
WPI

# IMGD 5100: Immersive HCI

## Augmented Reality

**Robert W. Lindeman**
Associate Professor
Interactive Media & Game Development
Department of Computer Science
Worcester Polytechnic Institute
gogo@wpi.edu

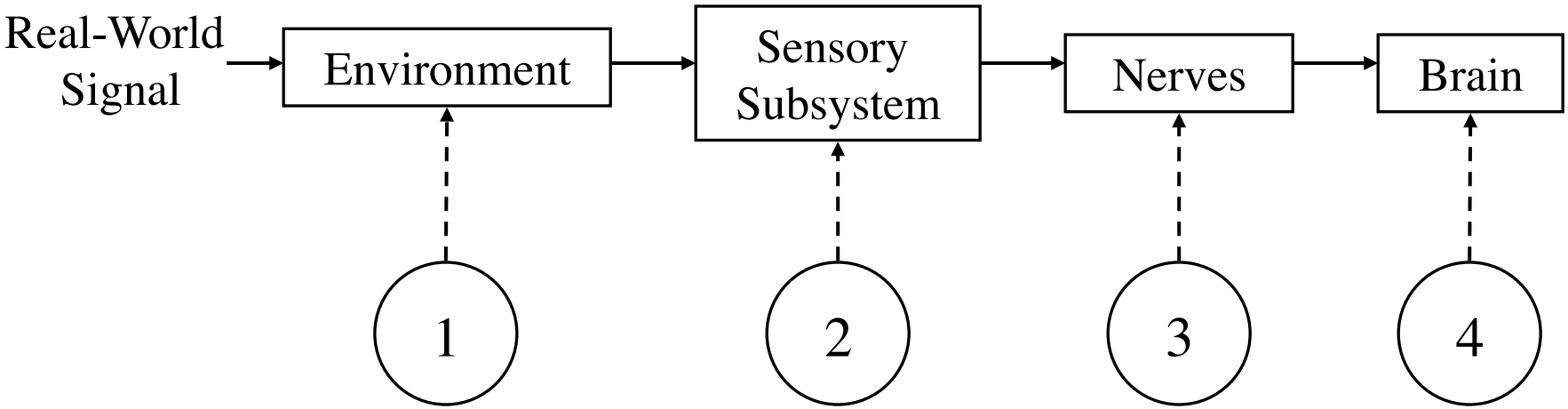
WPI

# Motivation

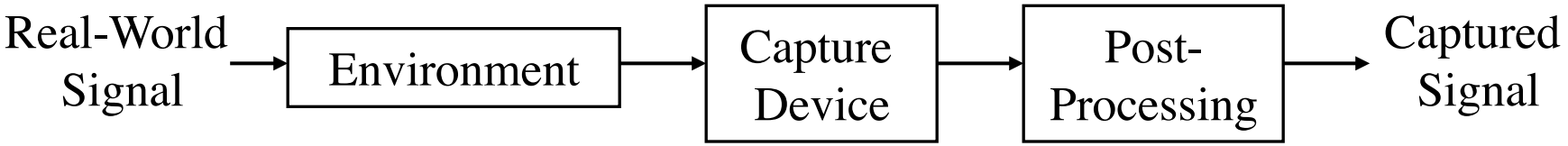
- Augmented Reality
  - Mixing of real-world (RW) and computer-generated (CG) stimuli
  - Graphical overlays on the real world
  - Adding information to real experiences
- Much work on visual sense
- Can be extended to auditory sense
  - Other senses?
- For the user to merge RW and CG, attributes must be matched
  - Visual: Lighting & shadows, level of fidelity
  - Audio: CG and RW sound occlusion and reflection

