WAVE – An Audio Virtual Environment

Leonel Valbom[1], Adérito Marcos[2] and Computer Graphics Centre[3]

ABSTRACT

This paper outlines the basis and gives a description of the project WAVE that is starting in the Department of Information Systems of the University of Minho in co-operation with a research group in the Computer Graphics Centre – ZGDV, Guimarâes.
The project aims to set up an immersive environment of virtual reality, where the music, sound and audio (3D or not) plays an important role in a virtual musical/sound instrument for performances, education, entertainment or experimentation.
This virtual environment uses a hardware system containing two batons with wireless sensors to convert movements into MIDI messages for interaction and a computer to process graphics, MIDI and sound.
The application uses three different performance workspaces one of which includes a synthesiser, another a mixer with recording/sequencing features and the third one a 3D spatial sound display. The workspaces can be used one at a time but they are always running processes in the background all of the time. In each different workspace there is a control zone at the left hand side that can change the workspace parameters. The baton in the left hand of the user changes the levels of the parameters of the workspace and sometimes it assists the musical/sound performance made by the baton in the right hand in the workspace.
This virtual instrument seeks to allow simple gestures to generate musical phrases and to control the behaviour of sounds with a great degree of freedom, thus giving the user the power to manage the expressive channels and the built-in artificial intelligence techniques. This paper will address, therefore, the main concepts and architecture supporting the WAVE platform/system.

WAVE – Ένας Εικονικός Ηχητικός Χώρος

ΠΕΡΙΛΗΨΗ

Στο άρθρο που ακολουθεί περιγράφονται οι μεθοδολογικές πτυχές της εφαρμογής διάφορων εδραιωμένων τεχνικών στο χώρο της Εικονικής Πραγματικότητας, προκειμένου να εμπλουτιστούν και να επεκταθούν οι δυνατότητες μιας πλατφόρμας αναπαράστασης εικονικών κόσμων. Το άρθρο περιγράφει τις αντίστοιχες τεχνικές και προχωράει στην ανάλυση προβλημάτων και λεπτομερειών υλοποίησης. Τέλος, παρέχονται παραδείγματα και αποτελέσματα εφαρμογής για καθεμία από τις τεχνικές που περιγράφεται.

Το άρθρο αυτό σκιαγραφεί το υπόβαθρο και δίνει μια περιγραφή για το έργο WAVE και την αρχιτεκτονική του υλοποιημένου συστήματος, το οποίο ξεκίνησε στο Τμήμα Πληροφοριακών Συστημάτων του Πανεπιστημίου του Minho, σε συνεργασία με μια ερευνητική ομάδα από το Computer Graphics Centre – ZGDV στο Guimarâes. Σκοπός του έργου είναι να στηθεί ένα εικονικό περιβάλλον εμβύθισης, στο οποίο η μουσική και ο ήχος (3D ή όχι) παίζουν σημαντικό ρόλο στην υλοποίηση ενός εικονικού μουσικού οργάνου για παραστάσεις, εκπαίδευση, ψυχαγωγία και πειραματισμό. Το εικονικό περιβάλλον χρησιμοποιεί ένα σύστημα που περιλαμβάνει δύο ράβδους με ασύρματους αισθητήρες που μετατρέπουν κινήσεις σε σήματα MIDI για την αλληλεπίδραση του χρήστη με αυτό, καθώς και έναν υπολογιστή για την επεξεργασία των γραφικών, των σημάτων MIDI και του ήχου.
Η εφαρμογή χρησιμοποιεί τρεις διαφορετικούς τρόπους έκφρασης/αλληλεπίδρασης. Ο ένας περιλαμβάνει έναν ψηφιακό συνθέτη μουσικής, ο δεύτερος ένα μείκτη με δυνατότητες εγγραφής και καταγραφής ακολουθιών τόνων και ο τελευταίος ένα σύστημα τρισδιάστατης αναπαράστασης ήχου. Τα διαφορετικά περιβάλλοντα εργασίας μπορούν να χρησιμοποιηθούν εναλλάξ, αλλά είναι συνεχώς διαθέσιμα και ενεργά. Σε κάθε διαφορετικό τρόπο εργασίας, υπάρχει μια ζώνη ελέγχου στα αριστερά, η οποία μπορεί να αλλάζει τις παραμέτρους του περιβάλλοντος. Η ράβδος στο αριστερό χέρι του χρήστη ελέγχει τη στάθμη των παραμέτρων και σε ορισμένες περιπτώσεις βοηθά στην παραγωγή του ήχου ή της μουσικής, για τα οποία είναι υπεύθυνο το δεξί χέρι.
Αυτό το εικονικό όργανο αποσκοπεί στη δημιουργία μουσικών φράσεων από απλές χειρονομίες, καθώς και στον έλεγχο της συμπεριφοράς των ήχων με μεγάλο βαθμό ελευθερίας, δίνοντας έτσι τη δυνατότητα στο χρήστη να διαχειρίζεται τα κανάλια εκφράσεων και τους ενσωματωμένους αλγόριθμους τεχνητής νοημοσύνης του συστήματος.

