



**HAL**  
open science

## From biological signals to music

Burak Arslan, Andrew Brouse, Julien Castet, Jehan-Julien Filatriau, Remy Léhembre, Quentin Noirhomme, Cédric Simon

► **To cite this version:**

Burak Arslan, Andrew Brouse, Julien Castet, Jehan-Julien Filatriau, Remy Léhembre, et al.. From biological signals to music. 2nd International Conference on Enactive Interfaces, Nov 2005, Genoa, Italy. hal-00524797

**HAL Id: hal-00524797**

**<https://hal.science/hal-00524797>**

Submitted on 8 Oct 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# From Biological Signals to Music

Arslan, B.<sup>\*</sup> Brouse, A.<sup>\*\*</sup> Castet, J.<sup>†</sup> Filatriau, J.J.<sup>‡</sup> Lehembre, R.<sup>‡</sup> Noirhomme, Q.<sup>‡</sup> Simon, C.<sup>‡</sup>

(<sup>\*</sup>) TCTS, Faculté Polytechnique de Mons, Belgium (<sup>\*\*</sup>) CMR, University of Plymouth, UK

(<sup>†</sup>) INP, Grenoble, France (<sup>‡</sup>) TELE, Université catholique de Louvain, Belgium

burak.arslan@tcts.fpms.ac.be, andrew.brouse@plymouth.ac.uk, julien.castet@free.fr,

(filatriau, lehembre, noirhomme, simon)@tele.ucl.ac.be

## Abstract

*This project proposes to use the analysis of physiological signals, such as the electroencephalogram (EEG), electromyogram (EMG) electrocardiogram (ECG) and electro-oculogram (EOG), to control sound synthesis algorithms in order to build a biologically driven musical instrument. This project took place during the eNTERFACE'05 summer workshop in Mons, Belgium. Over four weeks specialists from the fields of brain-computer interfaces and sound synthesis worked together to produce biologically controlled musical instruments playable in real-time.*

## 1. Introduction

Advances in computer science and specifically in Human-Computer Interaction (HCI) have now enabled musical performance using sensor-based instruments in real-time computer synthesis systems [1]. Musicians can now use positional, cardiac, muscle and other sensor data to control sound synthesis [2, 3].

Simultaneously, advances in Brain-Computer Interface (BCI) research have proven that cerebral patterns can be used as a source of communication and control [4]. Indeed, cerebral and conventional sensors can be used together with the object of producing a 'body-music' controlled according to the musician's cognitive and proprioceptive processes. Research is already being done toward integrating BCI and sound synthesis [5, 6].

One salient approach aims to sonify the data derived from physiological processes by mapping the data directly to sound synthesis parameters [7, 8, 9]. Another approach aims to build a musical interface where inference based on complex feature extraction enables the musician to intentionally control sound production [6].

In the following, we present: a short history of biologically-controlled instruments; the architecture we designed to acquire, process and play music based on biological signals; strategies for signal acquisition; a discussion of signal processing techniques; the sound synthesis implementation and the instruments we built; and conclude with a presentation of some future directions.

## 2 History

Brainwaves are a form of bioelectricity, or electrical phenomena in animals or plants. Human brainwaves were first measured in 1924 by Hans Berger. He termed these electrical measurements the electroencephalogram (EEG), which means literally 'brain electricity writing'. Berger first published his brainwave results in 1929 as "Über das Elektrenkephalogramm des Menschen" [10]. The English translation did not appear until 1969. His results were verified by Adrian and Matthews in 1934 who also attempted to listen to the brainwave signals via an amplified speaker [11].

This was the first attempt to sonify human brainwaves for auditory display. The first instance of the intentional use of brainwaves to generate music did not occur until 1965, when Alvin Lucier [12], who had begun working with physicist Edmond Dewan, composed a piece of music using brainwaves as the sole generative source. *Music for Solo Performer* was presented, with encouragement from John Cage, at the Rose Art Museum of Brandeis University in 1965.

In the late 1960s, Richard Teitelbaum was a member of the innovative Rome-based live electronic music group *Musica Elettronica Viva* (MEV). In performances of *Spacecraft* (1967) he used various biological signals including brain (EEG) and cardiac (ECG) signals as control sources for electronic synthesizers. Over the next few years, Teitelbaum continued to use EEG and other

biological signals in his compositions and experiments as triggers for the nascent Moog electronic synthesiser.

Then, in the late 1960s, another composer, David Rosenboom, began to use EEG signals to generate music. In 1970-71 Rosenboom composed and performed *Ecology of the Skin*, in which ten live EEG performer-participants interactively generated immersive sonic/visual environments using custom-made electronic circuits. Around the same time, Rosenboom founded the Laboratory of Experimental Aesthetics at York University in Toronto, which encouraged pioneering collaborations between scientists and artists. For the better part of the 1970s, the laboratory undertook experimentation and research into the artistic possibilities of brainwaves and other biological signals in cybernetic biofeedback artistic systems. Many artists and musicians visited and worked at the facility during this time including John Cage, David Behrman, LaMonte Young, and Marian Zazeela. Some of the results of the work at this lab were published in the book “Biofeedback and the Arts” [13]. A more recent monograph by Rosenboom, “Extended Musical Interface with the Human Nervous System” [14], remains the definitive theoretical aesthetic document in this area.

