An instrument of sound and visual creation driven by biological signals

Andrew Brouse, Jean-Julien Filatriau, Kosta Gaitanis, Rémy Lehembre, Benoît Macq, Eduardo Miranda, and Alexandre Zénon

***Abstract*— Recent advances in new technologies offer a large range of innovative instruments for designing and processing sounds. This paper reports on the results of a project that took place during the eNTERFACE06 summer workshop in Dubrovnik, Croatia. During four weeks, researchers from the fields of brain-computer interfaces and sound synthesis worked together to explore multiple ways of mapping analysed physiological signals to sound and image synthesis parameters in order to build biologically-driven musical instruments. A reusable flexible framework for bio-musical applications has been developed and validated using three experimental prototypes, from whence emerged some worthwhile perspectives on future research.**

***Index Terms*— EEG, EMG, brain-computer interfaces, digital musical instruments, mapping**

I. INTRODUCTION

MUSIC and more generally artistic creation has often drawn inspiration from the possibilities offered by technology. For example, the invention of the piano was a key event in the emergence of romantic music. More recently, the electric guitar and synthesizer have allowed elements of Jazz to move towards Pop Music. Digital signal processing and multimedia computers have enabled the creation of an overwhelming gamut of new sounds. More recently, work has begun to discover ways to control these new sounds with the ultimate goal of creating new musical instruments which are playable in real-time.

The present contribution is focused on the development of new musical instruments activated by the electrical signals of the brain (EEG) and of the muscles (EMG). We are exploring features of bio-signals by mapping them to parameters of computer-generated sounds. This work is the continuation of a project initiated last year during the first eNTERFACE workshop in Mons, Belgium [1] [2]. In our previous work we used inverse methods and left/right cortical activity differentiation - as in classical Brain to Computer Interfaces (BCI) [3] - to design the mapping between physiological signals and sound synthesis parameters. We felt, however, that the 'musification' of biological signals could benefit by using the richness of the raw brain and muscle signals rather than just relying on the results of analyses. Hence, we took the opportunity of this workshop to explore new fission/fusion strategies by conducting three experiments:

- The first application was the sonification of EEG signals, using a vocal model, which could be either used as a musical instrument or as a diagnostic tool
- The second experiment was more directed to musical applications and interactive performance, and is aimed to generating visual and sonic textures controlled by the results of EEG spectral analysis.
- The last application was a tentative attempt to extend the hyper-instrument paradigm by building a physiologically enhanced didgeridoo that relies on wearable sensor technology [4].

This report is composed of five main sections: a history of music and sonification controlled by biological signals; a theoretical framework which exposits the fission/fusion of biological signals in musical applications; a description of the hardware and software architecture of our platform dedicated to musification of biological signals; a section which details the different EEG analysis methods we have implemented; and finally detailed descriptions of the experiments with some possibilities for improvement for each.

II. HISTORY AND THEORY OF SONIFICATION OF BIOLOGICAL SIGNALS

Whereas the use of biological signals to control music systems has a long and rich history dating back at least 40 years [5], the contemporary notion of sonification of biological data for auditory display is relatively recent, the first articulated writings beginning to appear around 1994 [6]. Sonifications as evidence or as objects of scientific knowledge also present fascinating opportunities to interrogate notions of scientific truth and ontology. In fact, the practice of using sound as a tool for medical diagnosis for example, dates back more than 150 years to the development of the stethoscope by René Laennec [7] and the attendant practice of mediate auscultation. As listening to the body is one of the most basic skills in a standard medical education, trained doctors are thus highly sensitive to sound and its implications for diagnosis. Simultaneously, over the past 150 years or so, scientific measurement equipment has become increasingly sophisticated and precise. The possibility of making highly precise measurements of phenomena has - until recently however - been almost exclusively destined for visual display. That is, the results of

This report, as well as the source code for the software developed during the project, is available online from the eNTERFACE'05 web site: www.enterface.net.

This research was partly funded by SIMILAR, the European Network of Excellence on Multimodal Interfaces, during the eNTERFACE05 Workshop in Mons, Belgium.

R. Lehembre was supported by a grant from the Belgian NSF(FRIA).

Andrew Brouse and Eduardo Miranda are with the Interdisciplinary Centre for Computer Music Research, University of Plymouth, U.K.

Jean-Julien Filatriau, Rémy Lehembre and Benoît Macq are with the Communications and Remote Sensing Laboratory, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Alexandre Zénon is with the Neurophysiology Laboratory, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

these sophisticated measurements has been, to a very large extent, primarily expressed in visual terms: as graphs, line traces, charts, histograms, waterfall charts etc., either on paper or some similar support, or on a luminous display such as a CRT or TFT computer screen. Recently, the notion of auditory display of scientific or other information has become current. Auditory display has several advantages over visual display especially for critical applications largely due to the ways in which our auditory perceptual apparatus passes information to the brain. By using salient characteristics of sound, such as rhythm, duration, pitch, timbre and harmonic/enharmonic content, it is possible to rapidly and accurately express complex, multimodal information in a manner which can be quickly and accurately grasped by a trained listener. Our auditory apparatus is capable of distinguishing very subtle differences in simultaneous, complex auditory streams and it can do this very quickly and accurately [8]. Work has already been done in the sonification of biological signals such as EEG - notably by Gottfried Meyer-Kress and his early work in EEG sonification - which has been furthered by a workshop at ICAD2004 entitled ”Listening to the Mind Listening” and even more recently by a workshop and paper given at ICAD2006. In a related field, Mark Ballora did pioneering work in the sonification of the cardiac rhythms related to the diagnosis of conditions such as sleep apnea [9]. In most of the preceding cases, however, the sonifications were performed ”offline”, that is, not in real-time. The goal of this part of the project is to develop a real-time system for sonification of biological data. Previous efforts along these lines have led to very specific solutions with particular hardware and software components which have proved hard to re-use and not sufficiently flexible for diverse applications. Our goal, thus, is to begin work upon a flexible, re-usable, open-source framework for the generalized sonification of biological signals. This platform would provide a stable, re-usable, flexible and comprehensive environment for the sonification of human biological data for auditory display. This display would be useful for doctors, scientists, researchers and clinicians in the study and diagnosis of normal and abnormal indicators. Much as this is primarily a tool for scientific research, it is also envisioned as a useful tool for music technologists, composers and performers in the realisation of musical forms which are driven by measured biological phenomena. It is felt that a stable platform for such musical research is as useful in the musical sphere as it is in the scientific one. In fact, an historical survey of biologically driven music, such as brainwave music, shows periods of intense, productive activity followed by quiescent lulls where very little happens. It is felt that these lulls are due in part to a lack of appropriate tools and techniques for consistent and repeatable musical realisation and thus, little opportunity for practices and mastery of bio-instruments such as brainwave music.