1. University of Minho, Department of Information Systems, Azurém Campus, 4800–058 Guimarães, Portugal
lvalbom@ccg.pt
2. University of Minho, Department of Information Systems, Azurém Campus, 4800–058 Guimarães, Portugal
marcos@dsiuminho.pt
3. R. Teixeira de Pascoais, 596, 4800–073 Guimar_es. Portugal
Aderito.Marcos@ccg.p

Since the end of the 1980s, sound technologies have played an important role in our society, not only on the educational area and entertainment, but also in musical and sound activities, where digital audio recording and the sound synthesis (among other things) had played an excellent role not only in the new conception and synthesis of palettes of timbres and dynamic, but also in the aesthetic evolution of musical domains. Throughout history the appearance and perfection of new musical instruments has been a constant activity.

Regarding the cycle of acoustic instruments new electronic musical instruments like the Theremin, the Ondes de Martenot (1928), the Pianorad (1926) and the Givelet started to appear at the beginning of 20th century. The first two of these were the most successful in the musical community, with particular prominence for the Theremin due to its innovative technique whereby the sound was initiated with the free movement of the hands of the interpreter without any mechanical associated linkages.

By the end of the 20th century there appeared a series of new instruments, using common technologies similar to those used in Virtual Reality.

One of the instruments (or interface) featured by its contribution in the difficult attempt to exceed the levels of artistic expression demanded by musical interpreters was the Digital Baton, which was conceived during the 1990 decade in the MIT Media Laboratory.

The Digital Baton explores optical tracking in real time with an infra-red ray sensor in the exterior of a digital baton, five *Piezo-resistive* ribbons and three *orthogonal accelerometers* to send information of its the position, acceleration and orientation. Some musical/sound instruments and projects use different techniques, as for example, *scanning leisure to rangefinder* in order to measure the position of objects in a plane. Others use an optical system to cover, the *tracking* of the human movement.

Along with the creation of new musical instruments and differently used techniques of tracking, not to mention the evolution of graphical computation, other technological developments appeared to exert an influence (direct or indirect) on the habits of people and the construction of the Information Society in which Virtual Reality has started to take a prominent place. These developments are:

- digital audio recording
- the sound synthesis of new timbres
- the widespread use of technology
- the widespread consumption of information of all kinds distributed through networks or digital supports
- the accessibility to the tools of musical creation and experimentation

- the possibility of the incorporation of interaction into the systems
- the possibility of using new musical languages and music patterns

In the last decades applications for the manipulation of music and sound information such as the sequencer recorder, editors and creation technologies have been developed that are capable of simulating a complete recording studio in a single home computer. These digital tools have become accessible thereby allowing qualitative experiments of techniques using the facilities of professional equipment for the pursuit of improvements in the quality of services and the graphical representation of the digital sound information.

Independently of the tracking technologies that have been explored in the virtual instruments, it has become necessary to evaluate different ways of controlling interpreters, of opening up new expressive musical channels and also of incorporating some degree of intelligence in the computer graphic interface to allow that simple musical gestures can originate and mould complex musical phrases. Simultaneously, it is important to complement these new instruments with a good visual feedback, including multiple visualisations of musical and sound information in order to allow the efficient management/interpretation of the parameters of the expressive channels while at the same time incorporating intelligent automatisms in the system.