In France, scientist Roger Lafosse was doing research into brainwave systems and proposed, along with musique concrète pioneer Pierre Henry, a sophisticated live performance system known as Corticalart (art from the cerebral cortex). In a series of free performances done in 1971, along with generated electronic sounds, one saw a television image of Henry in dark sunglasses with electrodes hanging from his head, projected so that the content of his brainwaves changed the colour of the image according to his brainwave patterns.

Starting in the early 1970s, Jacques Vidal, a computer science researcher at UCLA, simultaneously began working to develop the first direct brain-computer interface (BCI) system using an IBM mainframe computer and other custom data acquisition equipment. In 1973, he published “Toward Direct Brain-Computer Communication” [15] based on this work.

In 1990 Jonathan Wolpaw et al [16] at Albany developed a system to allow a user to exercise rudimentary control over a computer cursor via the alpha band of their EEG spectrum. Around the same time, Christoph Guger and Gert Pfurtscheller also began researching and developing BCI systems along similar lines in Graz, Austria [19].

In the early 1990s two scientists, Benjamin Knapp and Hugh Lusted [17], began working on a human-computer interface called the BioMuse. It permitted a human to control certain computer functions via bioelectric signals. In 1992, Atau Tanaka [1] was commis-

sioned by Knapp and Lusted to compose and perform music using the BioMuse as a controller. Tanaka continued to use the BioMuse, primarily as an EMG controller, in live performances throughout the 1990s. In 1996, Knapp and Lusted wrote an article for Scientific American about the BioMuse entitled “Controlling Computers with Neural Signals” [18].

In 2002, the principal BCI researchers in Albany and Graz published a comprehensive survey of the state of the art in BCI research, “Brain-computer interfaces for communication and control” [4]. Then, in 2004, an issue dedicated to the broad sweep of current BCI research was published in IEEE Biomedical Transactions [20].

### 3 Architecture

Our intention was to build a robust, reusable framework for biosignal capture and processing geared towards musical applications. To maintain flexibility, signal acquisition, processing and sound synthesis are performed on different physical machines linked via ethernet. Data are acquired via custom hardware which is linked to a host computer running a Matlab/Simulink [21] real-time blockset. Data are analysed before being sent - via OpenSoundControl [22] - to the visualisation, software sound synthesis and spatialisation nodes. The sound synthesis and spatialisation are performed using the Max/MSP [23] programming environment.

#### 3.1 Software

##### 3.1.1 Matlab and Simulink

We are using various biosignal analysis methods including the wavelet transform and spatial filters. All of the signal processing algorithms are written in Matlab [21]. Because signal acquisition from the EEG cap is done using custom C++ code, we must use a method in C++ to send the data stream to Matlab directly. We implemented our signal processing code as a Simulink [21] blockset using Level-2 M file S-functions with tuneable method parameters. This allows us to dynamically adapt to the incoming signals. Subsequently, we proceed with a real-time, adaptive analysis.

##### 3.1.2 Max/MSP

Max/MSP [23] is a software programming environment optimised for flexible real-time control of music systems. It was first developed at IRCAM by Miller Puckette as a simplified front end controller for the 4X series of mainframe music synthesis systems. It was fur-

ther developed as a commercial product by David Zicarelli [24] and others at Opcode Systems and Cycling 74 [?]. It is currently the most popular environment for programming of real-time interactive music performance systems.

There are other open-source environments which could be more interesting in the long-term especially in an academic context: Pure Data and jMax are both open-source work-alike software implementations which although not as mature as Max/MSP are nonetheless very usable. SuperCollider would be another potential open-sourced programming environment. It is also very powerful and expressive, if somewhat more arcane and difficult to program, largely due to its proprietary text-based programming paradigm.

## 3.2 Data Exchange

Data transmission between machines is implemented using UDP/IP protocol over ethernet. We chose this for best real-time performance. Reliability of UDP on an ethernet LAN is not an issue from experience. Specific musical messages were encoded using the OpenSoundControl [22] protocol which sits on top of UDP.

### 3.2.1 Open Sound Control (OSC)

OSC was conceived as a protocol for the real-time control of computer music synthesizers over modern heterogeneous networks. Its development was informed by shortcomings experienced with the established MIDI standard and the difficulties in developing a more flexible protocol for effective real-time control of expressive music synthesis.

OSC was first proposed by Matthew Wright and Adrian Freed in 1997, since which time it has become very widely implemented in software and hardware designs (although, its use is still not as widespread as MIDI). Although it can function in principle over any appropriate transport/physical layer such as WiFi, serial, USB etc., current implementations of OSC are optimised for UDP/IP transport over Fast Ethernet in a Local Area Network. For our project, we used OSC to transfer data from Matlab (running on a PC with either Linux or Windows OS) to Macintosh computers running Max/MSP.

## 4 Data Acquisition

ECG, EMG and EOG were captured on one computer with a multipurpose acquisition system and EEG was acquired on another system specialised for brainwave data capture.

