Sound rendering in Interactive Multimodal Systems

Federico Avanzini

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Sound Music Computing

Outline

- Background
  - Ecological Acoustics
  - Multimodal perception
- Auditory-visual rendering of egocentric distance
  - Binaural sound
  - Auditory cues for distance
  - Bimodal rendering: recent results, ongoing work
- Multimodal rendering of virtual objects
  - Synthesis of environmental sounds
  - Multimodal rendering architectures
  - Haptic-audio: recent results, ongoing work

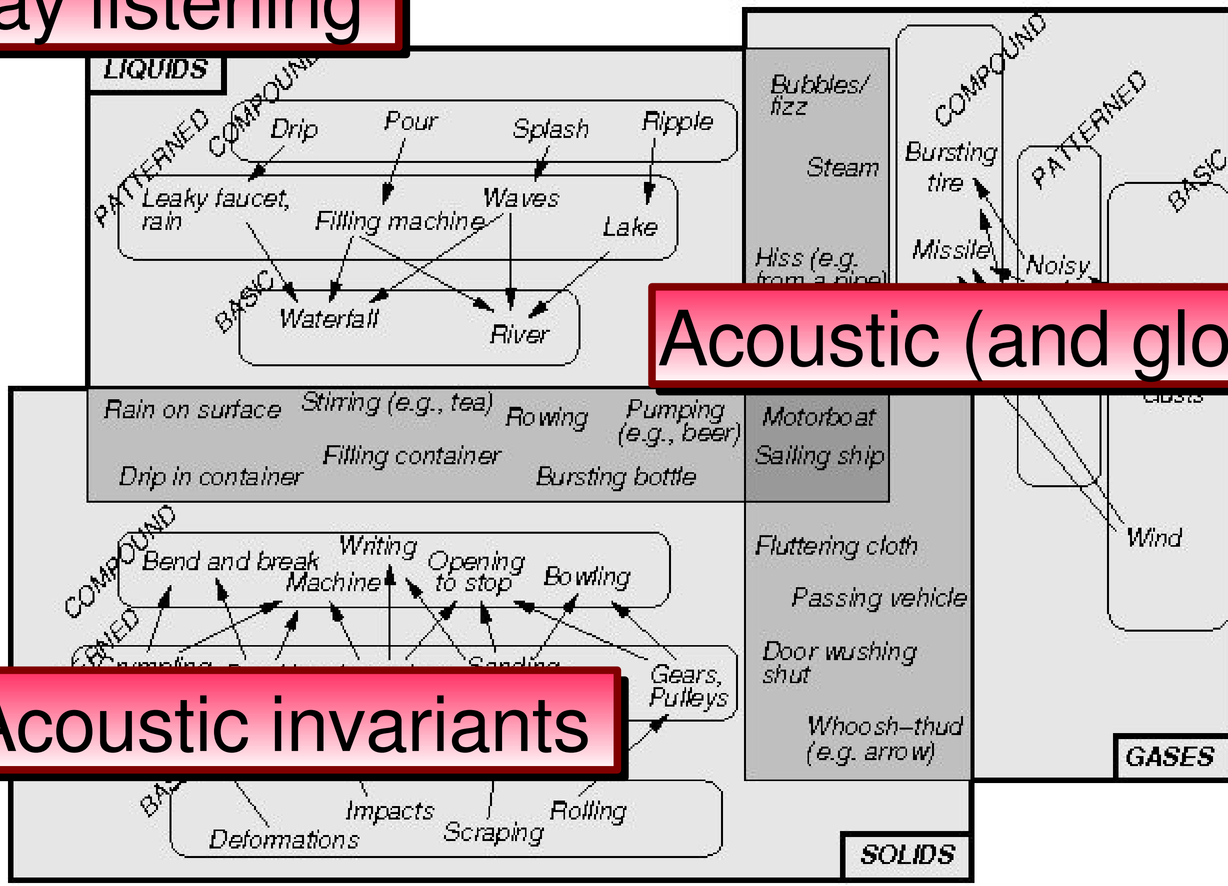
Ecological psychology, Ecological acoustics

