient instruments in Grid

The starting point is represented by all archaeological findings about the instrument such as fragments from excavations, written descriptions, pictures on ancient urns, etc. By applying the Physical Modelling Synthesis (PhM), ASTRA recreates computer replicas of some musical instruments and reproduces their sounds by simulating their behaviour as mechanical systems. The PhM Synthesis uses a set of equations and algorithms that describe the physical materials used in the ancient instruments to generate the physical sources of sound. For instance, to model the sound of a drum, a formula for how striking the drumhead injects energy into a two dimensional membrane would be created. Then, the properties of the membrane (mass density, stiffness, etc.), its coupling with the resonance of the cylindrical body of the drum, and the conditions at its boundaries (e.g., a rigid termination to the drum's body) would describe its movement over time and how it would sound. Similar modelling can be carried out in instruments such as violins, though replicating the slip-stick behaviour of the bow against the string, the width of the bow, the resonance and damping behaviour of the strings, the transfer of string vibrations through the bridge, and finally, the resonance of the soundboard in response to those vibrations. The computing model of the instrument reconstructed by ASTRA is parameterized by several quantitative factors. Some of them are constants that describe the physical materials and dimensions of the instrument while some others are time-dependent functions that describe the player's interaction with it such as plucking a string or covering tone holes. The "tuning" of some of these parameters is performed taking into account the sound spectra of a Monochord, an ancient musical instrument played by Pythagoras. The word "monochord" comes from the Greek and literally means "one string." In the Monochord, a single string is stretched over a sound box. The string is fixed at both ends while a moveable bridge alters the pitch. ASTRA researchers built a real Monochord in their laboratory and they verified that the model was correctly reproducing the sound of the physical instrument. In the last step, the refined computing model is simulated as a mechanical system in order to output the sound of the instrument. The simulation of the model is a key point in order to generate the sound of the instrument. It requires a large amount of computation because, in order to properly model the instrument, a lot of mathematical equations (differential equations of coupled oscillators) have to be solved. The computation also depends on the quality of the audio file one wants to obtain. In this scenario, Grid infrastructures , with their huge distributed computational and storage resources spread over large geographic regions worldwide, can effectively contribute to speeding up the simulation of the computing model and generate the sound of the instrument. With this approach, today ASTRA can generate digital sound libraries, using computing power wherever it is located, and share these results with the Cultural Heritage community. The most important advantage of this approach is to easily build sound libraries that students, researchers and museums can easily access and use. The digital sound libraries obtained with the use of the Grid, for instance, can be loaded into a keyboard and used by musicians to perform live concerts. Another main advantage of the ASTRA project is to bring experiences and knowledge together. Using e-Infrastructures, ASTRA allows archaeologists to have many interesting relics and findings, software developers, and historians to work together and create something unique.

2. The Implementation

2.1. Computational aspects of the Physical Modeling Synthesis used by ASTRA

Although physical modelling is not a new concept in acoustics and synthesis, having been implemented already in 1971 using the wave equations created by Hiller and Ruiz [1,2,3], its practical use has become feasible only recently, i.e. since large bandwidth research networks and sophisticated Grid middleware have started to interconnect thousands of computers in what are called e-Infrastructures. Physical modelling is a very computing-intensive technique, as the complex models of the musical instruments are solved by the numerical integration of coupled differential equations. For example, to have 100 samples (corresponding to an instrument with 50 strings plucked in 2 different ways) 400 CPU hours would be needed on a standard computer. Using Grid technology those computations could be run at the same time (as independent Grid jobs), automatically, on 100 different computers, having the results in just 4 CPU hours. This is particularly valuable, for example, when it is needed to create a large sound library. It would speed up by a factor of 100's the overall computing time. In order to exploit the computational power needed by the project and speed up the sound reconstructions, the ASTRA project moved from a local computer cluster towards distributed computing resources. First tests have been successfully conducted within the GILDA [4] (https://gilda.ct.infn.it) facility, the EGEE default training infrastructure, which brings together about 20 sites (not all on-line at the same time) in three different Continents, and later on the EUMEDGRID project (www.eumedgrid.org) Grid infrastructure which links computing resources across the Mediterranean region at up to 2.5 Gb/s through the GÉANT (www,geant2.net) and EUMEDCONNECT2 (www.eumedconnect2.net).research networks. The extremely high reliability of the GÉANT and EUMEDCONNECT2 networks, their performances in terms of number of institutions connected and overall throughput, made the ASTRA project possible. Thanks to EGEE and other projects, today ASTRA can rely on more than 500 CPUs and on the support of more than 20 different Grid sites in Europe, the Mediterranean, and other regions of the world.

