

The EyesWeb XMI Open Platform for Multimodal Interaction

Giovanna Varni

*InfoMus Lab-Casa Paganini Intl Centre of Excellence
DIST-University of Genoa, Italy*

EyesWeb XMI

What :

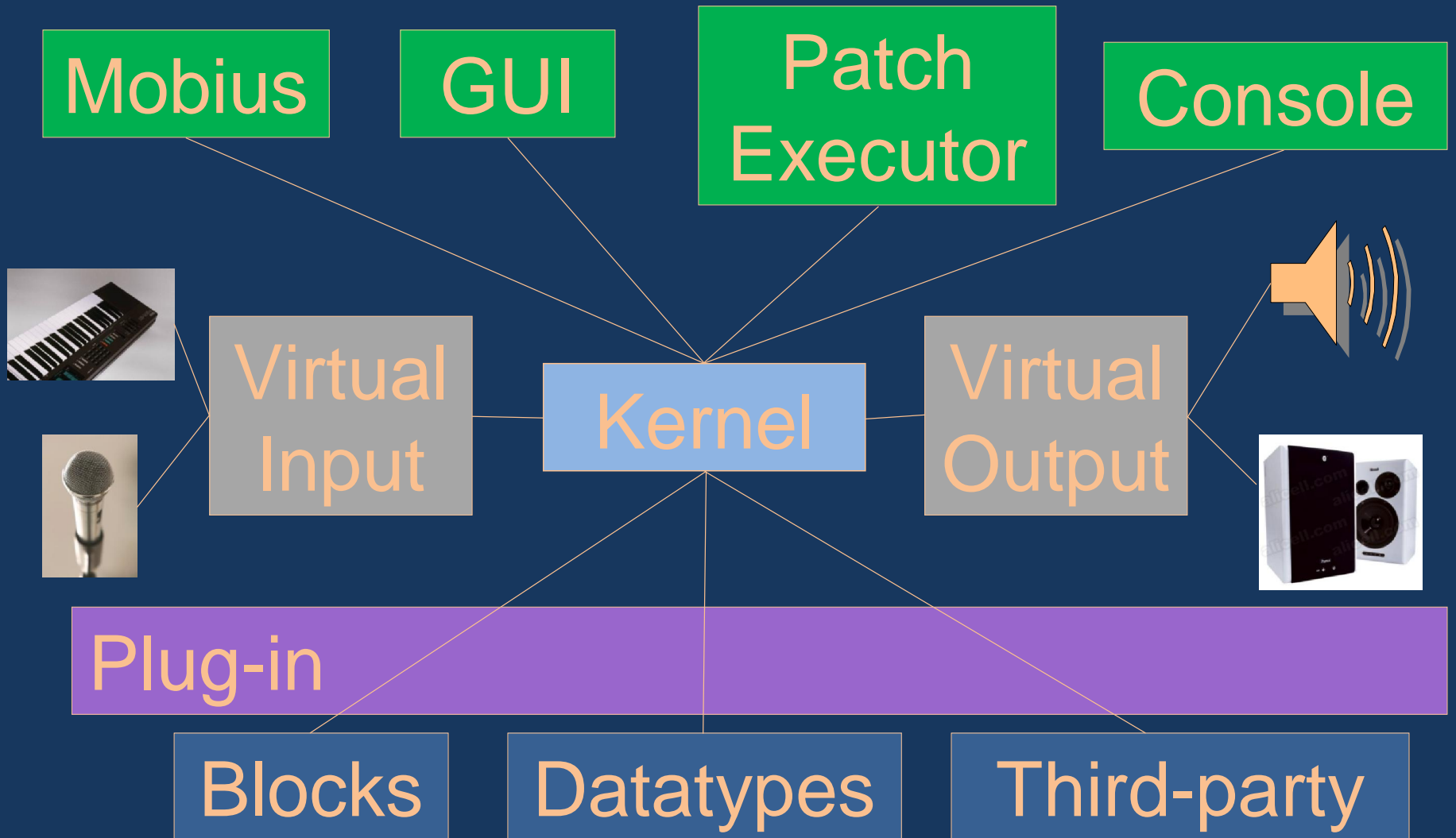
HW/SW platform for real-time multimodal processing