This virtual audio environment intends to offer to different kinds of users the possibility of interacting directly in the composition, improvisation, musical exploration or even simple entertainment in order to:

- incorporate themes such as sound, music and interpretation in a virtual audio environment;
- to develop a system that can be used for entertainment and artistic education while, at the same time, being a serious tool for composition, interpretation and experimentation;
- to implement a visual, musical virtual environment capable of awakening artistic/audio skills for users with different musical abilities, different ages and different musical preferences;
- to develop a high degree of interaction with the possibility of being used by several users simultaneously.

This virtual environment, that can go from entertainment, to education or even to *performance*, will be able to operate with different types and styles of music, in three-dimensional ambiences, in stereo, in mono, or interact simultaneously with real time sound synthesis conserving whenever possible, sound quality and creative freedom.

Description

The design of the virtual environment respects the sequence of tasks set out by Rovan (Rovan 2000) in a study of open-air controllers in order to allow a high degree of interactivity

1. *Think ahead (initial intention)*
2. *Initiate movement (performance gesture)*
3. *Gauge performance gesture through vision and proprioception*
4. *Hear resulting sound*
5. *Adjust gesture from vision and proprioception*
6. *Adjust intention from hearing*
7. *Repeat...*

The movements of the WAVE users will be captured by the tracking systems, interpreted and processed digitally triggering graphical/sound representations with sound feedback that will condition the users in their making of musical decisions in real time.

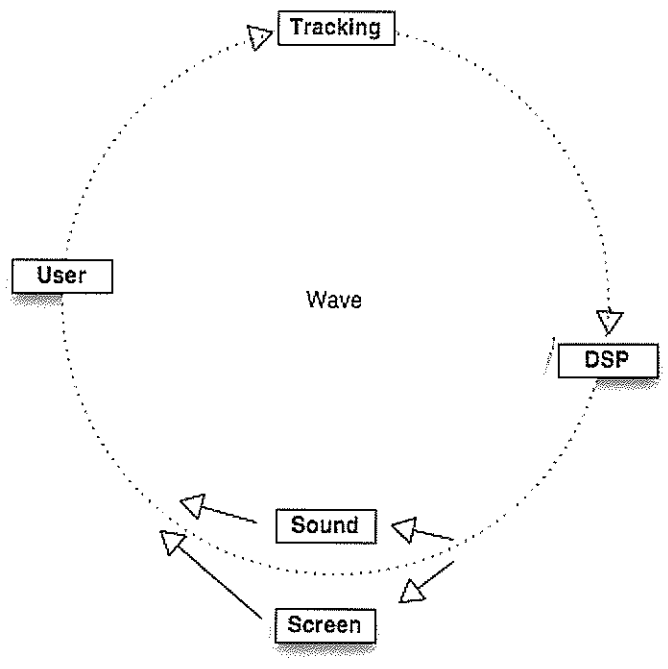

Fig. 1: Interaction in the System

As this system aims to be used by different types of users, it was necessary to endow it with some more advanced musical functions that enable the virtual environment to be used by musical performers. This instrument may be seen as a little demanding in respect of the control exerted by the user on the system and the level of rigour and precision of movements, however, any instrument is complex and needs great practice and devotion to reach a high level of control in performance. The conception of the system took into consideration the principles enunciated by Hunt and Kirk in that the systems of control in real time must have as attributes:

- *No fixed ordering to the human-computer dialogue.*
- *No single permitted set of options (e.g. choices from a menu) but rather a series of continuous controls.*
- *An instant response to the movements of the user.*
- *A physical and multi-parametric device as control mechanism which must be learned by the user until the actions become automatic.*
- *Further practice develops increased control intimacy and thus competence of operation.*
- *The human operator, once familiar with the system, is free to per form other cognitive activities whilst operating the system.*

In the WAVE system the perception of immersion is related to a large screen that must be seen with stereoscopic 3D glasses.

