# Audience

Audio Immersion Experience by Computer Emulation

**IMMERSIVE AUDIO FOR**
**COMPLETE VIRTUAL REALITY SYSTEMS**

Most current virtual reality systems do not possess an efficient mechanism for auralization and spatial sound projection, capable of creating or recreating a 3D sound field. Simulations in virtual environments have, for years, given priority to the scenes' visual experience production, and to enhancements to their qualitative aspects, such as texture, resolution, and depth sensation.
Particularly, the CAVE-like VR environments are the ones which would most benefit from auralization techniques, i.e., of the emulation of a three-dimensional real sound field of a given audiovisual scene, be it through loudspeakers or even using headphones, so as to allow the sensation of real immersion within the virtual environment, and guarantee coupling of 3D visual perception with 3D auditory perception.

## Complete Immersive Virtual Reality

VR systems where the auditory experience is matched to the visual experience are considered as complete virtual reality ones. If the VR environment is immersive, as occurs in the **Digital Cave** where users are immersed inside a virtual scenario and have freedom to move around, the experience of complete (audio and video) virtual reality is also said immersive.

Imagine yourself moving around exotic musical instruments playing in an ancient greek arena, or the experience of creating personal orchestras and listen to them not from the auditorium, but immersed in the scene, surrounding the instruments. Virtual audiovisual experiences like this could be explored and made feasible by means of complete immersive VR, in environments such as the Digital Cave.

It is desirable that the auralization system be based on loudspeaker (multichannel) paradigm so as to provide wider mobility and freedom for the users, who wouldn't need headphones. However, the creation of three-dimensional sound fields mathematically and/or physically correct is a complex task, which has gained a lot attention from the scientific community in the last years, being currently a research field very active and promising, specially after the considerable growth of interest due to the popularization of the surround systems for cinema (movies) and home-theaters, known as 5.1.

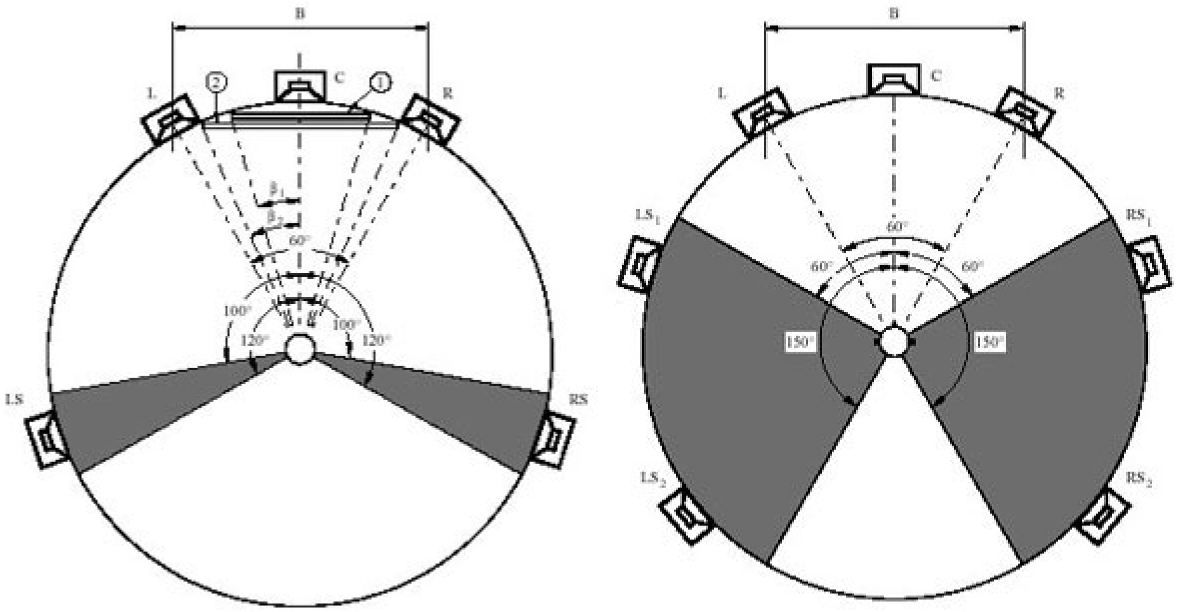


*Figure – 5.1 and 7.1 systems (standard ITU-R BS.775)*

There are several ways to develop sound spatialization solutions or auralization for virtual reality. Systems as a whole encompass from the design and development of the audio infrastructure and equipment installation, to the conception and design of applications for navigation in the 3D environment, the acoustic scene composition and control, and the choice and implementation of 3D audio coding formats, which could be also several.