Sound and Music Computing

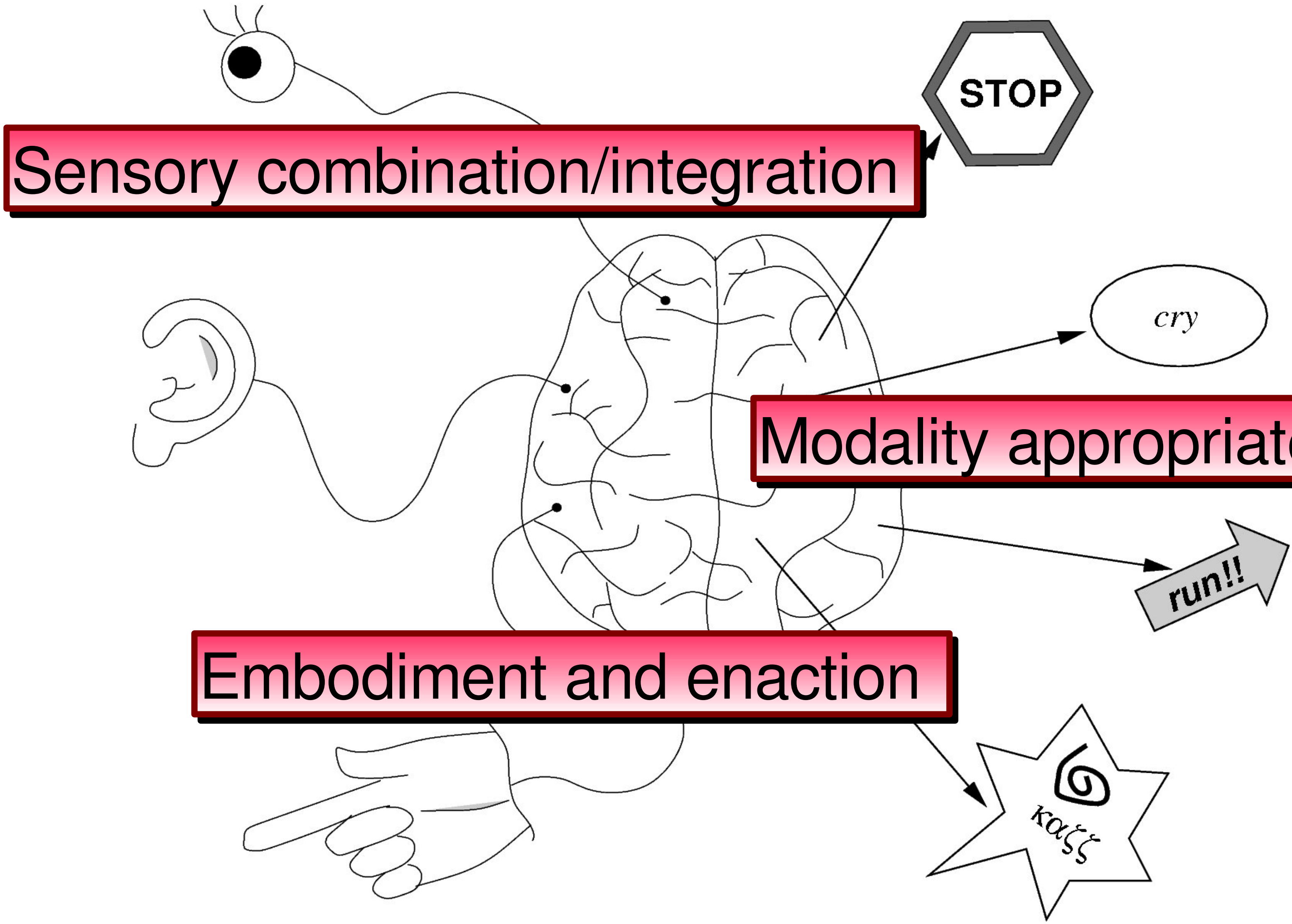
Everyday listening

Acoustic (and global) array

Acoustic invariants

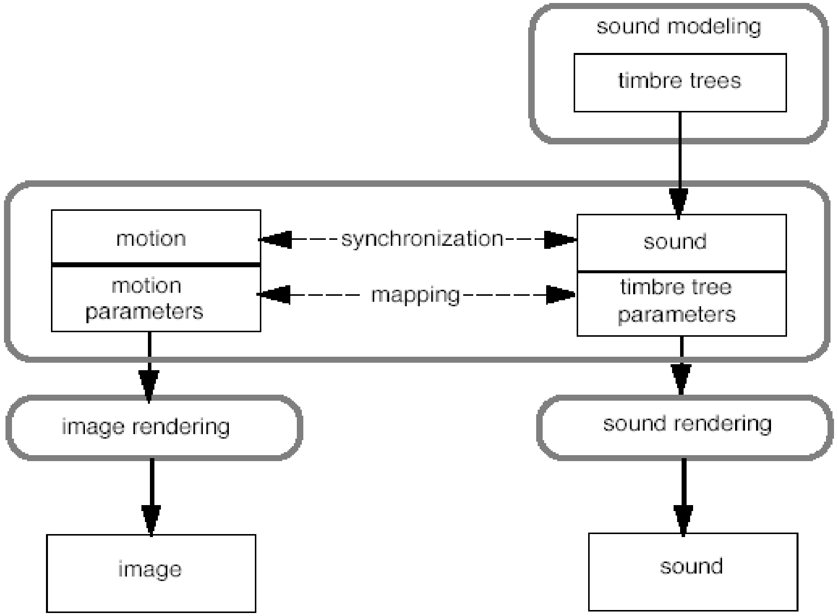


Multimodal perception



The sound rendering pipeline

- Modeling
  - Sound sources: resonating objects, interactions, ...
  - Physical modeling
- Syncronization
  - With other modalities: mappi from motion to sound params latency, ...
  - Physical modeling
- Rendering
  - Generate sound signals from models, in a given environment

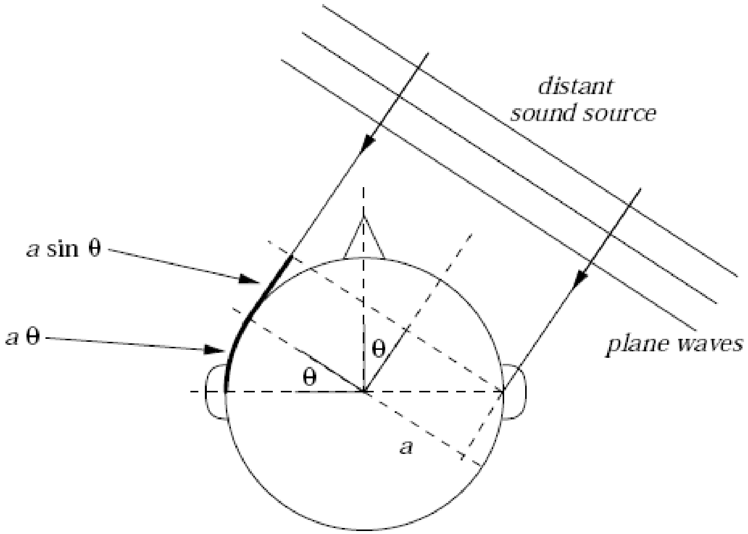


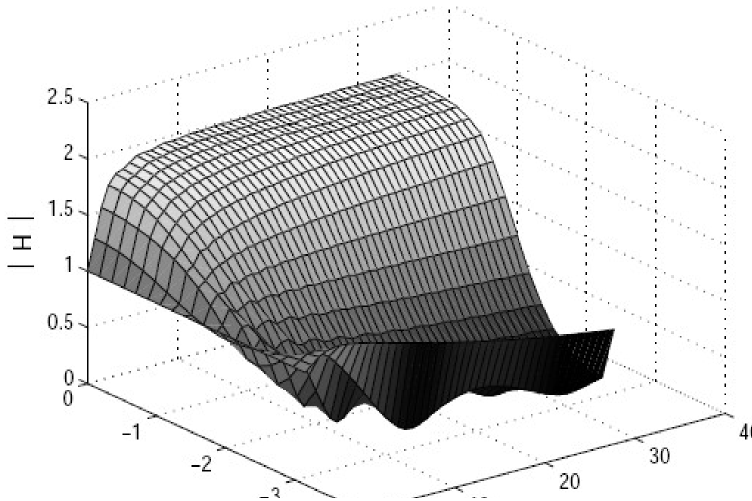
J. K. Hahn, et al. Presence 1998

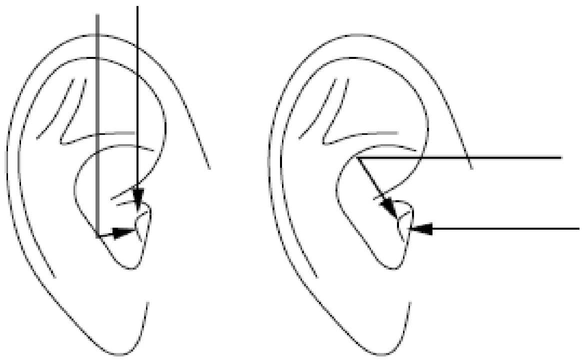
Sound and Music Computing

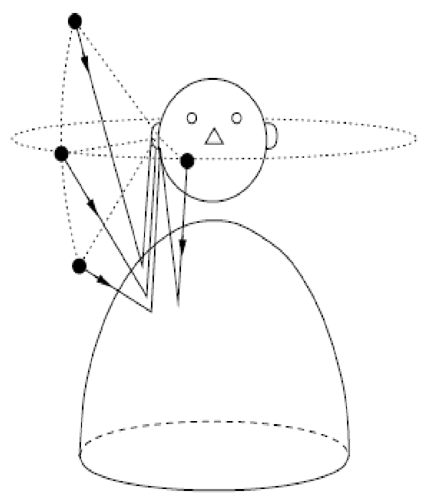
Auditory-visual rendering of egocentric distance