# Real-World Stimulus Paths

- Direct

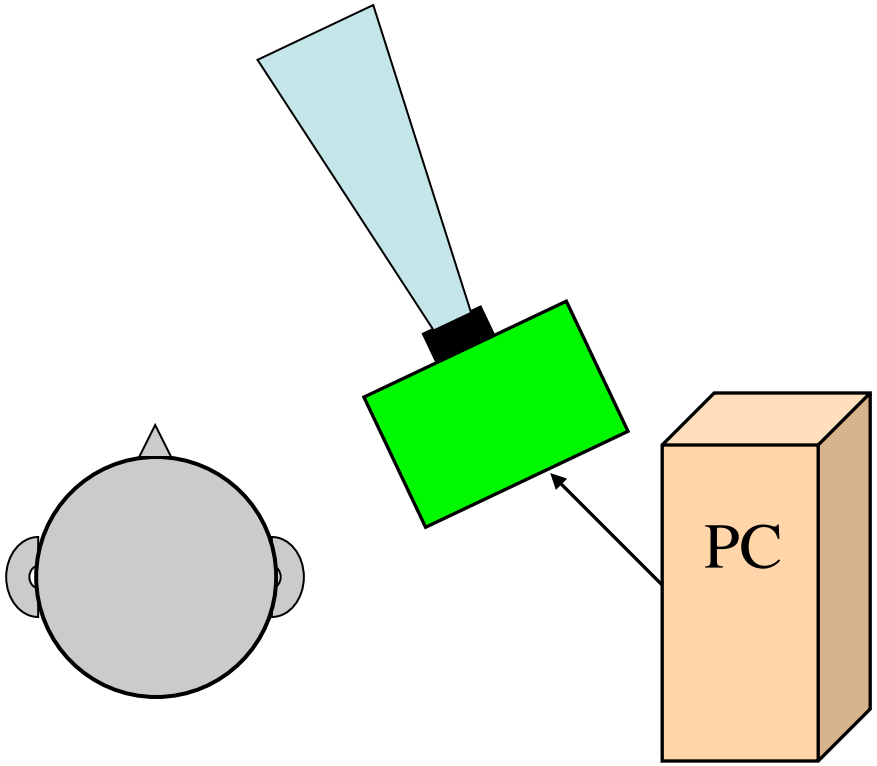



- Captured/Mediated



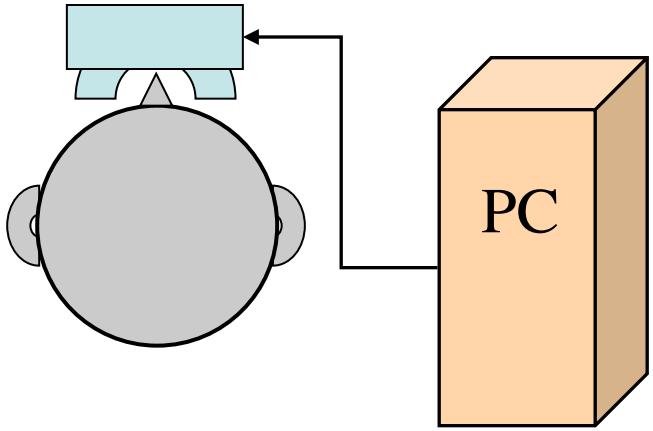
# Visual Sense

- Projection
  - Mixing in the environment (far)



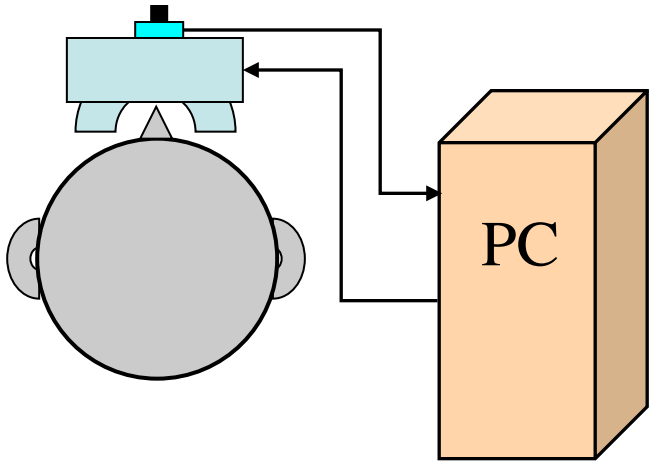
# Visual Sense (cont.)

- Optical-see-through AR
  - Mixing in the environment (near)



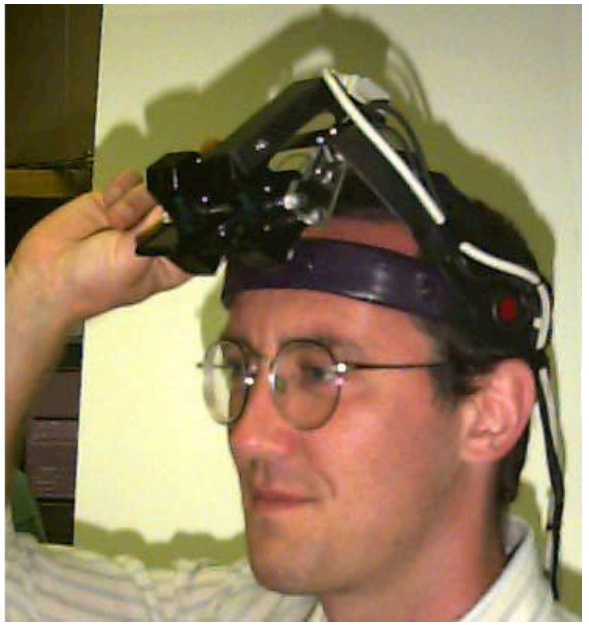
# Visual Sense (cont.)