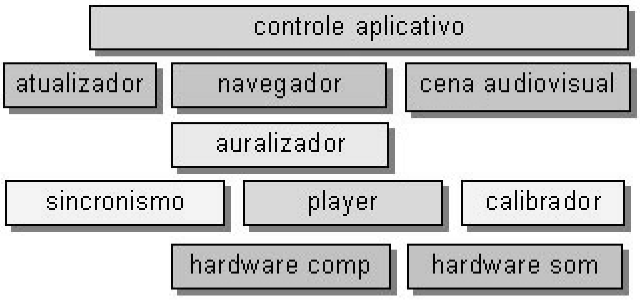


*Figure – Tasks Scenario*

## Immersive Audio Project

The on-going project for R&D of solutions for immersive audio in the Digital Cave looks for the implementation of a flexible and scalable system for 3D audio reproduction. Flexible in the sense of offering a minimum set of audio spatialization and auralization alternatives for applications having sound in several formats, from stereo/binaural up to the most sophisticated 3D multichannel audio coding under development. Scalable in the sense of allowing the associative usage of different formats, schemes or methods of sonorization, e.g. altering the spatial configuration (location) and number of loudspeakers depending on the auralization method, or employing complementary techniques, such as associating commercial reverb machines to provide selected types of envelopment, or either using different computer systems (hardware and software).

Objectives include since the implementation of decoding systems for the traditional formats and surround configurations (such as 5.1, 7.1, 10.2, DTS, THX, Dolby) up to the development and deployment of more sophisticated formats for 3D audio generation and reproduction, still restricted in the high-fidelity applications and research labs, and under development. For these formats, not only the reverb and surround envelopment are important, but also the location of sound objects in the 3D scene (perception of sound directivity) and the synthesis of the acoustic environment. In this category we include the Ambisonics and Wave-Field Synthesis formats.

The first, already known since the 70's when first coder/decoder proposals showed up, is based in an elegant mathematical formulation where both spatial and temporal sound information is coded into 4 vectors – x, y, z, and w (for 1st. order Ambisonics), as if it were projected onto cartesian axes. Initially conceived to register a real sound field recorded by soundfield microphones, this format can be adequately adapted to permit the computer simulation and creation of artificial sound fields. To reproduce the sound, the vectors are accordingly transformed in outputs for loudspeakers distributed in a known configuration around the audition area. A clear advantage of this method is in its flexibility to choose the number of loudspeakers one wants to use, and their positions around the audition area.

