OpenAUDIENCE

Sistema para Produção e Reprodução de Áudio Espacial

http://openaudience.incubadora.fapesp.br
regis@lsi.usp.br

Laboratório de Sistemas Integráveis Tecnológico

Apresentação

- Projeto OpenAUDIENCE
  - Objetivos, metas, características
  - Aplicações, mercados, parceiros
  - Arquitetura, conceitos e componentes
  - Desenvolvimentos em curso
  - Como colaborar
  - Exemplos de aplicação e uso
  - Próximos passos

Projeto OpenAUDIENCE

- O que é
  - Projeto para desenvolvimento colaborativo de software para áudio espacial multicanal
- Sucede ao projeto AUDIENCE com objetivos ampliados
- Hospedado na Incubadora FAPESP
- Objetivo principal
  - Desenvolvimento de software para produção, processamento, distribuição e auralização de cenas sonoras espaciais
- Metas:
  - Distribuição Beta, restrita a equipe de desenvovimento, base para sistemas dedicados
    - Closed distributions (for specific applications, clients) in PD support or other programming languages
  - Distribuição Livre, para uso de desenvolvedores de aplicações e disseminação como plataforma de auralização
    - ruled by BSD, and will have usage limitations and may contain embedded proprietary compiled externals

OpenAUDIENCE

- Objetivos gerais
  - prover soluções em produção e reprodução de áudio espacial, comunidades abertas e indústrias (multimídia, jogos, radiodifusão, internet, música profissional, etc)
  - divulgação da arquitetura e do sistema aberto para áudio 3D em desenvolvimento,
  - ampliação da fertilização cruzada entre as comunidades técnicas, científicas e artísticas das engenharias, música e ciência da computação,
  - identificação de pessoal capacitado e interessado em linhas de P&D em engenharia de áudio e tecnologias sonoras,
  - desenvolvimento colaborativo remoto de software,
  - uso do sistema em atividades de ensino e capacitação diversas usando software livre.

OpenAUDIENCE

- Software architecture based on AUDIENCE architecture (Faria 2005)
- Developers and researchers contribute by developing functional blocks in one or more layers, and algorithms for integration, control or interfaces
- Developers have access to the distribution software while active contributors

OpenAUDIENCE

- Deliverables
  - Software distributions for general spatial sound production and playing capable to export spatial audio in several formats and to be played in several configurations
- Academic contributors
  - POLI, UNICAMP
- Industrial partners
  - Lando Hi-Fi, Sankya Eletronica, Cabos Golden, Coding Technologies (Alemanha)
- Absolutely open to new researchers and application-interested groups!

OpenAUDIENCE Partners

- **Lando High-Fidelity**
  Speakers systems
  http://www.lando.com.br
- **Sankya**
  Amplifiers solutions
  http://www.sankya.com.br
- **Cabos Golden**
  Cabling and conectors
  http://www.cabosgolden.com.br
- **Coding Technologies**
  audio codecs
  http://www.codingtechnologies.com

Características principais

- Camadas funcionais com mínimo acoplamento
- Comunicação por mensagens
- Conceito de controle por camada
- Produção admite descrição da cena
- Renderização admite simulação acústica
- Codificação espaço-temporal, admite vários formatos de transmissão
- Decodificação e reprodução admite várias configurações de saída (modos e número de canais)

Some features

- Layer-oriented: controlls processes for spatial audio production, distribution and reproduction in 4 layers
- Open to many possible standards per functional layer
- Minimum coupling with visual systems, minimal coupling between layers,
- own sync strategies and process in every layer
- Each layer responsible to deliver its functionality, independent from the others.
- Many possible combination of layers and functions per layer to be employed into the encoder and in the decoder terminal (reconfigurable and flexible building of the engine).

Some features

- Permits to build spatial auralization machines (i.e. multi-object and spatial audio encoders, authoring tools, decoders, transcoders, and players)
- Layer self-controlling embedded
- Message passing feature to send and update real-time parameters and values
- Good candidate for reconfigurable audio coding architecture
- This spatial encoding/decoding method can be used together with MPS, and transcodes into this standard.

AUDIENCE advanced capabilities

- Permit further expansion, integrating new formats and techniques as functional units
- Can govern high-level tasks of spatial audio production such as
  - specifying in bitstream the formats/techniques used in the 4 main layers of processing and which were used to encode the payload
  - specifying (optional) how decoding blocks shall be put in line and connected (auralization engine buildup)
  - specifying valid formats and signals between layers
  - specifying a decoder which is a terminal SAOC engine, capable of analyzing the bitstream, calling the necessary functional units (AUDIENCE blocks), or providing the transcoding blocks necessary to remap sound according to its current capabilities
  - specifying a reconfigurable audio coding strategy

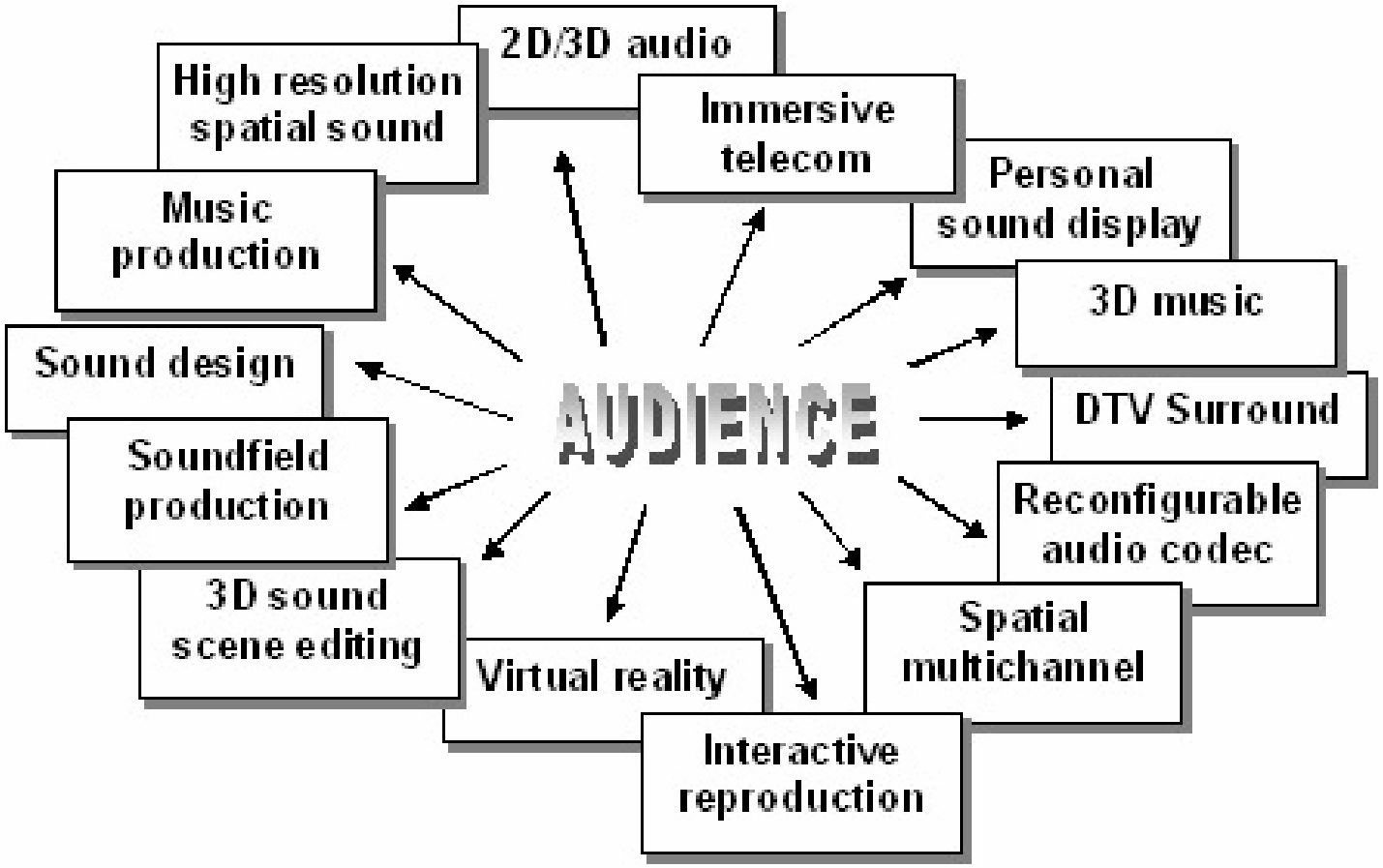
Application scenario

- Designed for spatial, multichannel and multi-object and multi-content scenarios
- Cover all described scenarios for MPEG SAOC, specially PAS (Personal Audio Services)
- Adequate for Hi-Fi distribution of spatial multichannel material (not only backward compatible downmix stereo)
- Accomodates general multichannel spatial encoding schemes, including 5.1 and Ambisonics
- Accomodates general compressions schemes including lossless

Áreas de Aplicações

- Produção sonora multicanal / espacial
- Codificação de áudio espaço-temporal, orientada a objetos
- Auralização 2D/3D, projeção de mundos virtuais sonoros e ambiências artificiais
- Espacialização sonora
- Áudio surround
- Comunicação multimodal
- Serviços de áudio personalizados

Áreas de Aplicações



Future Spatial Audio Markets

- Music/Phonographic industry
- powerful codecs, as AAC, OGG, lossless, etc.
- Real time distribution/consume, remote storage
- Spatial audio: 3D sound, MPEG Surround
- Music at the mobile
- Personalized Audio Services
- Hi-Fi Karaoke and Professional Music
- Cinema Digital 3D Authoring

Future Spatial Audio Markets

- Personalization of programs

No. of Object	3
Object Name	Voice (Announcer, Commentator), Environment (the audience)
Control	Loudness
Preset	1: Normal. 2: Mute mode (Vocal loudness zero). 3: More live mode (Loudness of environmental sound is increased)
User Control	Change the loudness of Announcer or Commentators voice

No. of Object	9
Object Name	Vocal, Violin 1, 2, Viola, Cello, Piano-L, R, Background-L, R
Control	Position, Distance, Loudness
Preset	1: Normal position(like figure) 2: Karaoke mode (Vocal loudness zero)
User Control	Change the loudness, position and distance of each object.

Aplicações em DTV

- Cobertura de diversos serviços e aplicações

Services	**Sound programs**	**Modes (channel configuration)**
main video sound program	mono stereo	1 (1.0) 2 channels (2.0)
audio description secondary language side audio (music, etc.)	surround (spatial sound)	2.1 channels (plus LFE/subwoofer) alternative modes (3.0, 4.0, 4.1, etc.)
separate voice and music additional sound services	simultaneous stereo and surround	5.1 (6 channels) 7.1 (8 channels)
audiovisual realism enhancement sound synthesis (terminal) 2D/3D sound scene capability	concurrent multiformat delivery (Surround formats)	advanced n.m modes (n+m channels) mixed modes (2.0 + 5.1, etc.)

Laboratório de Sistemas Integráveis Tecnológico

AUDIENCE System

Spatial Audio Production, Transmission and Reproduction

Audio Immersion Experience by Computer Emulation

AUDIENCE base architecture

- Layer-oriented
  - layer 1, responsible for scene compositions and description, production and interaction

AUDIENCE base architecture

- Layer-oriented
  - layer 1, responsible for scene compositions and description, production and interaction
  - layer 2, responsible for scene rendering (and acoustic simulation)