- Video-see-through AR
  - Mixing in the Computer



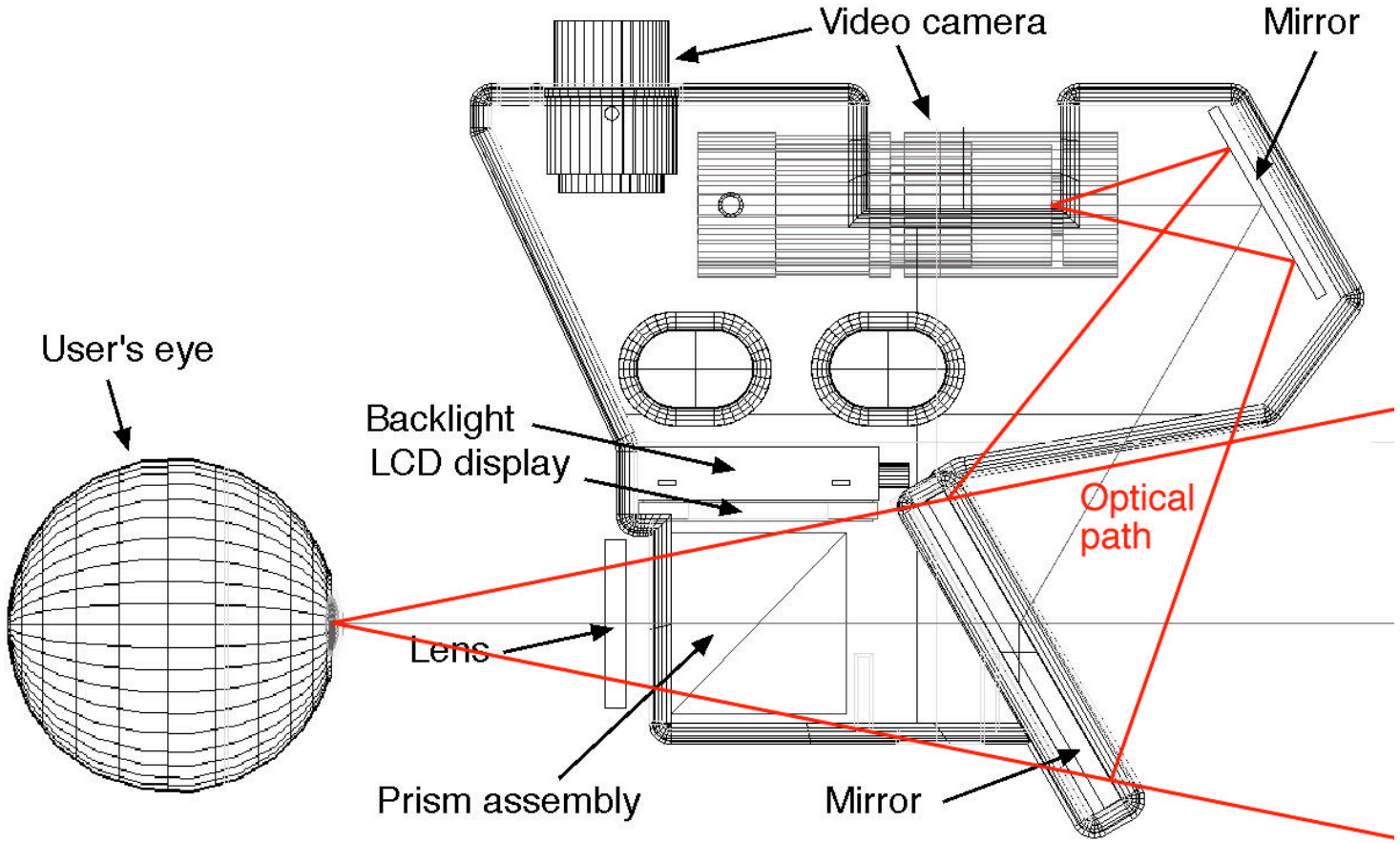
# Video-See-Through HMD





(Image: Fuchs, et al., Medical Image Computing and Computer-Assisted Intervention (MICCAI) ' 98, LNCS, 1998, Vol. 1496/1998, 934)

# Video-See-Through HMD (cont.)




(Image: Fuchs, et al., Medical Image Computing and Computer-Assisted Intervention (MICCAI) '98, LNCS, 1998, Vol. 1496/1998, 934)

# Video-See-Through HMD (cont.)

- NVIS: nVisor MH60-V (2010)

http://www.nvisinc.com/product2009.php?id=57

# Using Visual AR: SDKs

- ARToolKit
  - http://www.hitl.washington.edu/artoolkit/
  - Earliest usable kit
  - Now Open Source (free)
  - Commercial versions for iPhone & Android
    - http://www.artoolworks.com/
- Studierstube ES & Tracker
  - http://studierstube.icg.tu-graz.ac.at/handheld_ar/
  - *ES* sits on top of *Tracker*
  - Not free

# Using Visual AR: SDKs Examples

- ARToolKit
  - http://www.youtube.com/watch?v=5M-oAmBDcZk
  - (local clip)
- Studierstube ES
  - http://www.youtube.com/watch?v=JwluCuVKO9c
  - (local clips)

WPI

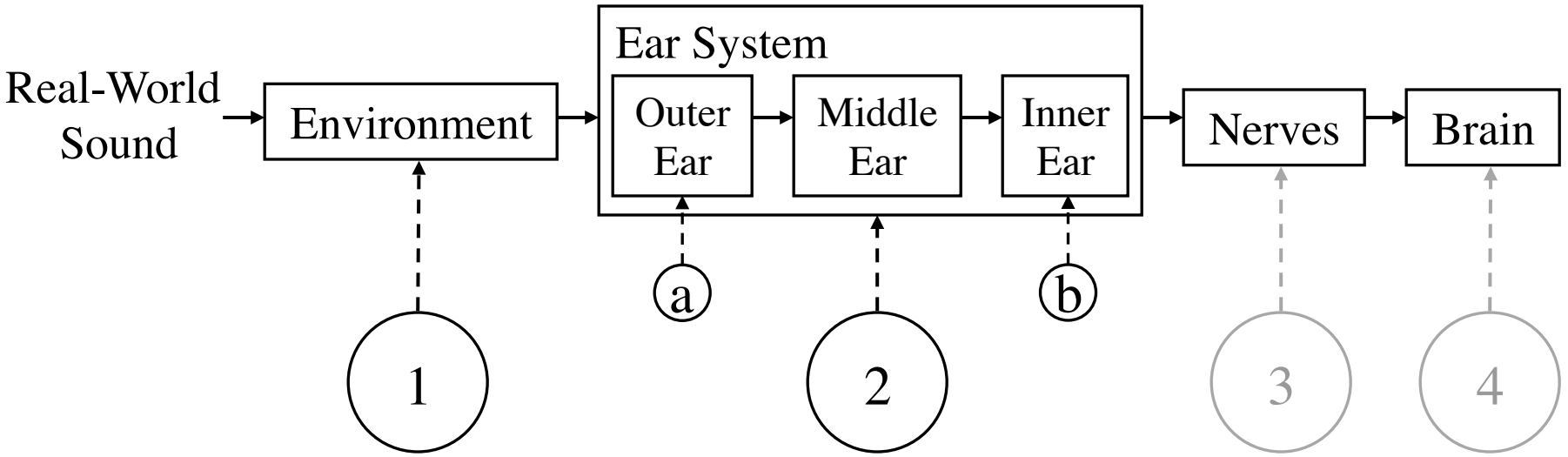
# Using Visual AR: Tools

- Google SketchUp + ARMedia Plugin
  - http://www.youtube.com/watch?v=wsQ-YGgVUT0
  - (local clip)
  - http://sketchup.google.com/
- Layar for mobile devices
  - http://www.layar.com/
  - Layering tool for layar browser
    - "Like HTML for AR"
  - (local clip)

# Using Visual AR: Tools (cont.)

- Cereal?
  - http://www.youtube.com/watch?v=jGdSslAJRwM
  - (local clip)
- Slot Cars?
  - http://www.youtube.com/watch?v=WMWEYqYPDfc
  - (local clip)
- Magic Tricks?
  - http://www.youtube.com/watch?v=Mk1xjbA-ISE
  - (local clip)
- Heads-up Display in Cars (play GE clip)
- Mobile AR (play Nokia clip)
- Mobile 3rd Party
  - http://news.bbc.co.uk/2/hi/technology/8193951.stm