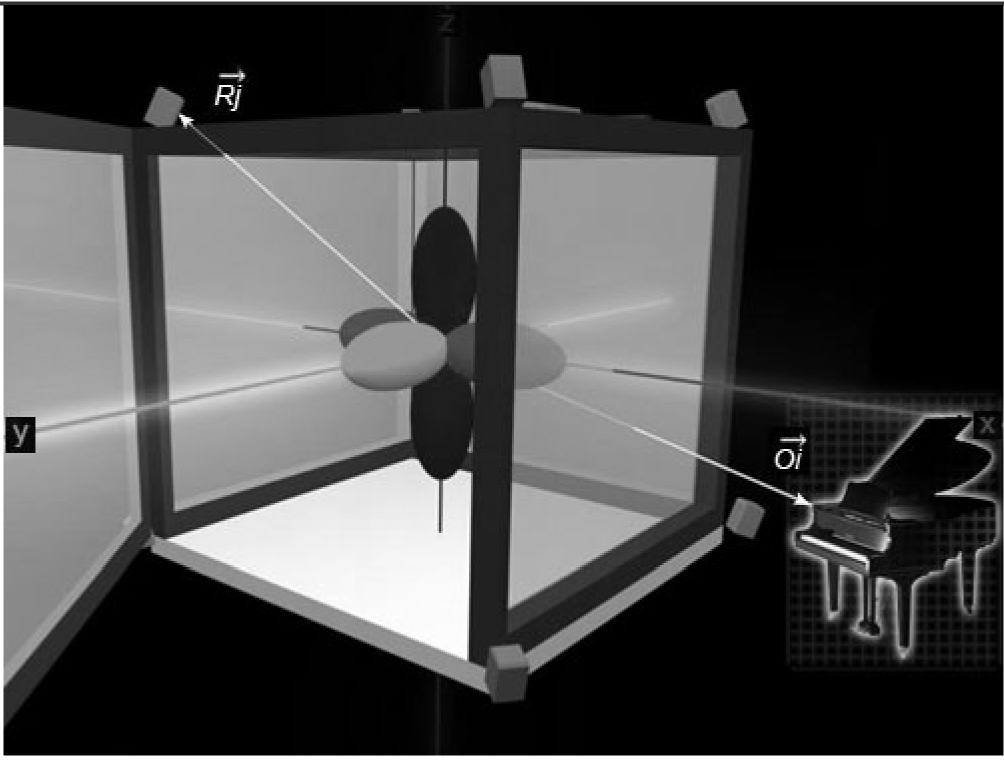

*Figure – Ambisonics scheme (both sound objects and loudspeakers have their location in the space parameterized by the system)*

The second format (Wave Field Synthesis – WFS) is more recent, and is based on a more computational complex formulation involving the physical modeling of the sound waves propagation in the environment, and takes into consideration a discretization of the physical space of the audition area. In this technique the main goal is to emulate a wave front that would be produced by real sound objects in a specific acoustic environment, through the usage of a densely distributed loudspeaker array around the audition area. An advantage of this technique is its higher tolerance to more listeners inside the audition area and its capability to induce the perception of sound fields located on a larger audition area, including the possibility of positioning objects in the middle of the area, around or among the listeners, and formed ahead of the loudspeakers. The wave front synthesis is a critical task, and may implicate in a high computational cost, which we intend to approach using cluster computing, distributing the computational tasks and calculus associated to the auralization process over the nodes of a computer cluster.

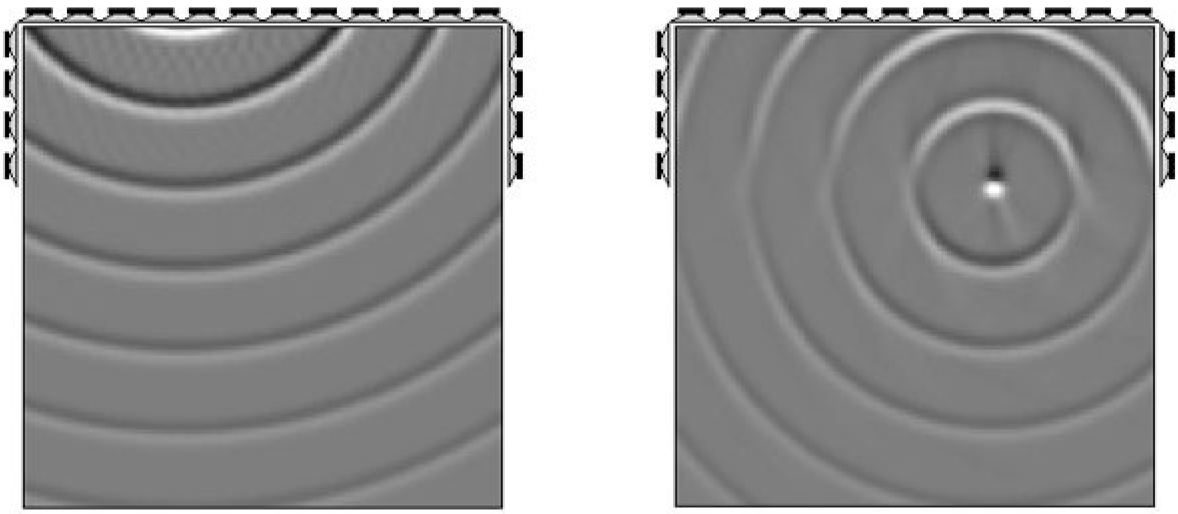

*Figure – Wave Field Synthesis (illustrating a sound object formed behind the speakers, and ahead of the speakers, inside the audition area)*

The immersive audio project in the Digital Cave considers activities addressing the development of software and also hardware (loudspeaker arrays and multichannel audio platforms) for the implementation mainly of these two 3D audio formats.