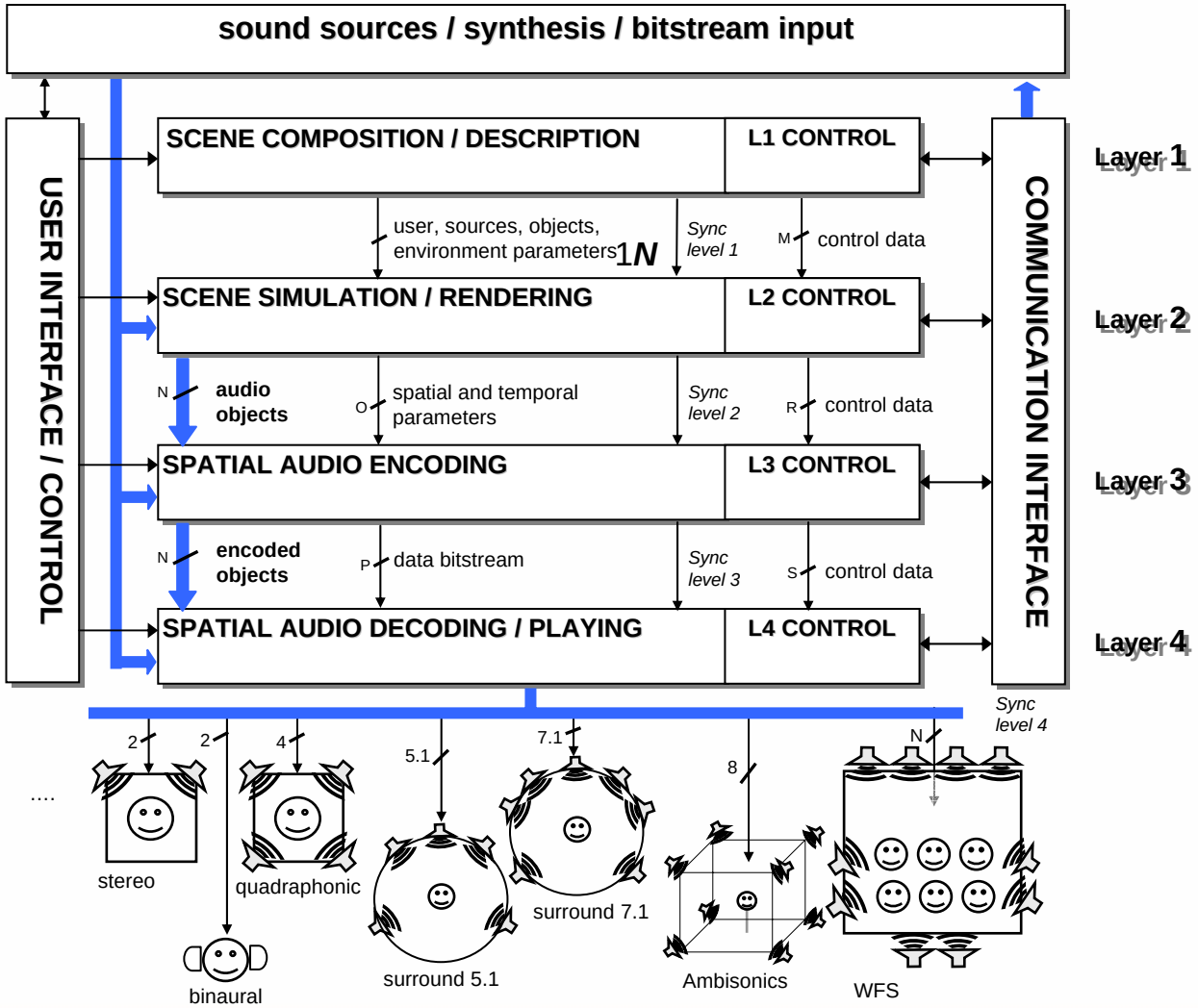
AUDIENCE base architecture

- Layer-oriented
  - layer 1, responsible for scene compositions and description, production and interaction
  - layer 2, responsible for scene rendering (and acoustic simulation)
  - layer 3, responsible for spatial audio encoding (encoding/compressing data) and can accommodate AAC, HE-AAC, MP3, MPS and other codecs

AUDIENCE base architecture

- Layer-oriented
  - layer 1, responsible for scene compositions and description, production and interaction
  - layer 2, responsible for scene rendering (and acoustic simulation)
  - layer 3, responsible for spatial audio encoding (encoding/compressing data) and can accommodate AAC, HE-AAC, MP3, MPS and other codecs
  - layer 4, responsible for decoding and playing (reproduction). This layer decodes/uncompresses data, to feed other layers or feed an output player

AUDIENCE base architecture



AUDIENCE architecture

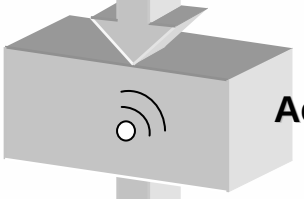
- Hierarchy
  - Technical-oriented
  - Human-oriented
- 1 layer = 1 functional group
- Layer independency
  - minimum coupling with visual systems, minimal coupling between layers
- Functional Units = Processing blocks
- Many possible combination of layers and functions per layer for the encoder and for the decoder terminal

AUDIENCE Auralization Phases

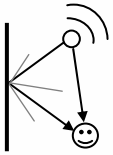
Mapping sound sources atributes, position, environment parameters

Acoustic scene description

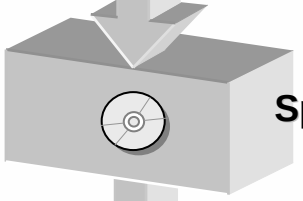
Acoustic propagation simulation (sound field rendering)



Acoustic simulation



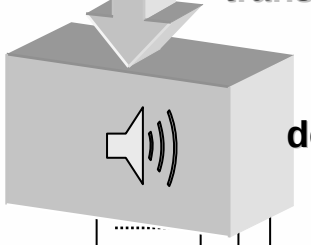
Processing and generation of spatial encoded sounds into a format for distribution



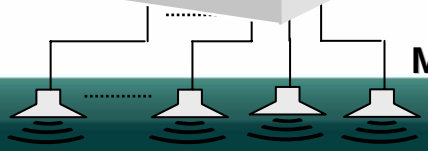
Spatial audio coding

transmission ..1010011011.. ..1001101101..

Decoding, mixing, filtering, generating N output signals, reproducing sound field

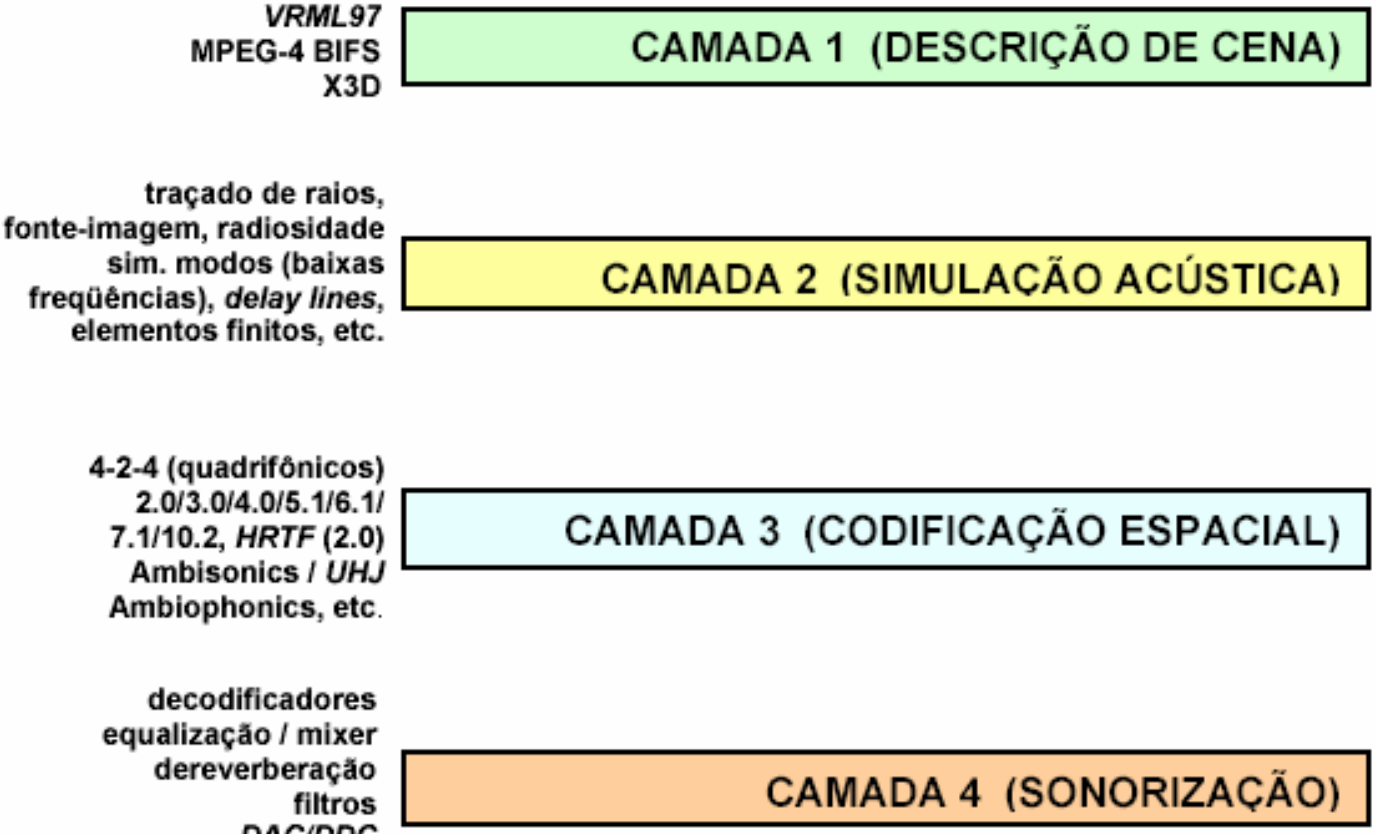


decoding



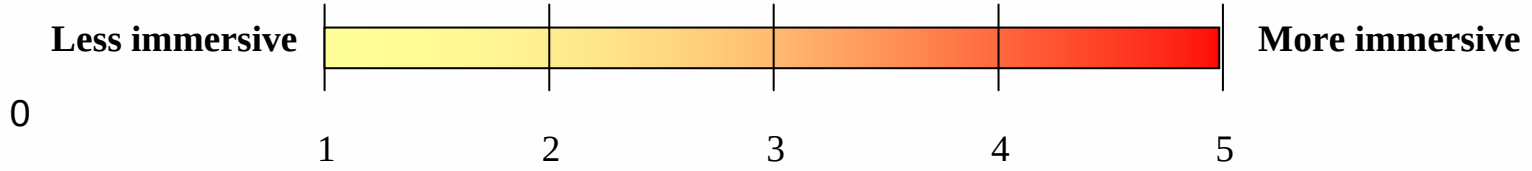
Multichannel playing

Auralization tools



Auralization level

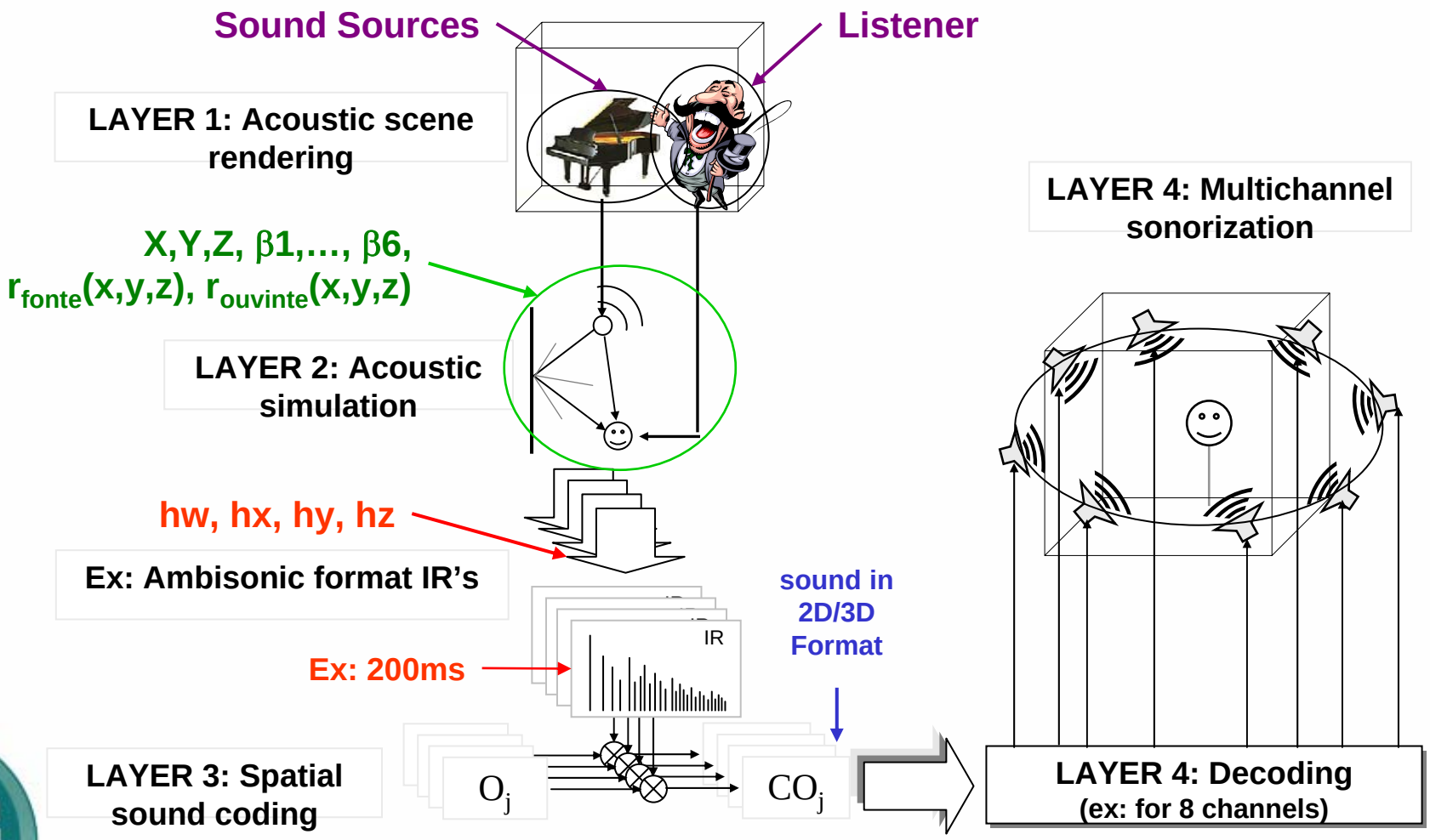
- Concept of Sound immersion: grading scale with 6 levels (Faria 2005)



level	Techniques/methods (examples)	perception (results)
0	monaural, anechoic, dry signal	no immersion
1	reverberation, echoes	spaciousness, ambience
2	panning (between speakers), stereophony, n.m (surround schemes)	sound direction, movements
3	amplitude panning, VBAP	correct positioning in limited regions
4	HRTF, auralization, periphony (Ambisonics, WFS, Ambiophonics)	stable 2D sound fields
5	HRTF, auralization, periphony (Ambisonics, WFS, Ambiophonics)	stable 3D sound fields, accurate distance and localization