2.2. Technical implementation of the Physical Modeling Synthesis

The modelling of our string instruments has been carried out using the Digital Waveguide Models (DWM). The Digital Waveguide Models, coined by Julius O. Smith III and introduced for the first time in 1985, are computational physical models made up of delay lines, digital filters, and often nonlinear elements which can be used to model the physical media through which acoustic waves propagate. Some features of these models are:

- Sampled acoustic travelling waves;
- Follow geometry and physical properties of a desired acoustic system;
- Efficient for nearly lossless distributed wave media (strings, tubes, rods, membranes, plates, vocal tract, etc.).

From a technical point of view, using the discrete form of d'Alembert's solution, the 1-D wave equation can be described by the following equation:

$$y(m,n) = y^{+}(m-n) + y^{-}(m+n)$$

where y^{+} is the right-going wave and y^{-} is the left-going wave. It can be seen from this representation that sampling the function y at a given point m and time n merely involves summing two delayed copies of its travelling waves. These travelling waves will reflect at boundaries such as the suspension points of vibrating strings or the open or closed ends of tubes. Hence, the waves travel along closed loops. Digital waveguide models therefore comprise digital delay lines to represent the geometry of the waveguide which are closed by recursion, digital filters to represent the frequency-dependent losses and mild dispersion in the medium, and often non-linear elements. Losses incurred throughout the medium are generally consolidated so that they can be calculated once at the terminations of a delay line, rather than many times in between. Waveguides such as acoustic tubes are three-dimensional but, because their lengths are often much greater than their cross-sectional area, it is reasonable and computationally efficient to model them as one dimensional waveguides. Membranes, as used in drums, may be modelled using two-dimensional waveguide meshes, and reverberation in three dimensional spaces may be modelled using three-dimensional meshes. Vibraphone bars, bells, singing bowls and other sounding solids (also called idiophones) can be modelled by a related method called "banded waveguides" where multiple band-limited digital waveguide elements are used to model the strongly dispersive behaviour of waves in solids.

2.3. String Modeling

In this sub-section we will describe how the DWM has been used for modelling the ideal plucked string. Let $y(t,x)$ denote the transverse displacement of an ideal vibrating string at time t, with x denoting the position along the length of the string. If we terminate a length L ideal string at $x=0$ and $x=L$, we then have the following boundary conditions:

$$y(t,0)\equiv 0 \qquad y(t,L)\equiv 0$$

where "$\equiv$" means "identically equal to" for all t's. The corresponding constraints on the sampled travelling waves are then

$$y^{+}(n)=-y^{-}(n)$$

$$y^{-}(n+N/2)=-y^{+}(n-N/2)$$

where N is the time in samples to propagate from one end of the string to the other and back, or the total "string loop" delay. The loop delay is also equal to twice the number of spatial samples along the string. A digital simulation diagram for the rigidly terminated ideal string is shown in fig 2. A "virtual pick-up" is shown at the arbitrary position $x=\xi$

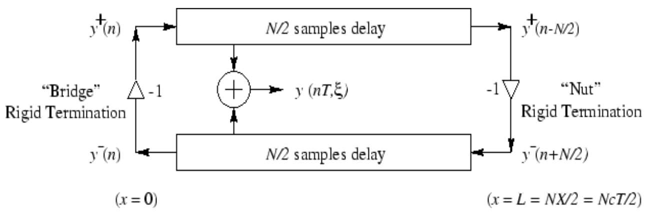

Fig. 2. The rigidly terminated ideal string with position output $x=\xi$

An ideal plucked string is defined as having an initial displacement and zero initial velocity. In the model, the delay lines should be initialized with displacement data corresponding to some arbitrary initial string shape.