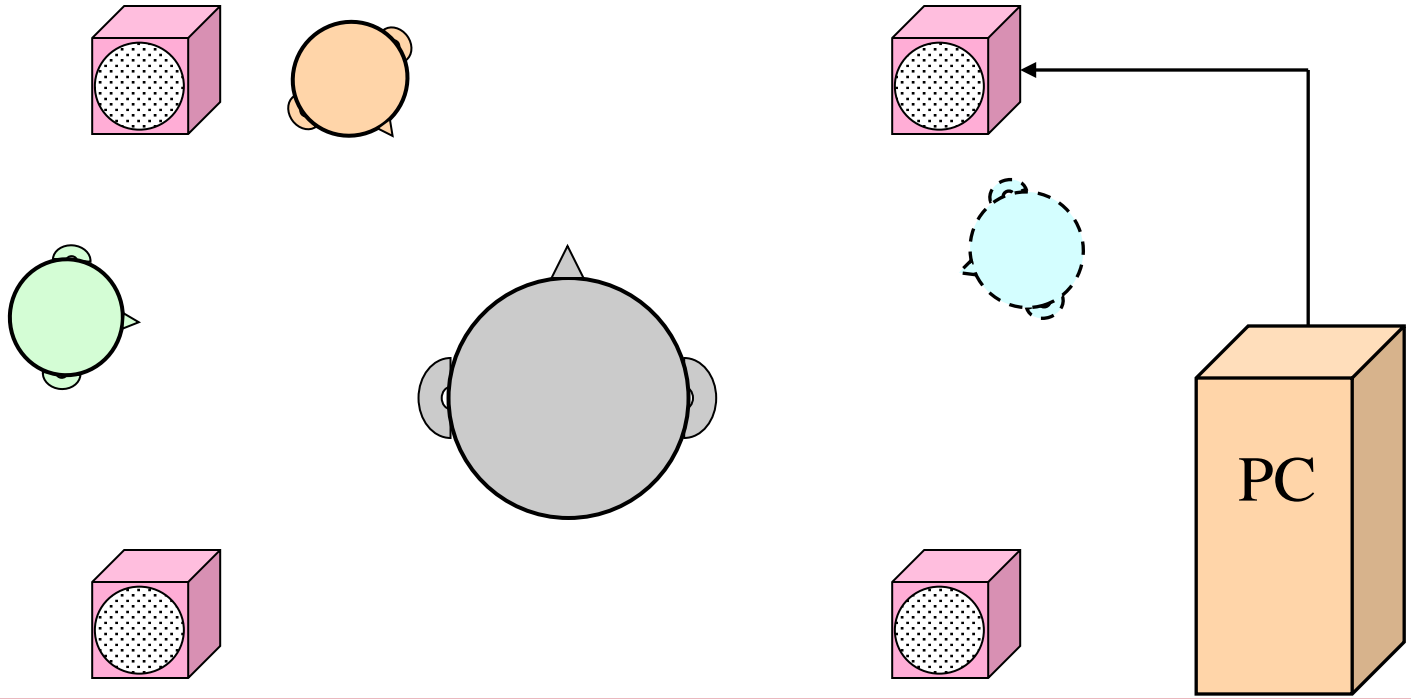
# Sound Paths & Mixing Points




- Typical VR/AR systems use speakers (1) or headphones (2a)
- Our approach performs the mixing at the cochlea (2b)

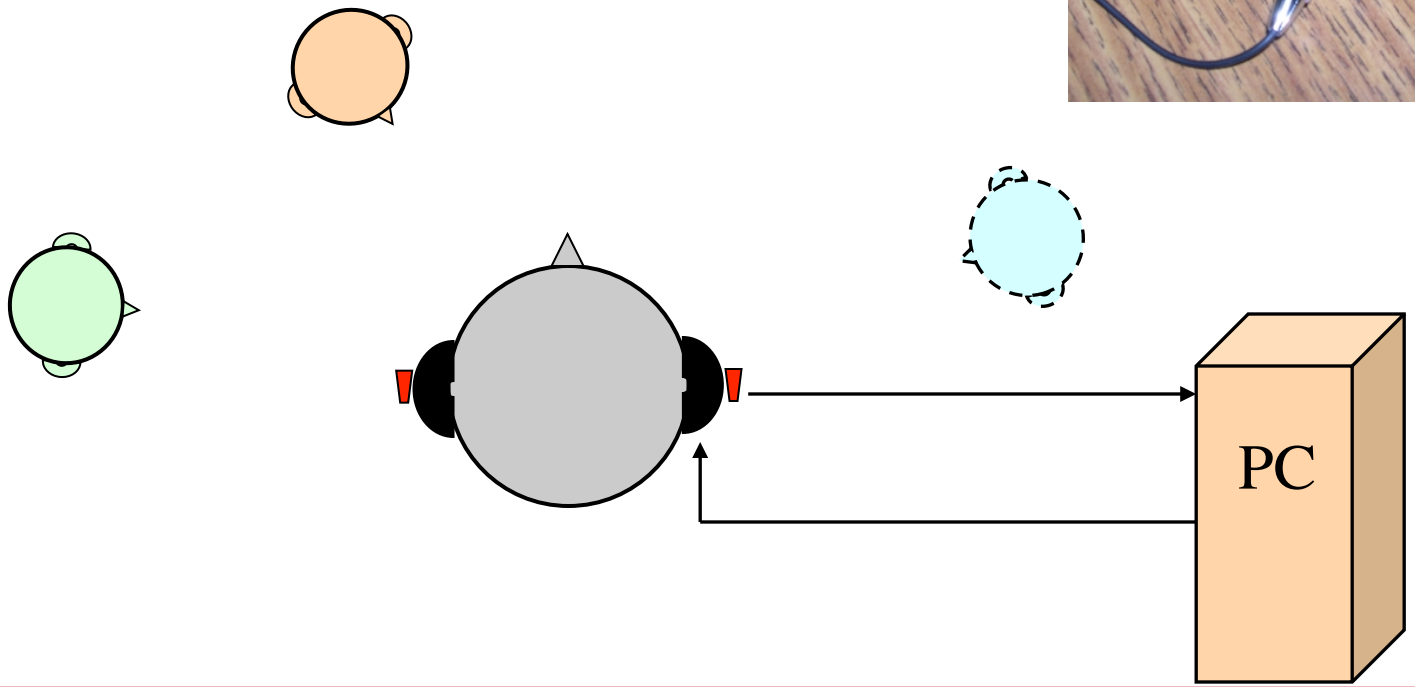
# Auditory Sense

- Acoustic-Hear-Through AR (Speakers)
  - Mixing in the environment (far)