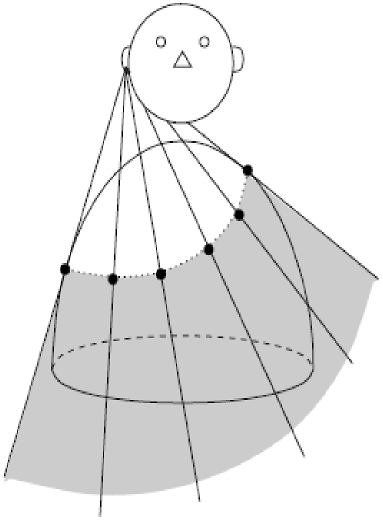
Binaural sound



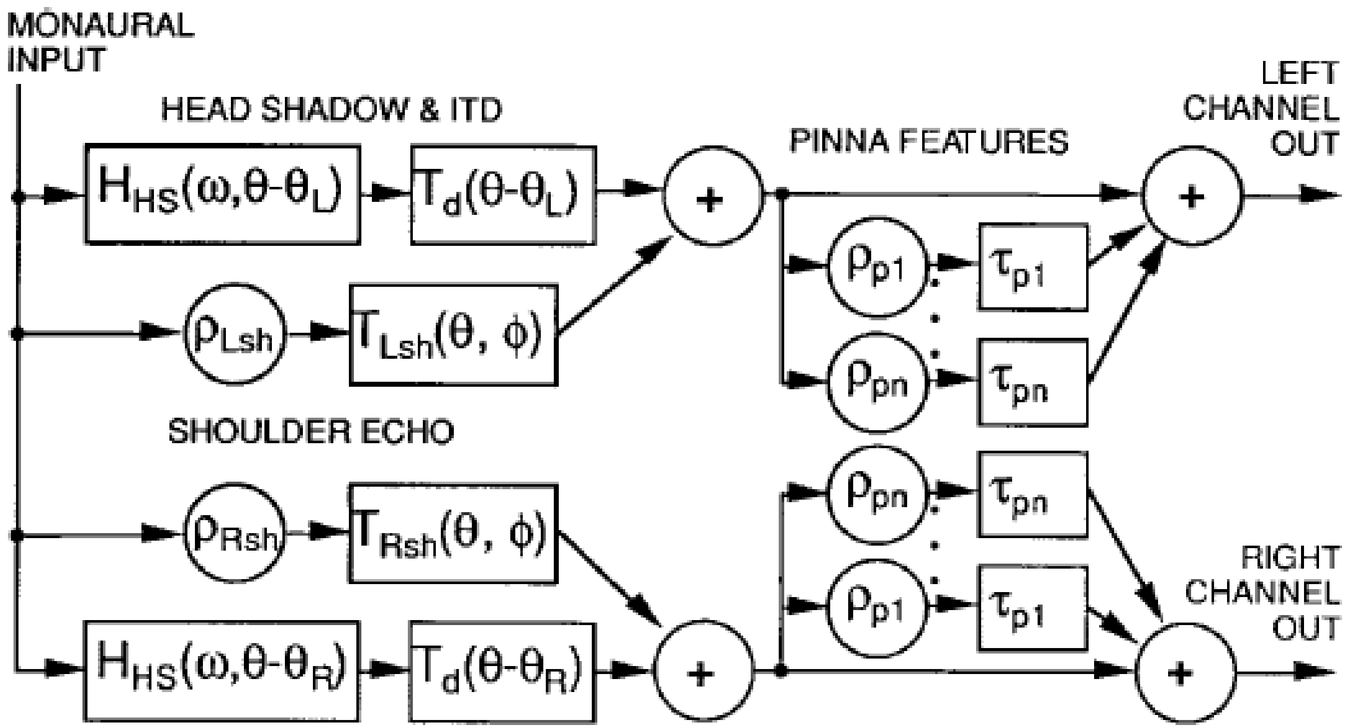








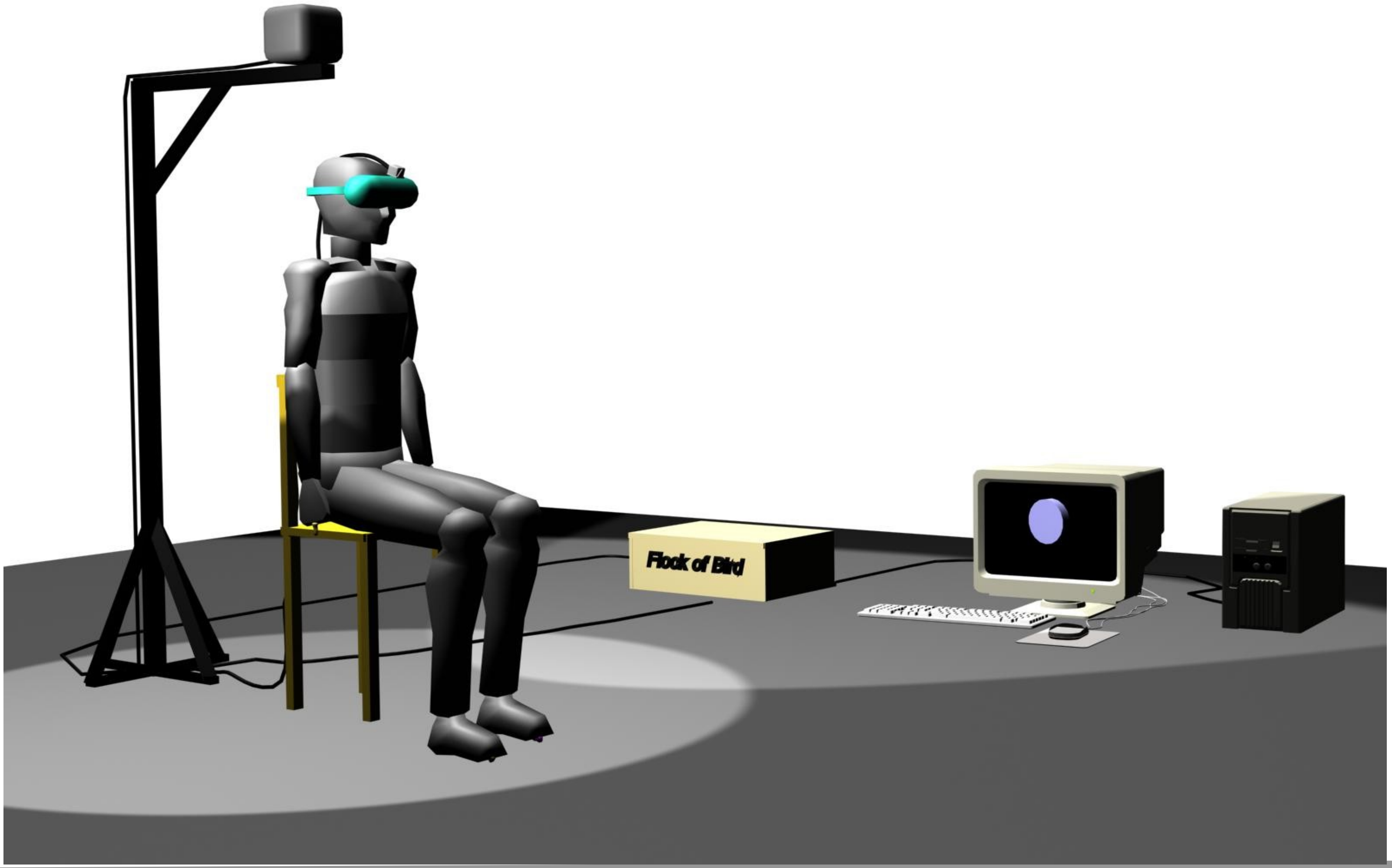
- Structural models (Duda and coworkers)