The implementation we have chosen for our simulation was derived from the Karplus-Strong observation that a delay line initialized with random numbers makes for a pleasant and bright plucked string tone. We modelled the time-dependent frequency losses using a moving average low pass filter, which implements in an efficient way the high frequency losses quite typical in real string instruments. Finally, the effect of the soundboard has been simulated analyzing the sound of tapping on a real soundboard with the same shape, measuring its impulse response and doing a convolution with the sounds coming out from the simulation.

From what concerns the computational implementation of the synthesis algorithm, the Synthesis ToolKit (STK, https://ccrma.stanford.edu/software/stk/index.html) has been used. The ToolKit, developed by Perry Cook and Gary

Scavone, is based on a set of open source audio signal processing and algorithmic synthesis classes written in C++ . STK was designed to facilitate rapid development of music synthesis and audio processing software with an emphasis on cross-platform functionality, real-time control, ease of use, and educational example code. The ToolKit is extremely portable, and it is completely user-extensible (all source included, no unusual libraries, and no hidden drivers) then it was an ideal candidate to design the simulation algorithms to be run on the Grid. The Synthesis ToolKit, which is free for educational purposes, can generate simultaneous SND (AU), WAV, AIFF, and MAT-file output sound file formats (as well as real-time sound output). The ToolKit and the wxWidgets C++ libraries (v. 2.8.0) have been used to create a GUI for ASTRA. Using this GUI the user can specify the sample frequency of the instrument, redirect the computation to an output file and, last but not least, specify a music score to be played.

2.4. Running ASTRA on the Grid

In this section we are going to describe the set-up of the Grid environment, first on GILDA and then on the EUMEDGRID infrastructure, created in order to allow non expert users to speed up the large scale production of sounds of ancient instruments. This task has been pursued by the Italian National Institute of Nuclear Physics (INFN) in close synergy with the Music Conservatories of Salerno and Parma. As first step, the two distributed Grid infrastructures have been successfully configured by installing, in the dedicated software areas, the package requested by ASTRA to perform the reconstruction process. The installation and the configuration of this software has been done using widely adopted tools such as *automake* and *autoconf*. A new software tag to reference the ASTRA software has also been published in the Information System in order to be used by the whole community of users. Another important task pursed by all the partners involved in the project has been the development of a web-based interface to help musicians and historians to execute the ASTRA algorithms without any knowledge of the middleware. In fact, there is a sort of “scientific gap” represented by the difficult protocols and the rather complex Command Line Interface that a normal user has to learn and practice to access Grids. In order to by-pass this problem, a web-based interface has been designed on top of the GENIUS Grid Portal [5,6,7,8]. This web portal, powered by EnginFrame (www.nice-software.com/web/nice/products/enginframe), provides end-users with a very intuitive interface which allows to access the distributed services and resources of a typical Grid infrastructure in a transparent and ubiquitous way.

2.5. The GENIUS architecture and the EnginFrame Java framework

EnginFrame is a web-based technology, developed by the Italian company Nice Srl (www.nice-software.com), that enables the access and the exploitation of Grid-enabled applications on e-Infrastructures. It allows users, gathered in a Virtual Organization (VO), to exploit the whole power of the Grid hiding all its complexity and providing an abstraction layer on top to the middleware. EnginFrame is also well known in the Grid world for being the technology which GENIUS (Grid Enable web eNvironment for site Independent User job Submission), one of the most known and used Grid portals of the European flagship project EGEE (www.eu-egee.org), is based on. Due to the modularity and flexibility of EnginFrame, the GENIUS portal can be easily customized to interact with several Grid services and/or VOs allowing scientists to access, execute and monitor their own applications by only using a simple web browser. The features exposed by EnginFrame through the GENIUS portal contributed to disseminate the achievements during demo sessions organized by ASTRA. For further information, see the website https://glite-demo.ct.infn.it.