XMI (for eXtended Multimodal Interaction)

How :

- Supporting research on multimodal Human-Computer Interaction
- Integration analysis coming from different sensorial modalities
- Integration novel paradigms for multimodal interaction
- Experiments on cross-modal techniques

EyesWeb XMI Architecture



EyesWeb XMI Architecture - Kernel

- Registration and organization of plugins into a coherent set of libraries
- Persistence of user designed applications
- Mapping of physical devices into logical ones
- Execution engine

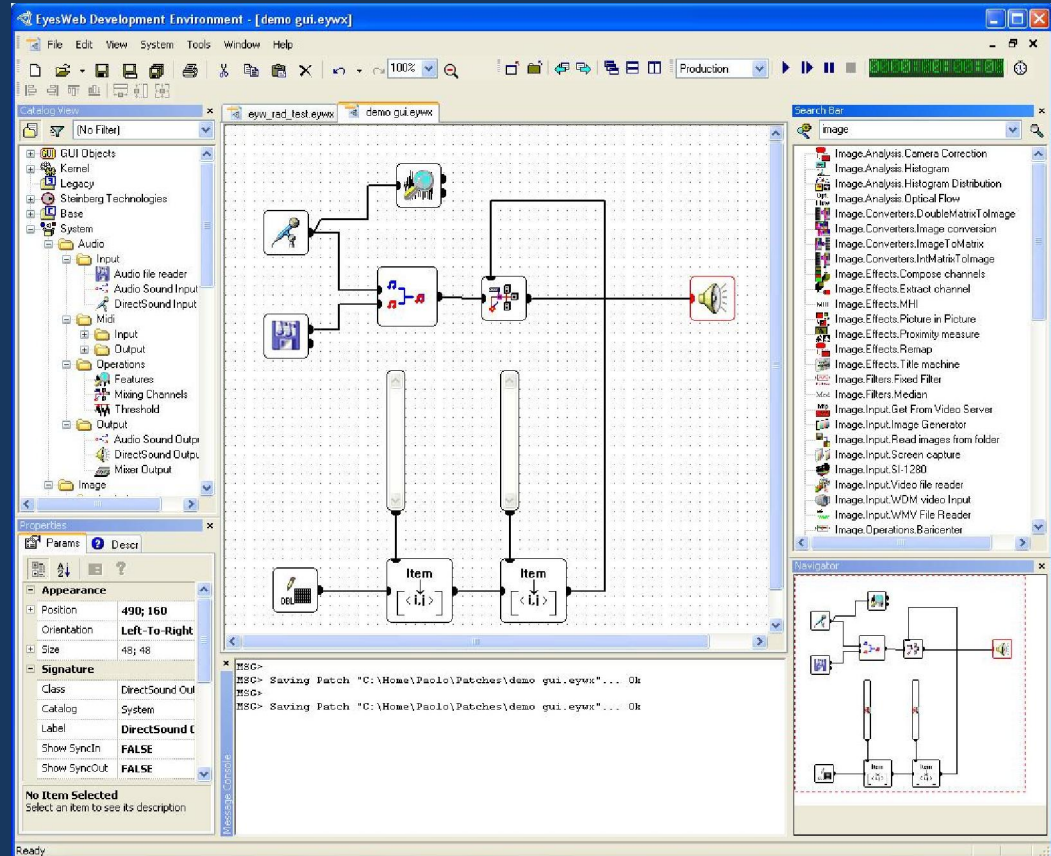
EyesWeb XMI – Execution Engine

- Management of the actual execution of patches
- Handling of:
 - data flow and synchronization
 - notification of events to the user interface and logging system
- Timestamping mechanism