Laboratório de Sistemas Integráveis Tecnológico

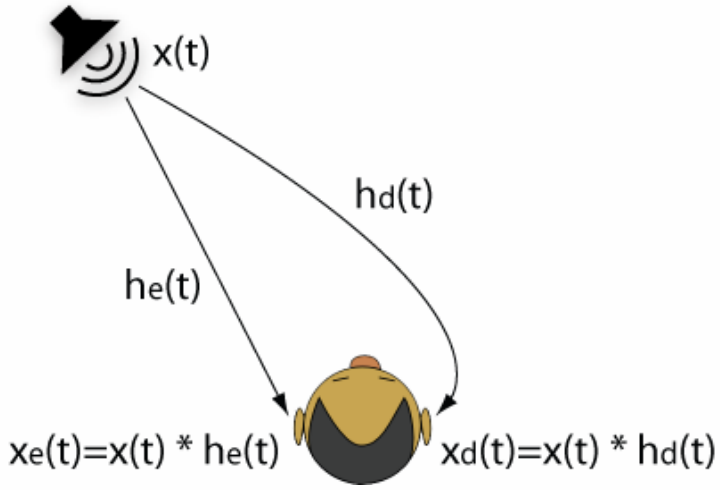
Auralization Methodology



Auralization techniques

- Bi-aural
- Dipolo Estéreo
- N.M (5.1, 7.1, etc)
- Ambiophonics
- VBAP
- WFS
- Ambisonics