## General System Block Diagram

Some mechanisms and/or communication protocols to allow control of the sound synthesis by the VR applications will be developed and investigated, so that applications can directly sound using supported auralization techniques. One proposal considers the usage of asynchronous channels to transmit commands and/or data updates concerning sound objects in the acoustic scene directly feeding the processes that control synthesis and auralization.

Usability and integrated audiovisual navigation control are fundamental issues for complete immersive virtual reality. Synchronization, scene description resource sharing, computational load allocation and interaction model are vital and even critical to obtain the final desired effect. These control and management requirements are necessary and considered in the whole system conception. Below, it is shown the reference block diagram for the implementation of an Ambisonics auralizator integrated to a VR application.

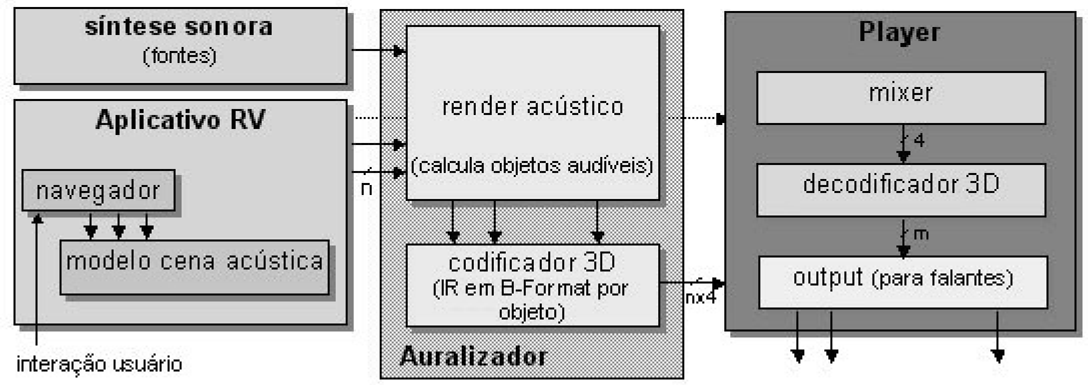


*Figure – VR Application + Ambisonics Auralizator (block diagram)*

**Blocks:**

- VR application (acoustic scene and 3D navigator setup)
- inter-process communication directives and audiovisual synchronization
- sound synthesis (wavetable and/or synthesizers)
- acoustic scene modeling (acoustical model)
- spatialization or auralization (coding in 3D format)
- decoding (output generation for loudspeakers)
- reproduction (amplification and distribution of audio "streams")

Some graphical programming languages to control synthesis and sonification processes may be used, for example MAX/MSP or PD (Pure Data, from Miller Puckette), as well as tools for scene description from MPEG-4 and Web3D (as X3D), as for example the 3D navigator **Jynx**, developed in the Digital Cave.

## Applicability

Auralization techniques find a huge number of possible applications in the areas of virtual reality, bi-directional multimedia and immersive television, allowing the enhancement of realism for the visual presentation or adding differential information that is context-dependant or relevant for the virtual environment.

Consider as an example an application design for aeronautics engineering, where designers navigate around an airplane model, interested in inspecting the air flow throughout the fuselage. The noise associated to the turbulence and different sound intensities in several points may add relevant acoustic information to the engineering/design process, bringing simulated reality to the reach of engineers.

In a insulation and internal noise control design for planes the audiovisual navigation within the virtual model will make possible to evaluate the sound quality and intensity in distinct spots around, helping to locate and assess localized problems, failures, and optimize the design tuning. A physically correct auralization of the sound field would e necessary for such activities and explorations in engineering projects.