EyesWeb XMI Architecture - GUI

Key words :

- Design patches
- Inplace execution
- Search/navigation

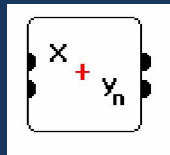


EyesWeb XMI Architecture – Block and Patches

Block

Conceptual unit

Example :

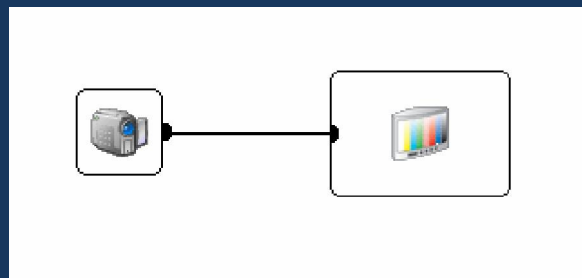


Properties	
Params	Descr
Position	490; 330
Orientation	Left-To-Right
Size	64; 64
Class	Arithmetic
Catalog	Kernel
Label	Arithmetic_1
Show SyncIn	FALSE
Show SyncOut	FALSE
<input type="checkbox"/> Active	TRUE
<input type="checkbox"/> Activate	
<input type="checkbox"/> Run level	Production
<input type="checkbox"/> SyncIn mode	Additive
<input type="checkbox"/> Operation type	Sum
<input type="checkbox"/> Number of inputs	2

Patch

Set of linked blocks

Example :



EyesWeb XMI – Patch Executor

- Scheduler of multiple patches
- Simple to use as it does not show the details of the patch

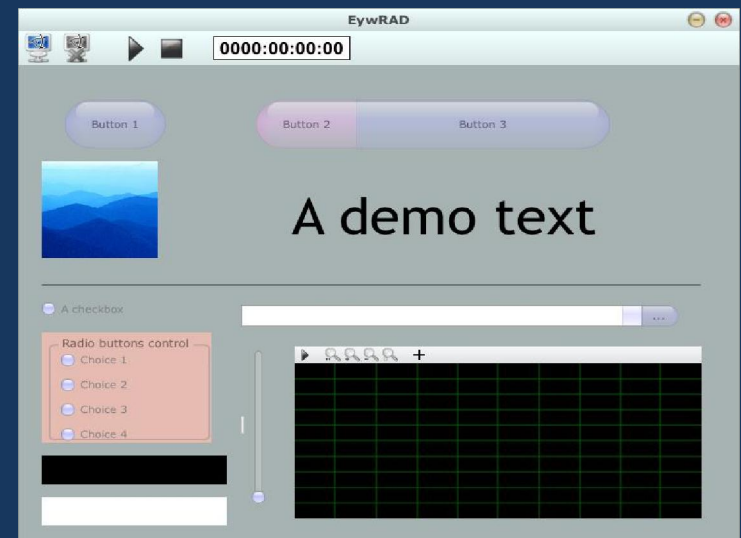
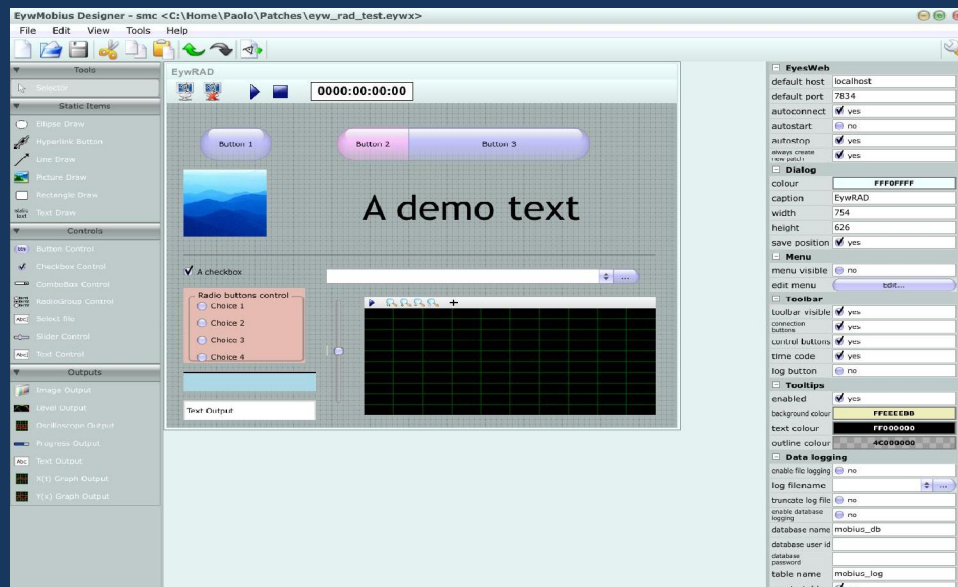
EyesWeb XMI – Console