## 4.1 EEG

EEG data are sampled at 64 Hz on 19 channels with a DTI cap. Data are filtered between 0.5 and 30 Hz. Electrodes are positioned following the 10-20 international system and Cz is used as reference. The subject sits in a comfortable chair and is asked to concentrate on different tasks. The recording is done in a normal working place: a noisy room with people working, talking and other ambient sounds. The environment is not free of electrical noise as there are many computers, speakers, monitors, microphones and lights nearby.

## 4.2 EMG, ECG and EOG

To record the EMG and ECG, three Biopac MP100 amplifiers were used. The amplification factor for the EMG was 5000 and the signals were filtered between 0.05-35 Hz. The microphone channel has a gain factor of 200 and DC-300 Hz bandwidth.

Another 2 channel amplifier is used to collect the EOG signals. This amplifier has gain factor of 4000 and 0.4-60Hz passband. For real time capabilities, these amplified signals are fed to a National Instruments DAQPad 6052e analog-digital converter card that uses the IEEE 1394 port.

Disposable ECG electrodes were used for both EOG and EMG recordings. The sounds were captured using the Biopac BSL contact microphone.

## 5 BioSignal Processing

We tested various parameter extraction techniques in search of those which could give us the most meaningful results.

We focused mostly on EEG signal processing as it is the richest and most complex bio-signal. The untrained musician normally has less conscious control over brain biosignals as opposed to other biosignals and therefore sophisticated signal processing was reserved for the EEG which needed more processing to produce useful results. The data acquisition program samples blocks of EMG or EOG data in 100 ms frames. Software then calculates the energy for the EOG and EMG channels, and sends this information to the related instruments. The heart sound itself is sent directly to the instruments to provide a rhythmic motif.

Two kinds of EEG analysis are done. The first one attempts to determine the user's intent based on techniques recently developed in the BCI community [4]. A second approach looks at the origin of the signal and at the activation of different brain areas. The performer has

less control over results in this case. There are more details on both of these EEG analysis methods at the end of this section.

## 5.1 Detection of Musical Intent

To detect different brain states we measured spatial distribution and temporal rhythms present.

Three main rhythms are of interest:

1. Alpha rhythm: usually between 8-12 Hz, this rhythm describes the state of awareness. If we calculate the energy of the signal using the occipital electrodes, we can evaluate the state of awareness of the musician. When he closes his eyes and relaxes the signal increases. When the eyes are open the signal is low.
2. Mu rhythm: This rhythm also ranges from 8 to 12 Hz but can vary from one person to another, sometimes between 12-16 Hz. The mu rhythm corresponds to motor tasks like moving the hands or legs, arms, etc. We use this rhythm to distinguish movements of the left or right hands.
3. Beta rhythm: Comprised of energy between 18-26 Hz. Beta is linked to motor tasks and higher cognitive functions.

The wavelet transform [25] is a technique of time-frequency analysis perfectly suited for task detection. Individual tasks can be detected by looking at specific frequency bands on specific electrodes.

This operation, implemented using sub-band filters, provides us with a filter bank tuned to the frequency ranges of interest. We tested our algorithm on two subjects with different kinds of wavelets: Meyer wavelet, 9-7 filters, bi-orthogonal spline wavelet, Symlet 8 and Daubechey 6 wavelets. We finally chose the symlet 8 which gave better overall results.

At the beginning we focused on eye blink detection and  $\alpha$  band power detection because both are easily controllable by the musician. We then wanted to try more complex tasks such as those used in the BCI community. These are movements and imaginations of movements, such as hand, foot or tongue movements, 3D spatial imagination or mathematical calculation. The main problem is that each BCI user must be trained to improve his control over the task signal. Therefore we decided to use only right and left hand movements first and not the more complex tasks which would have been harder to detect. Two other techniques were also used: Asymmetry Ratio and Spatial Decomposition.

### 5.1.1 Eye blinking and $\alpha$ band

Eye blinking is detected on Fp1 and Fp2 electrodes in the 1-8Hz frequency range by looking at increase of the band power. We process the signals from electrodes O1 and O2 -occipital electrodes- to extract the power of the alpha band.

### 5.1.2 Asymmetry Ratio

Consider we want to distinguish left from right hand movements. It is known that motor tasks activate the cortex area. Since the brain is divided in two hemispheres that control the two sides of the body it is possible to recognise when a person moves on the left or right side. Let C3 and C4 be the two electrodes positioned on the cortex, the asymmetry ratio can be written as:

$$\Gamma_{FB} = \frac{P_{C3,FB} - P_{C4,FB}}{P_{C3,FB} + P_{C4,FB}} \quad (1)$$

where  $P_{Cx,FB}$  is the power in a specified frequency band (FB), i.e. the mu frequency band. This ratio has values between 1 and -1. Thus it is positive when the power in the left hemisphere (right hand movements) is higher than the one in the right hemisphere (left hand movements) and vice-versa.

The asymmetry ratio gives good results but is not very flexible and cannot be used to distinguish more than two tasks. This is why it is necessary to search for more sophisticated methods which can process more than just two electrodes simultaneously.