3. The ASTRA Results

Since September 2008, and for the first time in centuries, the harp-like strings of the Epigonion, an instrument from Ancient Greece (see fig. 3), have been “virtually” plucked thanks to ASTRA. To achieve this, we used the advanced GÉANT and EUMEDCONNECT2 research networks to link high capacity computers together, sharing information to enable the computer-intensive modelling of musical sounds. In Figure 3 is shown the 3D structure of the Epigonion reconstructed by ASTRA in close synergy with some engineers from Parma. For this purpose, the following two additional packages: Blender (www.blender.com) and LuxRender (www.luxrender.com) have been installed on the Grid resources and used for the three-dimensional rendering of the instruments. The knowledge of the Epigonion was based on archaeological findings, historical pictures and literature. It was a wooden string instrument that musicians have likened the sound to something similar to a modern harp or a harpsichord. We have compiled the sounds of four Epigonion instruments to recreate a medieval musical piece, making this the first time that these instruments have been heard performing together. An important event to celebrate this result and communicate it to the general public was a concert held in Naples, Italy on the 14th of December 2008. It was the first time ever that an instrument of the past, reconstructed via compute-intensive modelling, performed alongside real instruments such as violins and flutes as well as voices. The Psalm “Laetatus sum” by Czech composer Jan Dismas Zelenka was performed by the Sonora Network Ensemble (www.scatolosonora.org) enriched with the digitally reconstructed sounds of the ASTRA Epigonion. Promoted as a concert produced in the ‘listening laboratory’, the performance aimed at paving the way for innovative musical research which enables a wide range of scientists such as archaeologists, network engineers and computer experts to collaborate with artists to revive lost sounds in real concerts (an excerpt of the concert is available at www.astraproject.org/files/concertdemo.mov). Other sound reconstruction examples can be downloaded from www.astraproject.org/download.html. Another notable step forward to improve the musical cultural heritage and brings history back to life has been reached on the 1st of December 2009 during the GÉANT launch event which took place in Stockholm, Sweden. During the meeting, researchers involved in the ASTRA project presented to all the community, as a world premiere, the sound of the Barbiton, an ancient Greek instrument similar to a double bass (see fig. 3). Accompanied by percussions, the instruments were played by the Lost Sounds Orchestra following melodies from a musical score written by Domenico Vicinanza (another co-author of this chapter). The melodies of the Barbiton, heard for the first time in 2,000 years, were sent 9,300 km away from the venue in Stockholm across the GÉANT and TEIN3 (www.tein3.net) research networks to provide music for dancers from the Arts Exchange of Asia, allowing them to simultaneously perform in Kuala Lumpur (Malaysia) at the ASEM (Asia-Europe Meeting) workshop’s Gala Dinner. The video of the performance can be downloaded from www.geant.net/Media_Centre/Media_Library/Media%20Library/ASTRA_musicians_09.mp4. The sound of percussions and the warm notes of the Barbiton worked perfectly with the background melody from the GÉANT and TEIN3 networks. Being the Barbiton a heavy, multi-stringed instrument, it produces a sound similar to a double bass and it was described in many poems and paintings of the time. It is known to have been played by the poets Alcaeus and Sappho. As well as for the Epigonion, also for the Barbiton a 3D rendering of the instrument has been created thanks the computational resources of the Grid.

