

Multimodal systems

Application domains
Examples

1

Application domains



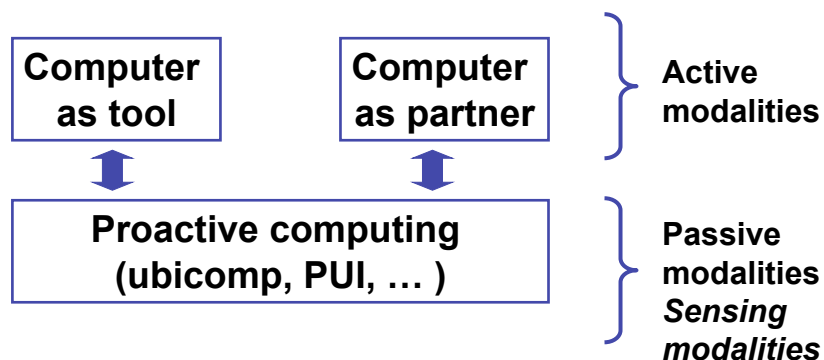
2

Three paradigms for multimodality

- **Computer as tool**
 - Multiple input modalities are used to enhance direct manipulation behavior of the system
- **Computer as partner**
 - The multiple modalities are used to increase the anthropomorphism of the user interface
- **Proactive computing (ubicmp, PUI, ...)**
 - The multiple modalities are used to sense the user and the environment

3

Three paradigms for multimodality



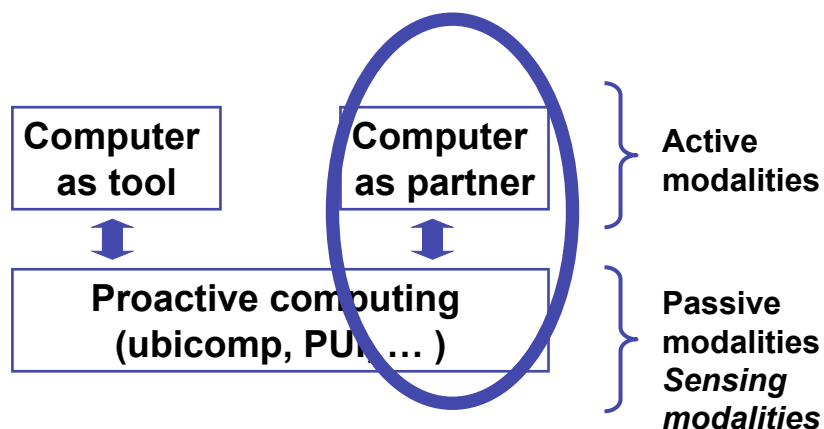
4

Active/Passive modalities

- Active modalities are used by the user to issue a command to the computer (e.g., a voice command)
- Passive modalities are used to capture relevant information for enhancing the realization of the task, information that is not explicitly expressed by the user to the computer such as eye tracking location/orientation tracking etc.
- Combination of active and passive modalities

5

Three paradigms for multimodality



6

Computer as partner Multimodality

- **Computer as partner**
- The multiple modalities are used to increase the anthropomorphism of the user interface
 - agent based conversational user interfaces
 - multimodal output is important: talking heads and other humanlike presentation modalities
 - speech recognition is a common input modality in these systems, and speech synthesis is used as an output modality



7

Computer as partner Multimodality

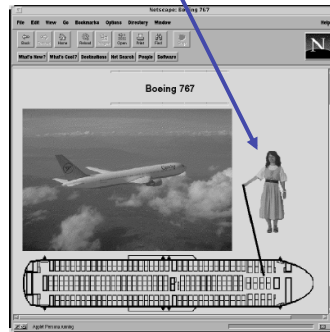
- **Computer as partner**
- Output modalities:
 - Talking avatar
 - Speech