### 5.1.3 Spatial Decomposition

Two spatial methods have proven to be accurate: The Common Spatial Patterns (CSP) and the Common Spatial Subspace Decomposition (CSSD) [26, ?]. We will shortly describe here the second one (CSSD): This method is based on the decomposition of the covariance matrix grouping two or more different tasks. It is important to highlight the fact that this method needs a learning phase where the user executes two tasks.

The first step is to compute the autocovariance matrix for each task. Given one signal  $X$  of dimension  $N \times T$  for  $N$  electrodes and  $T$  samples, we decompose  $X$  in  $X_A$  and  $X_B$ ,  $A$  and  $B$ . By using two different tasks, we can obtain the autocovariance matrix for each task:

$$R_A = X_A X_A^T \quad \text{and} \quad R_B = X_B X_B^T \quad (2)$$

We now extract the eigenvectors and eigenvalues from the  $R$  matrix that is the sum of  $R_A$  and  $R_B$ :

$$R = R_A + R_B = U_0 \lambda U_0^T \quad (3)$$

We can now calculate the spatial factors matrix  $W$  and the whitening matrix  $P$ :

$$P = \lambda^{-1/2} U_0^T \quad \text{and} \quad W = U_0 \lambda^{1/2} \quad (4)$$

If  $S_A = P R_A P^T$  and  $S_B = P R_B P^T$ , these matrices can be factorised:

$$S_A = U_A \Sigma_A U_A^T \quad S_B = U_B \Sigma_B U_B^T \quad (5)$$

Matrix  $U_A$  et  $U_B$  are equals and the sum of their eigenvalue is equal to 1,  $\Sigma_A + \Sigma_B = I$ .  $\Sigma_A$  et  $\Sigma_B$  can be written thus:

$$\Sigma_A = \text{diag} \left[ \underbrace{1 \dots 1}_{m_a} \quad \underbrace{\sigma_1 \dots \sigma_{m_c}}_{m_c} \quad \underbrace{0 \dots 0}_{m_b} \right] \quad (6)$$

$$\Sigma_B = \text{diag} \left[ \underbrace{0 \dots 0}_{m_a} \quad \underbrace{\delta_1 \dots \delta_{m_c}}_{m_c} \quad \underbrace{1 \dots 1}_{m_b} \right] \quad (7)$$

Taking the first  $m_a$  eigenvector from  $U$ , we obtain  $U_a$  and we can now compute the spatial filters and the spatial factors:

$$S P_a = W U_a \quad (8)$$

$$S F_a = U_a^T P \quad (9)$$

We proceed identically for the second task, taking care this time with the last  $m_b$  eigenvectors. Specific signal components of each task can then be extracted easily by multiplying the signal with the corresponding spatial filters and factors. For the task A it gives:

$$\hat{X}_a = S P_a S F_a X \quad (10)$$

A support vector machine (SVM) with a radial basis function was used as a classifier.

### 5.1.4 Results

The detection of eye blinking during off-line and real-time analysis was higher than 95%, with a 0.5s time window. For hand movement classification with spatial decomposition, we chose to use a 2s time window. A smaller window significantly decreases the classification accuracy. The CSSD algorithm needs more training data to achieve a good classification rate so we decided to use 200 samples of both right hand and left hand movements, each sample being a 2s time window. Thus, we used an off-line session to train the algorithm. However each time we used the EEG cap for a new session, the electrode locations on the subject's head changed. Performing a training session one time and a test session another time gave poor results so we decided to develop new code in order to do both training and testing in one

session. This had to be done quite quickly to ensure the user's comfort.

We achieved an average of 90% good classifications during off-line analysis, and 75% good classifications during real-time recording. Real-time recording accuracy was a bit less than expected. (This was probably due to a less-than-ideal environment - with electrical and other noise - which is not conducive to accurate EEG signal capture and analysis.) The asymmetry ratio gave somewhat poorer results.

## 5.2 Spatial Filters

EEG is a measure of electrical activities of the brain as measured from the external skull area. Different brain processes can activate different areas. Discovering which areas are active is difficult as many different source configurations can lead to the same EEG recording. Noise in the data further complicates this problem.

In the following, we present the methods - based on forward and inverse problems - and the hypothesis we propose to solve the problem in real time.

### 5.2.1 Forward Problem, head model and solution space

If  $X$  is a  $N \times 1$  vector containing the recorded potential with  $N$  representing the number of electrodes.  $S$  is an  $M \times 1$  vector of the true source current with  $M$  the unknown number of sources.  $G$  is the leadfield matrix which links the source location and orientation to the electrodes location.  $G$  depends of the head model.  $n$  is the noise. We can write

$$X = G S + n \quad (11)$$

$X$  and  $S$  can be extended to more than one dimension to take time into account.  $S$  can either represent few dipoles (dipole model) with  $M \leq N$  or represent the full head (image model - one dipole per voxel) with  $M \gg N$ . In the following we will use the latter model.

The forward problem is to try and find the potentials  $X$  on the scalp surface knowing the active brain sources  $S$ . This approach is far simpler than the inverse approach and its solution is the basis of all Inverse problem solutions.

