ASRU Panel:
Recording Situated Human Communications

Panelists:

John Hansen, Univ. of Texas, Dallas

Kazuya Takeda, Nagoya University

Alex Acero, Microsoft

Brandon Roy, MIT Media Lab

Moderator & Panelist:

Sharon Oviatt, Incaa Designs

Multi-Person Multimodal Data Collection Combining Audio, Visual & Pen Input

Sharon Oviatt
oviatt@incaadesigns.org

ASRU Workshop
December 2007

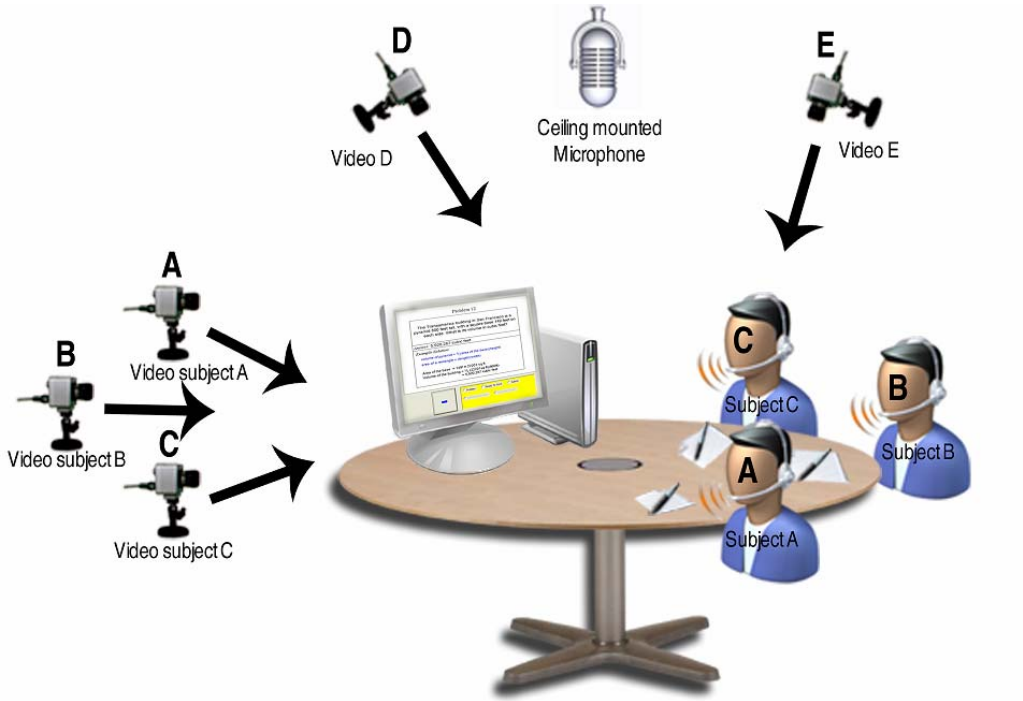
Goals, Characteristics & Amount of Data Collected

- Develop infrastructure for research & prototyping of novel collaborative multimodal systems, including computer-mediated group meetings *(simulation software, data collection & analysis software)*
- Capture of multiple audio, visual, and pen heterogeneous data streams *(10-12)*, synchronized, time-stamped & high resolution
- Ability to observe & respond to media streams in real-time during data collection for rapid prototyping *(both single & dual-wizard methods)*
- Minimally obtrusive recording devices, so natural communication & work patterns are preserved
- Supported 2 studies involving highly realistic & engaged meetings focused on user performance *(e.g., peer tutoring to learn geometry)*
- > 36 hours of multimodal multi-person data collected