III. Fission and Fusion of bio-signals

A. Our proposed framework

We proposed to model the design of musical instruments or sonifications as a fission-fusion process. Our theoretical framework is shown in Fig. 1. The central issue is the fission from each of the given input modalities (EEG and/or EMG in this case) into salient features channels. These features channels are then fused into commands which activate different aspects of the related sound and image synthesis processes. The process of fission of commands into the output feature channels which are then fused back into the global output signal is also seen as part of the fission-fusion process. This process can be likened to the attendant processes of analysis and resynthesis which are so central to digital signal processing.

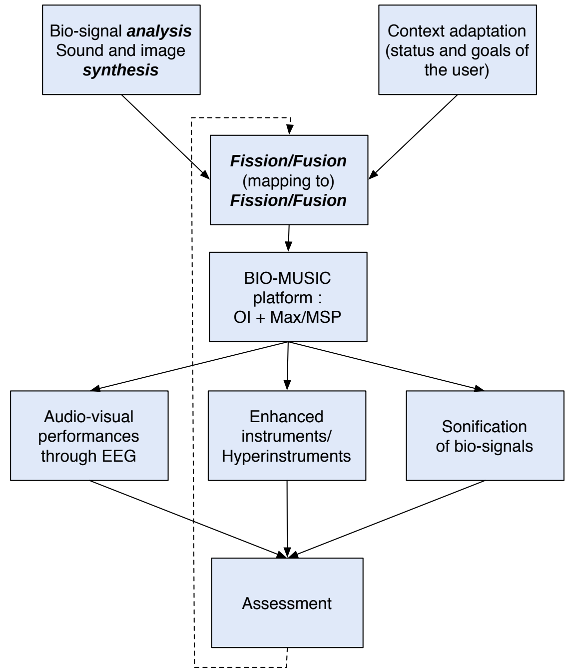

Fig. 1. Framework for the design of biologically-driven musical instruments

B. Mapping

In the literature on digital musical instruments [10], the term mapping refers to the transformations performed upon real-time data received from controllers and sensors into control parameters that drive sound synthesis processes. One of our objectives during this workshop was to design consistent mappings between biological signal features and sound synthesis parameters in order to create biologically-driven musical instruments and sonifications.

C. Usability measurements

The three systems will be improved based upon: assessments of usability and aesthetics by musicians, aesthetic judgements by audiences, and quality of discrimination between relevant EEG patterns in the case of sonification for diagnostic purposes.

IV. The platform architecture

A. Towards an open source system

Our aim in the long term is to produce an entirely open source platform dedicated to the real-time analysis of EEG

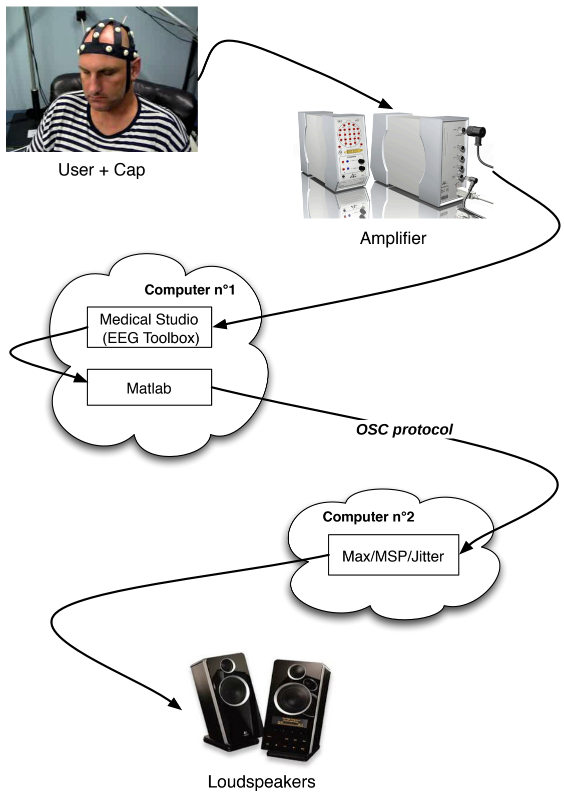

Fig. 2. Architecture of our bio-music platform

signals and other biosignals for musical applications. Since our work is multidisciplinary it involves using resources from different fields of study and thus different software packages are needed. During this workshop we used Matlab for the analysis of EEG signals and Max/Msp/Jitter for the sound and image synthesis. We plan to shift our development toward open source software like Octave [11], Python [12] and PureData [13] in the future. In the following, we describe the architecture of our system (Fig. 2) in a bottom-up way, from hardware data acquisition to software implementation.

B. Hardware

1) EEG equipment: EEG signals are recorded with a *dti* [14] cap containing 18 electrodes located according to the 10/20 international positioning system. The signals are amplified with an biosignal amplifier provided by *dti* with a gain of 10^6 and a default sampling rate of 128Hz. Due to limitations in real-time signal processing, we sampled at 64Hz. Once captured, the data is then bandpass filtered between 0.5 and 30 Hz to remove extraneous signals. C_z was used as a reference electrode while P_z was taken for the ground.

2) EMG equipment: For EMG signals, we worked with the same equipment but changed the gain to 1000 since EMG signals have much larger amplitudes than EEGs. Disposable electrodes were used, 3 per muscle, with one as a reference and placed near a bone (i.e. elbow or knee), a second was posed along the muscle (belly-bone junction), the third, taken as ground, was via a conductive bracelet worn by the user.

C. Software

Our platform is currently implemented via four software packages running on two computers which manage the specific tasks required by the global application :

1) MedicalStudio-EEGToolbox: Acquisition and visualisation is done using EEGToolbox, a plugin written in C++ for MedicalStudio [15], an open source software platform for medical data analysis and display which runs under Linux. This toolbox saves the data and can also send it using UDP to another computer running Simulink under Windows. The connection between the biosignal amplifier and the computer running Linux is made with a usb cable.