Auditory-visual rendering of egocentric distance

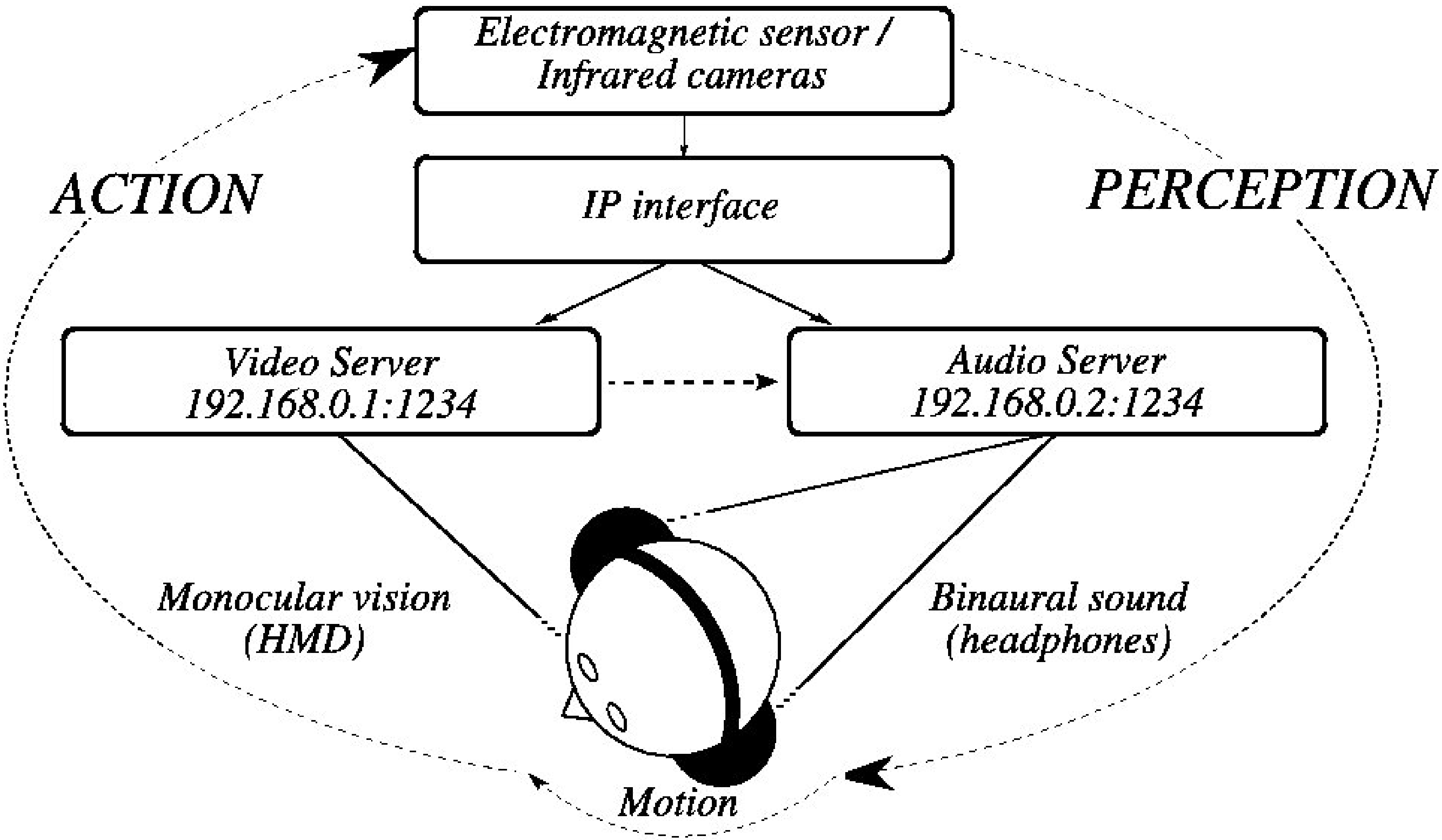
- Estimation of egocentric distance
  - Is an “enactive” task
  - Verbal judgements vs. action-driven judgments
- The problem of distance rendering is much studied in VR applications
- Auditory distance rendering is hard, auditory cues to distance are ambiguous
  - Static cues: intensity, R/D ratio, familiarity, distance-dependent spectral cues, auditory parallax
  - Dynamic cues: motion parallax, “acoustic tau”