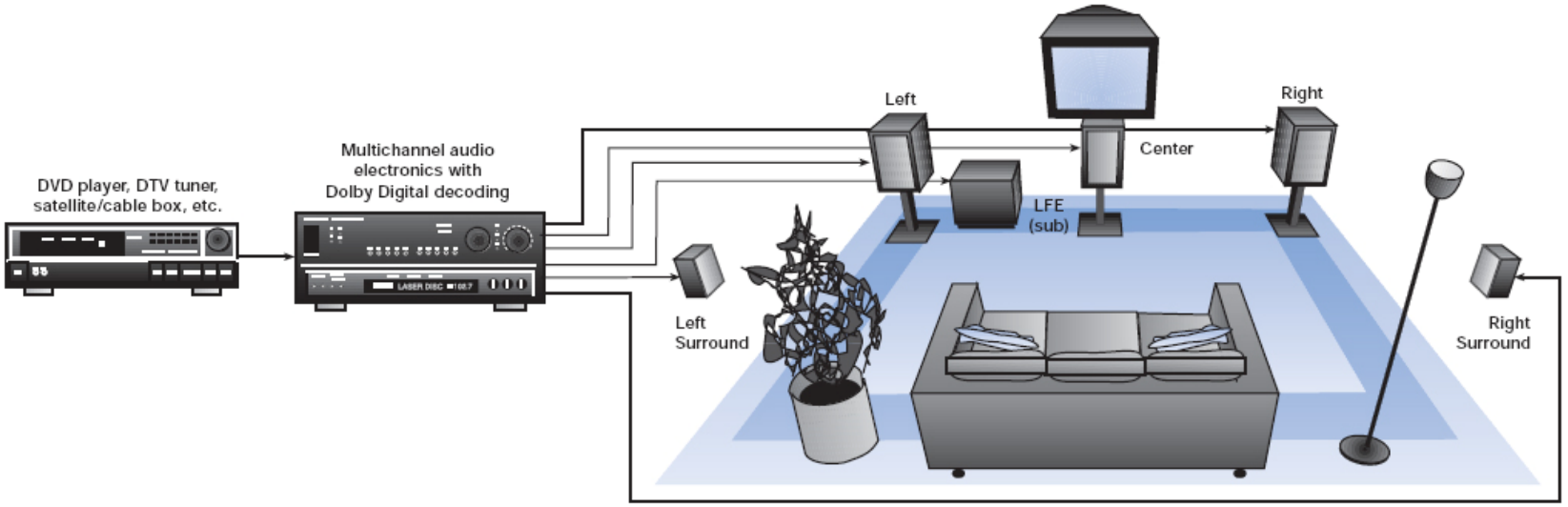
Bi-aural



Funções de transferências dos ouvidos esquerdo e direito

Dummy Head KU100, da Neumann

N.M

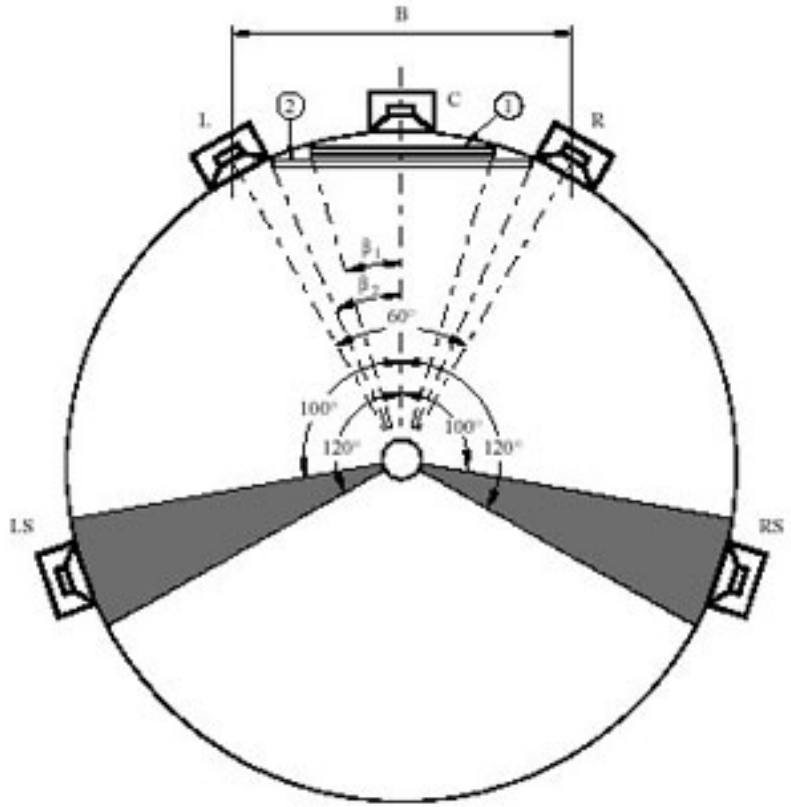


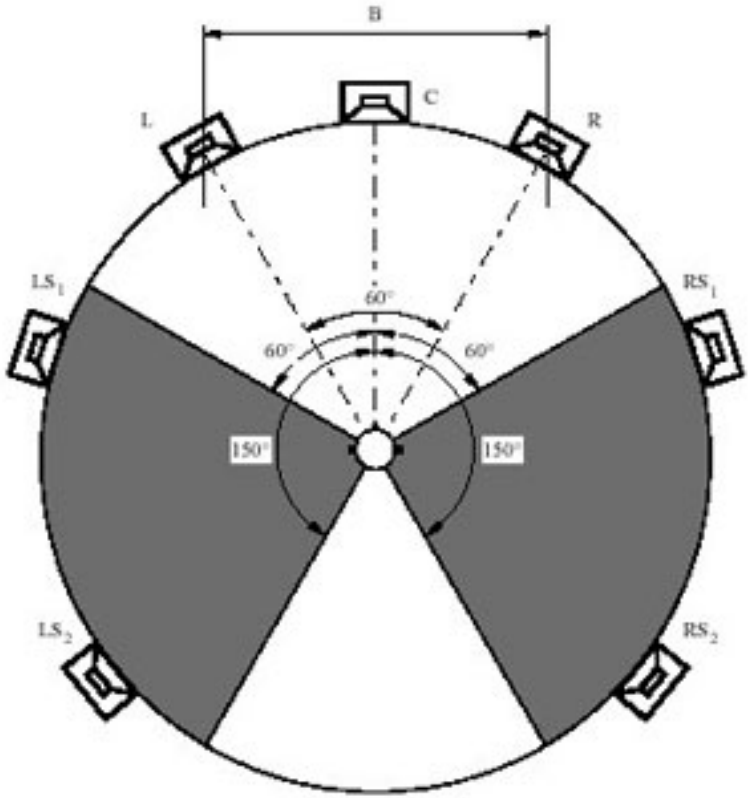
5.1 = 3/2/1 e 7.1 = 3/4/1

Dolby / DTS

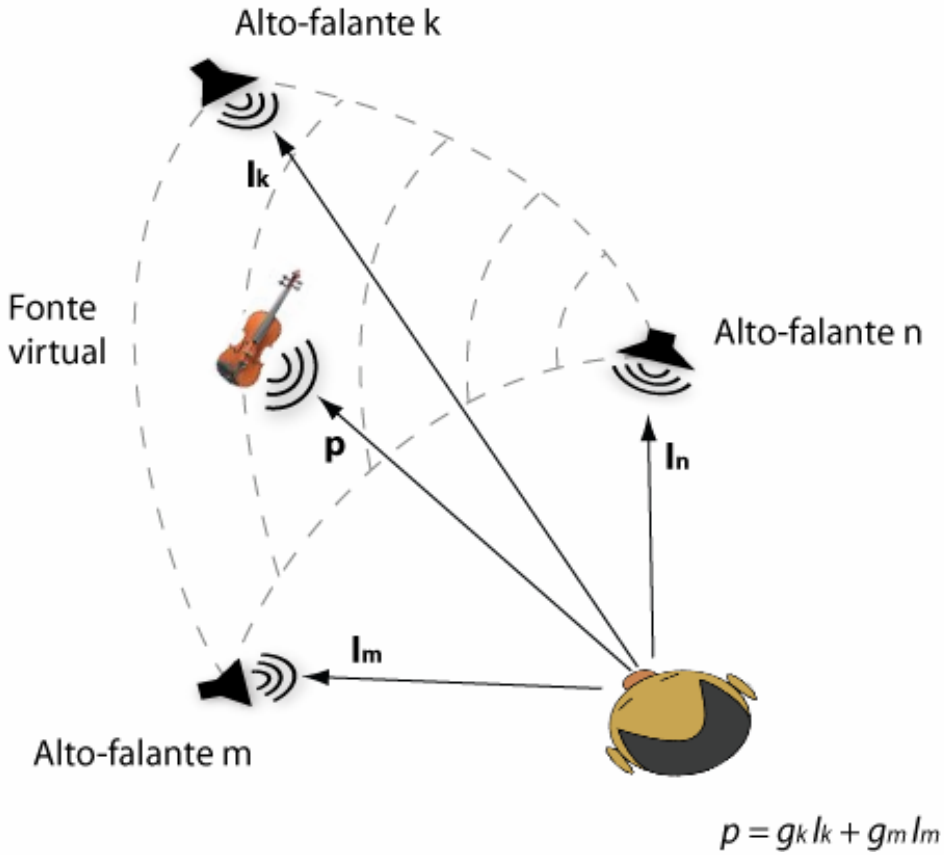
5.1/7.1

- ITU-R BS.775.1





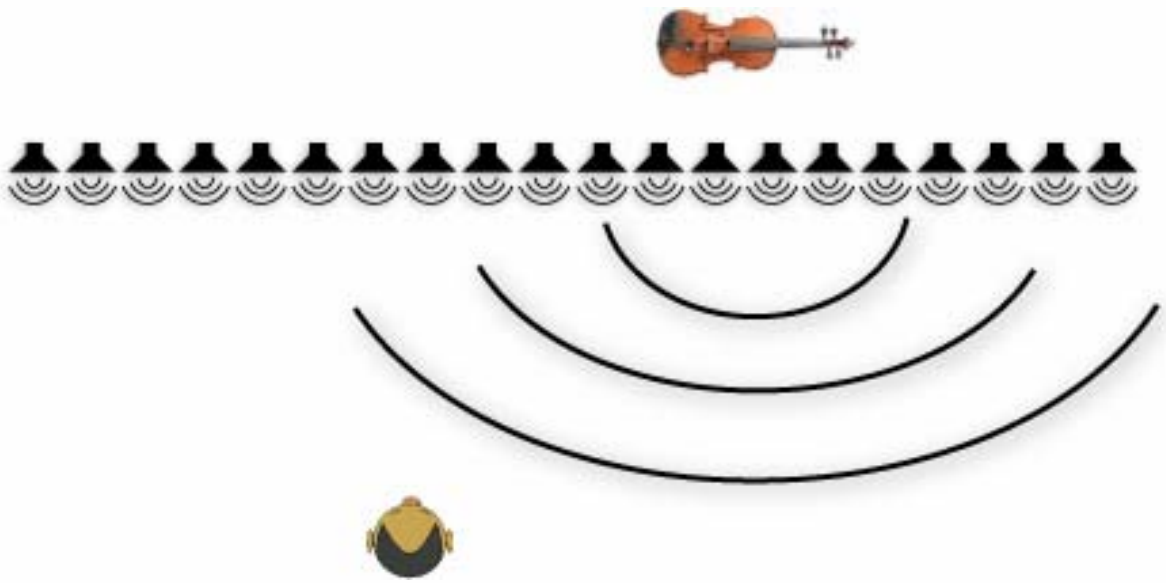
VBAP



$$p = g_k l_k + g_m l_m + g_n l_n$$

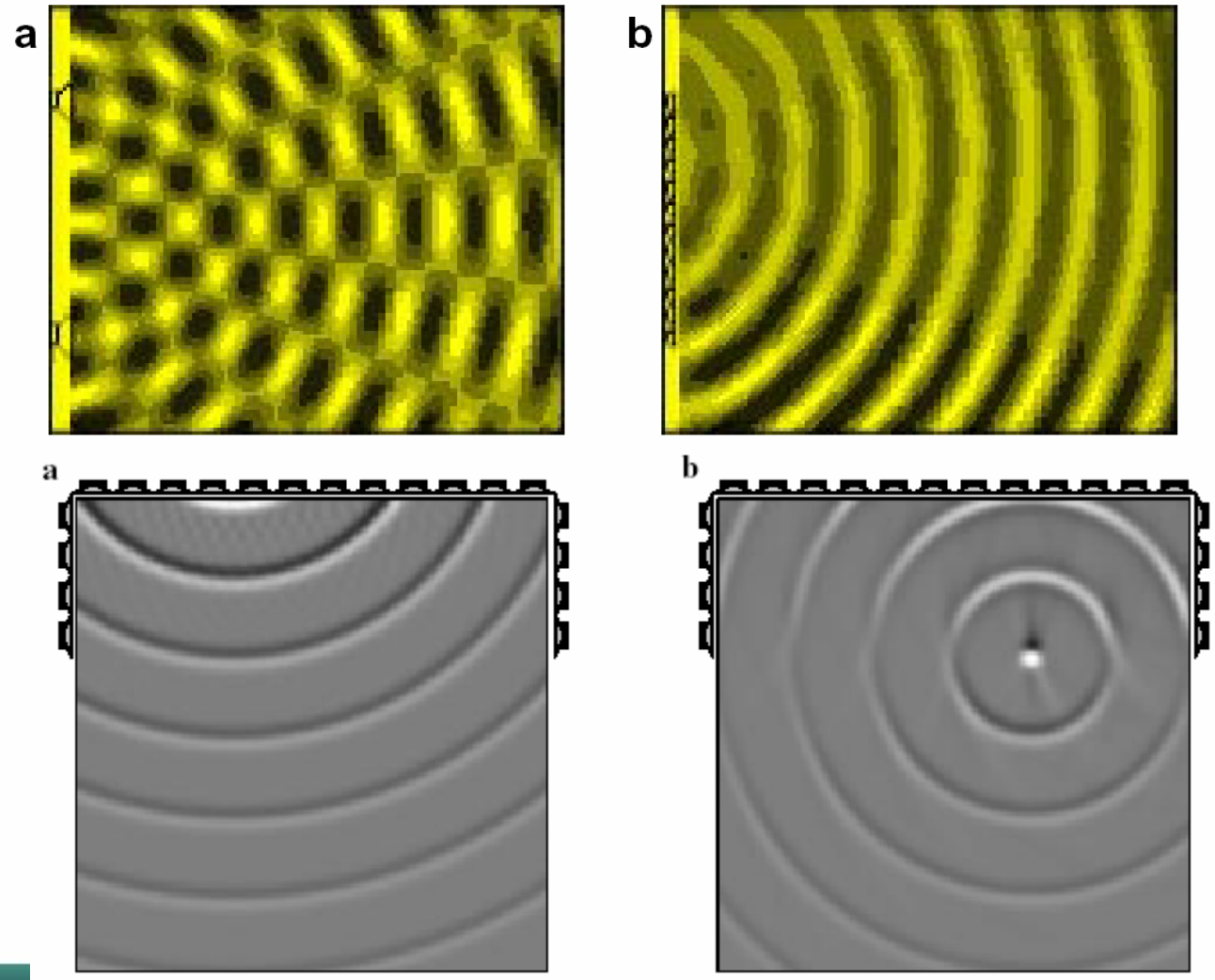
Vetores e posicionamento de alto-falantes no SVPA

Wave Field Synthesis

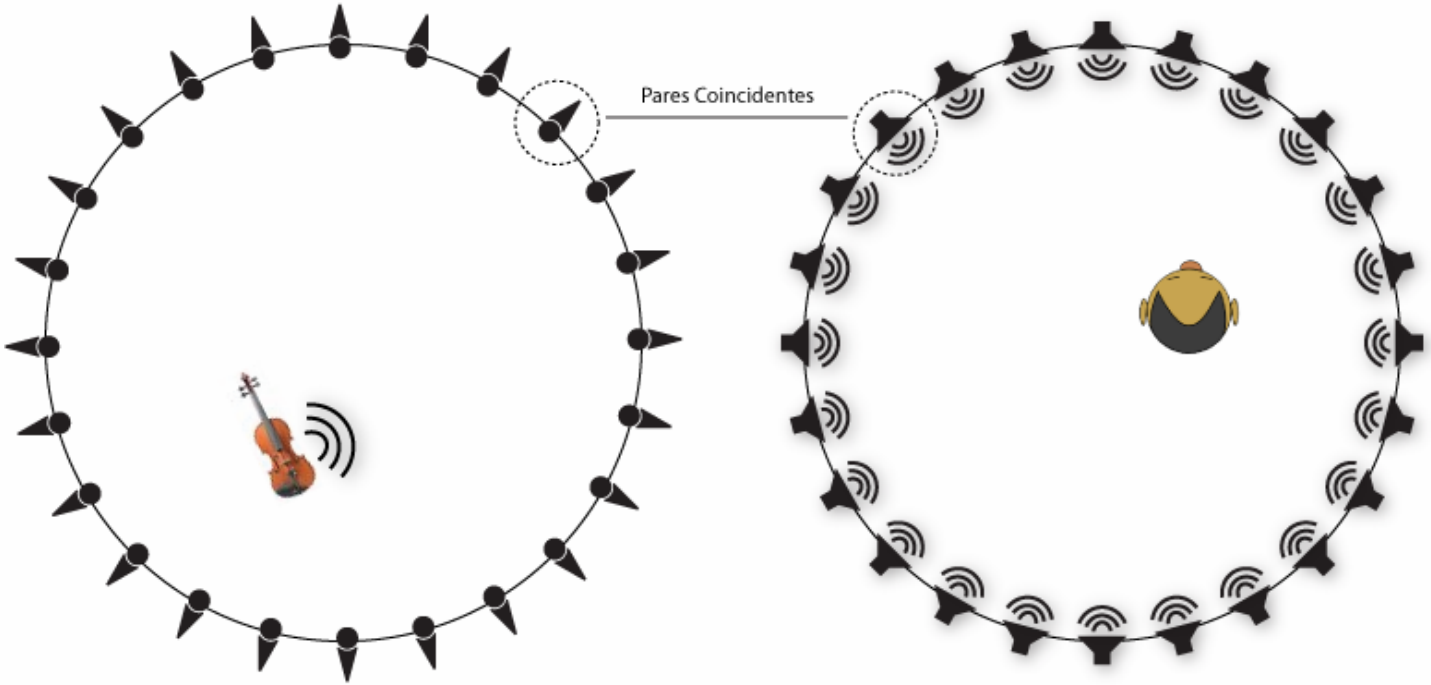


Montagem WFS frontal, na qual uma matriz de alto-falantes reproduz a onda original

Wave Field Synthesis

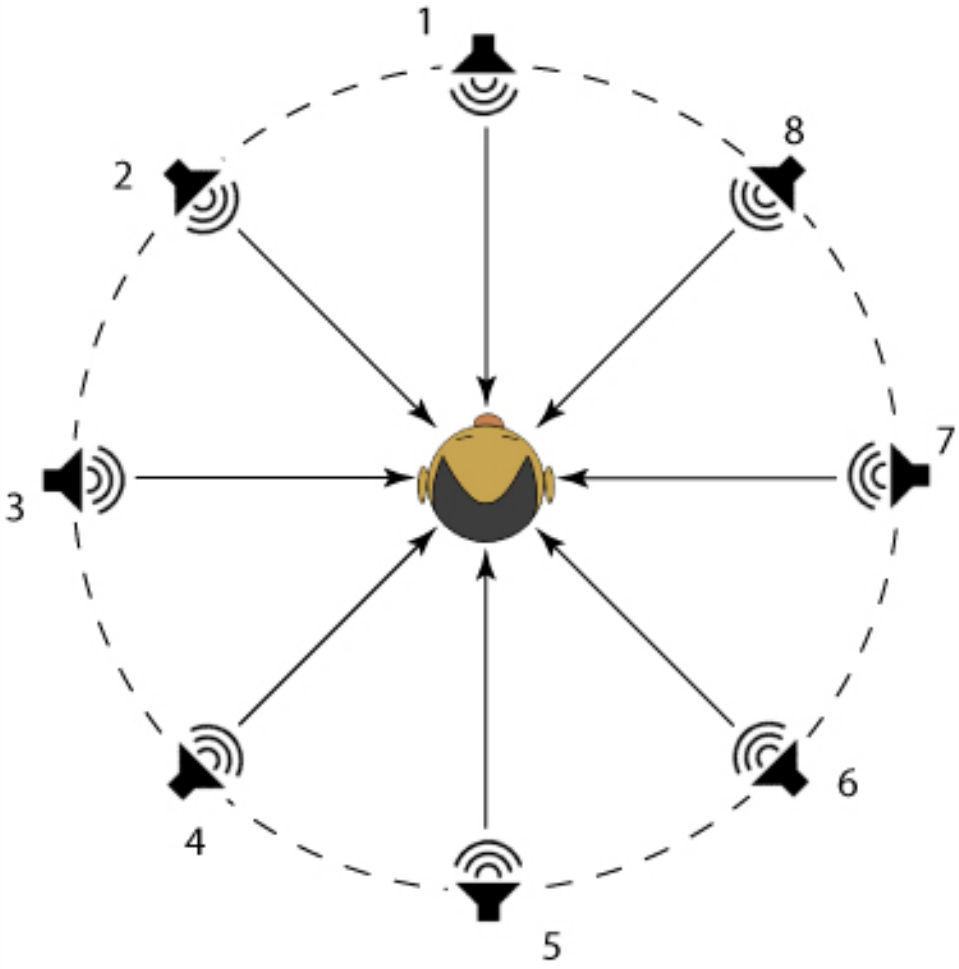


Wave Field Synthesis



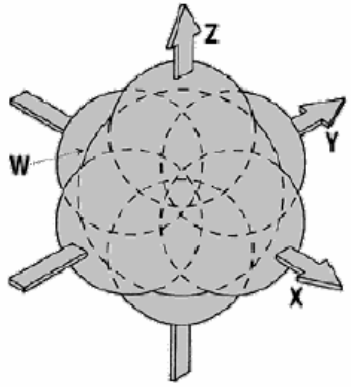
Montagem WFS circular para gravação e reprodução, com uma matriz densa de microfones e alto-falantes