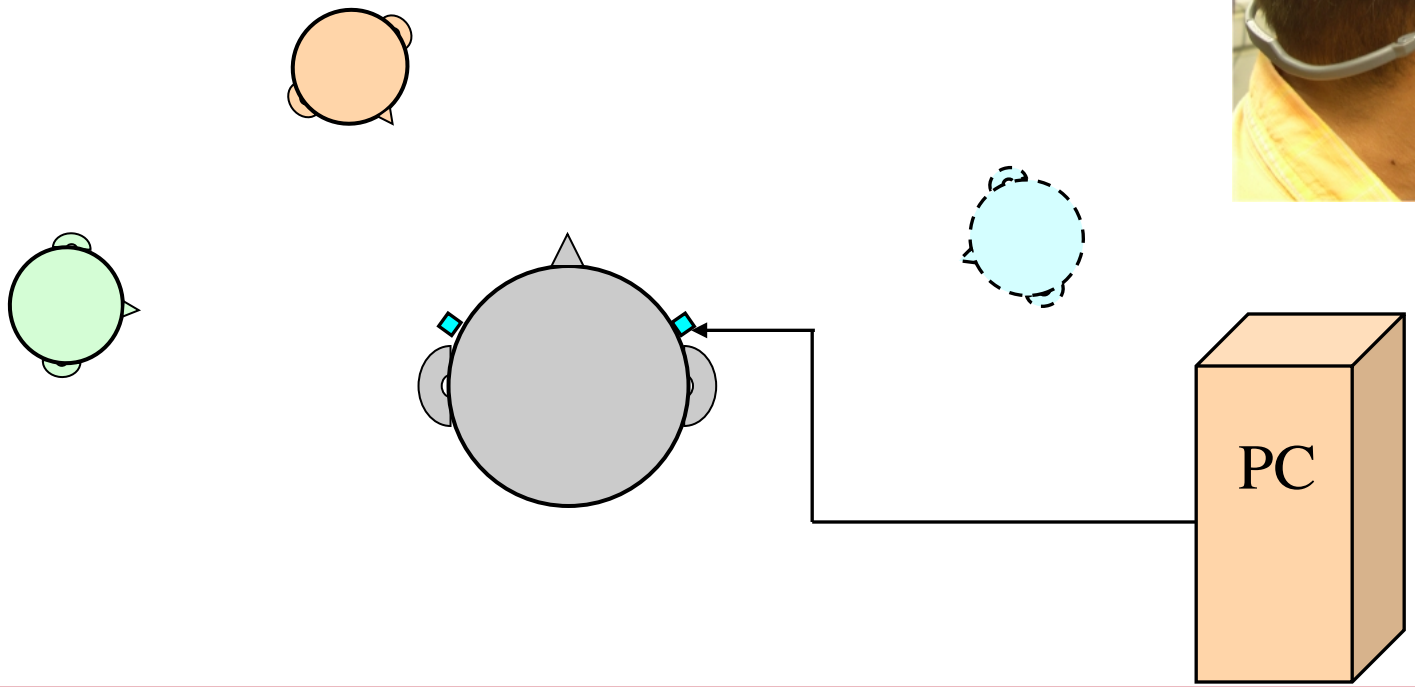
# Auditory Sense (cont.)

- Mic-Through AR
  - Mixing in the computer



# Auditory Sense (cont.)

- Hear-Through AR
  - Bone conduction
  - Mixing at the sensory subsystem



# Bone-Conduction Example

- The sound of your own voice is a combination of:
  - Sound reaching your ears through the air
  - Vibrations reaching your cochlea though your head
- Example
  - Sound heard through the air
  - Sound heard through the head
  - Combined sound