Experimental setup

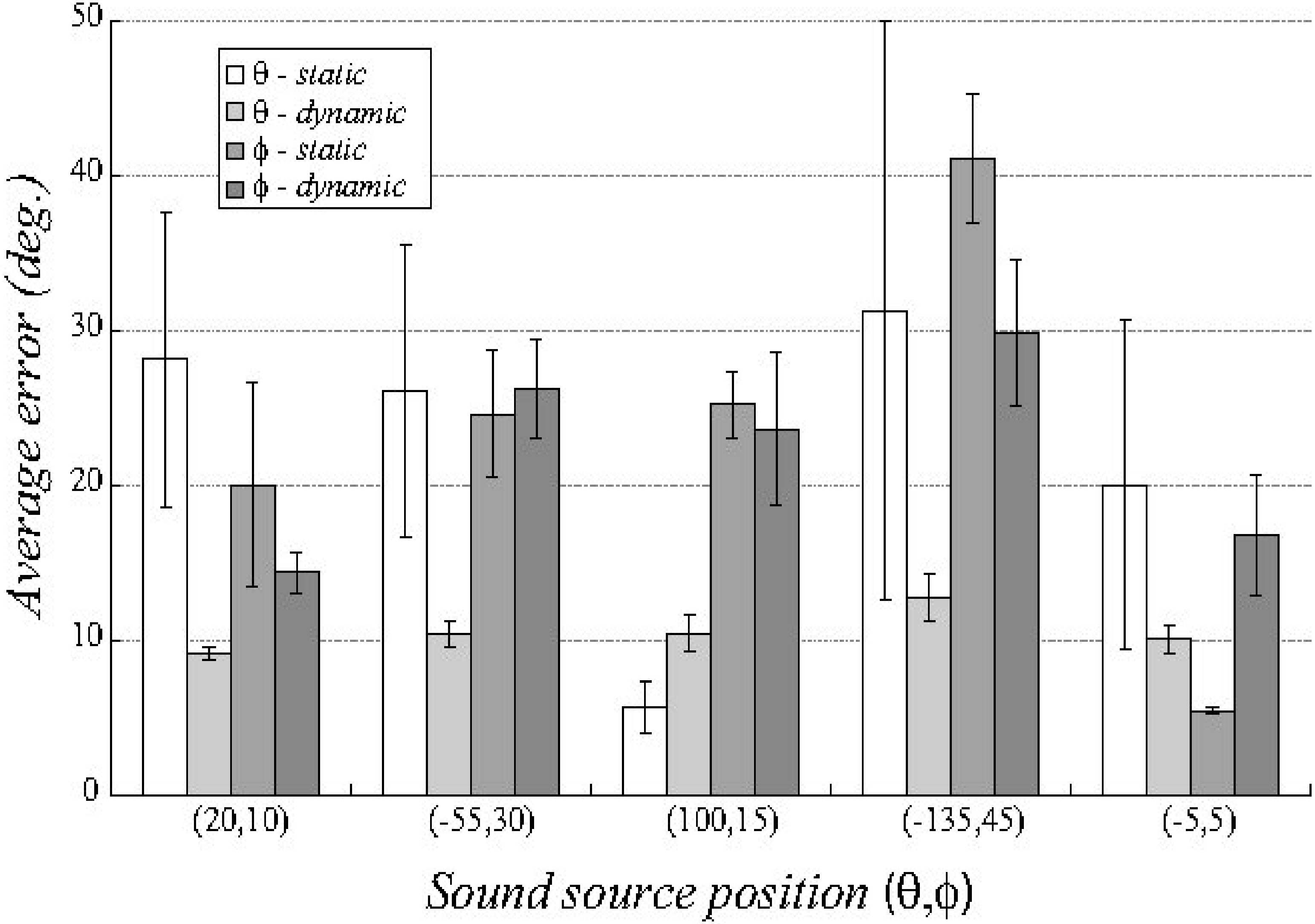


Sound and Music Computing

The rendering system



Preliminary exp. results: azimuth & elevation



Preliminary exp results: distance

- 10 subjects
- 9 stimuli (9 distances) presented in 3 conditions (visual only, audio only, audio-visual), and in active/passive settings
- Addition of audio improves distance estimation considerably
  - Even in “passive” settings (??)

Sound and Music Computing

Multimodal rendering of virtual objects

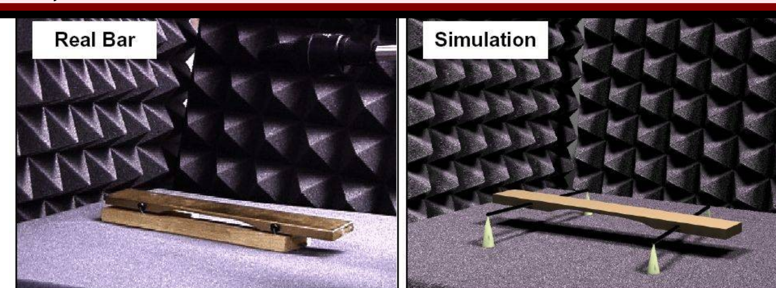
Sound Synthesis

- Contact sounds

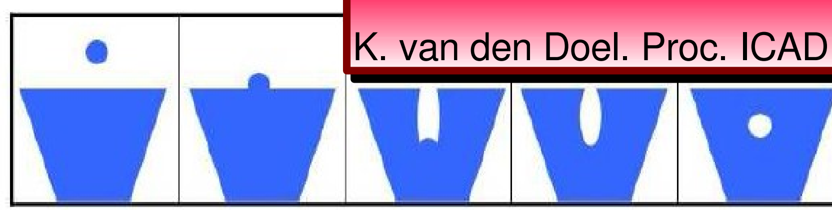
K. van den Doel, P. G. Kry, and D. K. Pai. Proc. ACM SIGGRAPH 2001

  - Modal synthesis
  - Finite Element Models
  - Other structures (e.g., WGs) are less used

J. F. O Brien, P. R. Cook, and G. Essl. Proc. ACM SIGGRAPH 2001

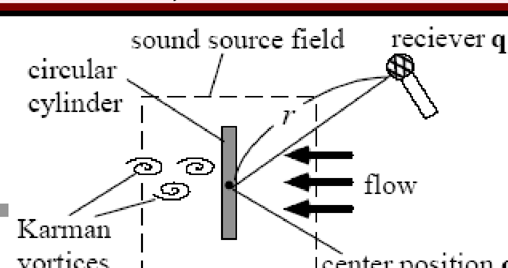


K. van den Doel. Proc. ICAD 2004



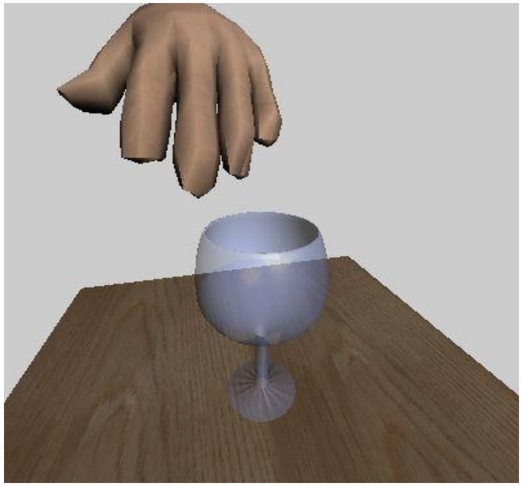
- Other typologies
  - Liquids
  - Aerodynamic sounds
  - More research needed

Y. Dobashi, T. Yamamoto, and T. Nishita Proc. ACM SIGGRAPH 2003



Sound models

- Basic sound models
  - Modal resonators
  - Non-linear interactions (impact, friction)
- Building blocks for *patterned* and *compound* environmental sounds
  - Toward a hierarchical taxonomy of sound models
- Link to ecological dimensions
  - Mappings from sound features to ecological dimensions are known
  - Need for mappings from physical parameters to relevant sound features



Sound and Music Computing

Physical models and ecological mappings

- Modal synthesis

$$\ddot{q}_n(t) + g_n\dot{q}_n(t) + \omega_n^2 q_n(t) = \frac{1}{m_n} f_A(t)$$

  - modal frequencies depend on *shape* and geometry
  - *material* determines decay characteristics
  - amplitudes of partials depend on strike location
- Impact force (Hunt-Crossley - *ASME J. Applied Mech.* 1975)