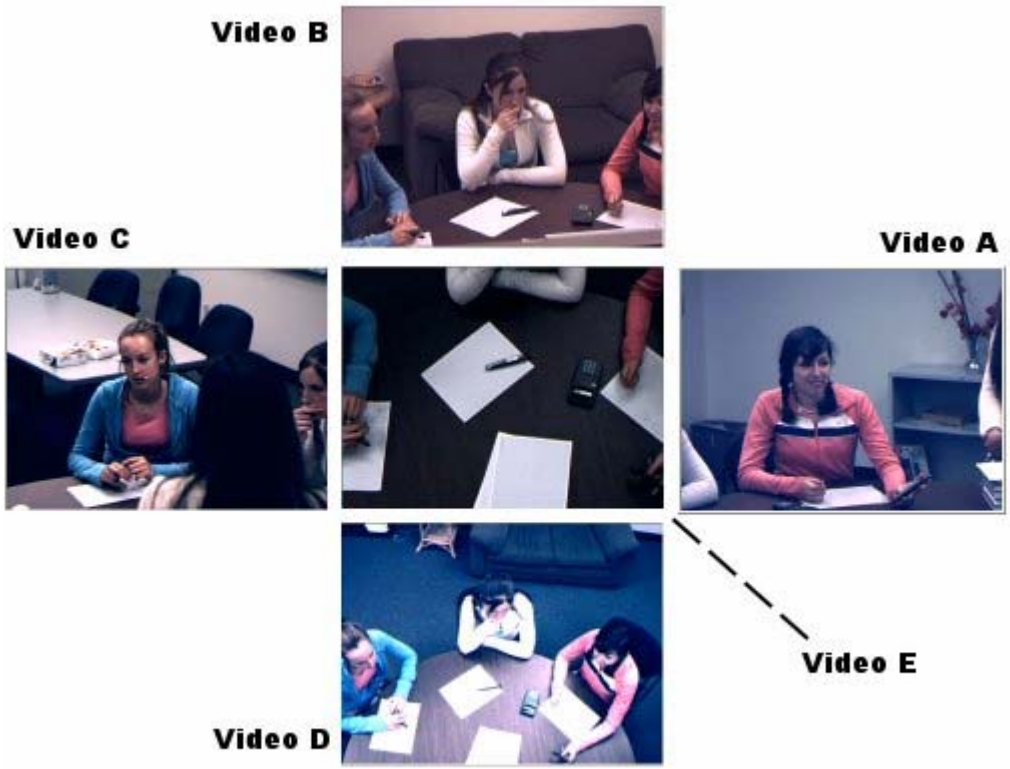
Research Topics Explored in Simulation Studies

- Prototyping of new interfaces involving speech, digital paper & pen, vision & their multimodal combination
- Modeling of human performance & cognitive load during extended problem solving & collaboration
- Design of educational interfaces with tutorial assistants for math peer tutoring
- Design of implicit user-adaptive interfaces for field use, collaborative use & educational activities