Mauldin & Scordilis, 2004

# Research Questions

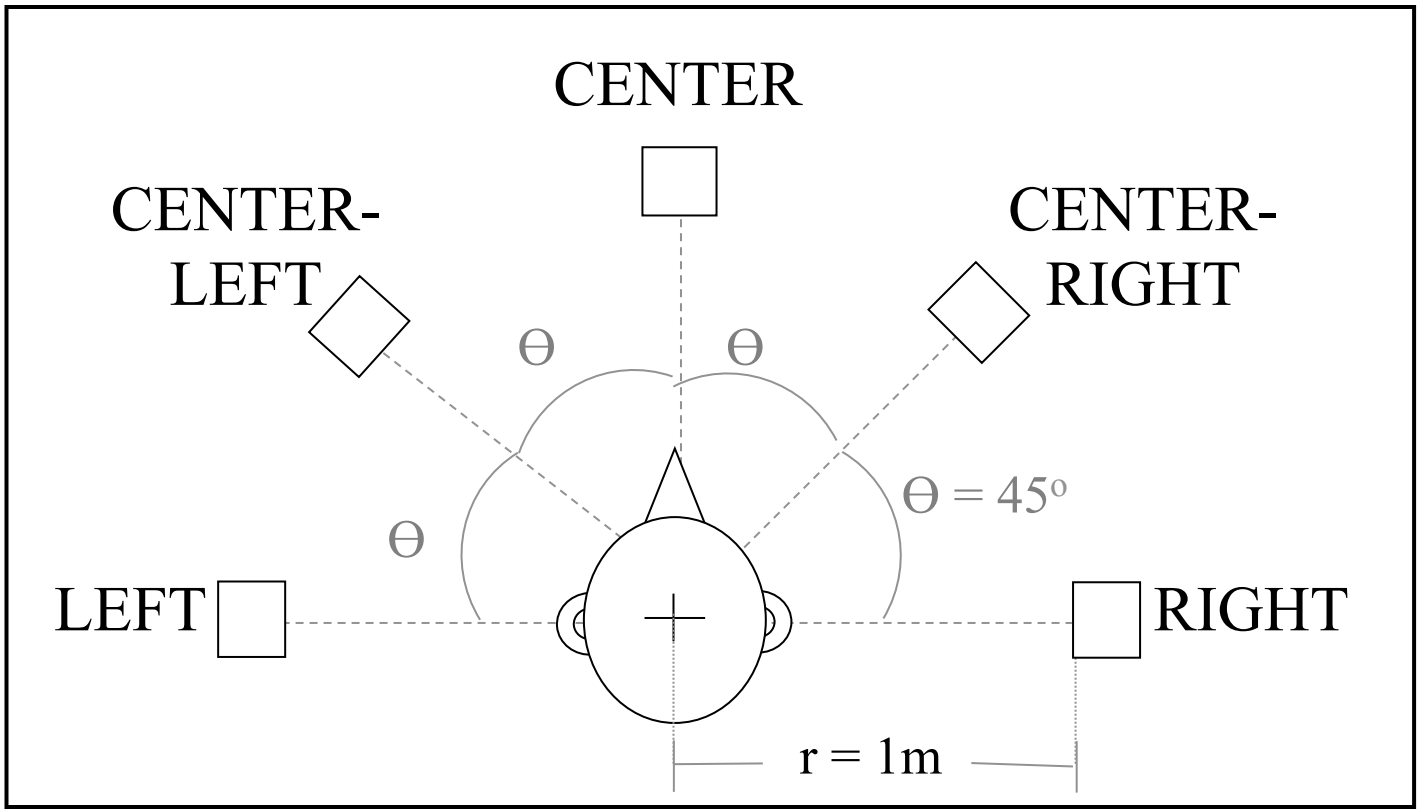
- How well can people localize sound using bone conduction?
- What types of sound works best?
  - Ambient sound
  - Spoken voice
  - Sound FX
  - Music
- We looked at basic sounds (sine waves) of various frequencies
  - Stationary and moving sounds

# Design of the User Study

- 24 Computer science students (22 male)
- 3 Main treatments (Audio Devices)
  - Speakers, Headphones, Bone-Conduction Device
- Each subject performed 63 trials with each device
  - 3 Frequencies
    - Low (200Hz), Medium (500Hz), High (1kHz)
  - 7 sound samples (5 sound locations + 2 directions)
    - Left, Center-Left, Center, Center-Right, Right
    - Moving, right-to-left moving
  - 3 repetitions of each combination
  - 3 * 7 * 3 = 63

# User Study

- Physical/Virtual sound locations



WPI

# User Study (cont.)

- Each sample was played for 1 second
- Subjects wore a blindfold
- No HRTFs used
- Subjects had to identify location/direction