Applications turned to enhancements or to the conception of new multimedia services are continuously demanding more realism and quality in the audio and video media presentation. For some applications, such as the development of new generations of teleconference, videophone or interactive television, there is an enormous interest in 3D models, which can reconstruct a remote audiovisual reality, which is by its turn also three-dimensional.

Systems oriented to telepresence or sophisticated multi-user immersive teleconference already consider the utilization of holography or stereoscopy techniques, not only for transmitting with incremented realism, but also addressing the production of a more fluid or continuous visual interface linking remote points. Joint stereoscopy and 3D auralization allow a fusion of the remote parts in a more uniform and gradual manner, or can transport users to a new environment, where visual and acoustic 3D properties may be shared.

New generations of systems for telepresence, where capturing multiple projections is already a reality, could also benefit of adopting a 3D audio capture, taking into consideration the sound directional properties. With fast networks (e.g. gigabit Ethernet), with multichannel multimedia compression techniques and adequate multiprojetion techniques it turns possible to investigate the limits and exercise creative capacity for conception of sophisticated 3D multimedia transmission and rendering.

A "soundfield" microphone coupled to a 3D audio coder, such as Ambisonics, may be employed to capture and transmit a 3D sound field from a place to another remote one. Sound channels may be properly coded and integrated to video/media bitstreams, opening many interaction modes, conception possibilities for new services and applications, benefiting from the fusion of audio and video 3D.

Also in the digital television (DTV) new applications addressing transmission and local rendering of 3D (virtual) acoustic environments are already feasible with recent “surround”-like systems (5.1 channels) and for some DTV standards. However, they still consist of high cost applications and less explored in respect of content production, interaction and realism enhancement.

We consider as a future goal in application design for the CAVE the prototyping of an integrated application for capture, transmission and rendering for both 3D audio and video, which could be ported to complete immersive virtual reality applications, as well as to advanced digital TV applications.

## Implementation Phases

phase I : Infrastructure (Aug/2004 to Mar/2005) : **[ Concluded ]**

- basic infrastructure (acquisition of audio server and multichannel soundcards; installation of speakers, amplifiers and cabling; loudspeaker support system; microphones and headphones)
- basic hardware/software installations (installation of audio server, gear rack, boards and patches; installation of basic software, drivers and libraries; tests)
- development software installation (DSP; C/C++ compilers; software for edition, sequencing, control and route of digital audio; patch-oriented development platforms; multichannel setups; graphical music programming; codecs and VST plugins; software and libraries for management and inter-node communication; libraries and gear for sound synthesis and acoustics; general apps; tests)
- acoustic setup (hw/sw integration; preliminary tests; acoustic measurements; patch setups and general calibration)
- acoustic treatment (acquisition/installation of additional equipment for acoustic conditioning; adaptations;)

phase II : Ambisonics I (Aug/2004 to July/2005) : **[ Concluded ]**

- research and development concerning a system for Ambisonic coding/decoding
- development of navigators and user interfaces integrated to VR tools in the Digital Cave
- integration design of software components for inter-process and inter-node communication (message/command passing, streaming, control, synchronization)
- development and/or integration of modules for description, simulation and rendering acoustic scenes (acoustic model)
- specific applications (virtual acoustic environment modeling in specific applications)
- design and first implementation of the "AUDIENCE Spatial Sound Package" for PD.