8

Computer as partner Multimodality

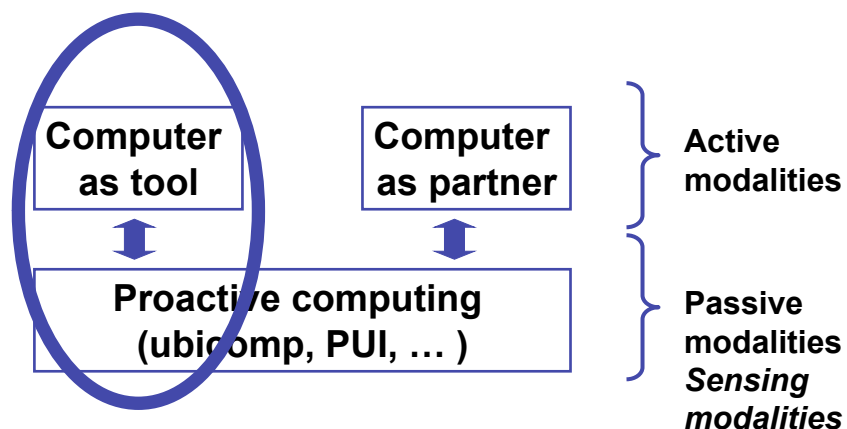
- **Computer as partner**
 - Output modality: posture language



E. André 96 Persona

9

Three paradigms for multimodality



10

Computer as tool Multimodality

- Computer as tool
- The user is responsible for initiating the actions
- Multiple input/output modalities are used to enhance direct manipulation behavior of the system
 - Interaction modalities

11

Computer as tool Augmented Reality / Augmented Virtuality



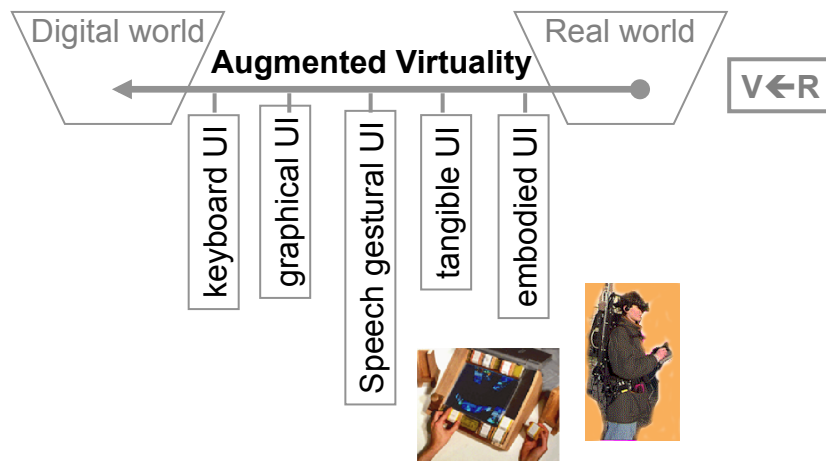
Purpose of the task = real world



Purpose of the task = computer

12

Augmented virtuality



13

Once upon a time HCI ...



- Ivan Sutherland 1963 (PhD thesis MIT)

- Sketchpad

- Drawing tool
- Optical pen and buttons
- Direct manipulation
- Icons
- Zoom
- Copy/Paste

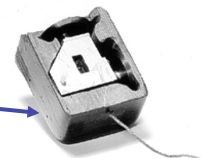
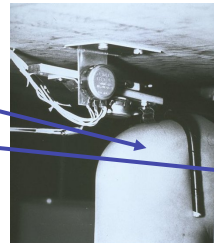


14

Once upon a time HCI



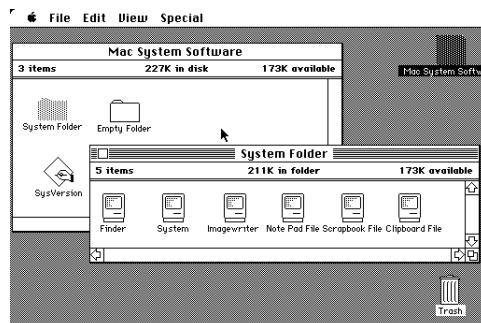
- Douglas Engelbart
- 1968 NSL oN Line System
- Augment/NSL
 - Text edition
 - Video conference
 - Two dimensional screen
 - Device on knee
 - Mouse