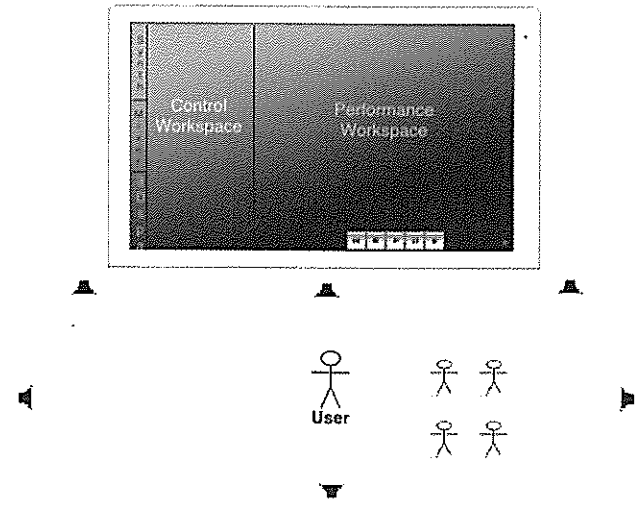

Fig. 2: System Layout

The main *screen* is composed of three different parts. The zone on the left side, which is also the narrowest, allows the choice of the type of interaction or available module in the second (B) of the three vertical zones of the screen. Selecting a different tab, it is possible to choose to interact with a synthesizer (synth), a sound mixer (mixer) or a 3D audio simulator (3D Spac).

The third zone (C) of the screen, which has about two thirds of the total surface of visualisation, has a direct relationship with the movements of the *baton* in the right hand of the user and with the second system of tracking. This zone has also a graphical representation of transport controls that are used in common sound reproduction systems (⏮ ■ ▶ ⏸).

In this zone, the user can also play sounds, which trigger multiple visualisations of scales, frequencies or musical patterns.

The B zone contains the real time control parameters of the synthesizer, mixer or spacialiser.

The synthesizer looks like a front panel of the analogue synthesizers with oscillators, resonances, filters, and some automatic effects such as different kinds of arpeggios, harmonies and scales. The performer is able to control sound synthesis parameters that are not normally controlled in its full potential. This type of control possibilities can be used in real time, creating new sounds that change in and within time and eventually other new sounds not yet found in a studio.

The mixer with eight tracks seems to be a common mixer but there are some differences. In the first insert it can manage sound files or MIDI files to play along with the performance.

Another feature of this mixer is that it can handle the surround position of the track.

These facilities allow it to play different sound or MIDI files at the same time during the performance of the user, while at the same time adding effects to the sound files played and to the sounds used by the user in the performance.

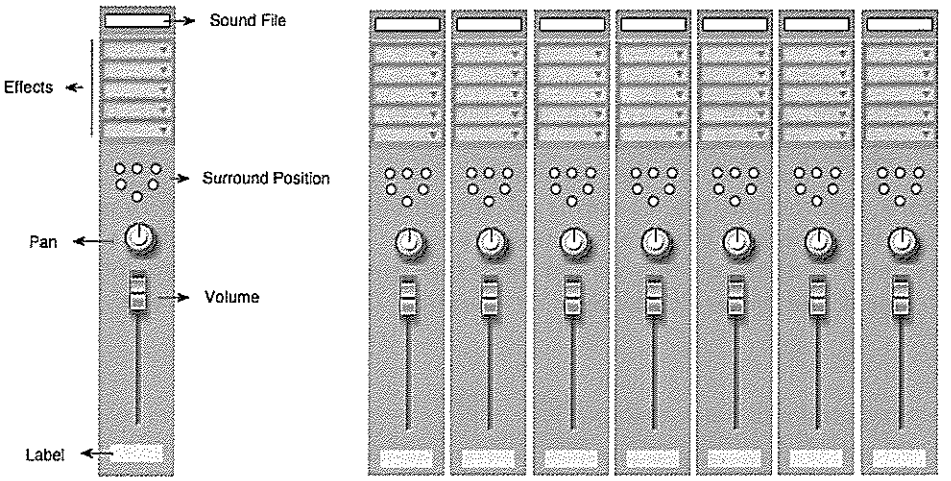

Fig. 3: Detail and Sound Mixer

The spacialiser shows in zone B of the screen the position of the sound speakers in an open space and allows their displacement in real time.