The leadfield  $G$  is based on the Maxwell equations. A finite element model based on the true subject head can be use as lead field but we prefer to use a 4-spheres approximation of the head. It is not subject dependent and less computationally expensive. A simple method consists of seeing the multi-shell model as a composition of single-shells -much as Fourier uses functions as sums

of sinusoid [27]. The potential  $v$  measured at electrode position  $r$  from a dipole  $q$  in position  $r_q$  is

$$v(r, r_q, q) \approx v^1(r, \mu_1 r_q, \lambda_1 q) + v^1(r, \mu_2 r_q, \lambda_2 q) + v^1(r, \mu_3 r_q, \lambda_3 q) \quad (12)$$

$\lambda_i$  and  $\mu_i$  are called Berg's parameters [27]. They have been empirically computed to approximate three and four-shell head model solution.

When we are looking for the location and orientation of the source, a better approach consists of separating the non-linear search for the location and the linear one for the orientation. The EEG scalar potential can then be seen as a product  $v(r) = k^t(r, r_q)q$  with  $k(r, r_q)$  a  $3 \times 1$  vector. Therefore each single shell potential can be computed as [28]

$$v^1(r) = ((c_1 - c_2(r \cdot r_q))r_q + c_2 \|r_q\|^2 r) \cdot q$$

with

$$c_1 \equiv \frac{1}{4\pi\sigma \|r_q\|^2} \left( 2 \frac{d \cdot r_q}{\|d\|^3} + \frac{1}{\|d\|} - \frac{1}{\|r\|} \right) \quad (13)$$

$$c_2 \equiv \frac{1}{4\pi\sigma \|r_q\|^2} \left( \frac{2}{\|d\|^3} + \frac{\|d\| + \|r\|}{\|r\| F(r, r_q)} \right) \quad (14)$$

$$F(r, r_q) = \|d\|(\|r\|\|d\| + \|r\|^2 - (r_q \cdot r)) \quad (15)$$

The brain source space is limited to 361 dipoles located on an half-sphere just below the cortex in a perpendicular orientation to the cortex. This is done because the activity we are looking at is concentrated on the cortex, the activity recorded by the EEG is mainly cortical activity and the limitation of the source space considerably reduces the computation time.

## 5.2.2 Inverse Problem

The inverse problem can be formulated as a Bayesian inference problem [29]

$$p(S|X) = \frac{p(X|S)p(S)}{p(X)} \quad (16)$$

where  $p(x)$  stands for probability distribution of  $x$ . We thus look for the sources with the maximum probability. Since  $p(X)$  is independent of  $S$  it can be considered as an normalizing constant and can be omitted.  $p(S)$  is the prior probability distribution of  $S$  and represents the prior knowledge we have about the data. This is modified by the data through the posterior probability distribution  $p(X|S)$ . This probability is linked to the noise. We assume the noise is gaussian, with zero mean and covariance matrix  $C_n$

$$\ln p(X|S) = (X - GS)^t C_n^{-1} (X - GS) \quad (17)$$

where  $^t$  stands for transpose. If the noise is white, we can rewrite equation (17) as

$$\ln p(X|S) = \|X - GS\|^2 \quad (18)$$

In case of zero mean gaussian prior  $p(S)$  with variance  $C_S$ , the problem becomes

$$\begin{aligned} & \operatorname{argmax}(\ln p(S|X)) \\ &= \operatorname{argmax}(\ln p(X|S) + \ln p(S)) \\ &= \operatorname{argmax}((X - GS)^t C_n^{-1} (X - GS) + \lambda S^t C_S S) \end{aligned}$$

where the parameter  $\lambda$  gives the influence of the prior information. And the solution is

$$\hat{S} = G^t C_n^{-1} (G^t C_n^{-1} G + \lambda C_S^{-1})^{-1} X \quad (19)$$

For a full review of a method to solve the Inverse Problem see [29, ?, 30].

Methods based on different priors were tested. Priors ranged from the simplest - no prior information - to classical prior such as the laplacian and to a specific covariance matrix. The well-know LORETA approach [30] showed the best results on our test set. The LORETA [30] looks for a maximally smooth solution. Therefore a laplacian is used as a prior. In (19)  $C_s$  is a laplacian on the solution space and  $C_n$  is the identity matrix.

To enable real time computation, leadfield and prior matrices in (19) are pre-computed. Then we only multiply the pre-computed matrix with the acquired signal. Computation time is less than 0.01s on a typical personal computer.

## 5.2.3 Results and Application

In the present case of a BCMI, the result can be used for three potential applications: the visualisation process, a pre-filtering step and a processing step.

The current of the 361 dipoles derived using the inverse method is directly used in the visualisation process. The current on every point of the half-sphere is interpolated from the dipole currents. The result is projected on a screen.

# 6 Sound Synthesis

## 6.1 Introduction

At the end of the workshop, a performance of music was presented with two bio-musicians and various equipment and technicians on stage orchestrating a live bio-music performance before a large audience. The first instrument was a midi instrument based on additive

synthesis and controlled by the musician's electroencephalogram along with an infrared sensor. The second instrument, driven by electromyograms of the second bio-musician, processed recorded accordion samples using granulation and filtering effects. Furthermore, biological signals managed the spatialized diffusion over eight loudspeakers of the sound produced by two musicians. We here present details of each of these instruments.