15

Once upon a time HCI

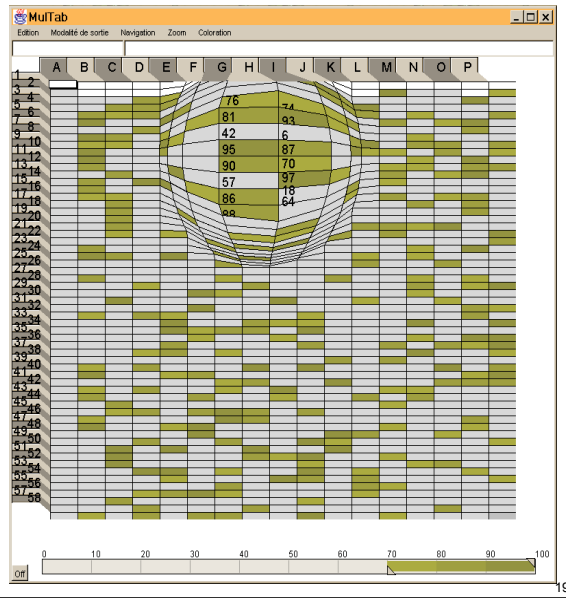
- Macintosh 1984
 - Direct manipulation
 - Mouse and Keyboard



16

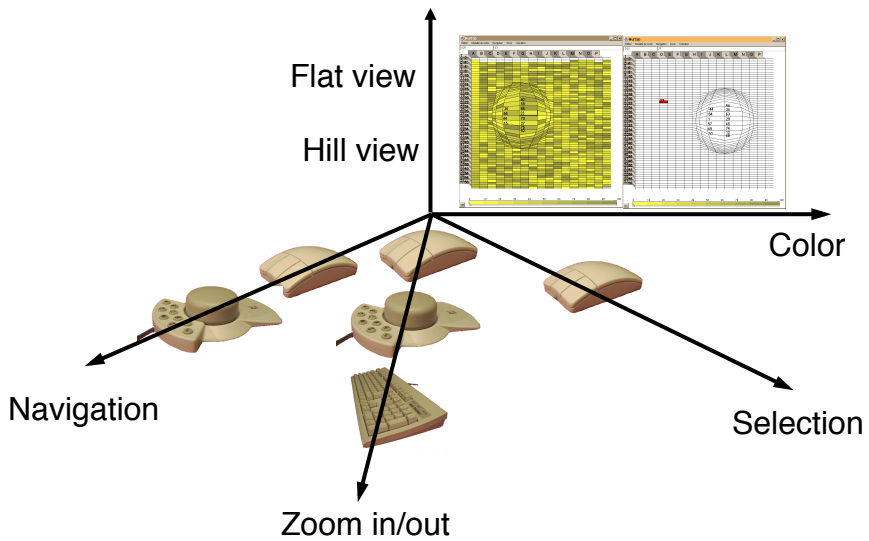
Graphical UI

- MULTAB
- LIG-IIHM



19

Experimental evaluation: MULTAB



20

Graphical UI

- MERL table
 - *Video: TableRondeVernier.mpg*



21

Graphical UI

- Multi-surface interaction



22

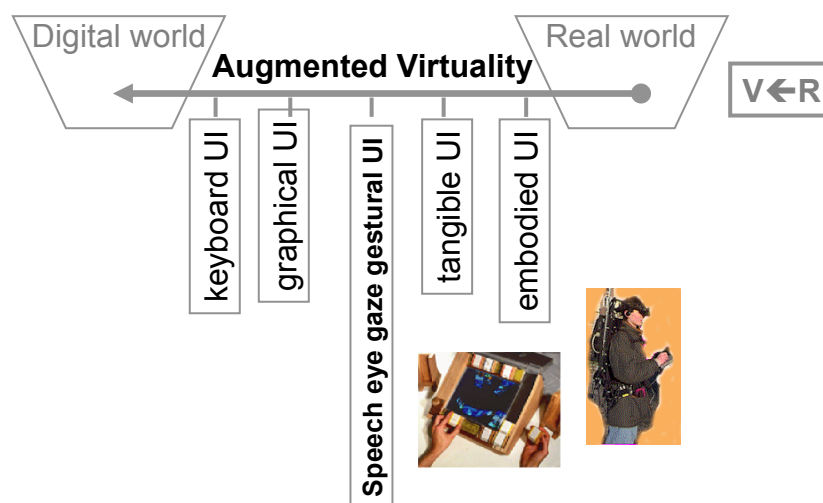
Manipulation and stereo

- Input modalities
 - 3D gesture
 - Speech
- Sensing modalities
 - Head tracker
 - Eye gaze tracker



23

Augmented virtuality



24

Meditor: Multimode Text Editor

- MEDITOR: Y. Bellick LIMSI-Paris
- Combines keyboard, Braille terminal, a French text-to-speech synthesiser, and a speech recognition system
- Allows Blind people to perform simple Document editing tasks

25

Meditor



26

Meditor Commands

- To put a word in italic