Ambisonics



Ambisonics

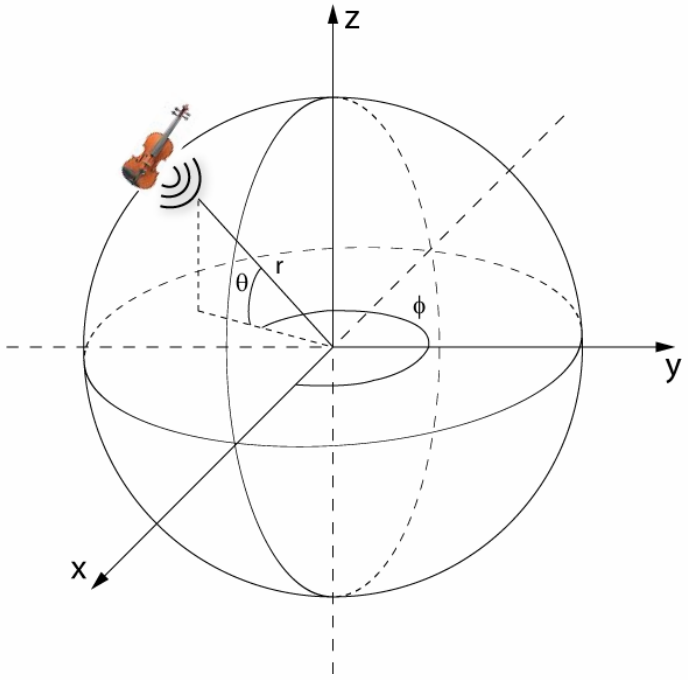
- Periphony: With-Height Sound Reproduction (Gerzon,1973)
- 3D sound field recording and reproduction
- Wave front reconstruction (Huygens)
- Encoding independent from decodin-
- Implementation flexibility
  - Define Ambisonics order
  - Define loudspeaker configuration



Ambisonics

- Encoding
  - Localization

$$W = \sqrt{2} \cdot S$$
$$X = \cos(\phi)\cos(\theta) \cdot S$$
$$Y = \sin(\phi)\cos(\theta) \cdot S$$
$$Z = \sin(\theta) \cdot S$$



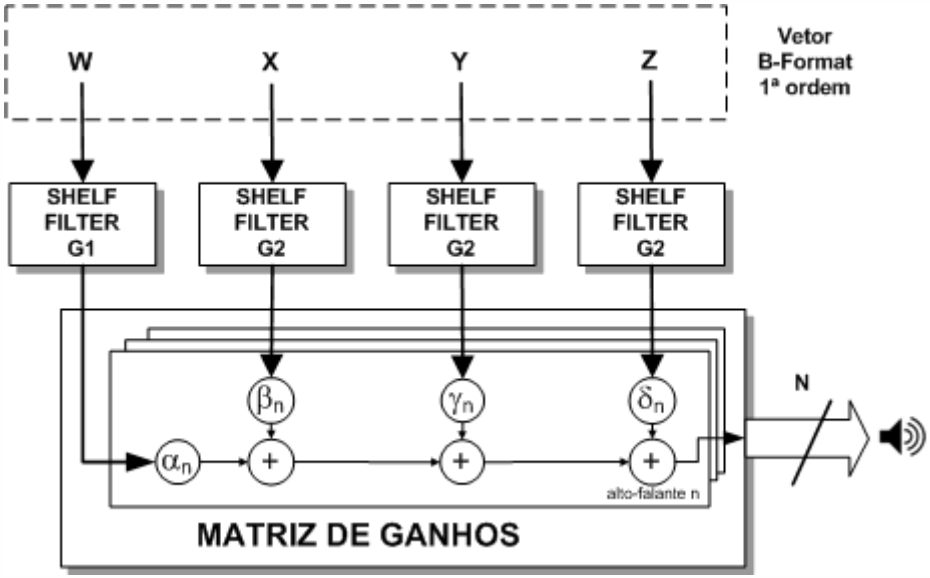
Mic Soundfield™

  - Distance > intensity
  - Delays

– Permit further manipulation of recorded sound field

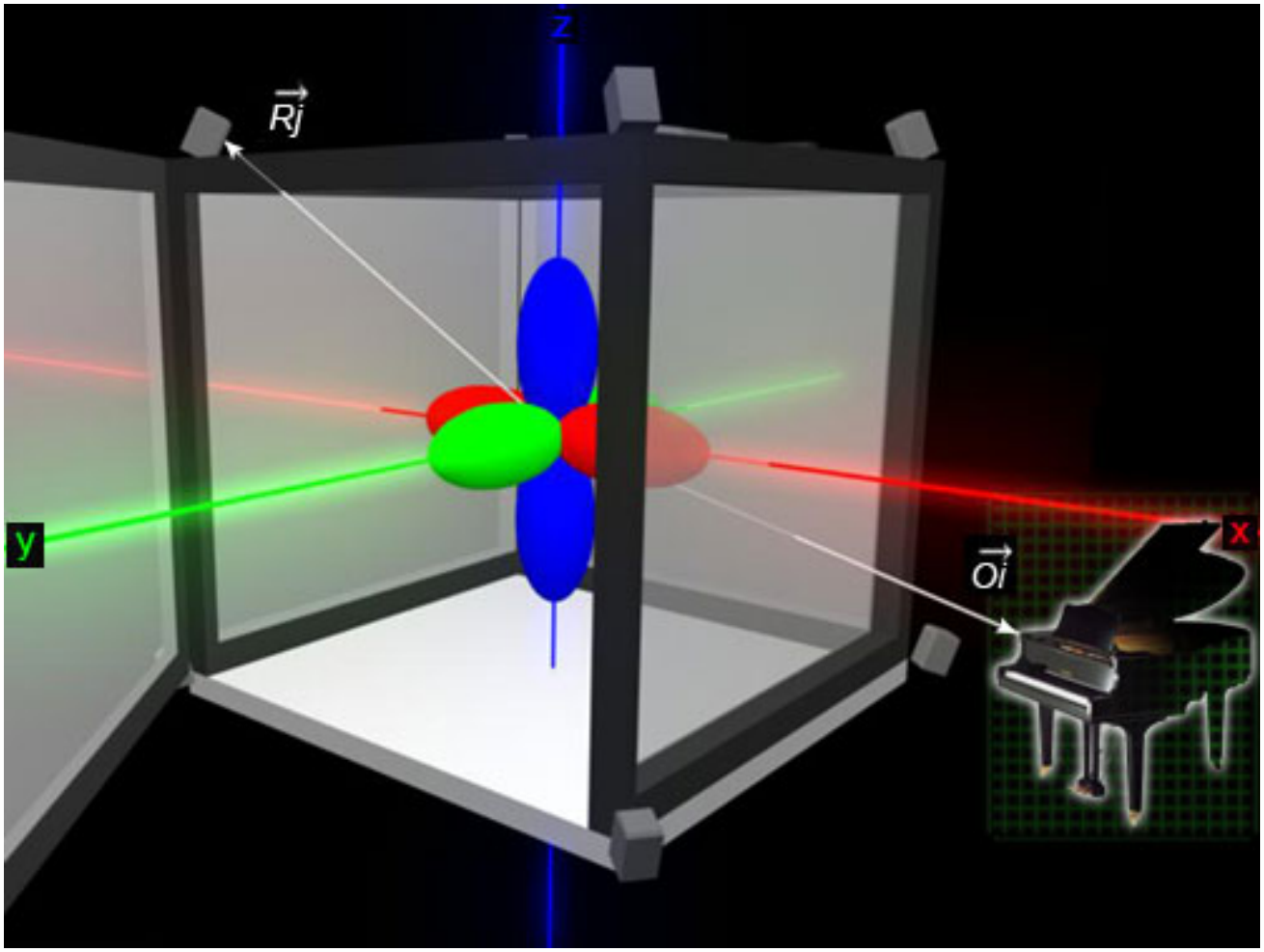
Ambisonics

- Decoding
  - Gain matrix
    - Regular configurations → tables
    - Irregular configurations → require complex solutions
  - Psychoacoustics filters