2) Matlab-Simulink: Matlab was chosen for easy code generation. Although we had developed the previous year a simulink program, we switched to Matlab in order to spare a computer. This way, the acquisition and signal analysis is made on the same Linux running computer. Simulink could not be used because it suffers from different bugs under Linux that makes it hard to use.

3) Max/Msp/Jitter[16]: : Max/MSP is a graphical development environment dedicated to real-time interactive applications. In use worldwide for over fifteen years by performers, composers or artists, Max/MSP is a combinaison of Max software for the control of musical applications through MIDI protocol, and MSP, an add-on package for Max enabling the manipulation of digital audio signals in real-time. Jitter is an other additionnal library for Max environment, offering a large range of real-time image and video processing tools.

4) OpenSoundControl (OSC) : A link between Matlab and MaxMsp: In order to transfer data between softwares, we used the OSC protocol [17] which sits on top of the User Datagram Protocol (UDP). It allows a fast and reliable data exchange since we work in a local area network. Packets are sent with a header containing the name of the corresponding data as well as the size of the packet. This makes it very useful since the receiving program can easily manage the arriving packets. The maximum size for the packets is 65536 bytes long. We were thus able to send the raw EEG signal and many features computed with matlab to Max/Msp allowing a maximum flexibility (An excerpt from the code is detailed in App. I).

V. Fission of bio-signals

A. Introduction

We worked with two different bio-signals, EEG and EMG. We describe in this section how to operate a fission of these signals in order to extract relevant features. Let us present briefly these two kind of signals:

- The EEG signal is a rich and complex reflection of neuronal electric activity that takes place in the brain. Since the first electroencephalogram recording made by Berger in 1929, different waves have been described corresponding to several frequency bands. Although these waves are well known, their frequencies and amplitudes are not directly under subject's control, but only reflects very general states of the brain. Therefore, using a simple frequency analysis as input to the sound synthesizer will

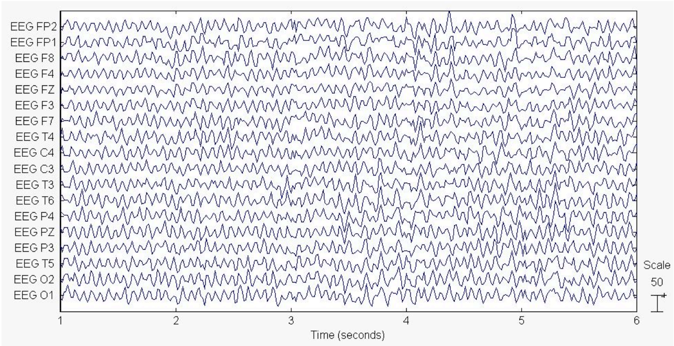

Fig. 3. Recorded EEG, a 13Hz alpha rythm can be observed. The user was in a drowsy state after a heavy lunch and had taken an espresso

not allow enough controllability. Other, more complex signal properties can reveal more useful. On an other hand, EEGs have a very good time resolution of $\simeq 1ms$ unbeaten by recent methods (fmri..). This property is very valuable for the purpose of musical instrument control and should be taken care of. Finally, since electrodes are placed at different locations, it is important to take into account the spatial information.

- The EMG signal is produced by the electrical potential generated by muscle cells. The increase in contraction strength of the muscle is associated with an increase in the number of cells that produce electrical potentials (depolarisation), and hence an increase in signal amplitude. This signal contains two main waves, a low frequency wave that describes the movement, and a higher frequency wave that includes more precise information on the electrical activity of the muscles. Due to hardware limitations, we focused on the the low frequency band (i.e. the envelope of the signal). The higher frequencies could be used in a further version of the project to take advantage of their higher temporal resolution.

B. EEG fission according to frequency bands

We describe here the partition of the EEG into frequency bands

- Delta (0.5-4 Hz): This wave has first been discovered by W. Gray Walter in 1936 with a patient that had a tumor. Thus in the awake, it is quite alarming to present the slow characteristic waveform of the delta rythm. However, for a sleeping person, high amplitude delta waves are normally present in the EEG. For our application it appears evident that this rythm will not be of great use unless we create a composition for sleeping performers!
- Theta (4-8 Hz): Scientists still debate whereas theta activity is relevant to an early drowsiness state or if it reflects some kinds of mental activity. Nonetheless it is a faster rythm than delta and could be linked to brain activities such as memory [18], or can be modulated by visual stimulation (ref).
- Alpha (8-12 Hz): Alpha rythm is a leading indicator of subject's relaxation. Alpha synchronization (leading to amplitude increase) occurs in the absence of any visual stimulation, as for example, when the user closes his eyes. In contrast, any visual stimulation lead to posterior alpha desynchronization. Therefore it is a good tool for our application since it can be used as a switch. Alpha waves could also be associated to conscious visual perception [19].
- Beta (12-24 Hz): Extending over a large bandwidth, the beta activity reflects intense activity such as listening, taking decisions, or more generally, arousal. It is a dominant rythm in the normal adult awake EEG.
- Mu rhythm: This rythm, as the alpha rythm is between 8 and 12 Hz but is specific to imaginary or real movements [20]. It is located in the motor cortex and is contralateral to the movement i.e. for a left hand movement, a desynchronization will appear in the right hemisphere.

Let us recall that the frequencies given above are not strict but subject dependent. The control of these waves by the subject can only been achieved following extensive training. As a consequence, it is difficult to produce a controllable EEG driven musical instrument on the basis of the amplitudes of these signals alone. However we can derive a few indicators from a spectral analysis:

1) Frequency Values: A Fast Fourier Transform was used to compute the frequency. We used a 1 sec window to compute a 32 points transform

2) Spectral Entropy: The spectral entropy, a measure widely used showing the complexity of a signal, is computed in order to detect salient rhythms. It is given by:

$$H_{sp} = -\sum_f p_f ln(p_f) \tag{1}$$

where p_f is the probability density function (PDF) that represents the normalization of the power given at frequency f regarding the total power spectrum:

$$p_f = \frac{s_f}{\sum_f s_f} \text{ with } f \in \aleph^+ \text{ and } f \leq 32 \tag{2}$$

3) Spectral Edge: The spectral edge is the frequency under which 95% of the spectral energy can be found This value gives an indication of where the signal is concentrated.