- The user says "*italic*" while clicking on any character of the word on the Braille terminal.
- To place a character into an exponent position
- The user says "*character exponent*" while clicking on the corresponding character.
- To delete a part of the text
- 1) The user says "*begin selection*" while clicking on the first character of the string to be deleted, 2) then says "*end selection*" while clicking on the last character, and 3) says "*delete*" to complete the command. The message feedback "*selection deleted*" is then generated by the speech-synthesizer.

27

Speech + gesture

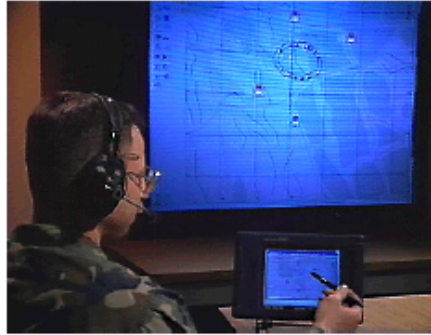
- VoicePaint LIG-IIHM
 - Graphical editor
 - Mouse + speech
 - Change colors using speech while drawing using the mouse

28

Quickset

Spoken and gestural interaction

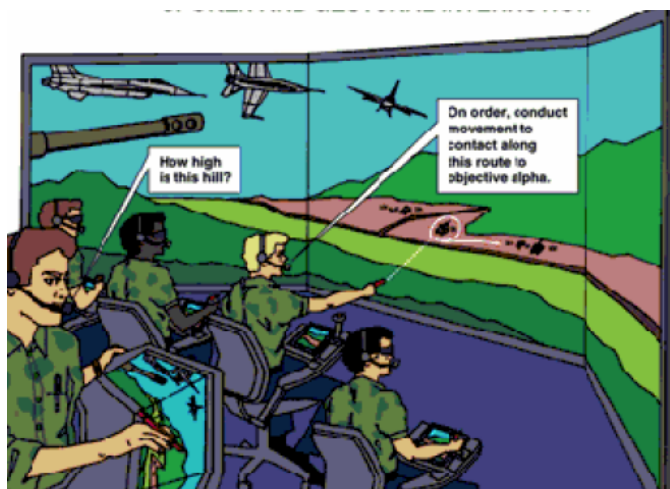
- Speech
- Pen input
 - Pointing (selection)
 - Gesture recognition