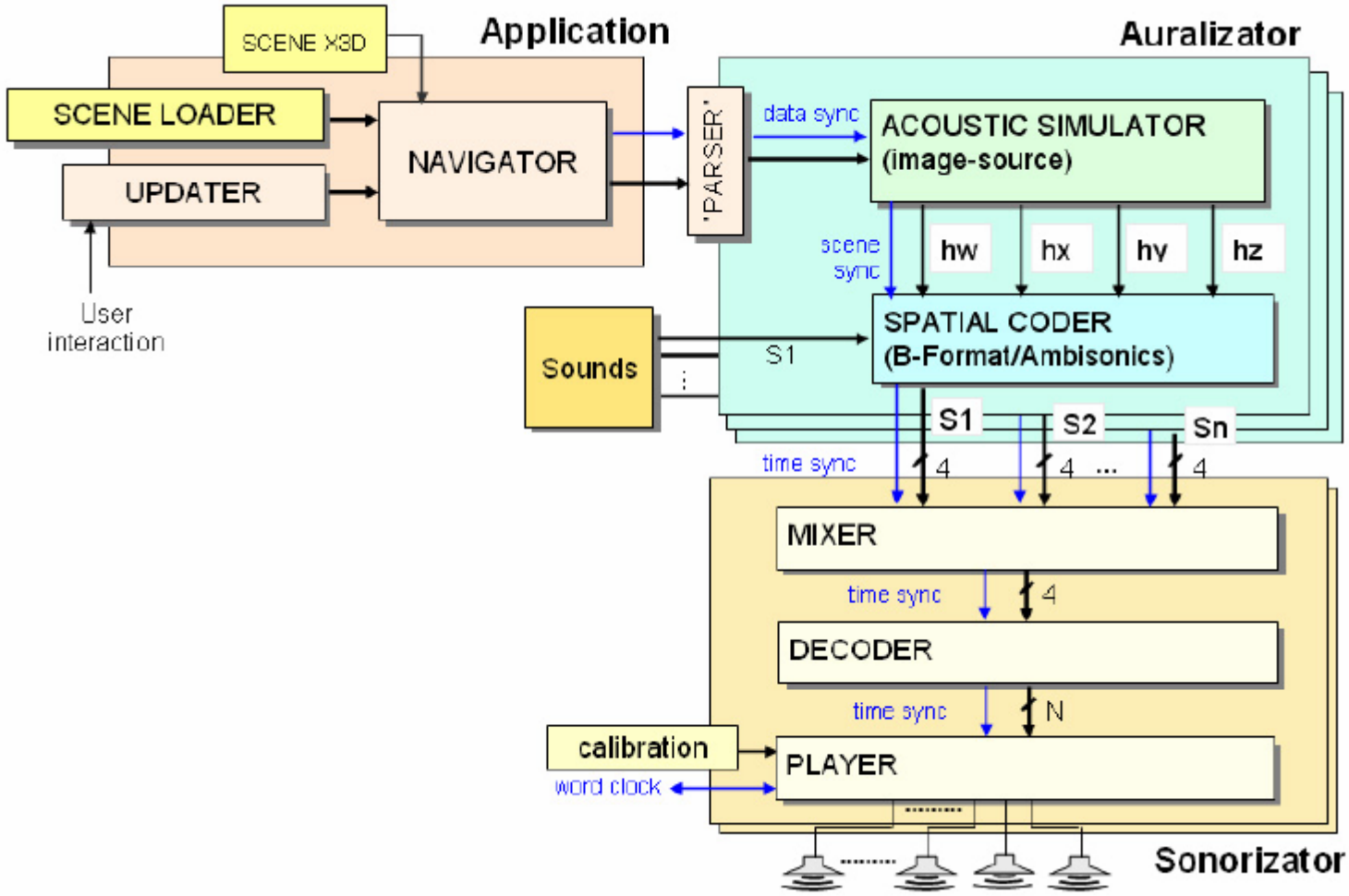


1st order Ambisonics decoder

Ambisonics

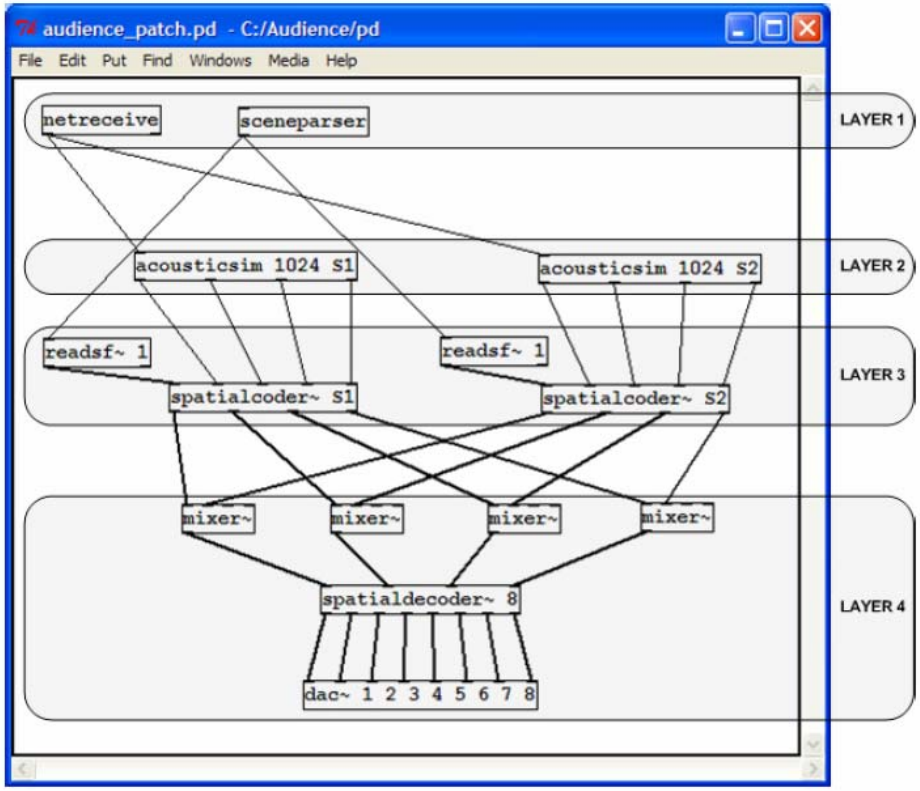


Building an Auralizator



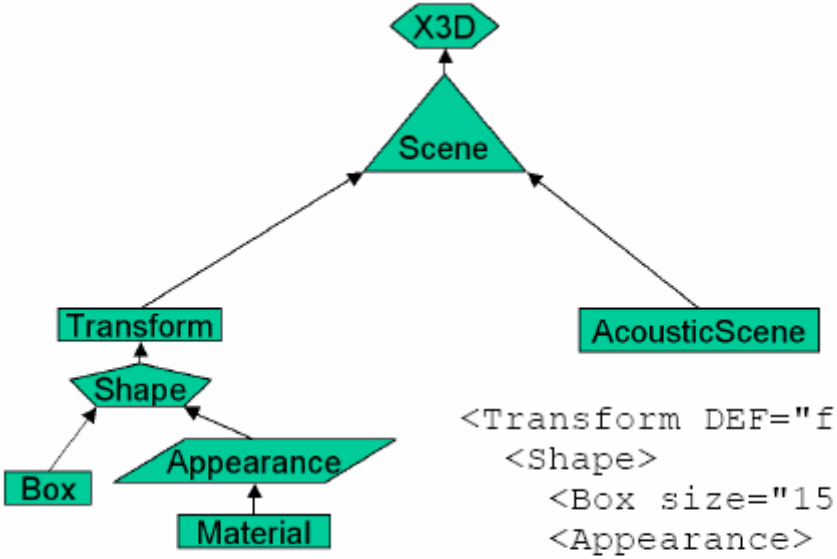
Building an Auralizator

- Underline software system: Pd
  - real-time graphical programming environment for audio, music and graphical processing by Miller Puckette, following MAX/MSP
- AUDIENCE Pd Patch:
  - Layer-oriented components
  - Flexible command passing and audio routing



Layer 1: Acoustic Scene

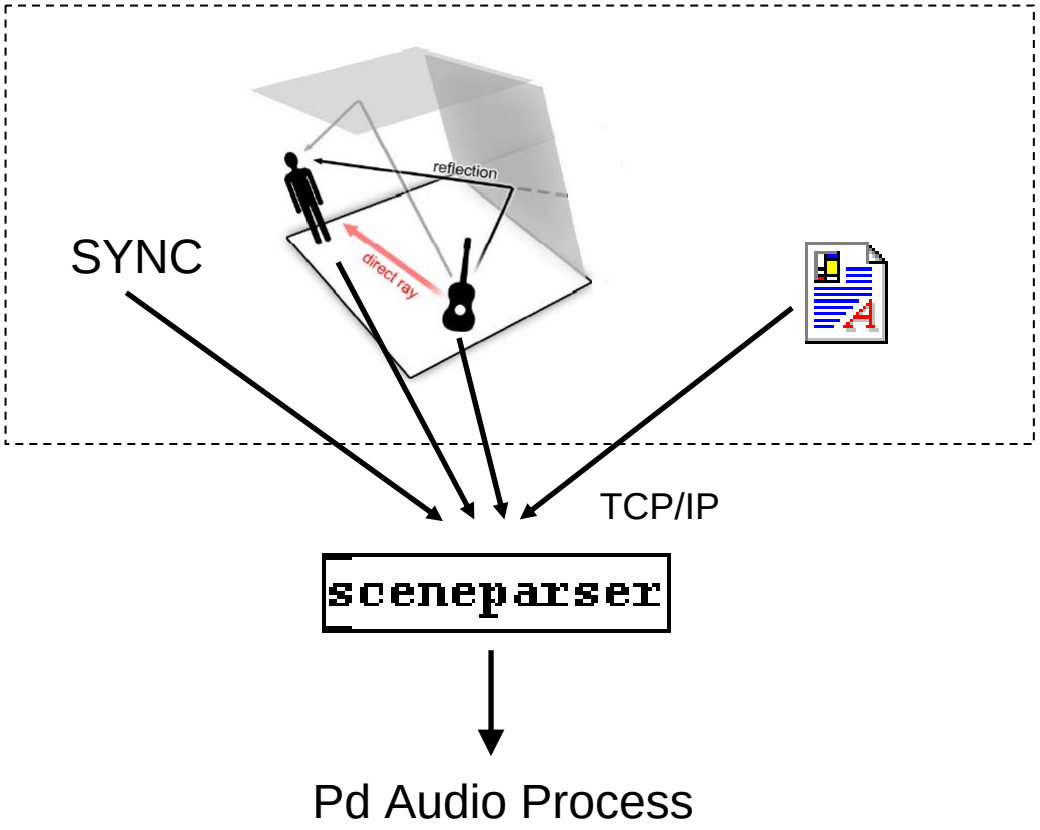
- Acoustic Scene Description:
  - X3D audio nodes primitives
    - New nodes and attributes added for audio




```
<Transform DEF="floor" center="7.5 0.05 10">
  <Shape>
    <Box size="15 0.1 20"/>
    <Appearance>
      <Material diffuseColor="1 1 1"/>
    </Appearance>
  </Shape>
  <Sound>
    <AcousticMaterial coefref=".8" freqref="1000"/>
  </Sound>
</Transform>
```

Layer 1: Acoustic Scene

- Acoustic Scene Parsing

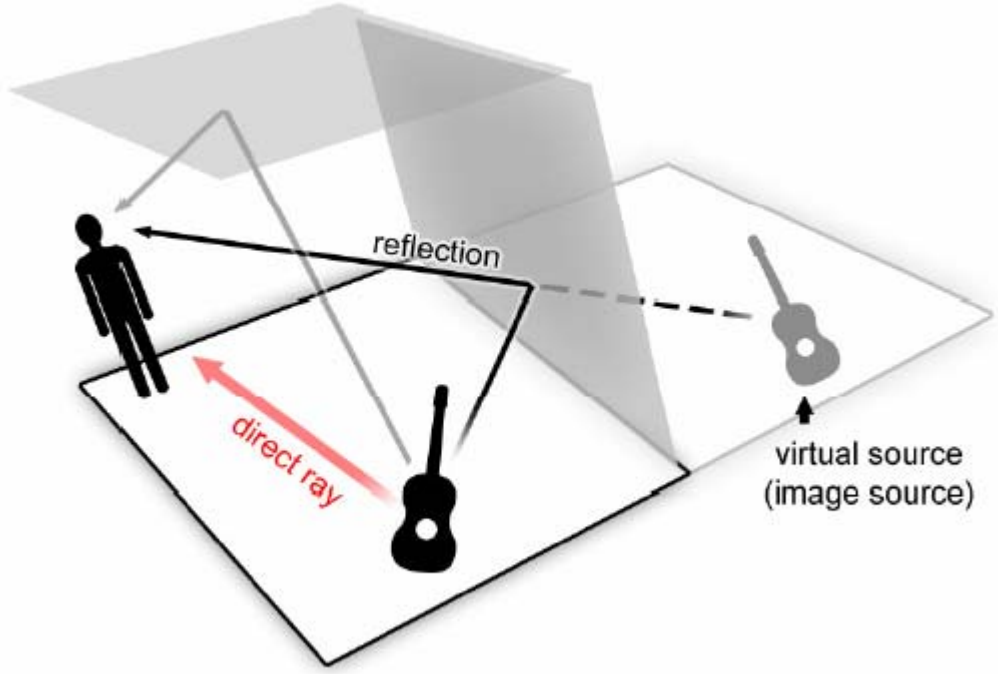


Layer 1: Acoustic Scene

- Acoustic Scene Parsing
  - Component receives user localization, sound source localization, sound file, synchronization by TCP/IP from the application (example: virtual reality browser, 3D visual processing software, etc)
  - Output: user localization, sound source localization, environment parameters, sound file attributes
  - Processing: receives information and pass it to the layer 2 and 3 for 3D audio processing and coding

Layer 2: Acoustic Render

- Acoustic Simulator
  - First implemented model: Allen's image-source technique for rectangular spaces [Allen, 1979]



Layer 2: Acoustic Render

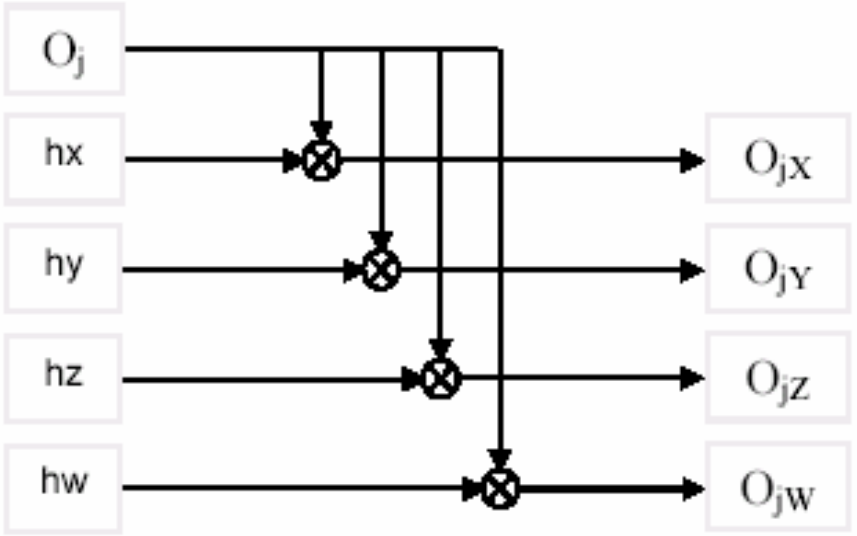
- Different techniques can derive a Pd layer-2 component and be combined for large bandwidth simulation (low and high frequencies)
- Next desired techniques to be prototyped in the auralizator include ray-based and wave-based methods:
  - Borish (1984) Arbitrary polyhedra virtual rooms
  - Radiosity methods
  - FEM (finite element methods)
  - BEM (boundary element methods)

Layer 2: Acoustic Render

- Acoustic Simulator component
  - Parameters: size of the impulse response
  - Inputs: room dimensions, sound source localization, listener localization, walls' absorption coefficients
  - Outputs: impulse responses (ex: W, X, Y and Z IRs), reflectograms, intermediary channel set processed
  - Processing: uses wave-based or ray-based acoustic renderers, such as Allen`s image source technique for rectangular spaces
- New component: RenderingMatrix generator for 5.1 scene
  - Parameters: normalization index
  - Inputs: sound source localization, listener localization
  - Outputs: intermediary 5.1 channel set processed
  - Processing: uses VBAP concepts for 2D rendering onto 5.1 speakers

Layer 3: Spatial Audio Encode

- Spatial Audio Coding
- For Ambisonics:
  - exports B-Format signals: X, Y, Z, W
  - Encoding process: convolution

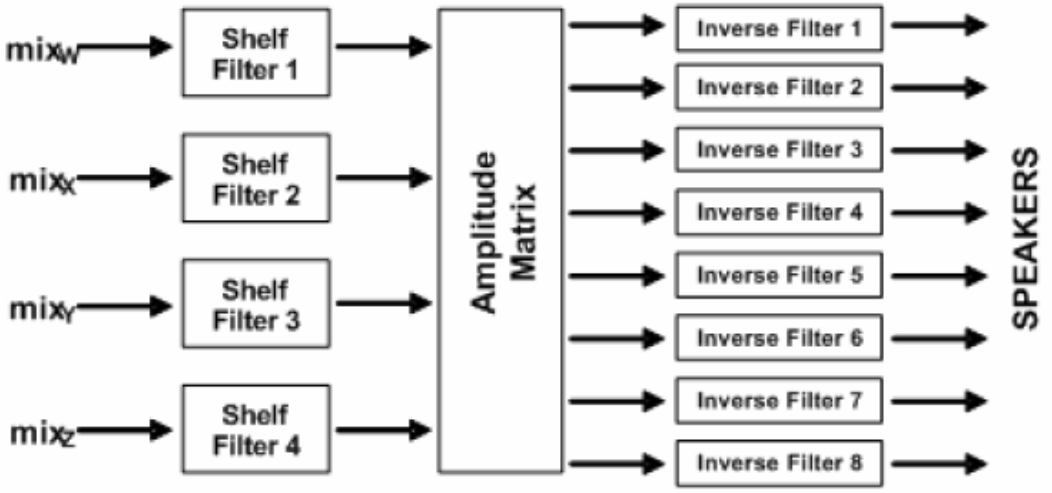


Layer 3: Spatial Audio Encode

- Spatial Audio Coder components
  - Inputs: sounds in mono, B-Format impulse responses (W, X, Y, Z), sound source reflectograms, intermediary channel set processed, inter-channel correlation data
  - Outputs: B-Format sound, MPS sound, SAOC sound, AAC bitstreams, etc.
  - Processing: using the overlap-add convolution algorithm, encodes the mono sound in the B-Format using the impulse responses; uses channel correlation for efficient coding, etc.