4) Asymmetry ratio: In order to detect when the user makes left or right side movement, we use a very simple tool that computes the normalized difference between the power contained in the mu rythm of two electrodes located in the left and right motor cortex (i.e. C_3 and $C4$):

$$\Gamma_{[8-12Hz]} = \frac{C_{3,[8-12Hz]} - C_{4,[8-12Hz]}}{C_{3,[8-12Hz]} + C_{4,[8-12Hz]}} \tag{3}$$

This ratio has values between -1 and 1, the sign indicating the side of the body that was moved

C. EEG fission according to signal spatialization

As mentionned above, taking into account the position of the electrodes is extremely important in EEG analysis. Two similar methods, the Common Spatial Subspace Decomposition (CSSD) [21] and the Common Spatial Patterns (CSP) [22], extract information from the most relevant electrodes.

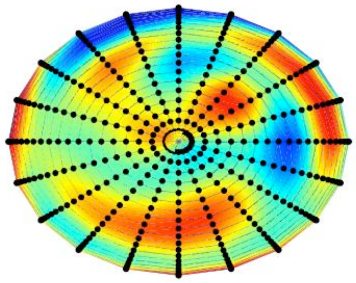

Fig. 4. Inverse Problem visualisation : The black dots are the location of the sources

These methods are known to be the most accurate in the BCI community. However they imply offline preprocessing and low variability between sessions which in our case is seen as a limitation. Indeed, our aim is to produce live music in different environments thus rendering a training session obsolete. An other method that is starting to gain success is based on the principle that the EEG signals are generated by sources (i.e. assemblies of neuronal cells that when combined produce a sufficiently strong current that can be measured at the surface of the scalp) and that the propagation of electrical currents through brain tissues can be modeled with Maxwell's equations. Therefore using a model of the brain it is possible to reconstruct the activity of sources and gain access to the spatial location of brain processes. Besides, this method offers a visualisation of the activity. Having described this method in [1], we will briefly resume the main steps of this method:

1) Head Model: We used a four spheres head approximation based on [23] ,[24] and [25]. Each layer represent, the brain itself, the cephalo-rachidian liquid, the cranial box and the scalp. There are 400 dipoles (Fig. 4) distributed over the cortex (the surface of the first sphere). As an approximation, deep sources are not taken into account. The potential measured on the n electrodes, ϕ, is linked to the value of the m sources, φ, according to the lead field matrix G and additionnal noise η:

$$\phi = G\varphi + \eta \tag{4}$$

The lead field matrix is computed once for a given head model and remains constant further on. Knowing ϕ from the recording, we wish to find φ. Unfortunately this so-called inverse problem is an ill-posed problem since the number of unknowns is much greater than the data at hand. Following is a short description of the inverse problem

2) Inverse Problem: Solving Eq. 4 can be done using a bayesian formalism :

$$P(\varphi|\phi) = \frac{P(\phi|\varphi)P(\varphi)}{P(\phi)} \tag{5}$$

where:

- $P(\varphi|\phi)$ stands for the *a posteriori* probability to have the source distribution φ matching ϕ
- $P(\phi|\varphi)$ is the *likelihood* i.e the probability to have the given data according to the sources. It depends on the quality of the recording and on the head model
- $P(\varphi)$ is the *a priori* knowledge we have about the sources.
- $P(\phi)$ is a normalizing probability that can be neglected

Finding the best solution to the inverse problem comes down to maximizing the *a posteriori* probability. This can be achieved in various ways as different methods have been proposed during the past 15 years [26] [27]. We implemented the LORETA algorithm because it gives a maximally smooth solution.

3) Features: Four features are derived from the solution of the inverse problem and are sent to the sound processing unit. To compute these features, we divide the source space in four subspace representing the frontal, occipital, left and right sensori-motor parts of the brain. This decomposition is based on the fact that the frontal zone is associated with memory and cognitive processes while the occipital region is linked with visualization. Left and right motor-cortex side are associated with left and right limbs movement. This is a very simplistic view of the brain but is adopted as a first approximation.

D. Further work : EEG fission in 3D

We discussed earlier the importance of taking into account the spectral, spatial and temporal information of the EEG. We studied some techniques of spectral information retrieval and a technique to improve the spatial resolution. We could in a future approach combine the inverse problem and spectral methods. Another approach would be to work with spherical harmonics using an interpolation of the electrodes on a half-sphere. Finally, including temporal constraints in the IP could improve the obtained solutions.

VI. EXPERIMENTS

A. Sonification (Vocalisation) of EEG

The current implementation of sonification uses a source-filter voice synthesis model developed by Nicolas D'Alessandro and others [28] which in this case has been tuned to emulate the multi-phonic singing chants typically produced by Tibetan Gyuto monks or by Tuvan traditional folk singers. The voice synthesis model as delivered, exposes a limited set of functionalities with given ranges. In the interests of proper encapsulation and OO design, we respect these givens and will work with them. In this case the controller mappings used the F1-4 formant frequencies whilst the F0 was not directly controlled. Additionally, parameters representing "tension", "hoarseness", "chest/head balance" and "fry" were also controllable. Any available mapped data source (alpha, beta, theta, mu etc.) can be used as a controller for any of the synthesis parameter. It was found that the formant frequencies were best controlled by signals which do not change too quickly or vary too greatly. A facility is available to control the positioning of any generated sound source with respect either to a stereo sound field or to a 5.1 quasi-surround sound field.

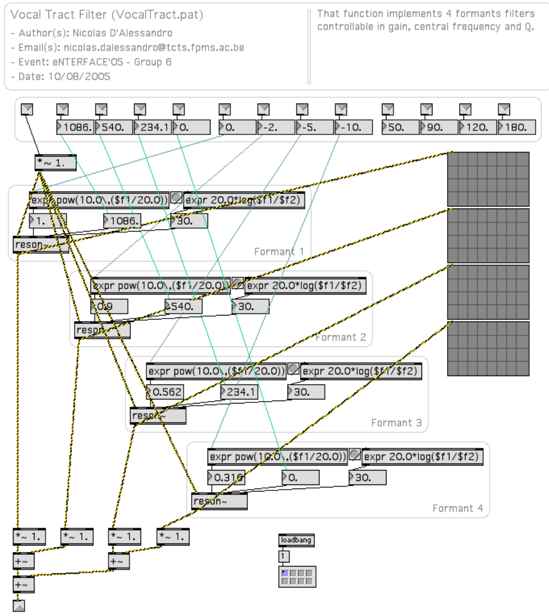

Fig. 5. Vocal Tract Filter realised by Nicolas d'Alessandro et al. It implements four formants controllable in gain, central frequency and Q.