### 6.1.1 Sound Synthesis

Artificial synthesis of sound is the creation, using electronic and/or computational means, of complex waveforms, which, when passed through a sound reproduction system can either mimic a real musical instrument or represent the virtual projection of an imagined musical instrument. This technique was first developed using digital computers in the late 1950s and early 1960s by Max Matthews at Bell Labs. It does have antecedents, however, in the musique concrète experiments of Pierre Schaeffer and Pierre Henry and in the TelHarmonium of Thaddeus Cahill amongst others. The theory and techniques of sound synthesis are now widely developed and are treated in depth in many well-known sources.

The chosen software environment, Max/MSP, makes available a wide palette of sound synthesis techniques including: additive, subtractive, frequency modulation, granular etc. With the addition of 3rd party code libraries (externals) Max/MSP can also be used for more sophisticated techniques such as physical modelling synthesis.

### 6.1.2 Mapping

The commonly used term mapping refers, in the instance of virtual musical instruments, to mathematical transformations which are applied to real-time data received from controllers or sensors so that they may be used as effective control for sound synthesis parameters. This mapping can consist of a number of different mathematical and statistical techniques. To effectively implement a mapping strategy, one must well understand both the ranges and behaviours of the controllers or sensors and the nature of the data stream produced along with the synthesis parameters which are to be controlled.

A useful way of thinking about mapping is to consider its origin in the art of making cartographic maps of the natural world. Mapping thus is forming a flat, virtual representation of a curved, spherical real world which enables that real world to be effectively navigated. Implicit in this is the process of transformation or projection which is necessary to form the virtual representation.

Thus, to effectively perform a musically satisfying mapping, we must understand well the nature of our data sources (sensors and controllers) and the nature of the sounds and music we want to produce (including intrinsic properties and techniques of sound synthesis, sampling, filtering and DSP)

This poses significant problems in the case of biologically controlled instruments in that it is not possible to have an unambiguous interpretation of the meanings of biological signals whether direct or derived. There is some current research in cognitive neuroscience which may indicate directions for understanding and interpreting the musical significance of encephalographic signals, but this is just beginning.

A simple example of a mapping is the alpha rhythm spectral energy to musical intensity. It is well known that strong energy in the frequency band (8-12 Hz) indicates a state of unfocused relaxation without visual attention in the subject. This has commonly been used as a primary controller in EEG-based musical instruments - such as in Alvin Lucier's "Music for Solo Performer" - where strong Alpha EEG directly translate into increased sound intensity and temporal density. If this is not the desired effect then consideration has to be given to how to transform the given data to produce the desired sound or music.

## 6.2 Instrument 1 : an interface between brain and sound

For this instrument, we used the following controls

- right or left body movement (Mu bandwidth)
- eyes open or closed (Alpha bandwidth)
- average brain activity (Alpha bandwidth)

This MAX/MSP patch is based upon these parameters. The sound synthesis is done with a plug-in from Absynth which is software controlled via the MIDI protocol. This patch creates MIDI events which control this synthesis. This synthesis is in particular composed of three oscillators, three Low Frequency Oscillators, and three notch filters. There are two kinds of note trigger:

- a cycle of seven notes
- a trigger of single note

Pitch is not controlled continuously.

Regarding the first kind of note trigger, the cycle of notes begins when the artist opens his eyes for the first



time. Right or left body movements can control the direction of cycle rotation and the panning of the result. The resultant succession of notes is subjected of two randomised variations of the note durations and the delta time between each note.

In the second note trigger, alpha bandwidth is converted to a number between 0 and 3, which is then divided into three parts:

- 0 to 1 : this part is divided into five sections, one note is attributed to each section and the time properties are given by the dynamics of the alpha variations
- 1 to 2 : represents the variation of the Low Frequency Oscillator (LFO) frequency
- over 2 : the sound is stopped

The EEG analysis for these controls happens over time. To have an instantaneous controller, an infrared sensor controller was added. Based on the distance between his hand and the sensor, the artist can control:

- the rotation speed of the cycle, using the right hand
- the frequency of the two other LFO, using the left hand

The performer decides the harmony before playing, which, in the case of live performance, has proved to be a good solution.

### 6.2.1 Results

The aim of this work was to create an instrument controlled by electroencephalogram signals. Musical relationships are usually linked with gestures, yet, here no physical interaction is present. Further, the possibility for complex interactions between a traditional musical instrument, like a guitar, and the performer, retains a great power. To be interesting from an artistic point of view, a musical instrument must provide a large expressive palette to the artist.

The relationship between the artist and the music acts in two directions: the musician interacts with sound production by means of his EEGs but the produced sound also interacts via a feedback influence on the mental state of the musician. Future work could turn toward the biofeedback potential for influencing sound.

## 6.3 Instrument 2 : Real-time granulation and filtering on accordion samples

In the second instrument, sound synthesis is based on the real-time granulation and filtering of recorded accordion samples. During the demonstration, the musician

starts his performance by playing and recording a few seconds of accordion which he will then process in real-time. Sound processing was controlled by means of data extracted from electromyograms (EMGs) in measuring muscle contractions in both arms of the musician.