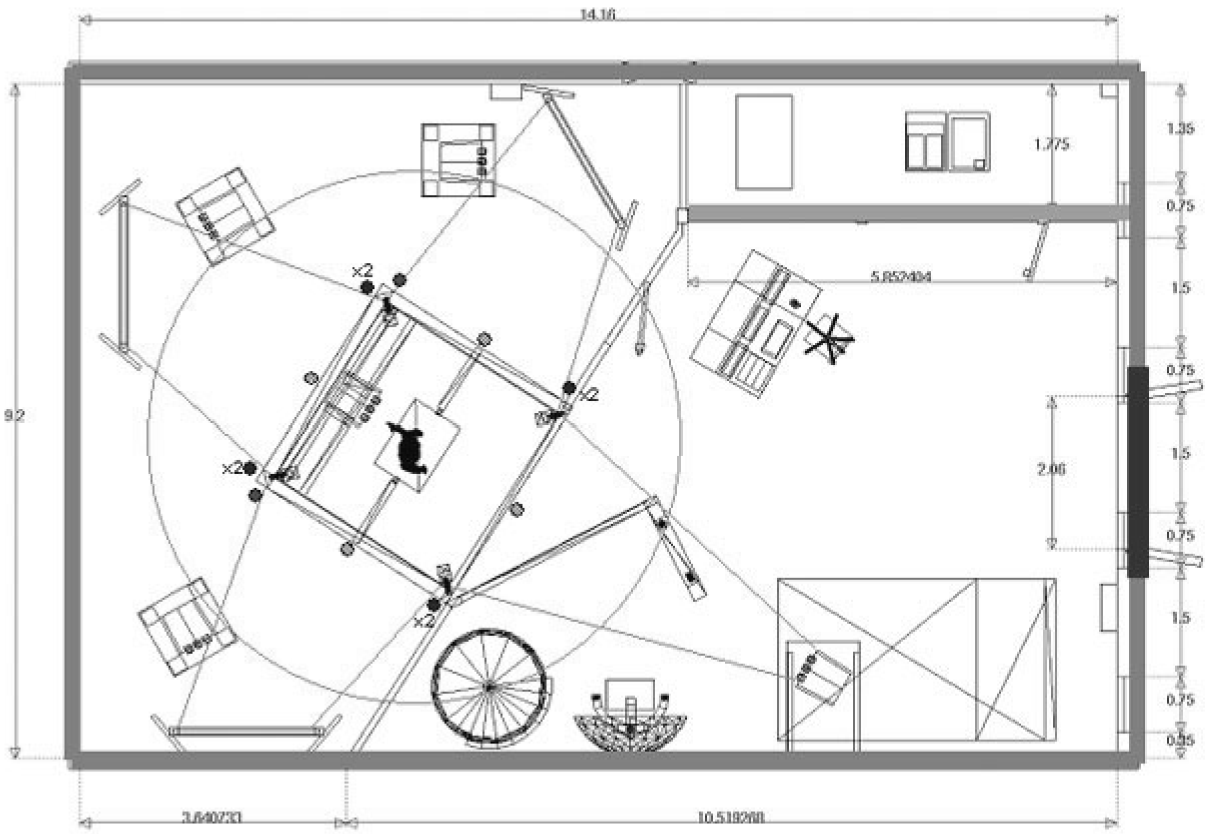


*Figure – Digital Cave and loudspeaker array for Ambisonics (preliminary design)*

phase III: Higher Sound Immersion Levels (Apr/2005 to Dec/2006)

- design and development of a system for higher sound immersion levels, exploiting higher order Ambisonics and WFS coding/decoding (partnership with other universities and/or institutions)
- release of AUDIENCE Spatial Sound Package for PD (v.1.0)

- development of loudspeaker array systems and multichannel reproduction systems (partnership with companies)
- development of a WFS interactive auralization control system (WFS auralizator) orienting to advanced applications in virtual reality and home-theater environments
- adaptation and improvement of acoustic models for higher sound immersion levels.
- development and/or adaptation of multichannel soundboards, circuitry (hardware) and software for driving dense loudspeaker arrays
- specific applications (high coupling with visualization and realistic audiovisual rendering; Voyager's holodeck is taking off!)