Results: Due to the highly prototypical nature of this platform, no extensive testing was done and it is thus not possible to provide comprehensive analyses of the relative success or failure or suitability of this platform for any current intended usage. Test that were made did indicate that basic functionality of modules and the software as a whole is intact and operational yet many improvements in precision, usability and flexibility are still lacking. Going forward it is envisioned that these characteristics will be ameliorated so that the platform will become a flexible, stable, consistent and useful tool for scientists, medical professionals and musicians in the future.

B. EEG driven audio-visual texture synthesizer

In this instrument we tried to link three modalities by exploiting results of EEG frequency analysis to control both visual and sonic textures synthesis modules (Fig. 6). This approach aimed to provide a visual feedback to the performer/audience enabling a better understanding of the fission/fusion process. Practically, the image synthesis module takes as input parameters data received from EEG analysis module, whereas sound synthesis parameters are extracted from both the output image and the results of EEG analysis. This strategy of linking synthesis processes should enable a strong correlation between resulting image and sound. Both synthesis modules have been implemented in Max/MSP environment, the image processing tasks relying on the specialized additional library Jitter. Following sections give more details on both image and sound synthesis modules.

1) Creation of the visual texture: The starting point of the creation of the visual texture is a space/frequency representation of cerebral activity: each second the EEG analysis module transmits to the visualization module a matrix containing

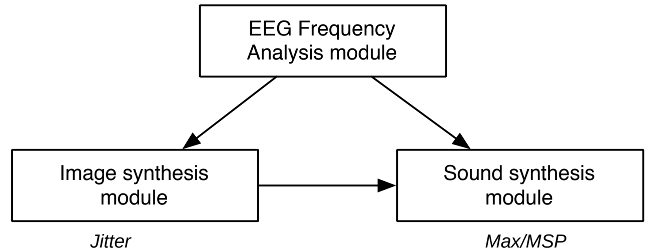

Fig. 6. General scheme of the instrument

the energy in the 32 bands of the spectrum of the signals measured by each of the 18 electrodes. A crossfading effect between consecutive matrixes is then achieved allowing to obtain a continuously and smoothly changing image. This moving image is then distorted: firstly a linear interpolation is done in order to blur the image. At this step of the process, the resulting image is a grayscale texture derived from the space/frequency representation of the EEG analysis (Fig. 7).

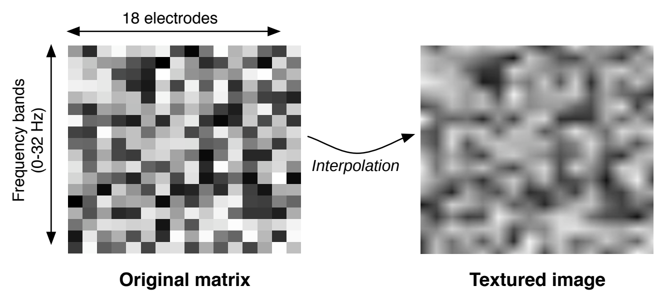

Fig. 7. Creation of a grayscale textured image from the space/frequency representation of brain activity

Then we apply a colorization process, based on color lookup tables, to remap grayscale into colored image. Lookup tables, also called transfer functions, are arrays of numbers where an input number is 'looked up' as an index in the table. The number stored at that index is then retrieved to replace the original number. In our case, we use lookup table to convert a monochrome into RGB value. In grayscale image, low-energy areas are represented in black and gradually whiten when energy increases. Our colorization process modifies the color associated to maximal energy, by defining a new color scale that will map in the resulting image high values, originally represented in white, to a new color defined by the result of EEG analysis image. The choice of the color, called C, associated to the maximum of energy, is driven by the distribution of energy between the alpha, beta and theta bands of the EEG signals. The three RGB components of this color, C_{Red}, C_{Green} and C_{Blue}, are thus weighted by the level of energy L_α, L_β, L_θ in the three frequency bands alpha, beta and theta respectively (Fig. 8). We obtain by this way a direct link between the color of the resulting image and the maximal energy frequency band of the EEG analysis.

The color lookup table is refreshed as soon as new values for alpha, beta, theta bands are received from EEG analysis module, i.e. one time per second, The transfer function used

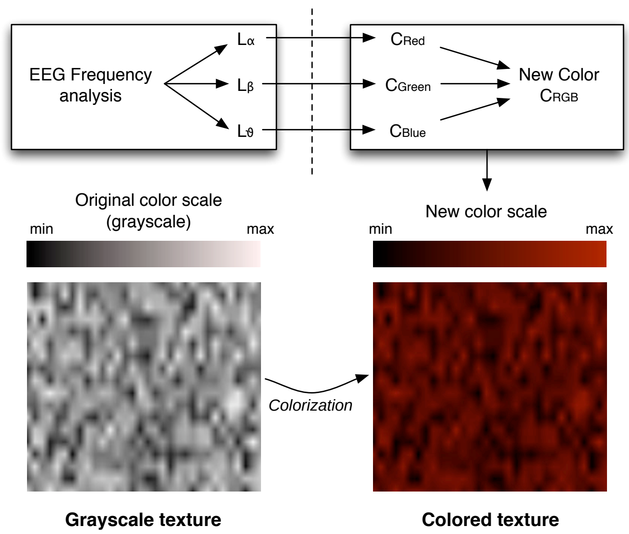

Fig. 8. Colorization of the texture following the distribution of energy in alpha, beta and theta bands

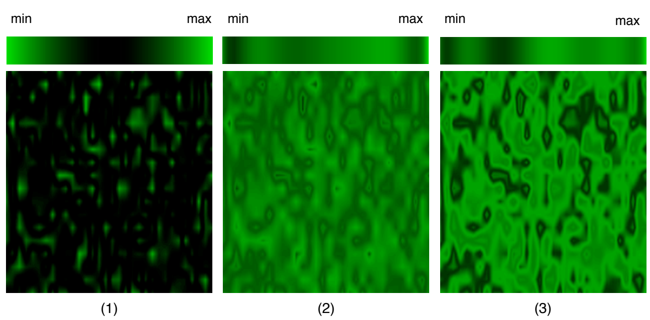

Fig. 9. Textures obtained from the same grayscale texture using different color transfer functions. Rightmost and leftmost images correspond to low and high level of entropy of the signal respectively.

for image in Fig. 8 is linear, but it is also possible to use non linear lookup tables, that give interesting effects on the resulting image and allow to obtain quite different types of visual textures, as shown in Fig. 9. In our instrument six predefined color transfer functions were available, and the choice among them was driven by entropy of the EEG signals, which is an indicator of state of relaxation of the subject, Mapping was done such a way that a dropping of the entropy results a more contrasted image.