29

Quickset

Spoken and gestural interaction



30

Quickset

Spoken and gestural interaction



31

Quickset - RASA: Multimodality and Augmented Reality

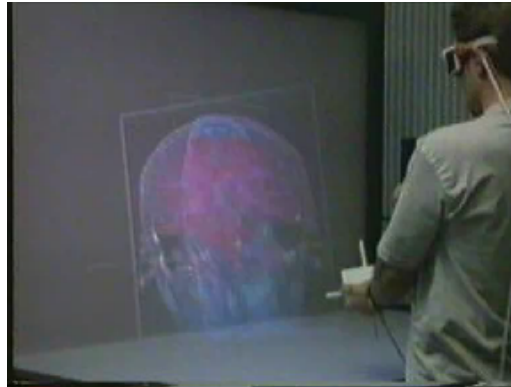
– Video DEMOMULTI/Rasa.mpg



32

Gestural UI

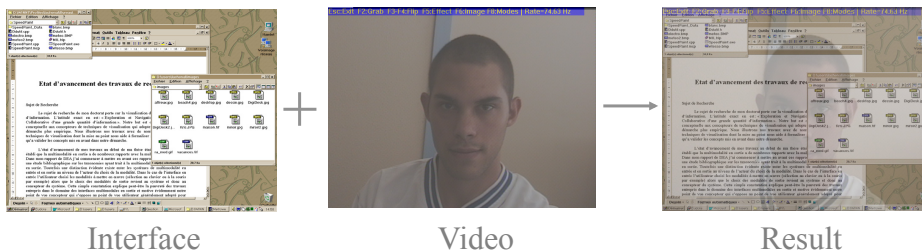
- 3D cubic mouse
 - Video CubicMouse.mov



33

Gestural UI

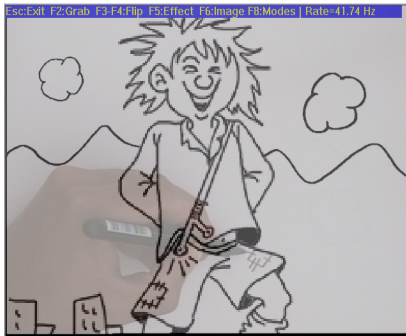
- Mirror Pixels
- Fusion of two images, one is the user interface, one is a video pointing to the user or to his/her hand



34

Gestural UI

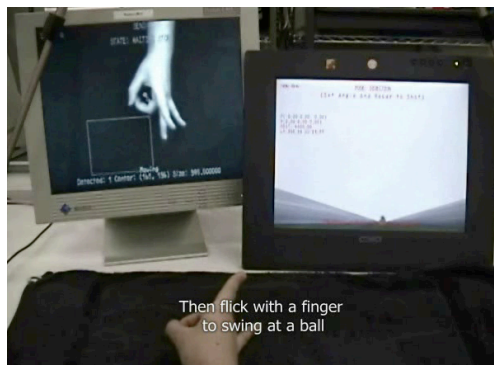
- Mirror Pixels: Drawing application
 - Input modality based on a pen
 - Video *PIXMIRROR/App_Dessin.mpg*



35

Gestural UI

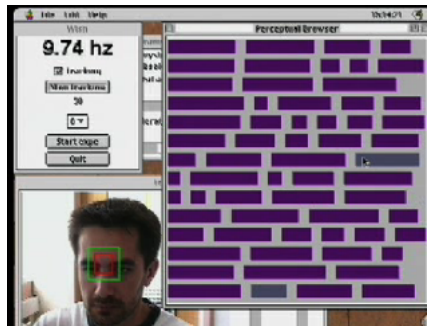
- Flicking gesture
 - Vision-based tracking
 - Video: *DEMOMULTI/BilleGestureVision-UIST06.mpg*



36

Gestural UI

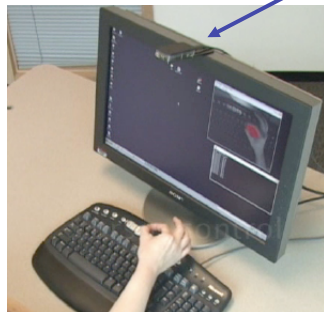
- Perceptual Browser LIG-IIHM
 - Two modalities
 - Head movement (vision based tracking) + mouse
 - *Video PBrowserMac.mpg*