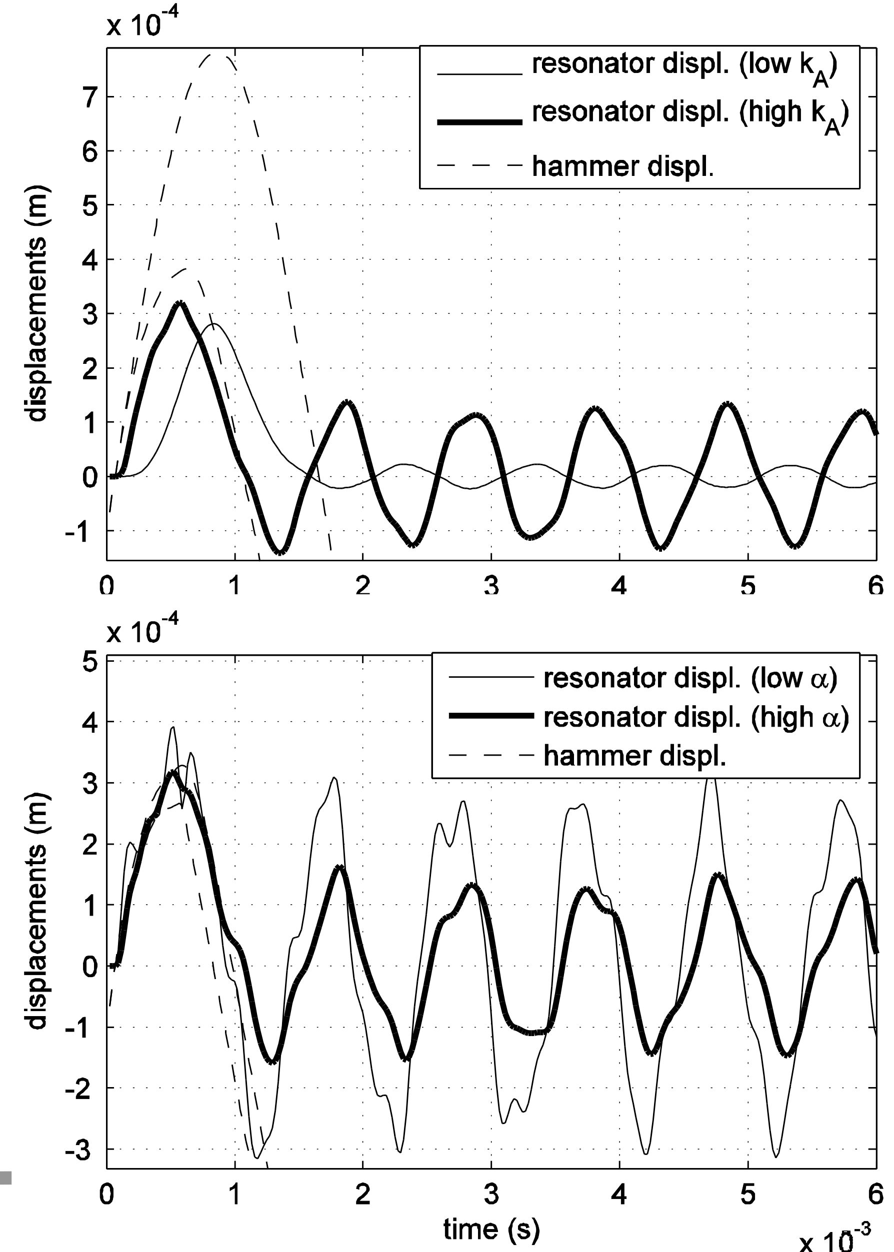
$$f_A(x(t), v(t)) = \begin{cases} k_A x(t)^\alpha + \lambda_A x(t)^\alpha \cdot v(t) & x \geq 0, \\ 0 & x < 0, \end{cases}$$

  - Parameters of contact force provide control over other ecological dimensions of the sound events
  - E.g., auditorily perceived stiffness.

Examples of attack transients

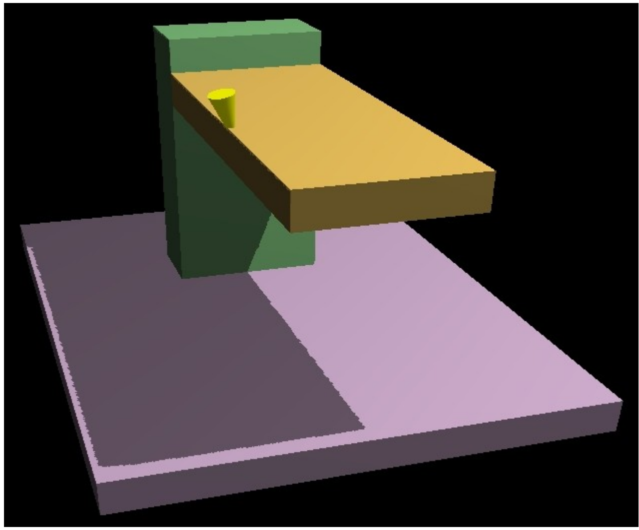
- Varying the force stiffness

- Varying the force exponent



Multimodal rendering of virtual objects

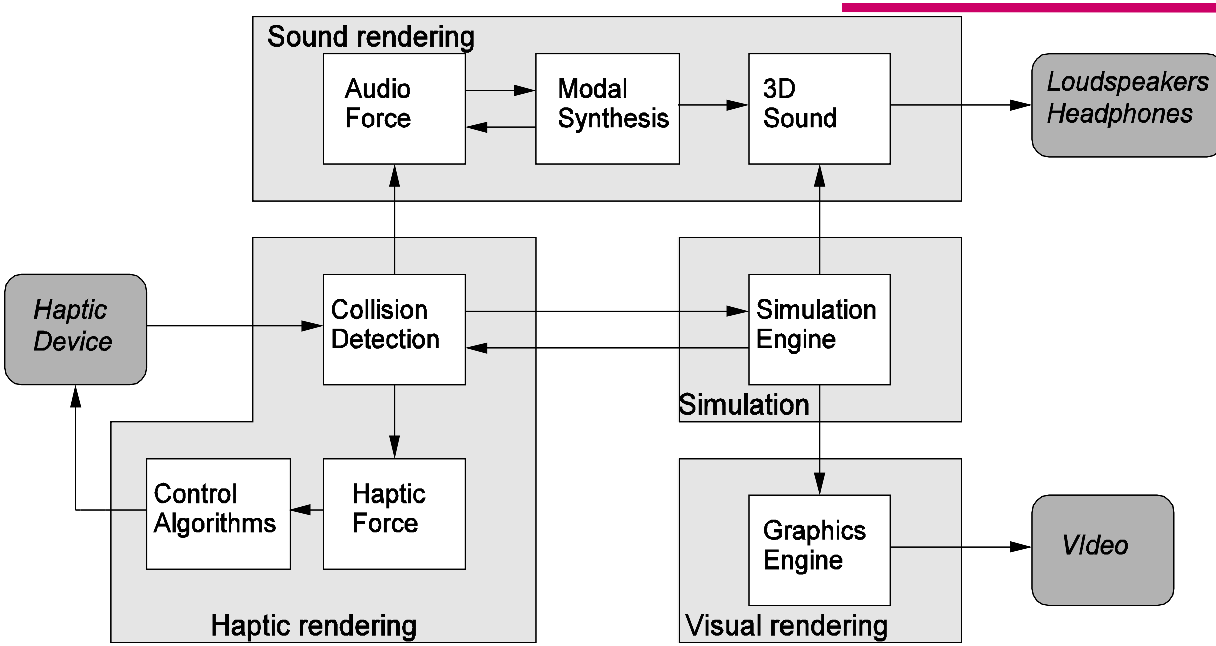
- Immersion
  - Properly designed/synchronized multimodal display provides greater immersion in a VE than a hi-fi visual display alone.
  - The range of object properties conveyed to the user can be enhanced