2) Translation in sonic texture: The translation of the visual texture created from EEG analysis into sound is based on one of the most popular technique of sound synthesis, the subtractive synthesis, widely used in musical applications such as analog synthesizers. The basic principle of subtractive synthesis is the use of complex waveforms, rich in harmonic or inharmonic information, which are then spectrally shaped by filters bank. In subtractive synthesis, the spectral envelope of the resulting sound is the product of the spectral envelope of the source with the frequency response of the filters bank (Fig. 10).

Here we used as audio source a pink noise, whose energy

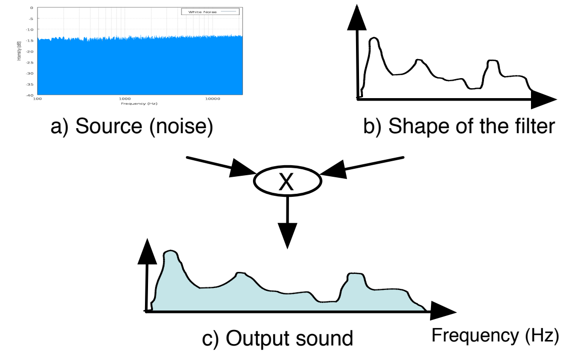

Fig. 10. Principle of subtractive synthesis

is geometrically distributed in the spectrum (constant energy per octave). The implementation of subtractive synthesis in the Max-MSP environment is based on the fffb object (fast fixed filter bank), that models a bank of 32 bandpass filters. This object takes as input a list of 32 values controlling the gain of each filter. In our instrument, this list is obtained from the visual texture created from EEG analysis by the following process (Fig. 11) : a sliding window extracts a sharp vertical band of the image (step 1), whose values are stored in a 1-D vector (step 2). This vector is then downsampled to obtain a list of 32 values (step 3) that will be used to drive gains of three filter banks (step 4). In order to musically enrich the resulting sound, we placed three filters bank in parallel, that resonances are differently distributed in the spectrum, implying each of the filters bank to produce its proper and discriminable timbre. Final synthesized sound is a mix of these three sounds whose loudness are respectively controlled by the level of energy in the alpha, beta and theta frequency bands extracted from EEG analysis (step 5), in a similar way of the weighting of RGB components of the final color in the colorization process of the visual texture. This enables a strong correlation between synthesized image and sound, both driven by the results of EEG frequency analysis. Videos demonstrating this instrument are available online [29].

3) Results and future works: One aim of this work was to build a brain-computer interface linking image and sound synthesis processes to EEG analysis. We reached this objective by designing a subtractive synthesis instrument that spectral envelop is extracted from a visual texture resulting of EEG analysis. This approach enabled to establish a clear relation between output image and sound. In the future some main tracks of improvement should be investigated. Firstly it would be interesting to modify the space/frequency representation of brain activity that is the basis the creation of the visual texture. Indeed, a spherical representation relying on the localization of the electrodes on the scalp would be closer to the actual spatial brain activity. Concerning the image-to-sound translation, other sound synthesis techniques should be tested, such as additive or granular synthesis, in order to enhance the correlation between the synthesized visual and sonic textures. For this it would be interesting to exploit existing works in the fields of image sonification and auditory display [30]. Finally,

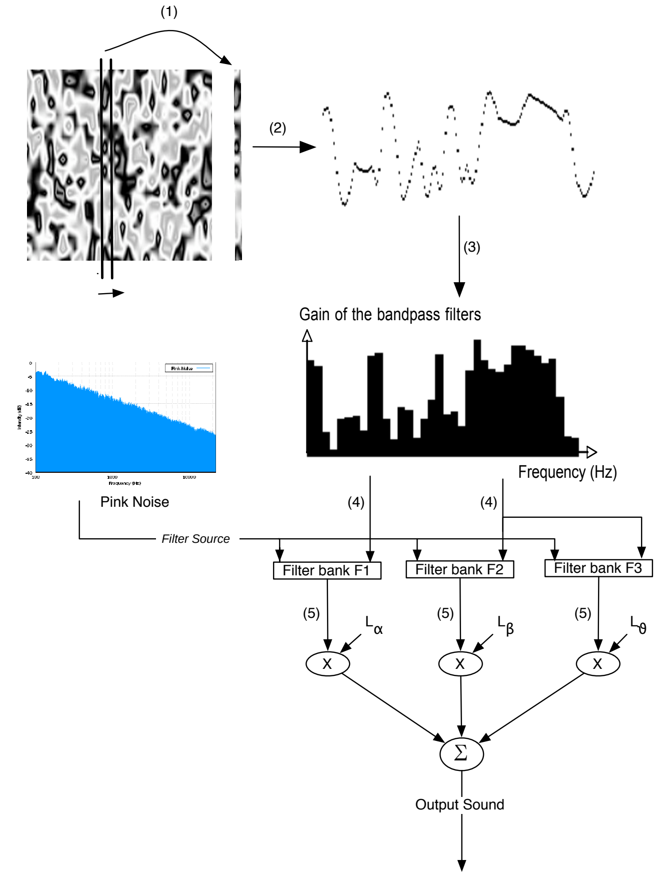

Fig. 11. Image-to-sound texture translation driven by EEG frequency analysis

we should keep working on the improvement of mapping between EEG analysis features and synthesis parameters. In this instrument, the user was actually not able to control the resulting image and sound, mainly because data we interpret as input parameters in the synthesis modules (spectral content of EEG signals) are hardly controllable by the human. In order to increase the playability of the instrument, it could be worth to add in the mapping easily controllable parameters such as EEG features linked to eye blinking. More generally the design of a mapping between EEG analysis results and synthesis parameters in such a brain-computer interface requires an explorative and inventive approach that could only be reached by intensive experimental sessions.

C. *EMG enhanced didgeridoo*

The third experiment we led during this workshop aimed to design an EMG-enhanced didgeridoo. The didgeridoo is an Australian traditional wind instrument, sometimes described as a wooden trumpet or a drone pipe. Because it is made up without keys, pitch produced by a didgeridoo is limited in a quite sharp range of frequencies, directly related to the dimensions of the instrument. In this experiment we tried to exploit EMG captors measuring contraction of muscles on one leg to enlarge the possibilities of the musician, especially in extending the range of pitch produced by the didgeridoo.