Fig. 3. 3D rendering of the Epigonion (a) and the Barbiton (b) with Blender and LuxRender (courtesy of F. Baghino and F. Ugozzoli, VisArc Studio, Parma, Italy)

4. Introduction to Data Audification and Melodisation

Current knowledge of volcanoes does not yet allow scientists to predict eruptions but since the sonification of volcano seismograms of Mount Etna (grid.ct.infn.it/etnasound/page4/page8/etna.aif) has been carried out on the GILDA training infrastructure, a notable step forward has been taken in this direction by the scientific community. A seismogram is a particular representation of the regular oscillations of a volcano. The sonification of these seismograms has recently become possible thanks to the new advanced features introduced by the data audification analysis (www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/Manual). Data audification is the representation of data by means of sound signals (typically, waveforms or melodies). It can be considered as the acoustic counterpart of data graphic visualization, i.e. a mathematical mapping of information from data sets to sounds. In the last few years this new way to represent data sets and patterns has acquired more and more interest in different disciplines such as science and engineering and education and training. Although most data analysis techniques are exclusively visual in nature, data presentation and exploration systems could benefit from the addition of sonification capabilities. The possibility to add sonification aspects and describe patterns or trends through sounds is particularly useful when dealing with complex high-dimensional data, or in data monitoring tasks where it is practically impossible to use the visual inspection. Indeed, sonification may give information about the inner structure of the represented data using the power of an abstract description. All kinds of regularities in the original data set will be reflected to the aural signal generated by the audification algorithm. The main advantage of this technique is that it make easier for users to recognize a change in sound with respect to a modification in something which has to be looked at. It is quite impossible, for instance, to distinguish a blinking light flashing 100 times per second from another one flashing at 200, 1000, or 10000 times per second, while it is quite easy to recognize periodic signals from 20 Hz to 20000 Hz. In this paper, this approach it has been used to approximate the volcano to a huge musical instrument which can produce a sound non detectable by the human ear as the frequency is too low. The larger the mountain the lower the frequencies emitted. Seismograms are the graphical representations of these infrasonic sounds. To hear these sounds, it is necessary to perform a special manipulation called *"pitch shift"*. In other words, the infrasonic frequencies are mapped to the ones that can be heard by the human ear. The hope is that by studying large amount of these sounds (which can be easily collected, categorized, stored, and retrieved thanks to the grid databases) it will be possible to learn more about the phase before eruptions. It will also potentially be possible to fully characterize eruptions from a sonic point of view to make more accurate predictions of their occurrences.

The melodisation of a data set, i.e. the creation of a melody starting from a series of data, is an interesting way to convert into aural signals almost any kind of information. One of the main reasons for converting a data set into a melody is the possibility to perform quantitative analyses on them by means of standard open source tools, like the seek for self similarities, self organization behaviour, recurrent patterns, quasi periodicities, "signature tunes" (which would help blind researchers to deal with data analysis as well). For the sake of simplicity we can imagine a set of m elements (a list or a discrete set of values) to be sonified. The sonification will provide a melody, a list of notes chosen among a set of n, to be drawn onto a music score. Without loosing generality, we can consider a standard, numerical, coding convention for the (well tempered) notes to be represented. In particular, we shall refer to the MIDI code, according to which the central "C" note corresponds to the integer 60, "C#" is 61, "D" is 62, and so on (any semitone shift will add or subtract 1 from the previous value). The lowest acceptable MIDI value is 0 and the higher is 127 so there are 128 possible notes that can be used to represent our data. Compared to other sonification methods, the MIDI sonification has the great advantage to code information into one of the most widely used musical formats. Almost all the applications for music analysis/sequencing are able to operate and play MIDI files in extremely customizable ways. Moreover, many programming languages, such as Java for example, have the possibility to load and use special libraries to efficiently manage them (see for example jMusic: http://jmusic.ci.qut.edu.au/).