- Command line-based interface
- Can runs as a Window service
- Supporting for EyesWeb Mobius

EyesWeb Mobius

Designed and developed by InfoMus Lab for
TRIL Centre

- Simple user interface designed for EyesWeb patches
- Runtime also available under Windows CE



EyesWeb XMI Features

- Support of a large number of external devices
- Exploitation of multi-processors/multi-core architectures
- Reduced footprint of Kernel
- Different run-levels (production/design)
- Synchronization

EyesWeb XMI – Supported technologies

- Communication : TCP, UDP, OSC, HTTP, COM/Serial
- Audio: ASIO, DirectSound, VST
- Video: FreeFrame plugins
- Other : MIDI
- Devices: High-speed cameras (e.g. Silicon Imaging), PTZ cameras, HTTP cameras, GPI devices

EyesWeb XMI – Libraries

Examples :

- Expressive Gesture Processing Library:
 - **MotionAnalysis**: motion trackers (e.g., LK), movement expressive cues (QoM, CI, ...)
 - **SpaceProcessing**: processing of (physical or abstract) spaces. Use of metaphors from physics: hysteresis, abstract potentials, superposition principle ...
 - **TrajectoryProcessing**: processing of 2D (physical or abstract) trajectories (e.g., kinematics, directness, ...)
 - **Machine learning techniques** (e.g., SVMs, clustering, neural networks, ...)

EyesWeb XMI – Libraries

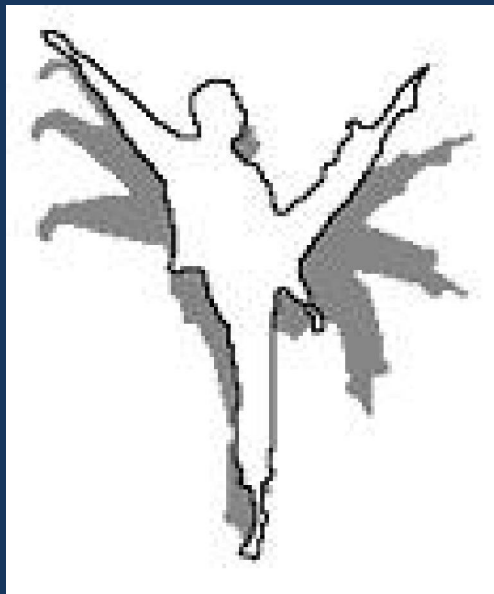
Expressive Gesture Processing Library.....

- Global measures depending on full body movement (e.g., body orientation, overall motion direction)
- Measures from psychological research (e.g., Boone & Cunningham's amount of upward movement)
- Cues from Rudolf Laban's Theory of Effort (e.g., directness, impulsiveness)
- Cues derived from analogies with audio analysis (e.g., Inter Onset Intervals, frequency analysis)
- Kinematical measures such as velocity, acceleration, average and peak velocity and acceleration.