- Physical models
  - Interactivity and ease in associating motion to sound control
  - Sound-rendering attributes can be incorporated into data structures that provide multimodal encoding of object properties

Sound and Music Computing

A multimodal rendering architecture

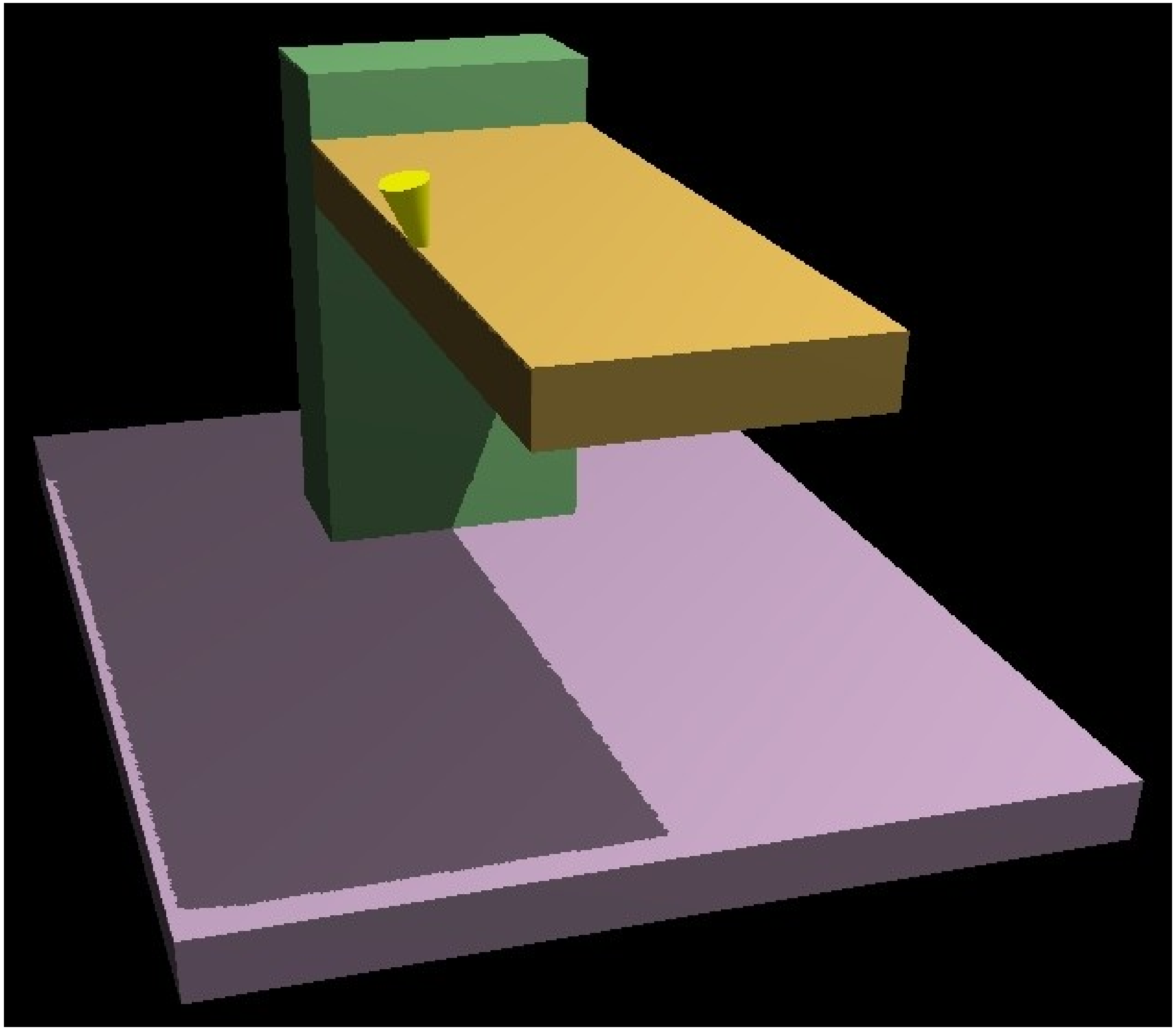


- Need for synchronization and low intermodal latency
- Sound computation triggered by contact detection
- Contact force computation embedded in the audio rendering thread
- Current (well... future...) work: integrate sound models into unified scene graph APIs

Bimodal perception of stiffness

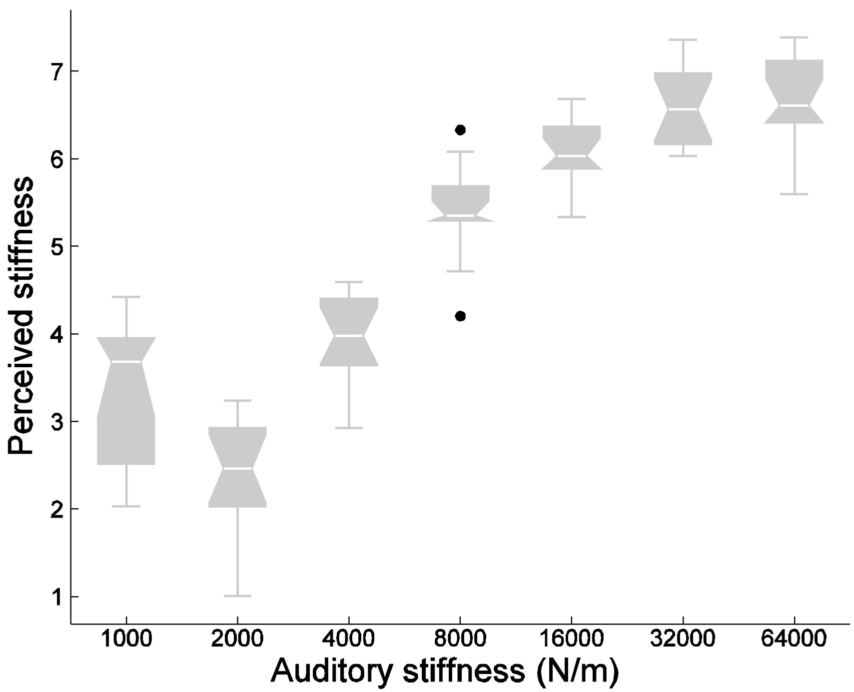
- Goal
  - Assess relative contributions of haptic and auditory information to bimodal judgments of *contact stiffness*
  - It's the prototypical haptic task!
- Stimuli
  - Graphic display does not change between conditions
  - In every condition same (400N/m) haptic stiffness
  - Auditory stiffness is the only parameter of the sound model which is varied between stimuli (7 levels)
- Procedure
  - 18 participants instructed to judge stiffness through absolute magnitude-estimation (scale from 1 to 8)