4.1. Sound production on the Grid

Even though the usefulness of sound representations was clearly accepted by the scientific community, its practical use has so far received limited attention because of the intensive computation usually required to produce the sounds. The Grid has made the application of this method possible. Digital audio usually deals with very high sampling rates. The standard value for CD quality audio signals is 44100 Hz, so to produce one second of audio data is necessary to compute 44100 values. One minute will take 60 x 44100 = 2646000 calculated samples. Geophysical data sonification with the use of Grid infrastructures has been carried out since 2003. Initially, the sonification process has been applied to data collected by a digital seismograph placed on the Etna volcano in Italy and recently to data collected on the Tungurahua (www.volcanodiscovery.com/en/tungurahua.html) volcano in Ecuador. A total of 60252 samples were analyzed. The goal was to learn more about eruption dynamics from patterns (and pattern evolution) in the sonograms. In the case of the sonification of Tungurahua, scientists from Italy and Ecuador involved in EGEE and EELA (www.eu-eela.eu) projects, collected the geophysical information on seismic movements which were then transformed into audible sound waves using the process of data sonification. In a second step, the sound waves were also "scored" as melodies. The resulting volcanic music can then be analyzed to look for patterns of behaviour and self similarities in eruption dynamics that can be used to predict the future activity of a volcano.

4.2. MIDI Sonification of Mt.Etna Volcano Seismograms

Mount Etna is the Europe tallest and largest active volcano (its volume is more than 350 km^3) and is one of the most active volcanoes on Earth. It is located along the eastern (Ionian) coast of Sicily and occupies a surface area of around 1200 km^2 with a perimeter exceeding 135 km. Its summit height varies frequently (from about 3315 to 3350 m) depending on the eruptive activity or minor collapses at the highest craters. The MIDI sonification of the Mt. Etna seismograms could be described in a simple way as follows.
Let us start from an empty score, as the one shown in fig. 4. We can then imagine to superimpose the seismograms to the score. How many musical notes or octaves are involved is one of the parameter of the sonification; one can take into account the 88 keys of the piano or the 128 MIDI notes, or some other subset like all the white keys and draw the notes according to the shape of the seismogram. The time scale is another parameter of the sonification. As we were discussing above, one can tight or enlarge the rhythmical structure to disentangle some peculiar phenomena. The score is hence ready to be played. The Java package developed for this kind of sonification works just in this way, starting from a discrete data set (the digitalized seismogram) to produce the MIDI score. The metro is customizable as well as the musical scale (i.e., the subset of 128 MIDI notes to be used) and the playing instrument. Following this procedure, we get a melody whose profile perfectly follows the shape and the behaviour of the seismograms, reproducing the oscillation amplitude variations and giving a quick and effective insight to their overall structural properties. Some examples of this sonification can be listened from the Mt. Etna volcano sonification web site (http://grid.ct.infn.it/etnasound/) where it is also possible to download the PDF scores of such sonifications.

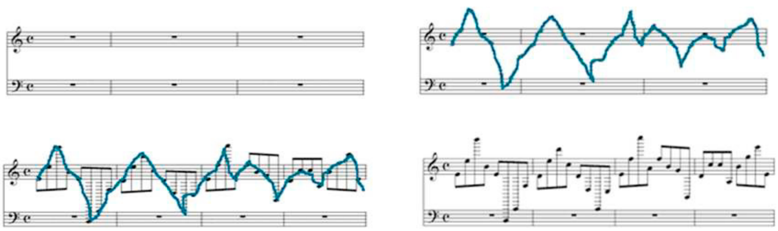

Fig. 4. (a) we start from an empty score; (b) we superimpose the seismogram; (c) we put notes according to the shape of the seismogram; (d) finally we get the melody

4.3. Sound production on the Grid

The first experiments involving sound production with INFN-Grid facilities started during the last months of 2003. In September 2003, CSound (www.csounds.com/whatis/index.html), a free and cross-platform acoustic compiler, developed by Barry Vercoe (MIT), was installed at INFN Catania and the software to interface it with the Grid middleware was written. The compiler, written in C, was tested within the new environment and since its beginning the test phase produced interesting results such as efficient use of the computing resources and customizable quality of the audio files. After a 4-5 months test phase, CSound was deployed on other Grid sites in