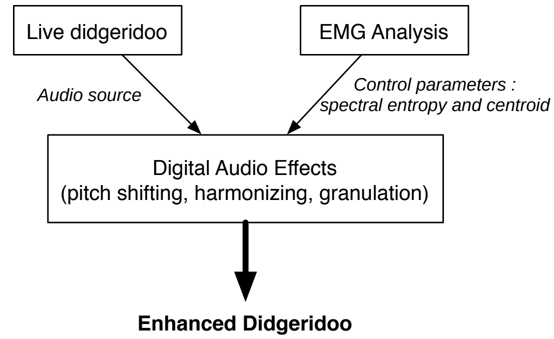

Fig. 12. General scheme of the enhanced didgeridoo

This instrument was running on two computers, one managing Medical Studio for the capture of EMG signal and the other one running Max-MSP for the implementation of digital audio effects. EMG signal, captured with Medical studio, was transferred to Max/MSP, where spectrum centroid, entropy and signal power around a frequency band of 8 Hz were computed. These resulting signals were differentially modulated by leg movements in such a way that the subject was able to control each of them, more or less independently. Two digital audio effects modules were thus designed: in the first one, entropy of the EMG signal, which was the most easily controllable parameter, was used to modify the cutoff frequency of a bandpass filter applied on the didjeridoo's sound. Spectrum centroid controlled a very slight pitch shifting (with a maximum ratio of 1.05) and power in the 8 Hz band controlled the cutoff frequency of a bandpass filter which was used in a feedback loop inside a granular synthesis process. In the second audio effect module, we used entropy of EMG signal to drive two simultaneous pitch shifting processes, one moving downward and another one moving upward. Videos demonstrating these experiments are available online [31]. These quite simple experiments demonstrated the musical potential of EMG-enhanced musical instruments: indeed mapping audio effects parameters with muscles contraction seems to get their control very intuitive and expressive. In the future we will pursue to investigate this field by testing more complex configuration of EMG-enhanced instrument, with multiple captors on several areas of human corpus (arms, neck), providing an actual measure of the physical activity of the musician. Similar experiments will be also carried out with other musical instruments (clarinet, accordion), taking into account the specificity of musical gestures associated to each instrument for the design of captors configuration and mapping strategy.

VII. Conclusion and further works

Building on the experience gained during the eNTERFACE'05 workshop, we have explored new horizons in biomusic. Last year we focused mainly on left and right hand movements thus working with limited inputs to the sound synthesis algorithms. Our current approach is to take maximum benefit of the richness of the EEG by extracting as many independant features as possible. We have adapted our architecture to enable multi-dimension data transfer between Matlab and MaxMsp. More sophisticated mapping could then be made under MaxMsp giving a higher correlation between sound and EEG analysis. The gap between art and science was filled by combining a relevant and aesthetic visual feedback. However the question of controlling the instrument remains open as the development of the interface itself did not leave enough time for a necessary training and assessment step. This underlines the amount of work left to achieve an instrument that could be in the future the equal of traditionnal instruments. We will focus in the future on two parts that seem important to us. First the migration of the platform to an open source software platform using OpentInterface will allow the sharing of our results and perhaps trigger new partnerships. The second and last step, but certainly not the least, will be to achieve a better control of the instrument itself. This means many training sessions over a long period of time. The authors are dedicated to pursuing their goal of achieving an entirely biological music.

APPENDIX I
OPEN SOUND CONTROL

In order to implement OSC, we use the freely available tcp-udp-ip toolbox for Matlab. The pnet function allows to create a packet and send data through a UDP connection. It is possible to include different headers in the same packet. For example, sending all the EEG raw information can be done with the following code:

```
% head of the message
header = ['F1','F2'...]
for j=1:nbrElectrodes
pnet(udp,'write','/header(j,:)');
% mandatory zero to finish the string
pnet(udp,'write',uint8(0));
...
% comma to start the type tag
pnet(udp,'write',',');
% number of float to write
for i=1:sizedata
pnet(udp,'write','f');
...
% data to send
pnet(udp,'write',single(data(i,j)));
```

The source code is available online on the enterface website.

APPENDIX II
DISCUSSIONS ON BIO-MUSIC

1) The BIO-MUSIC Platform: The Bio-Music platform (Fig. 13) is being prototyped using the Max/MSP graphical dataflow programming environment (similar to LabView) which allows for rapid development cycles and the possibility of making stand-alone applications directly. The functions of this prototype system are outlined below: Data Acquisition, Data Preprocessing, Intermediate Representation, Visual Mapping, Sonic Mapping, Visualisation and Sonification.

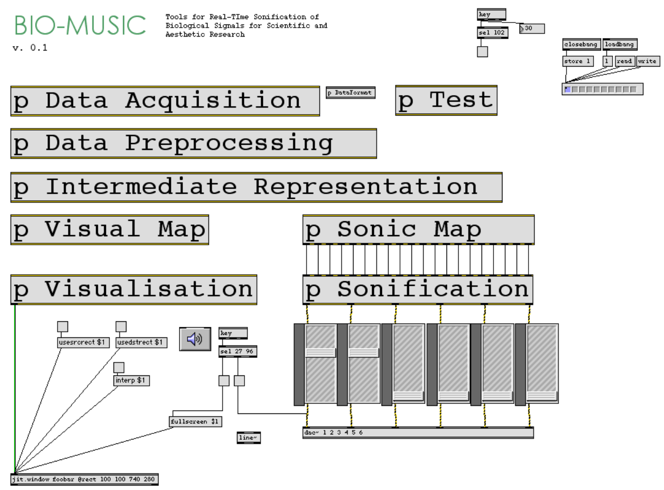

Fig. 13. Bio-music

2) Data Acquisition: The Data Acquisition module is charged with interfacing to various hardware which can be sources of live bio-data. In the current implementation, this involves receiving data over ethernet and UDP/IP packets which are formatted in the OpenSoundControl specification. For the meantime, this is considered to be an acceptable abstraction of the hardware-¿software interface. In future, a facility to load executable code libraries which could generically interface to any given hardware as desired. It is envisioned that this functionality might be provided by the OpenInterface software project at a future date. It is also possible in the Max/MSP environment to load ”external” code libraries which encapsulate the functionality of executable code within the Max environment. By this method, a different “external” would be needed for every particular hardware interface. Some simplification might be gained by the specification of only USB external devices and further adoption of any Data Acquisition over USB standards which are extant or pending. That said, most major manufacturers of Data Acquisition hardware provide standard libraries and SDKs to aid in such development. In any case, the current UDP/OSC network model for data acquisition will remain.