### 6.3.1 Granulation

Granulation techniques split an original sound into very small acoustic events called grains and reproduce them in high densities of several hundred or thousand grains per second. A lot of transformations on the original sound are made possible with this technique and a large range of very strange timbres can be obtained. In our instrument, three granulation parameters were driven by the performer : the grain size, the pitch shifting, and the pitch shifting variation (that controls the random variations of pitch).

In terms of mapping, the performer selected the synthesis parameter he wanted to vary thanks to an additional midi foot controller and this parameter was then modulated according to the contraction of his arm muscles, measured as electromyograms. The contraction of left arm muscles allowed choosing either to increase or decrease the selected parameter, whereas the variation of the parameters were directly linked to right arm muscle tension.

### 6.3.2 Flanging

In addition to granulation, a flanging effect was implemented in our instrument. Flanging is created by mixing a signal with a slightly delayed copy of itself, where the length of the delay, less than 10 ms, is constantly changing. The performer had the ability to modulate several flanging parameters (depth, feedback gain) separately via his arm muscle contractions much as was done to control the granulation parameters.

### 6.3.3 Balance dry/wet sounds

During the performance, the musician had the possibility to control the intensities of dry and wet sounds with the contraction of his left and right arm respectively. This control gave the musician the ability to cross-fade original sound with the processed one by means of very expressive gestures.

### 6.3.4 Results

Very interesting sonic textures, near or far from original accordion sound, have been created by this instrument. Granulation gave the sensation of clouds of sound,

whereas, very strange sounds, reinforced by spatialisation effects on eight loudspeakers, were obtained using certain parameter configurations of the flange effect.

As with any traditional musical instrument, the first thing going forward will be to practice the instrument in order to properly learn it. These training sessions will aim to improve the mapping between sound parameters and gestures. Data gloves and EMGs measuring muscles contraction in other body parts (legs, shoulders, neck), along with new kinds of sound processing could bring interesting results.

## 6.4 Spatialization and Localization

The human perception of the location of sound sources within a given sound environment are due to a complex series of cues which have evolved according to the physical behaviour of sound in real spaces. These cues can include: intensity, including right-left balance, relative phase, early reflections and reverberation, Doppler shift, timbral shift and many other factors which are actively studied by researchers in auditory perception.

The term ‘spatialisation’ refers to the creation of a virtual sound space using electronic techniques (analogue or digital) and sound reproduction equipment (amplifiers and speakers) to either mimic the sound-spatial characteristics of some real space or present a virtual representation of an imaginary space reproduced via electronic means. The term ‘localisation’ refers to the placement of a given sound source within a given spatialised virtual sound environment using the techniques of spatialisation. Given the greatly increased real-time computational power available in today’s personal computers, it is now possible to perform complex and subtle spatialisation and localisation of sounds using multiple simultaneous channels of sound reproduction (four or more).

The implementation of a system for the localisation of individual sound sources and overall spatialisation in this project was based around an 8-channel sound reproduction system. Identical loudspeakers were placed equidistant about a centre point to form a circular pattern around the listening space. All speakers were elevated approximately to ear level.

Sounds were virtually placed within the azimuth of this 360 degree circular sound space by the use of mixing software which approximates an equal-power panning algorithm. The amplitude of each virtual sound source can be individually controlled. Artificial reverb can be added to each sound source individually in order to simulate auditory distance. Finally, each individual sound source can be placed at any azimuth and panned around the circle in any direction and at any speed.

Future implementations of this software will take into account some more subtle aspects of auditory localisation including timbral adjustments and Doppler effects.

## 6.5 Visualization

In a traditional concert setting, the visual aspect of watching the musicians play is an important part of the overall experience. With an EEG driven musical instrument, the musician must sit very still and immobile. By adding a visual dimension to this, we can enhance the spectator’s appreciation of the music.

We studied different ways of visualising the EEG and finally chose to present the signal projected on the brain cortex as explained in section 5.2. While the musician is playing, EEG data are processed once per second using the inverse solution approach and then averaged. A half sphere with the interpolation of the 361 solution is projected on the screen.

## 7 Conclusion

During this workshop, two musical instruments based on biological signals were developed. One was based on EEG and the other on EMG. We chose to make an intelligent musical instrument rather than to just sonify the data. The same biosignals were also used to spatialise the sound and visualise the biodata.

We have built an architecture for communication between data acquisition, signal processing and sound synthesis nodes. Our software is based on Matlab and Max/MSP and thus new signal processing and sound synthesis algorithms can be easily implemented.

The present paper reflects the work of seven people over four weeks. This work did not stop at the end of the workshop - it is ongoing and there is much still to be done. The signal processing and musical instrument designs can be improved. The musicians need to achieve better control of their instruments using biological signals. Mapping algorithms need to be improved and the software implementations must be made more robust. Going forward, the members of this team, together and individually, are committed to pursuing the dream of a Music which springs eternal from human biological signals.

## References

- [1] Tanaka, A., “Musical Performance Practice on Sensor-based Instruments”, *Trends in Gestural Control of Music*. (M. Battier and M. Wanderley eds.), IRCAM, 2000, pp.389-406.