Europe and, since the beginning of 2004, CSound is one of the new resources available on a relevant fraction of the continental infrastructure. Data sonification, sound synthesis and algorithmic composition have been investigated using couple of CSound files and Python (or sometimes Perl) scripts. Python scripts have been used to prepare the right data files during the tests. The sonification package has been written in Java and made it run on the INFN Grid. Sonified data were geophysical signals (Mt. Etna volcano activity) collected by a digital seismograph. The computing power needed to run such sonifications and the storage required to keep the results are really large. To create one minute of music coming from volcanoes, the algorithm has to process more than 2.6 millions of data, taking more than 120 MB of disk space and between 10 and 12 CPU hours on a state-of-the-art PC. This is why Grid is a natural platform for this kind of studies. Here follow the steps performed on Grid to let Mt. Etna volcano plays. First of all seismgraphical data have been recorded by a digital seismograph at a sampling frequency of about 100Hz. The total amount of processed data were 44 files, each of them taking almost 1MB and containing 27 minutes of recording of the Mt. Etna volcano activity. Here follows an example taken from on of these files:

```
ASCII files processed:
Starting time: 15/06/2001 00:03:39.920, Frequency: 100.1603 Hz, Samples: 168960
44 43 42 44 44 45 46 43 45 41 48 43 45 45 [..]
```

After a first scaling procedure (to properly arrange the samples in the [-1,1] interval) an array of samples is constructed. The waveform coded in the audio file will have exactly the same regularities, also recognizable thanks to the presence of some higher lines in the spectrum file. The order of magnitude of the frequency of quasi-regular phenomena is in the range 0-50 Hz, with a spectral envelope centred around 25-30 Hz. Starting from the scaled and re-sampled collection of values, two interpolation methods have been implemented: linear interpolation and sample-and-hold (SH) interpolation. Even if linear interpolation returns particularly "smoother" sound strings, nevertheless, for some analysis application, it could be useful to prefer the SH audio file where no external values are added to the original (scaled) data set. For each waveform coded, a Grid job, described by a special file containing (at least) the files to be transferred to the computing node (the Java source code Sonification.java), the data file to be sonified (etna.dat), and a list of files to be transferred back from the computing node to the PC when the computation has finished (audio file, ASCII files representing waveform and spectrum), has been submitted. The audio file collected back contains the sonification of the volcano activity. Moreover, the possibility, granted by the Grid, to have a unique (distributed) shared database for all the seismological data, accessible from the researchers from all over the world, will surely pave the way for a more effective and quick interaction among scientists (even coming from different domains). An interesting and worth mentioning “side effect” of the work done with seismic data audification was the world premiere which aimed to create a dance show based on volcanic sounds to raise public awareness about global warning, The event created a unique creative union between Science and the Cultural Heritage communities. For the first time ever, on the 14th of March 2009 a modern dance company performed to music generated from seismic data, recorded from four volcanoes across three continents. This unique event was facilitated by DANTE, the provider of high speed research and education networks, the two distributed computing projects EGEE and EELA, as well as the CityDance Ensemble (www.citydance.net), a prestigious company based in Washington, D.C.. The dance, titled “The Mountain”, and choreographed by Jason Garcia Ignacio, was part of CityDance Ensemble’s Carbon, a work-in-progress about climate change. The video about this performance is available for download at www.dante.net/upload/mpg/TheMountain.mpg. “The Mountain”’s choreography is based on the structure of melodies created out of seismic waves recorded from Mount Etna in Italy, Mount Tungurahua in Ecuador, and the Mountains Pinatubo and Mayon in the Philippines. Data were transformed into audible sound waves using the sonification technique described above by Domenico Vicinanza (the corresponding author of this paper) who also composed the music used in the dance performance.