SMI and Quantity of Motion

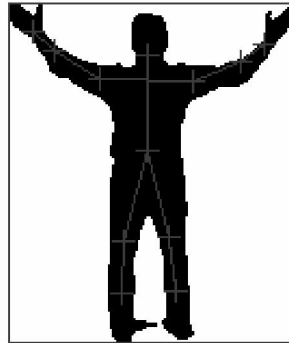
SIMs (Silhouette Motion Images) carry information on variations of the silhouette in the last few frames.

SIM's area (Quantity of Motion) is a measure of the detected amount of motion.



Contraction Index

A measure of how the subject's body uses the space surrounding it.



EyesWeb XMI – Libraries

- Expressive Cues Dynamics Analysis Library:

Focus on extraction of information about the dynamics of motion expressive cues

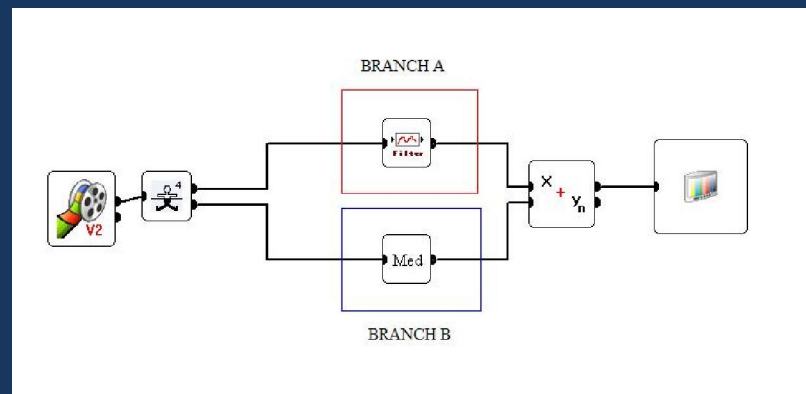
- initial and final slope
- initial and final slope of the main peak
- maximum value
- ratio between the maximum value and the duration of the main peak
- mean value
- ratio between the mean and the maximum value

EyesWeb XMI – Libraries

- ratio between the absolute maximum and the biggest following relative maximum
 - centroid of energy
 - distance between maximum value and centroid of energy
 - symmetry index
 - shift index of the main peak
 - number of peaks and number of peaks preceding the main one
 - ratio between the main peak duration and the whole profile duration
-
- **Tangible Acoustic Interface Library :**
 - a separate library from TAI-CHI EU IST
 - partially supported by the TAI-CHI EU IST Project (2004-2007)

EyesWeb XMI Features - Parallelism

Topological parallelism → parallel execution



Parallel execution of branches A and B!

Automatic synchronization at convergence points!

EyesWeb XMI - Synchronization

Multimodal processing.....

streams ability to handle several data

different clock sources

- New mechanism for synchronization:

Sync-in pins

Sync-out pins

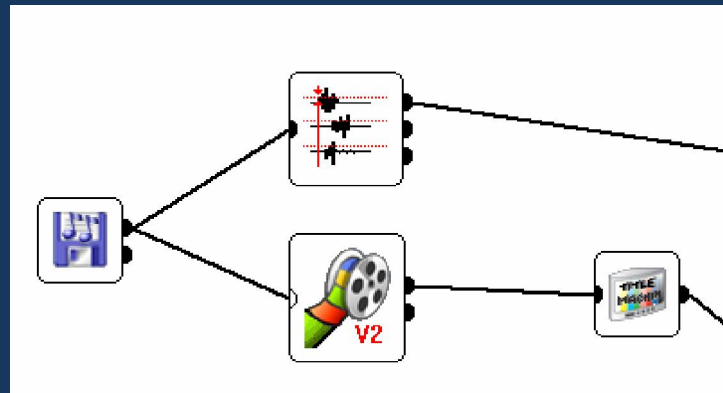
EyesWeb XMI – Sync-in pins

- Sync-in pins
 - make the clock source explicit
 - available for every block
 - exported on user's choice