- [2] Nagashima, Y., "Bio-Sensing Systems and Bio-Feedback Systems for Interactive Media Arts", *Conference on New Interfaces for Musical Expression (NIME03)*, Montreal, Canada, 2003, pp.48-53.
- [3] Knapp, R.B. and Tanaka, A., "Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing", *Conference on New Interfaces for Musical Expression (NIME02)*, Dublin, Ireland, 2002, pp.43-48.
- [4] Wolpaw, J. R. and Birbaumer, N. and McFarland, D. J. and Pfurtscheller, G. and Vaughan, T. M., "Brain-computer interfaces for communication and control", *Clinical Neurophysiology*, vol.113, 2002, p.767-791.
- [5] Brouse A., "Petit guide de la musique des ondes cérébrales", *Horizon0*, vol. 15, 2004. <http://www.horizonzero.ca/>
- [6] Miranda E. and Brouse A., "Toward Direct Brain-Computer Musical Interfaces", *Conference on New Interfaces for Musical Expression (NIME05)*, Vancouver, Canada, 2005.
- [7] Berger, J. and Lee K. and Yeo W.S., "Singing the mind listening", *Proceedings of the 2001 International Conference on Auditory Display*, Espoo, Finland, 2001.
- [8] Potard G. and Shiemer G., "Listening to the mind listening : sonification of the coherence matrix and power spectrum of eeg signals", *Proceedings of the 2004 International Conference on Auditory Display*, Sydney, Australia, 2004.
- [9] Dribus J., "The other ear : a musical sonification of EEG data", *Proceedings of the 2004 International Conference on Auditory Display*, Sidney Australia, 2004.
- [10] Berger H., "Über das elektroencephalogramm des menschen", *Arch. f.Psychiat*, vol. 87, pp.527-570, 1929.
- [11] Adrian, E. and Matthews, B. The Berger Rhythm: Potential Changes from the Occipital Lobes in Man, *Brain* 57, No. 4, 355385 (1934).
- [12] Lucier A. and Douglas S., *Chambers*, Middletown: Wesleyan University Press, 1980.
- [13] Rosenboom D. ed., *Biofeedback and the arts : results of early experiments*, Vancouver: Aesthetic Research Centre of Canada, 1976.
- [14] Rosenboom D., *Extended Musical Interface with the Human Nervous System*, Berkeley, CA: International Society for the Arts, Science and Technology, 1990.
- [15] Vidal, J., "Toward Direct Brain-Computer Communication", in L.J. Mullins, ed., *Annual Review of Biophysics and Bioengineering* (Palo Alto, CA: Annual Reviews, 1993) pp. 157180.
- [16] Wolpaw, J.R., McFarland, D.J., Neat, G. W., and Forneris, C.A., "An EEG-based brain-computer interface for cursor control," *Electroencephalogr Clin Neurophysiol*, vol. 78, pp. 252-9, 1991.
- [17] Knapp B. and Lusted H., "A Bioelectric Controller for Computer Music Applications.", *Computer Music Journal*, 14(1) pp. 42-47. 1990.
- [18] Knapp B. and Lusted H., "Controlling Computers with Neural Signals", *Scientific American*, pp. 82-87. October 1996.
- [19] Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoser, H., Schlgl, A., Obermaier, B., Pregenzer, M., "Current trends in Graz Brain-Computer Interface (BCI) research," *IEEE Trans Rehabil Eng*, vol. 8, pp. 216-9, 2000.
- [20] BCI special issue, *IEEE Transactions on Biomedical Engineering*, vol. 51, 2004.
- [21] The Mathworks. <http://www.mathworks.com/>
- [22] Wright, M. and Freed, A. *OpenSoundControl*. <http://www.cnmat.berkeley.edu/OpenSoundControl/>
- [23] Puckette, M. and Zicarelli, D. *Max/MSP* <http://www.cycling74.com/products/maxmsp.html>
- [24] Zicarelli D., "An extensible real-time signal processing environment for Max", *Proceedings of the International Computer Music Conference*, Ann Arbor, Michigan, 1998.
- [25] Mallat S., *A wavelet tour of signal processing*, Academic Press, 1998.
- [26] Wang, Y., Berg, P. and Scherg, M., Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: a simulation study, *Clinical Neurophysiology*, vol. 110, pp. 604614, 1999.
- [27] Berg, P., and Scherg, M. A fast method for forward computation of multiple-shell spherical head models, *Electroencephalography and clinical Neurophysiology*, vol. 90, pp. 5864, 1994.
- [28] Mosher, J.C., Leahy, R.M. and Lewis, P.S., "EEG and MEG: Forward solutions for inverse methods", *IEEE Transactions on Biomedical Engineering*, vol.46, 1999, pp.245-259.
- [29] Baillet, Sylvain and Mosher, John C. and Leahy, Richard M, "Electromagnetic brain mapping", *IEEE Signal processing magazine*, November 2001, pp.14-30.
- [30] Pascual-Marqui, R.D., "Review of methods for solving the EEG inverse problem", *International Journal of Bioelectromagnetism*, 1999, pp.75-86.