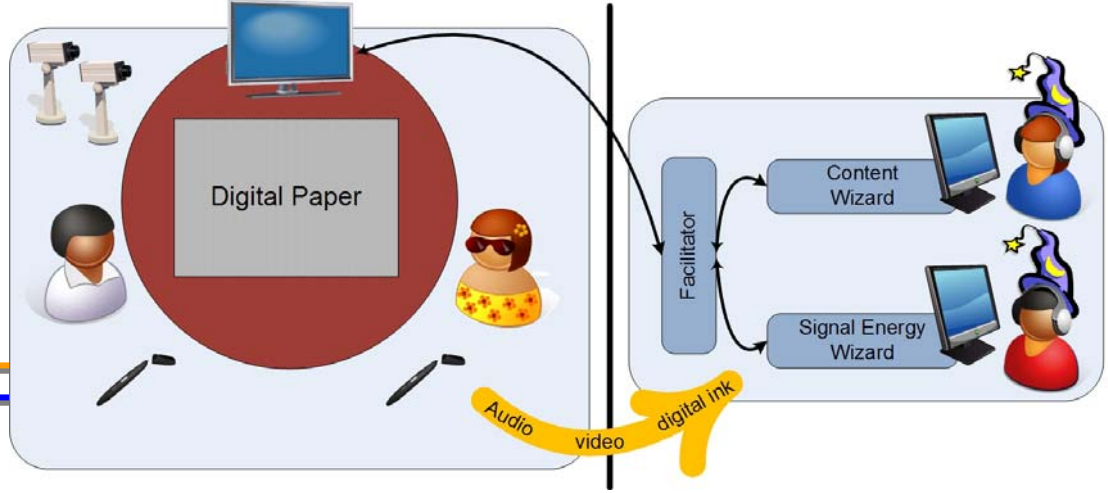
Data Collection Devices in Meeting Room



Synchronized Playback of Multiple High-Resolution Videos during One Moment of Interaction
(coordinated with playback of users' speech & pen input)



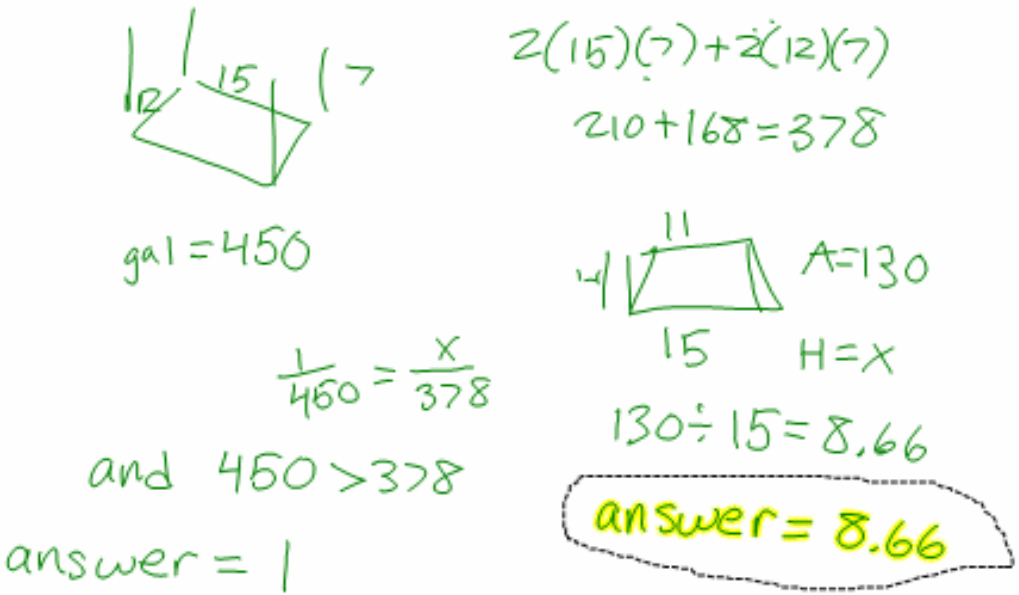
Example of Novel Dual-Wizard Simulation Method



- Content wizard: Responded to speech or pen constructions *when semantic content was compatible with tutorial system functions*
- Signal energy wizard: Responded when signal energy *(amplitude, pressure)* of real-time constructions during meeting met threshold
- Simulation involved real-time contingent learning paradigm to develop user-adaptive multimodal interface prototypes
- Signal detection methodology
- Wizard coordination supported by distributed agent architecture

Example of Real-Time Wizard Response Capabilities to Students' Pen Input during Simulation Studies

- Wizards saw real-time streamed digital ink from student pens while they worked & could:
  - -Pan, rotate & zoom their displays
  - -Encircle ink constructions to calculate pen signal features (*pressure)*



Data Availability & Future Directions

- Some data available to other external research partners, and some restricted to partners within CALO project (IRB issues)
- Our future data infrastructure will focus on emerging mobile interfaces:
  - Miniaturized & more flexible
  - Widely usable in different settings
  - More automatic data capture & playback
  - Support for larger volume data collections
  - More focused on multimodal information retrieval
  - Less labor intensive & expensive to conduct
- Community needs more mobile, multimodal & multi-person collaboration data