Layer 4: Decode & Reproduce

- Spatial Audio Decoder
  - Decodes to a number of speaker / configuration (cube, octagon etc)
    - Ex: Ambisonics 1st Order Decoder based on Richard Furse's calculations
  - Inverse filters to minimize the diffusion effects of the CAVE's projection screens



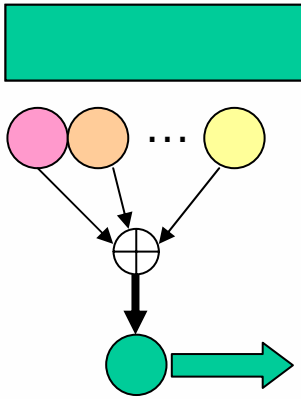
Layer 4: Decode & Reproduce

- Spatial Audio Decoder component
  - Parameters: speaker configuration type
  - Input: sound in Ambisonics 1st Order B-Format, in MPS, in AAC, etc.
  - Output: sound decoded for n speakers
  - Processing: inverse of the auralization technique encoding, decompressing, de-reverberation, normalization

Desenvolvimentos em curso

- Layer 1: interface 3D p/ produção de cena
- Layer 2: projeto de reverberadores
- Layer 3: MPS encoder, SAOC encoder, 5.1 encoder, projeto para WFS codec
- Layer 4: MPS, SAOC, 5.1 decoders
- Camadas auxiliares: I/O via RTA
- Documentação: *white-papers (ex:* “how-to-write OpenAUDIENCE objects”), FAQ, lista de blocos funcionais, lista de mensagens, helps

SAOC streams



StreamType1:
efficient combination of multiple objects
dwmx 2.0 or 1.0

StreamType2:
efficient association of multiple objects
bitstream

OpenAUDIENCE Software

- Plataforma: pd
- Desenvolvedores: acesso ao ambiente CVS
- Blocos auxiliares incluem conversores, formatadores de entrada e saída (ex: I/O multicanal) e patches agrupadores de funções
- Estrutura atual da distribuição
  - Bin (main layers executables and externals)
  - Doc
  - Patches
    - AUDIENCE (main patches components)
    - Projetos (specific project patches)
    - Demos
    - Tests
  - Src
  - Lib
  - Include

Metodologias

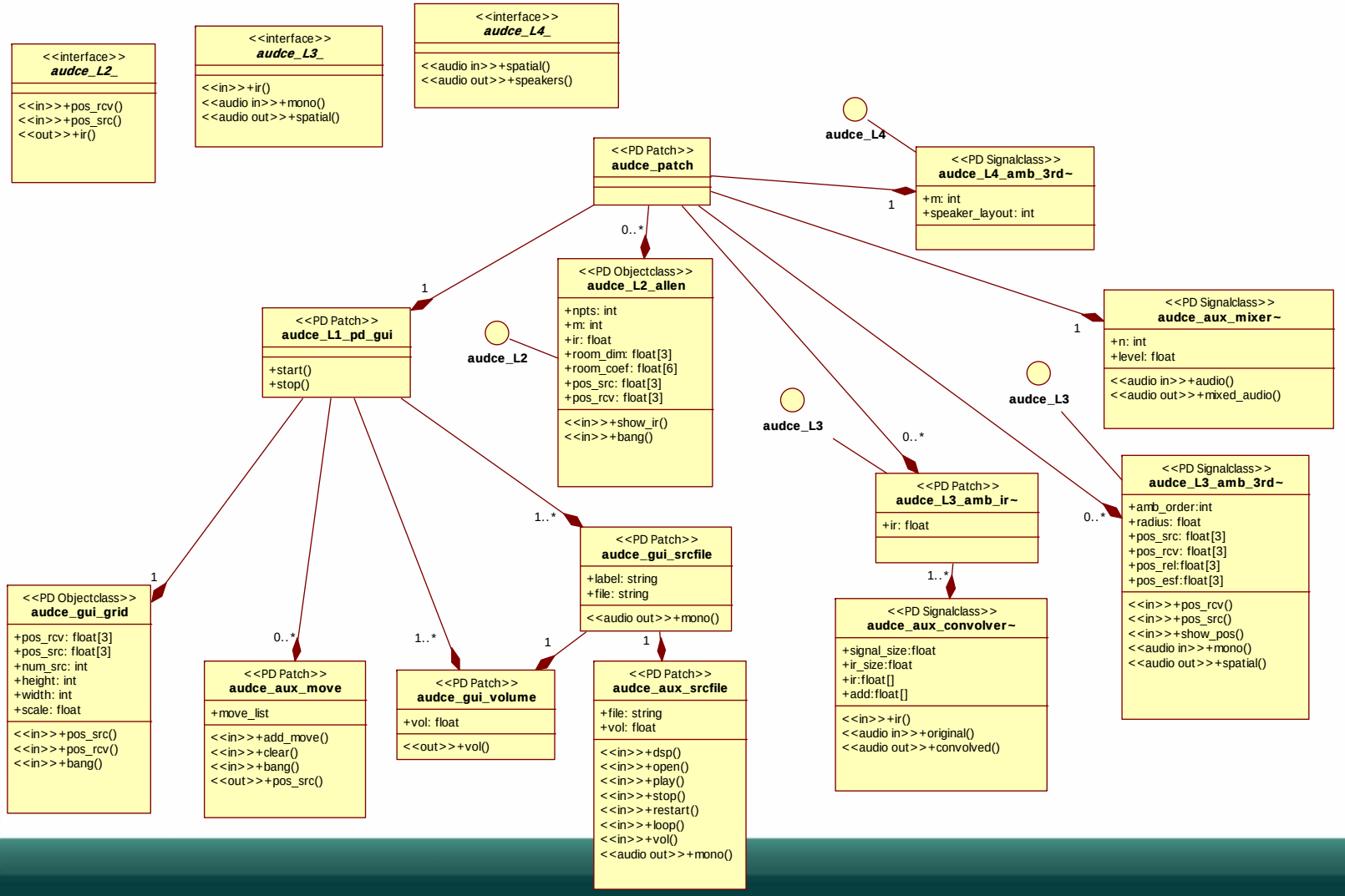
- Desenvolvimento colaborativo
  - Sete perfis de colaboração
  - Equipe local e colaboradores remotos
  - Identificação de uma camada ou aplicação de interesse levando a especificação de um bloco funcional ou patch de aplicação
- Engenharia de software
  - Identificação camada alvo, especificação formal, identificação de aplicabilidade e casos de uso
  - especificação de entradas e saídas, adequação a requisitos da camada, estrutura de chamada/uso, documentação/help, alcance e limitações,

Block design

- Functional block design considers:
  - Layer matching: functional requirements
  - Inter-block communication
    - Message passing
    - General form: attribute <new value>
  - Valid inter-layer signals
    - I/O design
    - Audio and Metadata data flows
  - Layer control
  - Software primitives
    - Class diagrams, methods, attributes, etc.
  - Naming convention
    - audce_Layer_Function

Laboratório de Sistemas Integráveis Tecnológico

AUDIENCE UML classes



Como colaborar

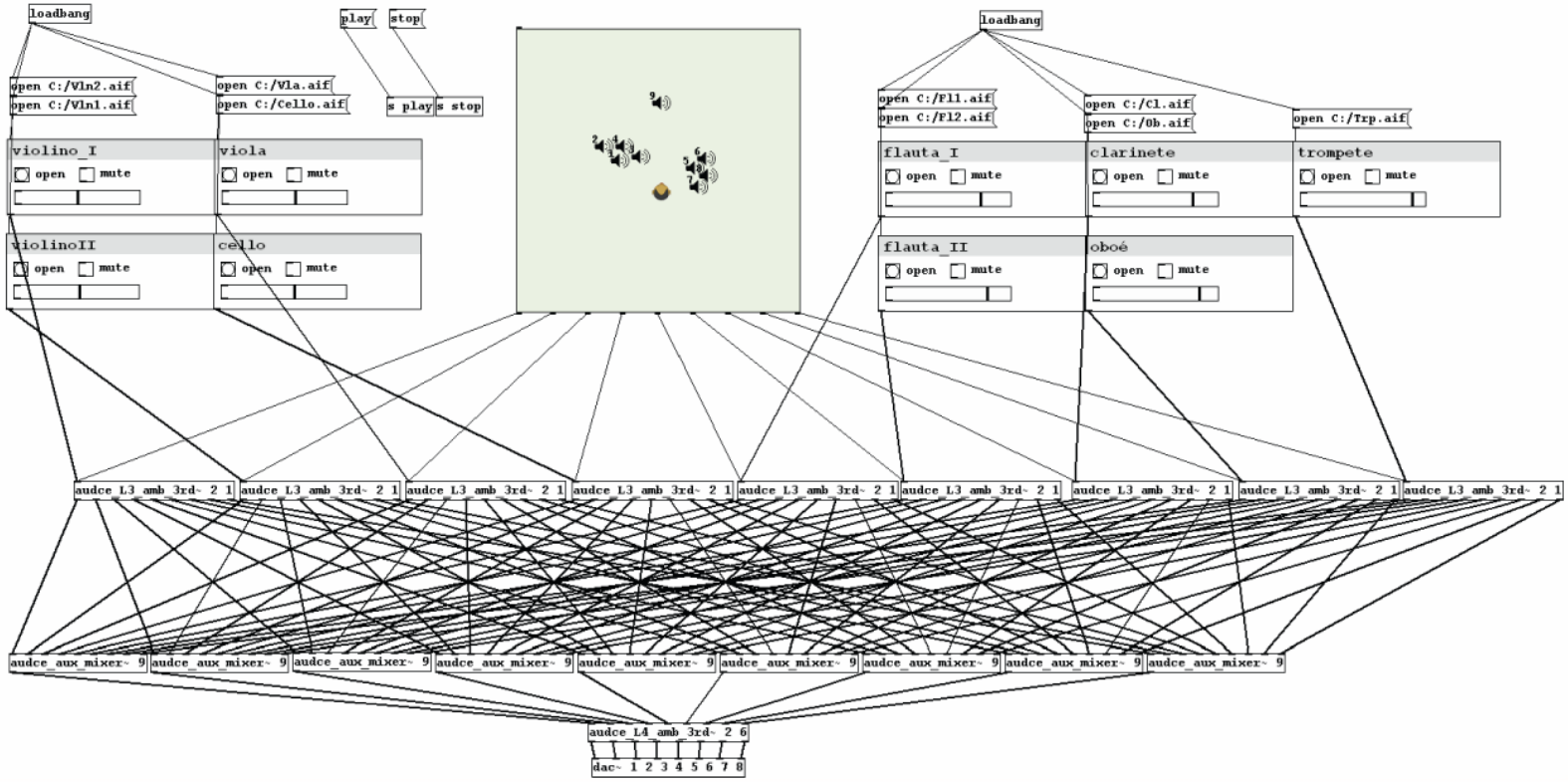
- Perfis de colaboração

Project developer	equipe básica de pesquisadores e desenvolvedores do projeto, responsável pela consolidação das distribuições e gerenciamento do acesso
Associate developer	colaborador desenvolvedor associado; colaborador pontual, remoto, responsável pelo desenvolvimento de uma funcionalidade no sistema
Doc writer	colaborador encarregado de acompanhar a documentação do software e do sistema
Support tech	colaborador encarregado de dar suporte técnico no portal a desenvolvedores remotos
Application developer	colaborador responsável pelo desenvolvimento de uma aplicação utilizando o software
Content producer	colaborador responsável pela produção de conteúdo sonoro para aplicações e testes
Portal support	colaborador encarregado de manutenção e atualização do portal

Exemplos de aplicações e uso

Music spatialization

- Virtual Stage: Charles Ives. “The unanswered question” (1906)
- 9 instruments in 3 clusters or free-field



Facilities using AUDIENCE

- CAVERNA Digital at University of Sao Paulo
- NEAC (Audio Engineering and Coding Center) at University of Sao Paulo

CAVE Multichannel Infrastructure

- Audio-dedicated node in the cluster
- 8 hi-fi speakers, in cubic array
- High quality and low noise multichannel amplifiers
- Hardware patch bay and multicable distribution
- Strategic speakers positioning to cope with visual projection

Laboratório de Sistemas Integráveis Tecnológico

Audio Infrastructure

AUDIO NODE
(CLUSTER)