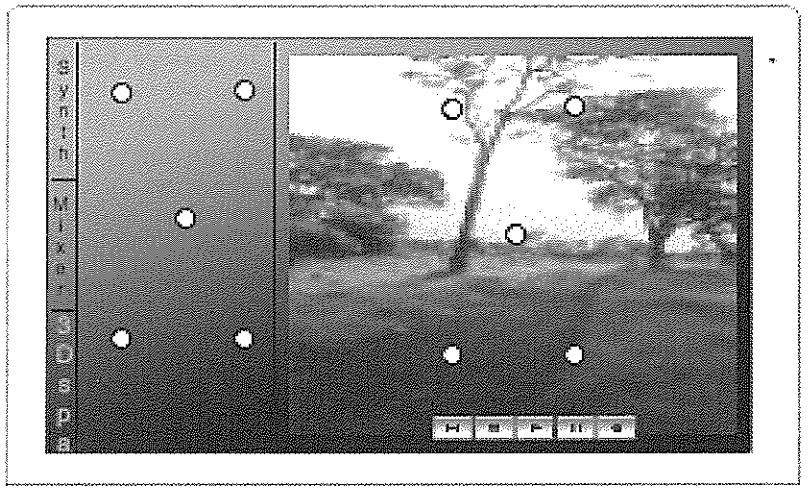

Fig. 4: 3D Spa Screen

In the zone C of screen, the module has a visualisation of a landscape or an acoustic room that can show the changes operated in real time in zone B. It is also possible to move a sound source through a landscape, while the system lets the listener hear noises of a country landscape.

The Zone C of the screen is of prime importance, because it is here that occurs the concerns of the three-dimensional interface, and it is also where the users concentrate their activity most of the time. It is in this zone of the screen that the user drives the spacialiser and triggers the notes, frequency or duration (performance screen).

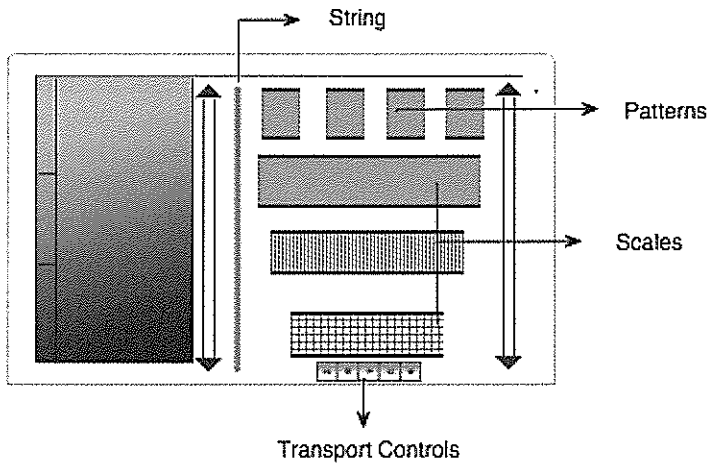

Fig. 5: Performance Screen

In the performance screen, where the user triggers the sound, there is:

- the visualisation of scales with different degrees of complexity (an initiated user will be able to use scales with five sounds, while an advanced user will be able to use scales with 12 sounds)
- one string to allow the use of portamentos
- a set of mini sequencers for the creation of small rhythmic/melodic patterns that can be triggered at any time

The tracking of the movements of the user is made through two systems of sensors; one of them is controlled by the movement of the arms of the user, as if the user was to leading an orchestra, while the second system detects the movements of the arms of the four users placed on fixed positions.