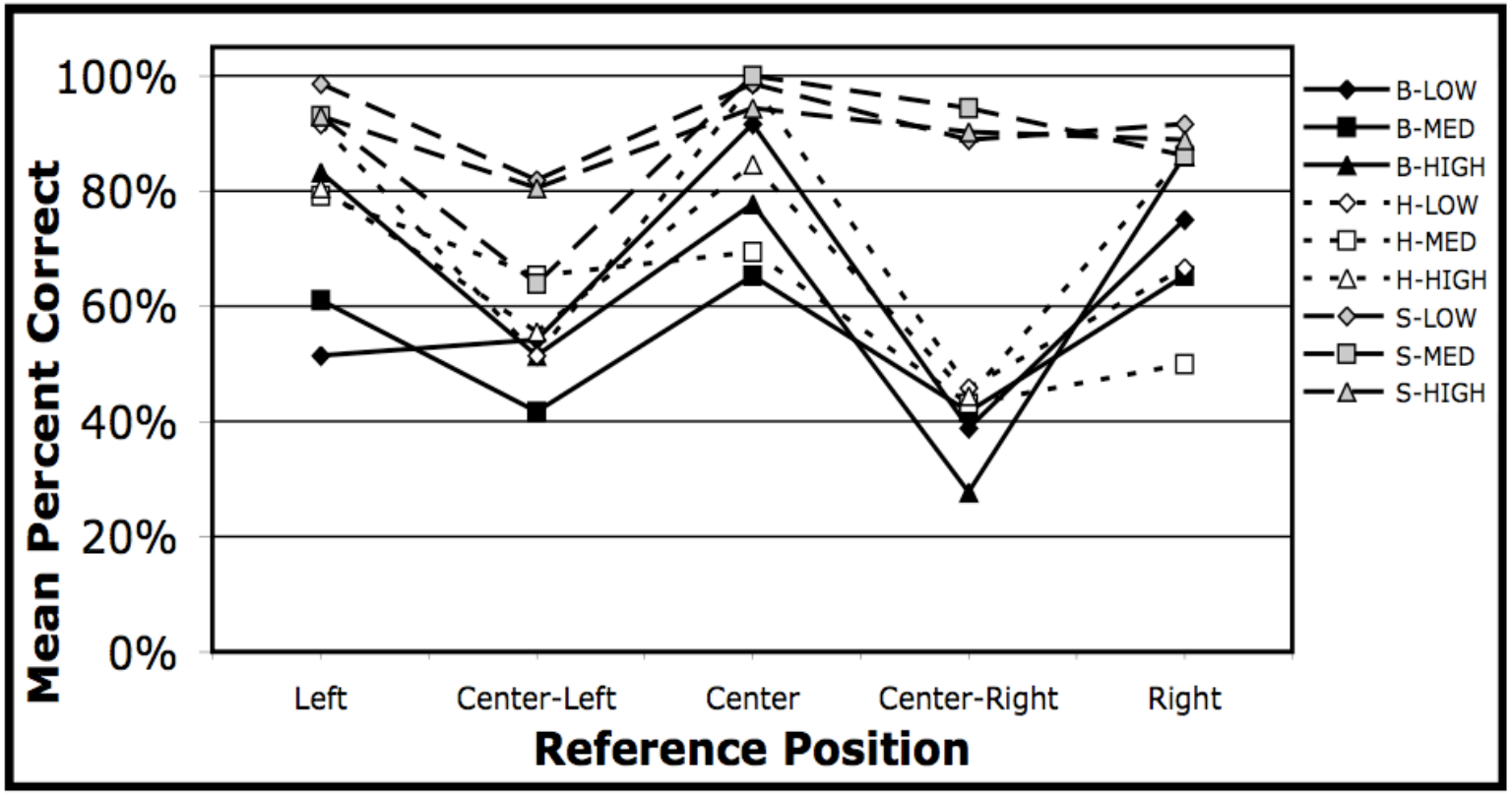
# Results

- Accuracy for **Stationary** Sounds
  - Speakers > headphones > bone conduction
  - High-Freq. == Low Freq., both > Medium Freq.
- Accuracy for **Moving** Sounds
  - Speakers == Bone conduction
  - Bone Conduction == Headphones
  - Speakers > headphones ($\alpha$ = .05)

| | Stationary | Moving |
|---|---|---|
| Audio Device | S H B | S B H |
| Frequency | HIGH LOW MED | *ns* |
| Interaction | *ns* | *ns* |

# Results (cont.)

- Problems with the "in-between" locations
  - Center-Left/Center-Right



WPI

# Analysis

- Real-world sound
  - High fidelity
  - Low control
- Computer-generated sound
  - Low(er) fidelity
  - Complete control
- Later mixing point = Closer to the brain
  - More personalized, but
  - More processing for transforming and mixing

# Analysis (cont.)

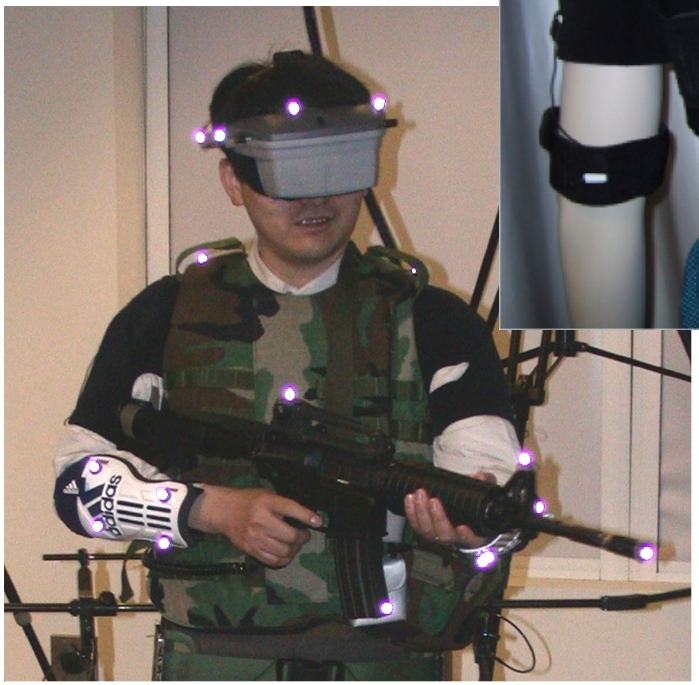
- Bone-conduction/headphone approaches
  - Require head tracking for CG sound
  - Require processing for spatialization (e.g., HRTF or BRTF)
- Speaker-based
  - Allows for shared experience (like projection systems in visual field)

# Haptic Sense

**Mixing in Computer (teleoperation) or in Environment**
(Immersion CyberGrasp)

**Mixing at Sensory Subsystem**
(Novint Falcon)

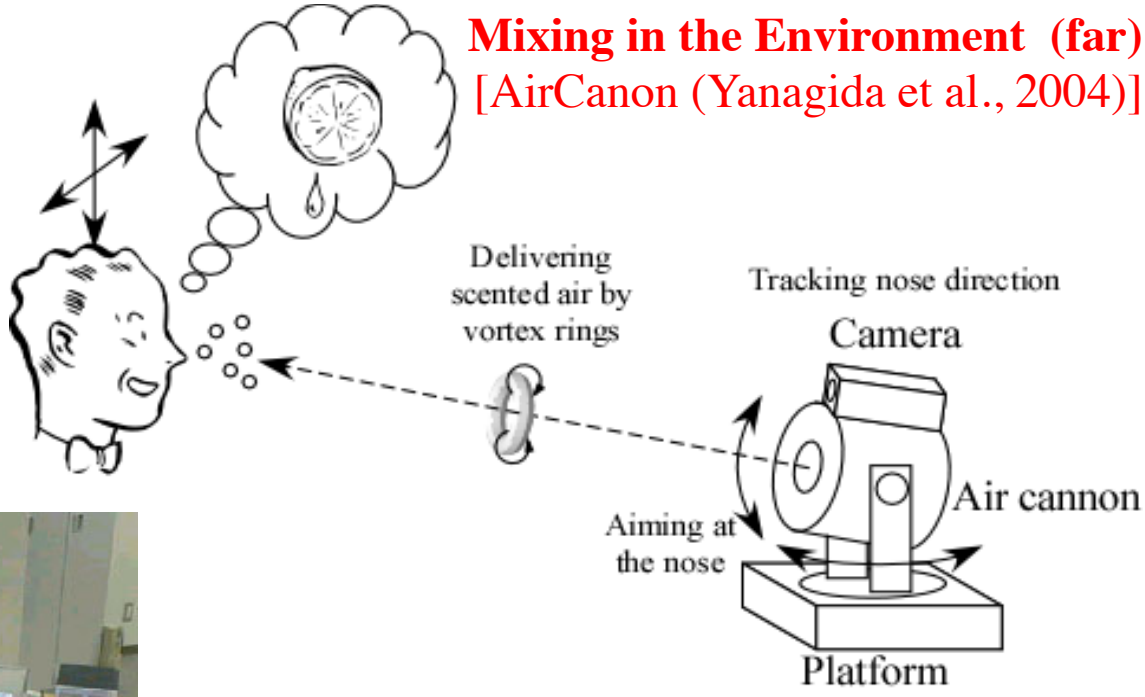
# Haptic Sense (cont.)