EyesWeb XMI – Sync-in pins

A clock signal may be connected to this pin block is activated according to the speed of this clock source



Example :

The *Video File Reader* block is synchronized with an audio signal which acts as clock generator

EyesWeb XMI – Sync-in pins

Noteworthy things :

- Any type of data can act as a clock signal
- Sync-in pin connection does not cause data transportation :
 - only information needed for synchronization is propagated!

connection to sync-in pin does not imply higher CPU usage

EyesWeb XMI – Sync-in pins

Sync-in pins operation mode :

- overriding of the usual EyesWeb activation rule
- when a signal is connected to the sync-in pin, the block is activated when the clock signal is raised
- More advanced techniques :
 - the sync signal can be mixed with the native clock of the block
 - the activation rule of the block instead of being overridden is integrated with the clock signal

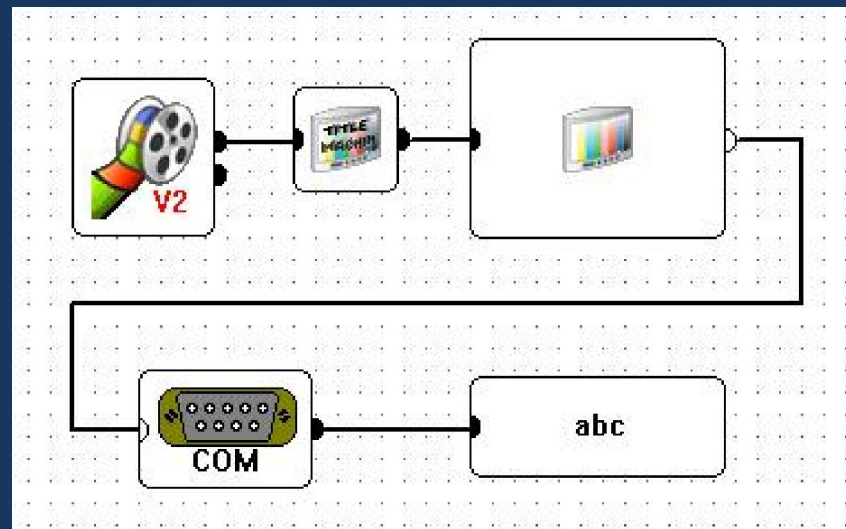
EyesWeb XMI – Sync-out pins

- The sync-out pin generates a clock signal, and this clock is raised each time the block is activated
- The clock signal generated by this pin is not directly related to the data stream generated
- The block might be a sink block



EyesWeb XMI Features – Sync pins

Example :



The *VideoDisplay* block, which has no outputs, exports a sync-out pin used as a clock source for a *SerialInput* block.

Casa Paganini - InfoMus Lab

- **Staff**
 - University personnel; PhD students; external staff (e.g. artists, visiting researchers); MS degree students
- **Activities:**
 - **Research:** national and intl contracts; EU Projects: Example
 - 5 FP IST: MEGA, CARE-HERE, COST ConGAS, TMR MOSART, ...
 - 6 FP IST: TAI CHI, NoE ENACTIVE and HUMAINE, S2S²...
 - **Education**
 - **Support to artistic production:** exploitation and validation of research outputs
 - **Industry contracts, applications, spin-off**

EyesWeb XMI Installations

- Science exhibition “Cimenti di Invenzione e Armonia”
(Casa Paganini, Genova, Italy, October 2006 - January 2007)
- The exhibition was part of “Festival della Scienza”, a huge international science festival held in Genova every year
- The exhibition was organized in a path in the auditorium and the museum rooms of Casa Paganini.