3) Data Preprocessing: The DAQ module passes the raw data to the preprocessing (signal conditioning) module which is responsible for normalising the ranges and characteristics of the various acquired data. Thus almost methematical or algorithmical transform can be applied to live data to change its range or behaviours. For example, if you wish, all incoming information can be converted to floating point numbers ranging from -1.0 to +1.0, which could be done either via manual adjustment or by an auto-adaptive process which would watch the incoming values and make the adjustments automatically.

4) Mapping: Once the incoming signals have been normalised as desired, the task is to map, or, to use a more appropriate term, project, those signals onto the sonification module in such a way that they produce a meaningful and useful result. This is the area where it is imagined researchers, scientists and even composers and musicians will spend most of their time working. Much of the functionality of the other modules is to standardise and simplify interfaces with data sources and sinks. To aid in the flexible and rapid ability of users to make fine-grained adjustments in this module, a facility is provided for MIDI continuous controller parameters from any kind of MIDI control surface to be used to adjust the mapping parameters of this module. In this case, the software was prototyped using the Behringer BCR2000, a sophisticated yet easy to use and relatively inexpensive control surface using rotary knobs to adjust continuous controllers. It should be noted that there are similar but independent modules for the mapping of data to visualisation and sonification processes respectively.

ACKNOWLEDGMENT

The authors would like to thank the eNTERFACE'06 organizing team from University of Zaghreb, Faculty of Electrical Engineering and Computing, for hosting with great success this second edition of the Similar Summer Workshop.

REFERENCES

[1] B. Arslan, A. Brouse, C. Simon, R. Lehembre, J. Castet, J.J. Filatriau, Q. Noirhomme , "A real time music synthesis environment driven with biological signals", ICASSP 2006.
[2] http://www.enterface.net/enterface05
[3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan , "Brain-computer interfaces for communication and control", Clinical Neurophysiology, vol.113, 2002, p.767-791.
[4] A. Kapur, E. Yang, A. Tindale and P. Driessen , "Wearable sensors for real-time musical signal processing", Proceedings of the Conference on New Interfaces for Musical Expression (NIME05), Vancouver, Canada, 2005.
[5] A. Brouse, Petit guide de la musique des ondes cérébrales, Horizon0, vol. 15, 2005.
[6] G. Kramer (ed.), Auditory Display : Sonification, Audification and Auditory Interfaces. Santa Fe Institute, 1994.
[7] J. Sterne, The audible past : cultural origins of sound reproduction, Duke University Press, 2003, pp. 99-136.
[8] A. Bregman, Auditory Scene Analysis : The Perceptual Organisation of Sound, MIT Press, 1990.
[9] M. Ballora, Data analysis through auditory display : applications in heart rate variability, Ph.D. Thesis, McGill University, 2000.
[10] D. Arfib, J.M. Couturier, L. Kessous and V. Verfaille , Mapping strategies between gesture control parameters and synthesis models parameters using perceptual spaces, Organised Sound 7(2), Cambridge University Press, pp. 135-152 .
[11] http://www.gnu.org/software/octave/
[12] http://www.python.org/
[13] http://pure-data.sourceforge.net/
[14] http://www.dti-be.com/
[15] MedicalStudio [Online]. Available: http://www.medicalstudio.org
[16] Max/MSP. [Online].
Available: http://www.cycling74.com/products/maxmsp.html
[17] OpenSoundControl.[Online].
Available: http://www.cnmat.berkeley.edu/OpenSoundControl/
[18] D. Osipova, A. Takashima, R. Oostenveld, G. Fernandez, E. Maris and E. Jensen, "'Theta and gamma oscillations predict encoding and retrieval of declarative memory.", J Neurosci, 2006, 26, 7523-7531
[19] C. Babiloni, F. Vecchio, A. Bultrini, G.L. Romani and P.M. Rossini, "Pre- and Poststimulus Alpha Rhythms Are Related to Conscious Visual Perception: a High-Resolution EEG Study.", Cereb Cortex, 2005
[20] G. Pfurtscheller, C. Brunner, A. Schlgl, and F.H.L. da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks.", Neuroimage, 2006, 31, 153-159
[21] Y. Wang, P. Berg and M. Scherg, "Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: a simulation study.", Clinical Neurophysiology, vol. 110, pp. 604614, 1999.
[22] Z. J. Koles and A. C. K. Soong, "EEG source localization: implementing the spatio-temporal decomposition approach.", Electroencephalogr. Clin. Neurophysiol., vol. 107, pp. 343-352, 1998.
[23] P. Berg and M. Scherg, "A fast method for forward computation of multiple-shell spherical head models.", Electroencephalography and clinical Neurophysiology, vol. 90, pp. 5864, 1994.
[24] J.C.Mosher, R.M. Leahy and P.S. Lewis, "EEG and MEG: Forward solutions for inverse methods", IEEE Transactions on Biomedical Engineering, vol.46, 1999, pp.245-259.
[25] S. Baillet, J.C. Mosher and R.M. Leahy, "Electromagnetic brain mapping", IEEE Signal processing magazine, November 2001, pp.14-30.
[26] C. Michel, M. Murray, G. Lantz, S. Gonzalez, L. Spinelli and R. Grave de Peralta, "EEG source imaging", Clinical Neurophysiology, vol.115, 2004, pp. 2195-2222.
[27] Pascual-Marqui and Roberto Domingo, "Review of methods for solving the EEG inverse problem", International Journal of Bioelectromagnetism, 1999, pp.75-86.
[28] C. d'Alessandro, N. d'Alessandro, S. Le Beux, J. Simko, F. Cetin and H. Pirker, "The speech conductor : gestural control and synthesis" In proceedings, eNTERFACE'05, Mons, Belgium
[29] http://www.tele.ucl.ac.be/ jjfil/EEGTexture
[30] W. S. Yeo and J. Berger., "Application of Image Sonification Methods to Music", in Proceedings of the 2005 International Computer Music Conference (ICMC 2005), Barcelona, Spain, September 2005.
[31] http://www.tele.ucl.ac.be/ jjfil/enhancedDidge