P10 (TRS) S/PDIF

IN 1 2 3 4 5 6 7 8

OUT

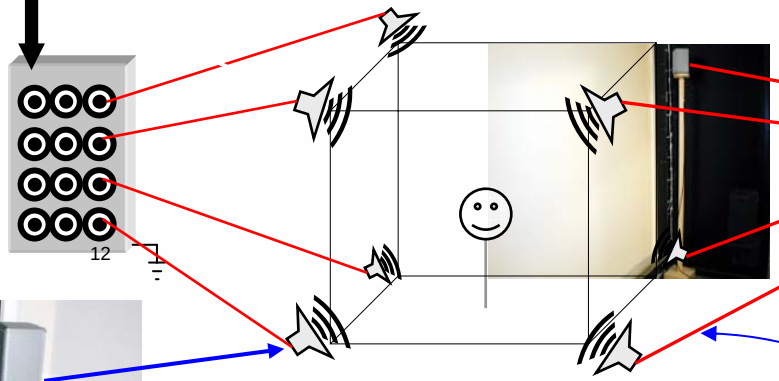
word clock (BNC)

MULTICHANNEL SOUNDCARD

RCA
Bornes

AMPLIFIERS

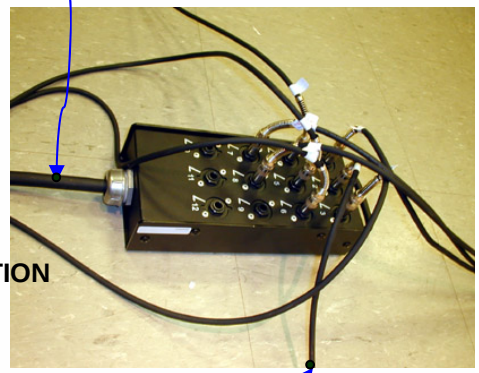
MULTICABLE



12

12

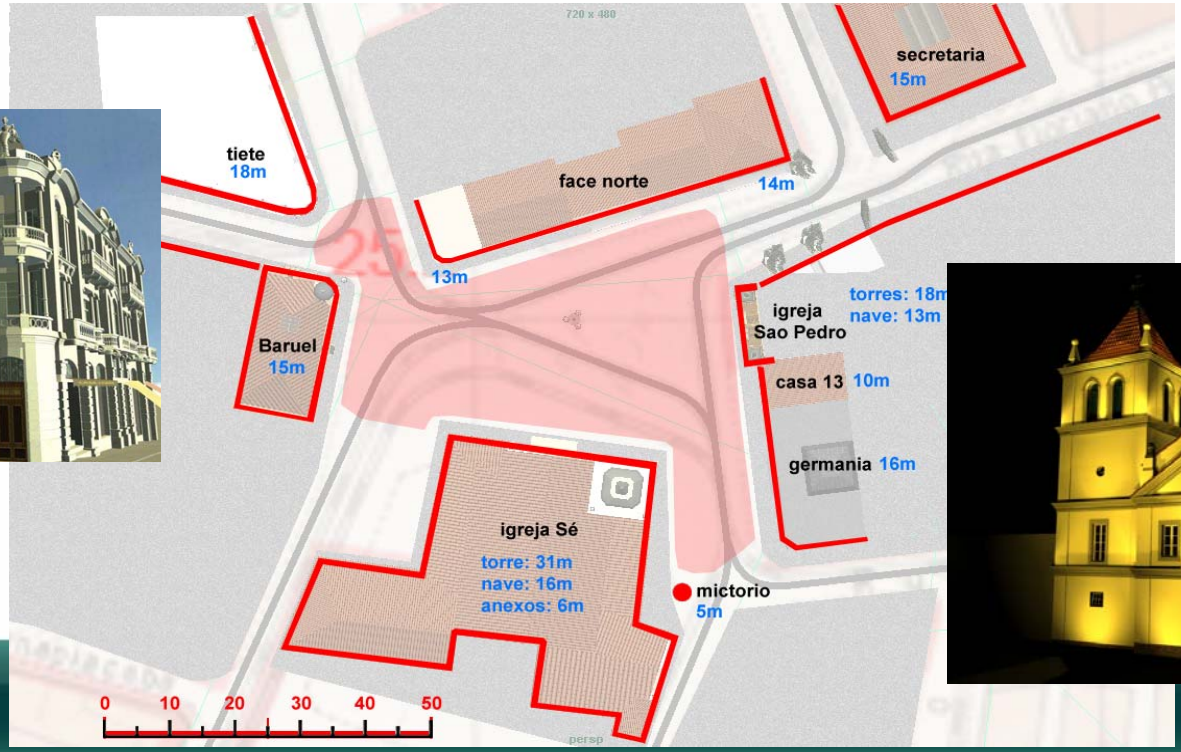
DISTRIBUTION PANELS



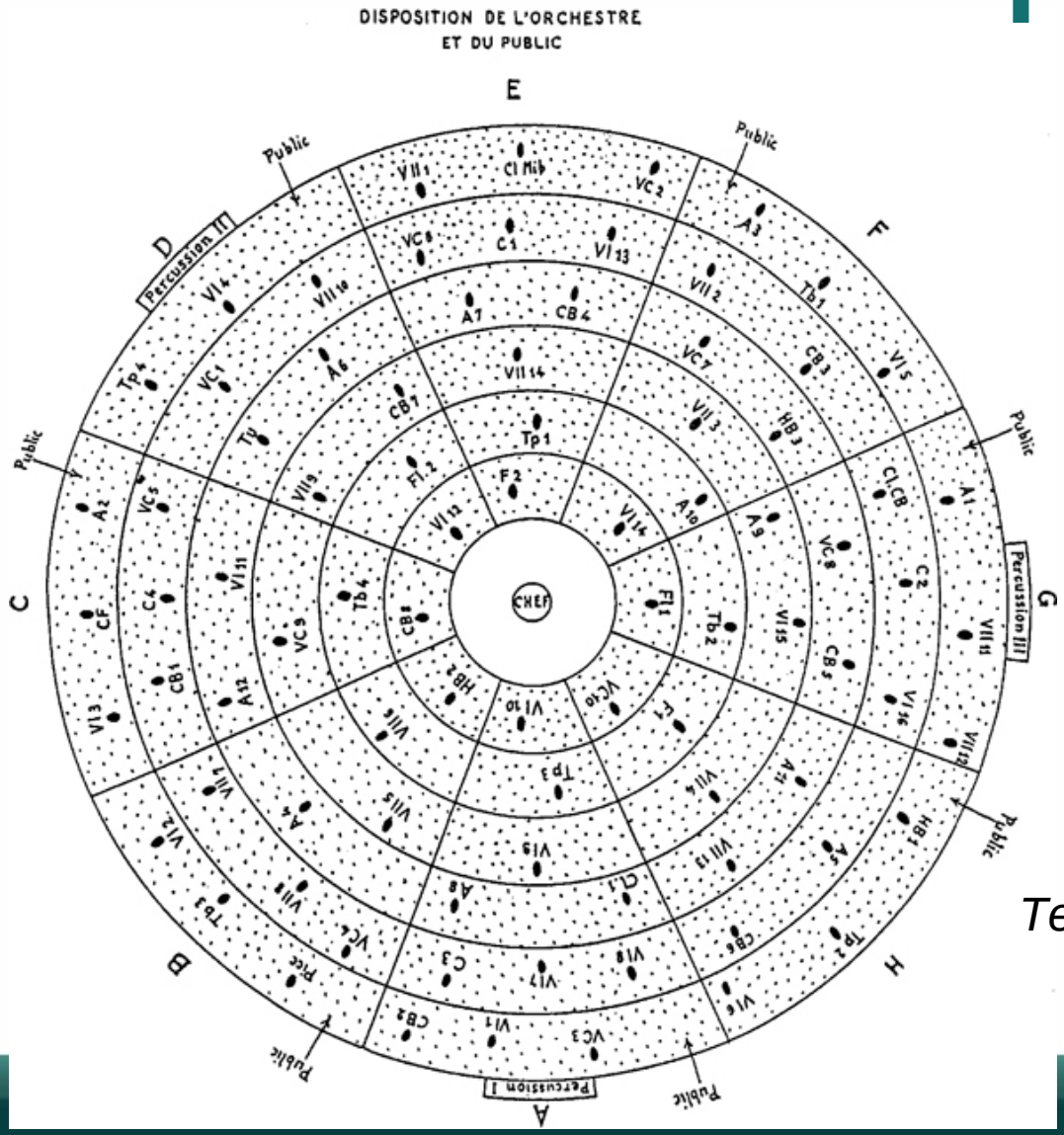
CAVERNA Digital

São Paulo Downtown 1913

- Virtual 3D city
- Sounds: cathedral, wind spots, people concentrations, birds in trees, background music following user
- Proximity loudness pre-calibrated for increased realism
- Partnership with Sao Paulo Historical Institutes



A desired future experiment



Terretektorth, I. Xenakis (1965-66)

Next research targets

- Wave Field Synthesis codecs for OpenAUDIENCE
- Loudspeakers clusters (hardware and software)
- Transport mechanisms for high channel density
  - Audio networks and Wireless solutions
- MPEG top-down applications combining
  - FTV (FreeView Point TV)
  - MAF (Multimedia Application Formats)
    - Portable applications Formats
  - Lossless codecs
  - Reconfigurable Audio Codecs
  - SMR (Symbolic Music Representation)
  - MPEG-21 (Digital item)
  - MPEG-7 Audio descriptors

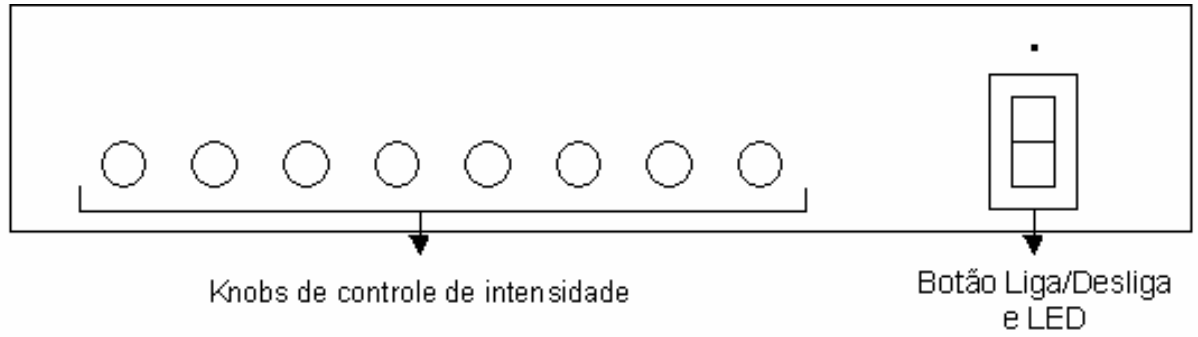
Laboratório de Sistemas Integráveis Tecnológico

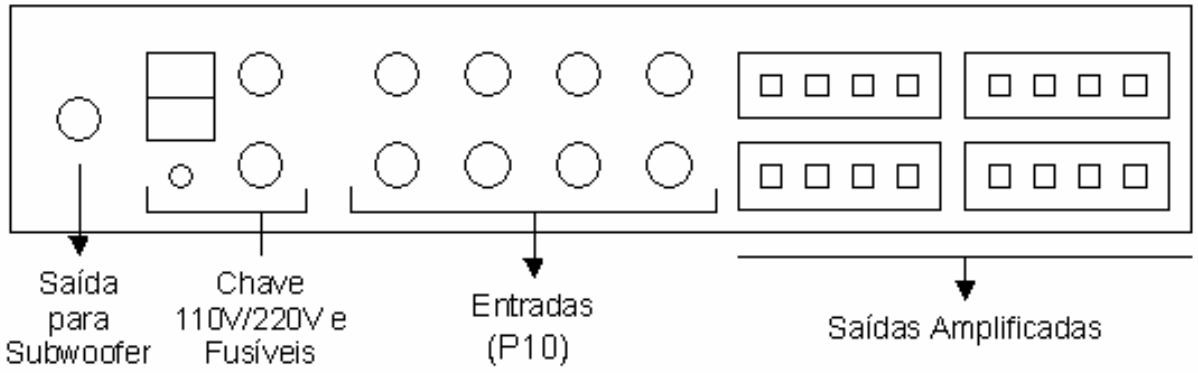
Other R&D areas and activities

- other r&d lines and activities includes
  - Support other centers in audio related technologies design and implantation of systems
  - Audio electronics design
    - Joint development of Multichannel Hw & Sw
    - Multichannel amplifiers prototyping

Multichannel Amplifiers Design

- Design and prototyping of multichannel eletronic amplifiers





Multichannel Amplifiers Design

- Novelties in conectorization
  - Input P10
  - Output “bornes” Non existent in market!

Laboratório de Sistemas Integráveis Tecnológico

Multichannel Amplifiers Design

- Assembled prototype, 30W/channel, 8 channel
- For 7.1 and octophonic loudspeakers setups
- Partnership with national industries (Sankya Eletrônica)

Laboratório de Sistemas Integráveis Tecnológico

Obrigado

Regis Rossi A. Faria
Núcleo de Engenharia de Áudio e Codificação
LSI – Laboratório de Sistemas Integráveis
Escola Politécnica da USP
regis@lsi.usp.br