An experimental session

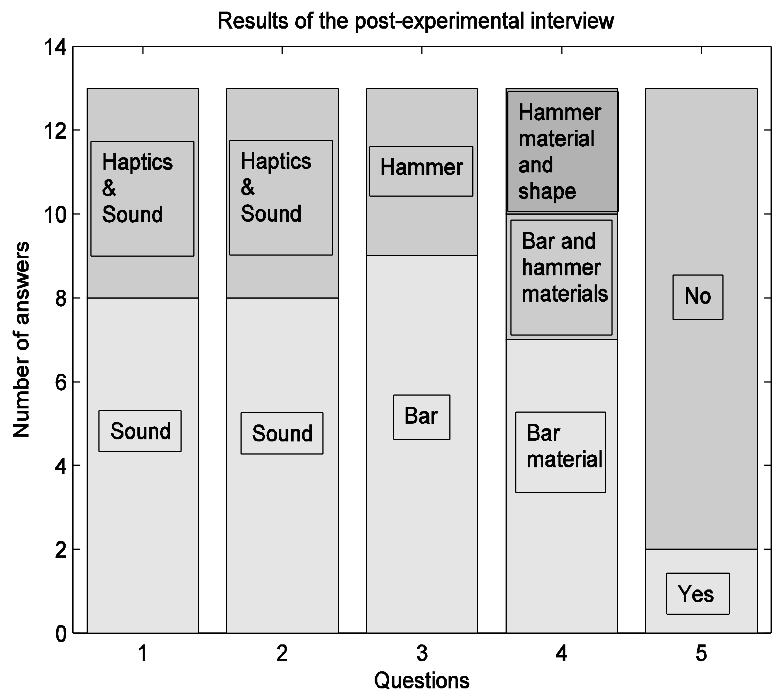


Results

- Magnitude estimates
  - Averaged across repetitions
  - Normalized with respect to grand mean across participants

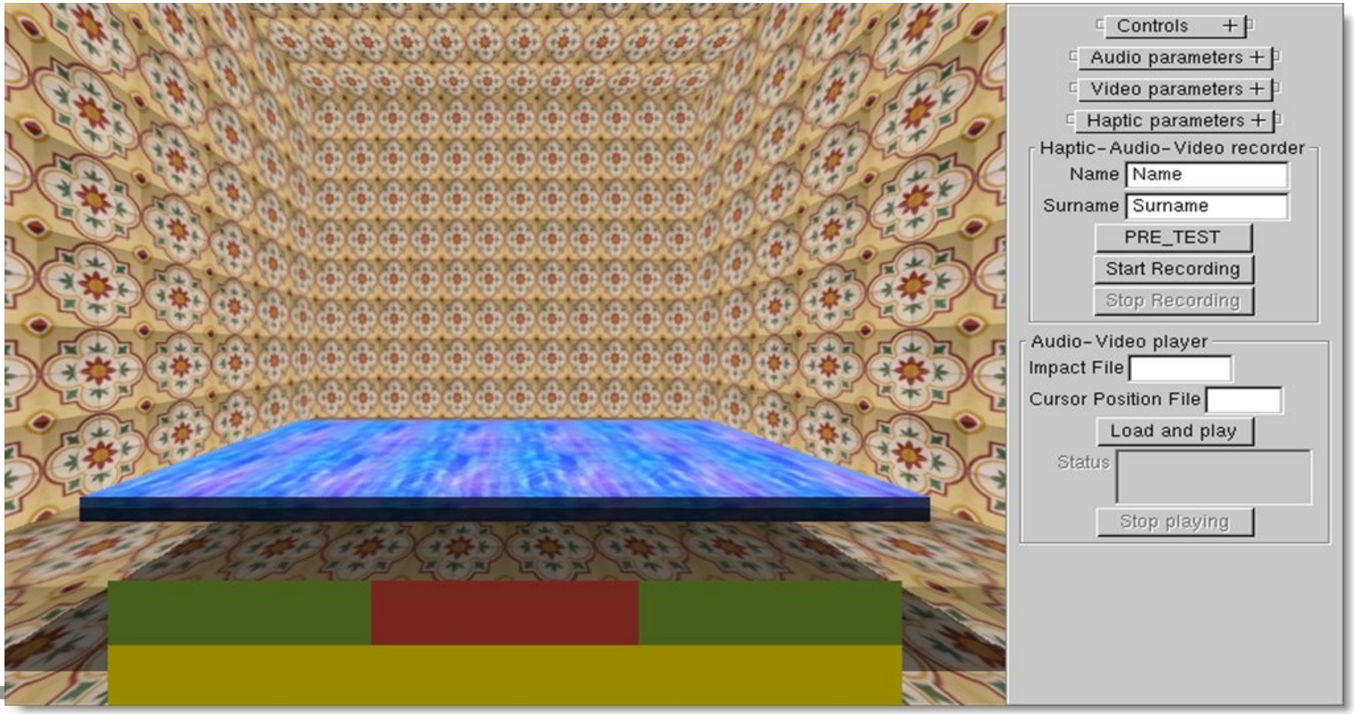


- Post-experimental interview
  - Auditory illusion: haptic stiffness was perceived to vary



Sensory integration in percussion performance

- Focus on performance
  - We avoid judgements based on verbal descriptions
  - We focus on knowledge generated through action
- Task
  - Strike an object with given impact force & given tempo



Sound and Music Computing

Experimental design

Training	Adaptation	Modification
Display A	*Performance criterion reached* / *Display A*	*Display B:* visual/haptic/auditory change

Real-time target force & tempo tracking

Feedback on force	No feedback on force	No feedback on force

Striking tempo feedback (metronome)

- Unimodal change: haptic k, auditory k, visual speed
- Performance variation induced by display change proportional to perceptual relevance of interested modality
- Possible extension: study performer kinematics