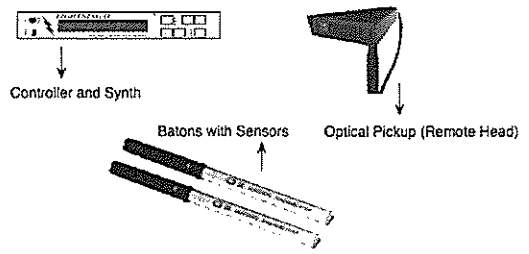

Fig. 6: Lightning II

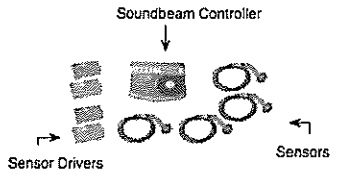

Fig. 7: SoundBeam

In the first system, while the right arm handles with the baton controls the expression and dynamics parameters, the left arm (also handling a baton) controls the control parameters or complements the activity of the right arm, so that the user feels free to concentrate on the movement of the melodic lines, chords, time, sound, phrases, volume, timbre, three-dimensional parameters or on the dynamics and the expressive contents of the musical materials.
The second tracking system uses ultra-sonic sensors and it allows four users to interact at the same time in the C part of the screen. This interaction is mainly with the pitch and/or frequency and the duration of sound.

The sound system of the virtual environment is Surround 5.1 Home Cinema to allow the spacialisation of the sound through the multiple columns. During the project, a different configuration will be tested, as a result of a proposal from the Audio Engineering Society, using a suspended speaker at the top of system.

The WAVE system is built using affordable standard hardware components without compromising the performance of the system.

The base of the system is two computers. The first one has a Linux SuSe operating system and this computer does the reception of signals, mapping and processing the signals and running the main application. It is also this computer that sends the graphics information to the video projector. The second computer, connected to the first one by Ethernet and MIDI, runs a Windows system that works with the audio and MIDI. These computers are both equipped with interface MIDI, Delta FireWire 410 and Revolution 7.1 sound cards, and the one that is connected to the video projector has a Nvidea series FX video card.

The system uses a passive projector with 100/120 MHz in conjunction with the use of wireless stereoscopic 3D glasses.

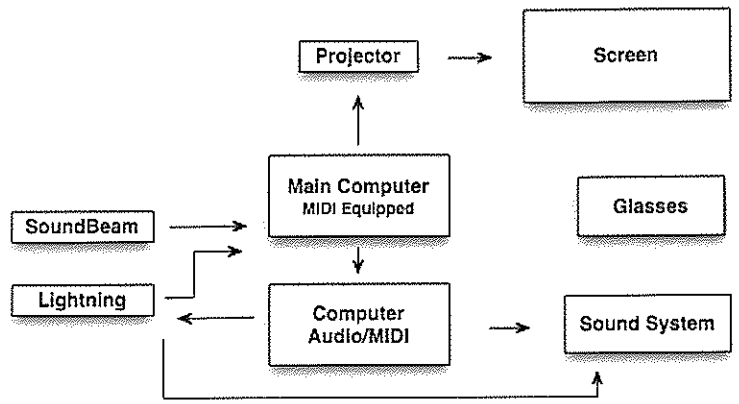

Fig. 8: Hardware connections

The screen for visualising the application is a Lusoscreen model with dimensions of 240 cm by 180 cm. This screen was chosen because it allows a good image contrast and it claims to be the only screen that has different reflection angles according to the two orthogonal axes. Its function varies as a result of the source position and incidence angle of the parasitic light, thus presenting a superior quality contrast in normal conditions.

WAVE uses a surround Home Cinema system made by Sony, which is sufficient enough to guarantee a good sound response.
The environment uses two different tracking systems; one is the Lightning II and the other is the SoundBeam.

LIGHTNING II is a MIDI controller system that senses the position and movement of handheld batons and transforms this information to MIDI. It uses the principle of optical triangulation and provides complete freedom of movement within a performance space of 3.6 metres height six metres depth. The included digital signal processor computes instantaneous velocity and acceleration and it performs detailed analyses of gestures allowing expressive control of the electronic musical instrumentation. The system has a digital synthesizer using Kurzweil Multimedia Audio Sample-playback System (MASS) technology and a digital sound effects processor that provides global delays and reverberation.

SoundBeam is a modular system with ultra-sonic sensors that convert physical movements, large or small, without physical contact and at distances of up to six metres, into digital MIDI messages. Also, SoundBeam can use switches with additional functions.