**Mixing in the Environment (Lindeman, VRST 2004)**

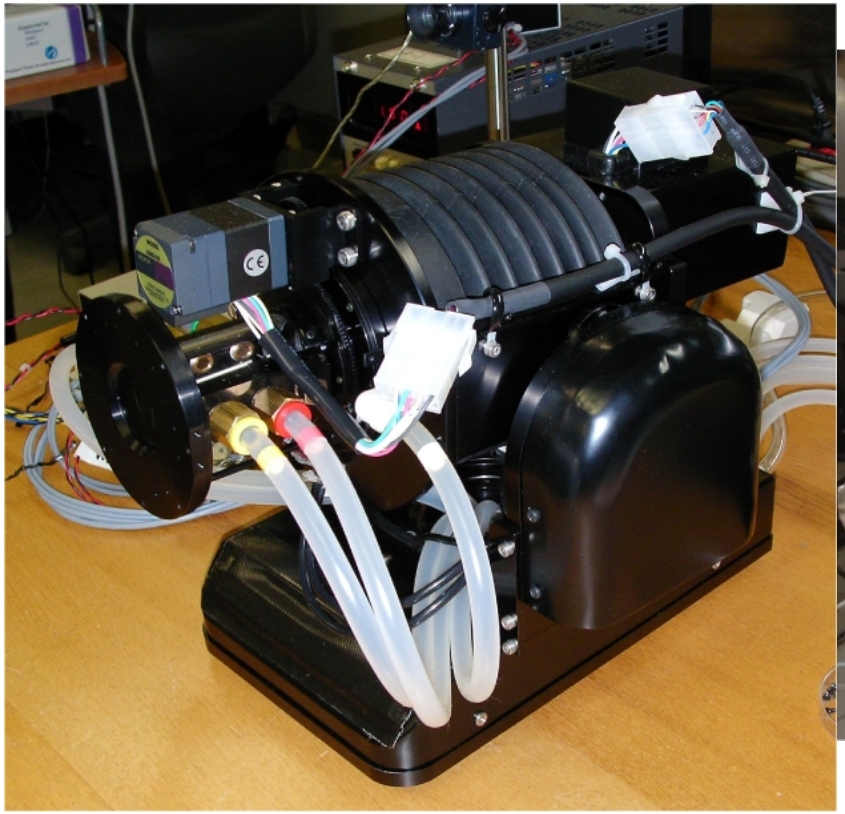
# Olfactory Sense

**Mixing in the Environment (far)**
[AirCanon (Yanagida et al., 2004)]

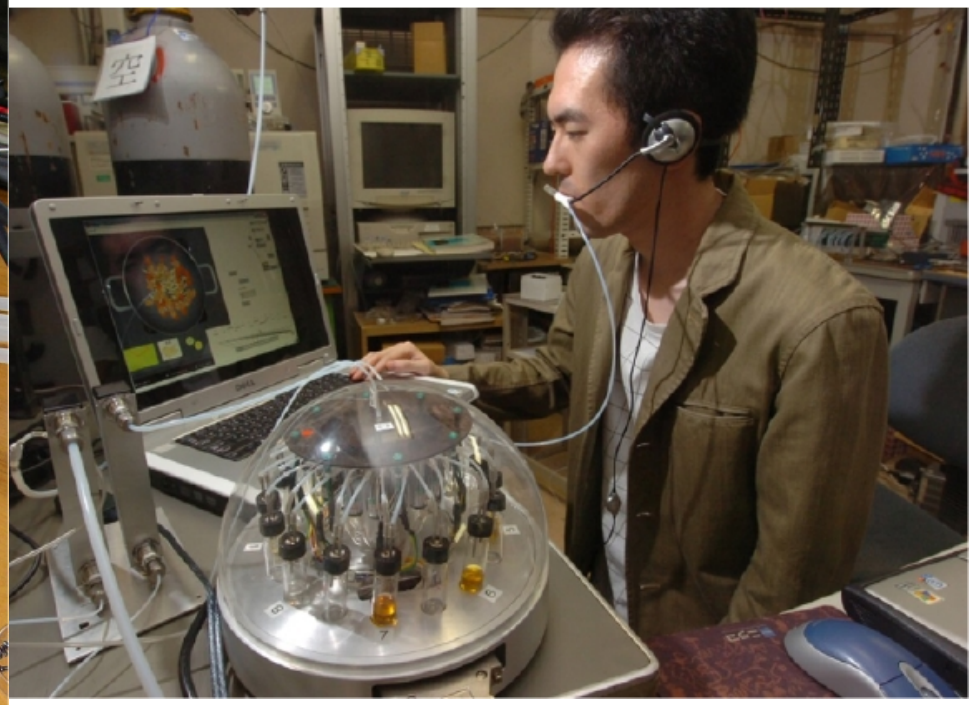



**Mixing in the Computer**
(Hirose et al. 1997)

# Olfactory Sense (cont.)



**Mixing in the Environment (mid)**
[AirCanon (Yanagida et al., 2004)]



**Mixing in the Environment (near)**
(Nakamoto & Min, 2007)

# Gustatory Sense

- Bite interface
  - Really haptics (near)

Iwata, 2004
(photos: Sid Fels)

# Gustatory Sense (cont.)

- Edible bits
- Straw-like interface
  - Mixing in the env.

(Nakamoto, 2007)


(Maynes-Aminzade, 2005)

# Gustatory Sense (cont.)

# Final Thoughts

- What about a 3D printer+robot arm?
- RW stimuli
  - High fidelity / low control
- CG stimuli
  - Low(er) fidelity / complete control
- Later mixing point = more "personal" stimuli
  - Closer to the brain
- Multi-sensory approaches are interesting
  - Compensate for weaknesses in one sense with another sense
  - Use speakers for environmental, bone-conduction for virtual characters