Tangible Acoustic Interfaces Room

TAI Google Earth application:

Interface and sensors

- 1 MDF surface
- 1 B/W video camera in front of the surface - 4 accelerometers

Techniques

- Movement of the hands of the user is tracked by means of multimodal integration of visual and acoustic continuous tracking techniques
- Kinematical and expressive features are extracted from user's gestures

Interaction

- The user can rotate and zoom in and out satellite images projected on the table

Sync-in pins were used to synchronize the results from the analysis performed on the audio and video stream

Tangible Acoustic Interfaces Room

TAI chair

Interface and sensors

- 1 chair
- 2 accelerometers

Techniques

- Acoustic localisation of touch position on the chair
- Processing of qualitative features of touch (e.g., with hard and quick tapping-like movements, or with light and smooth caress-like movements) to control sound generation and processing in real time

TAI chair was used in the music theatre opera “Un avatar del diavolo”, composer R.Doati, presented at La Biennale, Venezia, September 2005

TAI Room – Painting stand

- This application was composed of a painting stand with a white canvas, which was transformed in a TAI
- Gesture direction and impulsiveness was detected and used to control the playback of video sequences projected on the canvas. For example, a fast gesture toward the right triggers fast-forward reproduction of the video
- Expressive qualities of the performed gestures were mapped to local expressive variations of the reproduction (e.g., hesitant movement mapped to slower reproduction rate)

TAI Room – Connection of applications

- These applications were connected together and their activation was controlled through a small cell-phone like device
- Touching the TAI objects (the table, the chair, the painting stand) with the cell-phone provokes the recognition of the object from the noise produced by the touch of the cell-phone on the object and therefore causes the activation of the corresponding application. The cell-phone is also able to recognize other (passive) objects in the room (e.g. a metal handrail) from the noise caused by the touch/impact on the object itself

Orchestra Explorer

Physical navigation inside a virtual orchestra

- active exploration of the music piece the (virtual) orchestra is playing
- modification and moulding of the sound and music content in real-time.

By walking and moving on the stage, the user discovers each single instrument and can operate through her expressive gestures on the music piece the instrument is playing.

Orchestra Explorer

- Simultaneous playback and real-time processing of :
 - 26 mono audio tracks
 - expressive full-body movement and gesture
- Different tuning parameters for different physical places on the stage
- The whole application was run on a single workstation

EyesWeb XMI Installations

Three permanent installations at the Museum "Pavillon de la Roche d'Oëtre" (Roche d'Oëtre, Normandie).

The interactive installations, based on XMI and since summer 2006, concerns an "acoustic telescope", a "journey on time" using audiovisual, meteo and environment information, and an installation based on a large map of Normandie on the floor that can be navigated in real time by one or more users, who are tracked and can access and manipulate multimedia content on the regions and sites they visit by wandering on the map in the floor

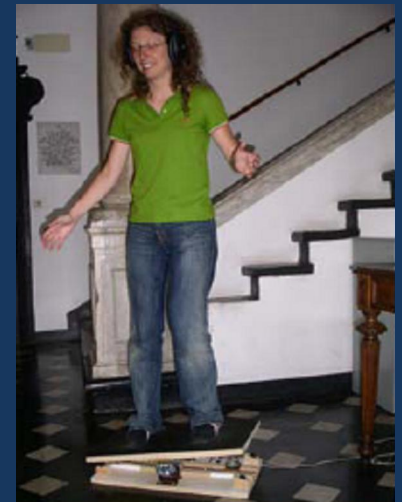
Enactive Interface



- Action-driven interfaces
- NOT symbolic or iconic knowledge
- BUT enactive knowledge :
 - natural and intuitive
 - based on experience
 - stored in the form of motor responses
 - acquired by the act of “doing”
 - closing the action-perception loop

ENACTIVE experiment on audio referents

- Investigation on dynamic auditory patterns:
 - control of human action
 - referents for orientation
 - perception level of movements
- Reaching target task
 - controlling standing medio-lateral body sway
 - dynamic acoustic information on orientation
 - achievement of specific acoustic patterns in acoustic stimuli



Links

- <http://www.eyesweb.org>
- <http://www.casapaganini.org>
- <http://www.infomus.org>

Contact

- info@infomus.org

....Questions?

.....Comments?

Thank you for your attention!