The 3D spacialiser Wave uses basic inter-aural cues combined with the digital reverb and delay of the Lightning II hardware and mixed

with the different audio signals carried to each speaker of the surround system. Wave does not use detailed frequency-dependence of normal HRTFs, because the users do not walk or move in the environment. Instead, they move their arms and hands, thus permitting that they can correct the direction of the sound source with dedicated controls within the screen.

From a general point of view there are some guidelines in the overall system like:

- The Wave starts with some defaults, allowing the user to play music with a pre-configured style without changes of parameters or with out any previous recordings.
- All the equivalent fields have real time actualisation.
- It implements the best possible graphical resolution on the screen.
- It gives a high degree of freedom to the performer.
- It implements a graphical function with reset/bypass/panic functions.

Of course, there are some limitations in this kind of systems. The body or the sensor positions are measured/sensed without reference to any physical contact. This fact causes various design problems that researchers are trying to solve with the use of an intense amount of visual feedback and with the creation of small music scales for inexperienced users.

Also an attempt is being made to develop strategies that allow the users to achieve consistent control and repeatable results without requiring long training and specialist skills.

Conclusions and Future Work

Up to the beginning of the twentieth century, music was normally written on paper and performed with mechanical musical instruments. Today, in the Information Society, there are new tools for making music and new ways of performing.

This paper presents a musical instrument in a virtual audio environment, based on affordable, standard hardware components and using three musical elements: pitch, timbre and space, while trying to give the user an environment of great freedom and at the same time some entertainment or experimentation.

Some improvements must be made in the interface and in the interpretation of the movements of the users.

In the final phase of the project, WAVE will be tested as an instrument of musical performance, as a tool for composition, as an entertainment instrument during cultural and artistic events and as a tool for developing educational components in the artistic domain and in corporal expression. At the same time, an attempt will be made to adapt

the virtual environment so that it can be used in research and development institutions, cultural, artistic or technological performances, artistic education and in electro acoustic performances.

References

1. Begault D, 2000, *3-D Sound for Virtual Reality and Multimedia*, Ames Research Center California 94035.

2. Chodura H., (IGD), "Virtual Music Reproduction", in the Proceedings of *Siggraph 1999*.

3. Dobrian Ch., in *Aesthetic Considerations in the Use of "Virtual" Music Instruments*.

4. Gardner B. and M. Keith, *Frequently asked questions about KEMAR HRTF data*, last revised 1/27/97.

5. Jordà S., "Interactive Music Systems for Everyone Exploring Visual Feedback as a way for Creating More Intuitive, Efficient and Learnable Instruments", in the Proceedings of the *Stockholm Music Acoustics Conference*, August 6-9, 2003 (SMAC 03), Stockholm, Sweden.

6. Marrin T., *Possibilities for the Digital Baton as a General-Purpose Gestural Interface*, CHI 97 Electronic Publications: Late-Breaking/Short Talks 1997.

7. Mulder A., "Virtual Musical Instruments: Accessing the Sound Synthesis Universe as a Performer", in the Proceedings of the *First Brazilian Symposium on Computers and Music*, 1994.

8. Paradiso J. A., F. Sparacino, "Optical Tracking for Music and Dance Performance", presented at the *Fourth Conference on Optical 3D Measurement Techniques*, ETH, Zurich, September 1997.

9. Rovan Joseph e Hayward Vincent, "Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance", in *Trends in Gestural Control of Music*, Wanderley, M., Battier, M. (eds.), Editions IRCAM, Paris, 2000.

10. Scipio Agostino Di, "Sound is the Interface" Sketches of a Constructivistic Ecosystemic View of Interactive Signal Processing, in the Proceedings of the *Colloquium on Musical Informatics*, Firenze 8-10 May 2003, CIM 2003.

11. Wanderley M. M., "Gestural Control of Music", *International Workshop - Human Supervision and Control in Engineering and Music*, Kassel, Germany Sept. 21-24, 2001.

2nd International Workshop on ICTs, Arts and Cultural Heritage Digital Art Technologies, Applications & Policy

2nd Session – Applied Technologies