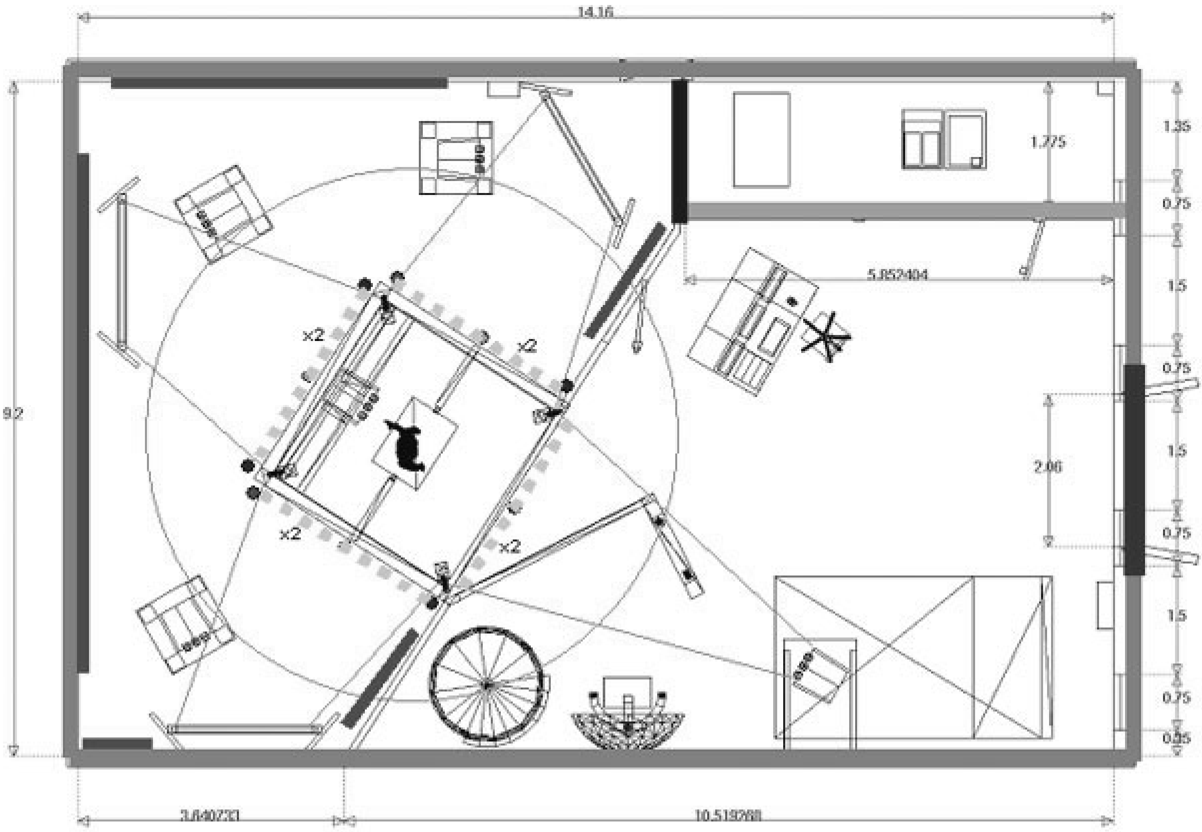


*Figure – Digital Cave and loudspeaker array for Wave Field Synthesis (preliminary design)*

phase IV : Low cost Auralizators and Commercial Systems (May/2006 to May/2007)

- activities concurrent with phase III
- acquisition and integration of commercial 3D/2D audiovisual codecs
- integration of surround decoders (partnership with companies)
- research and development of mechanisms for synchronization and control of low cost sound cards and chipsets for sound field generation applications
- proposing and prototyping low cost multichannel auralizators

## Team

**Regis Rossi A. Faria** ( PhD.EE )

Associate Researcher / R&D Manager

phone **(+55.11) 3091-5589**

email: **regis@lsi.usp.br**

**Leandro Ferrari Thomaz** ( EE, MSc. graduating )

Researcher at the Media Engineering Center and Digital Cave

phone **(+55.11) 3091-5589**

email: **lfthomaz@lsi.usp.br**

**Luiz Gustavo Berro** ( EE, undergraduate student )

Trainee

phone **(+55.11) 3091-5589**

email: **berro@lsi.usp.br**

**João A. Zuffo** ( PhD. EE )

LSI-EPUSP head

phone **(+55.11) 3091-5254**

email: **jazuffo@lsi.usp.br**

**Marcelo K. Zuffo** ( PhD. EE )

email: **mkzuffo@lsi.usp.br**

## Project Collaborators

Prof. Dr. **Sylvio Bistafa** (EP-USP, Dept. Mechanical Eng.)
Prof. Dr. **Fabio Kon** (IME-USP)
Prof. José Augusto Mannis (**CDMC-UNICAMP**)
Dr.Eng. **Márcio Avelar Gomes** (UNICAMP)

## Intitutional Partners