5. Conclusions

The benefits of the works presented in this paper are manifolds. ASTRA makes it possible to recreate the sound of instruments that would have previously been either too expensive or too difficult to manufacture manually. It also allows any model and its associated data to be accessed by other parties. This means that the research data can be shared around the world, making it a truly international project of immense value both to archaeologists and historians. The world premiere concert held in Naples (and, the Catania Concert “Musica@Fisica” held on the 26th of June 2009), clearly showcased the immense potential for creativity when two worlds come together: art/music on one side and science/technology on the other. With their huge computing power, essential for the sound modelling, Grid infrastructures like EGEE and EUMEDGRID, based on high-speed networks such as GÉANT and EUMEDCONNECT2, can act as catalysts of creativity. The success of the ASTRA project demonstrates, once again, how high speed networking technology can underpin research collaboration across a wide range of subjects and allow the academic world to work together across multiple locations. As its full name indicates, ASTRA hopes to bring a number of ancient instruments back to life, eventually creating a complete orchestra made up of yesteryear’s gear. One of the most relevant, long-term results is the birth of the Lost Sounds Orchestra (www.lostsoundorchestra.org), the ASTRA project orchestra: a unique ensemble made of reconstructed ancient instruments coming from the ASTRA research activities. It is the first orchestra in the world composed only by reconstructed instruments. The purpose is to involve musicians, artists, collaborating with real players, instrumentalists and singers, to enjoy together a heritage lost in the past. It plays sounds being lost in time due to instruments too complex both to build and play. The GILDA and EUMEDGRID infrastructures will continue to be heavily used to produce the huge amount of high-quality sound samples needed by the orchestra. For any further information or enquiry, readers can contact either info@astraproject.org or info@lostsoundorchestra.org.

Data sonification and melodisation are two applications aiming at converting data and information coming from scientific environments into audible signals (melodies, waveforms, etc). One of the most important outcomes is the possibility of spotting similarities and patterns in more efficient way. Fast Fourier Transforms and Wavelet analysis can be applied to sonified data in an easy and effective way, trying to correlate spectral evolutions and geophysics (changes in volcanoes’ behaviour). The successful tests done in Italy with Mt Etna and in Ecuador with Mt Tungurahua clearly demonstrate the interest of the scientific community in this new approach of sound analysis. Another relevant result is the possibility offered to blind researchers to perform data analysis and data mining listening to the data instead of looking at them on a screen. The data sonification experiments have been carried out in collaboration with the INFN Division of Catania and the TRAC (Technologies and Research for Contemporary Arts) project. For more information about the TRAC project, readers can contact gaetano.foti@ct.infn.it.

Acknowledgements

We gratefully acknowledge all the people who supported this work contributing with ideas, requirements and feedback. Authors thank in particular the support provided by the EELA, EUMEDGRID and EGEE-III projects. The GILDA Team is also acknowledged for the help provided in interfacing the application described in this abstract with the gLite (www.glite.org) grid middleware.

References

1. L. Hiller, P. Ruiz, Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Part 1 - Part 2 JAES (1971).
2. J. O. Smith, Physical modeling using digital waveguides. Computer Music Journal, 16(4), (1992) 74–91.
3. J. O. Smith, .Physical modeling synthesis update. Computer Music Journal, 20(2), (1996) 44-56
4. R. Barbera, L. N.. Ciuffo, et al. GILDA: Grid INFN Virtual Laboratory for Dissemination Activities. In: International Conference on Computer Supported Education, Lisboa. Conference Proceedings. Setúbal: INSTICC Press, (2) (2009) 273-278
5. G. Andronico, et al. GENIUS: a web portal for the grid. *Nuclear Instruments and Methods in Physics Research*, A 502, (2003) 433-436.
6. G. Andronico, et al. GENIUS: a simple and easy way to access computational and data grid. *Future Generation of Computing System, 19,* (2003) 805-813.
7. G. Andronico, et al. The GENIUS web portal: an easy way to access the grid. *Methods of Information in Medice,* 44, (2005) 210-220.
8. R. Barbera, et al. The GENIUS Grid Portal: Its Architecture, Improvements of Features ,and New Implementations about Authentication and Authorization. Paper presented at the 16th IEEE International Workshops on. doi: 10.1109/WETICE.2007.4407171 (2007)