37

Gestural UI

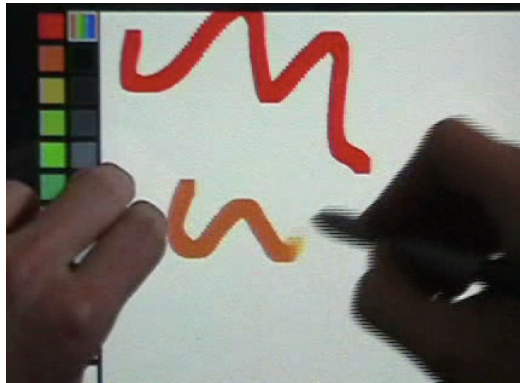
- Thumb and Fore-Finger Interface Microsoft
 - Modality based on vision-tracking
 - *Video DEMOMULTI/HandsOverKeyboard.wmv*



38

Two-handed interaction

- Drawing editor: Berkeley
 - Two modalities (one per hand)
 - Video: *2HandsInteraction.mov*



39

Two-handed interaction

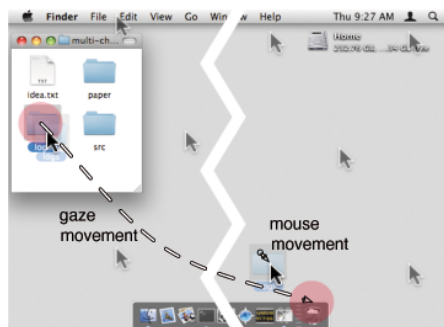
- NavGRAPH: LIG-IIHM
- NAVRNA: a system to visualize, explore and edit RNA Ribonucleic Acid
 - Video: *NAVGRAPHE/AVI_2.wmv*



40

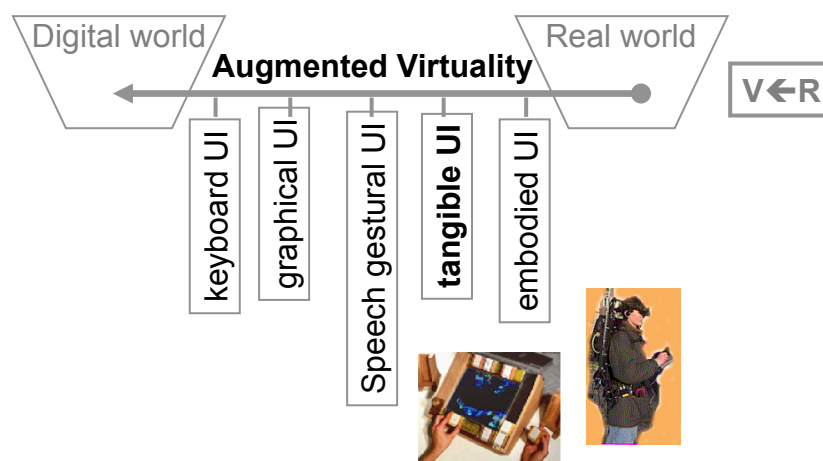
Eye gaze and mouse

- Rake cursor: LIG-IIHM
- Pointing tasks: two parallel input modalities
 - <http://iihm.imag.fr/blanch/projects/rake-cursor/>
 - Video: RakeCursor.mov



41

Augmented virtuality



42

Tangible UI

- AmbientRoom : MIT Medialab
 - Video: *AmbientRoom.mpg*



43

Tangible UI

- Lumino:
tangible blocks for tabletop computers
 - Video: *Lumino.mp4*

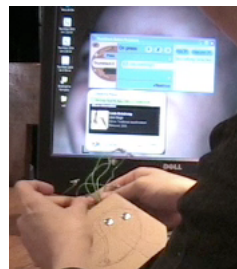


44

Tangible UI

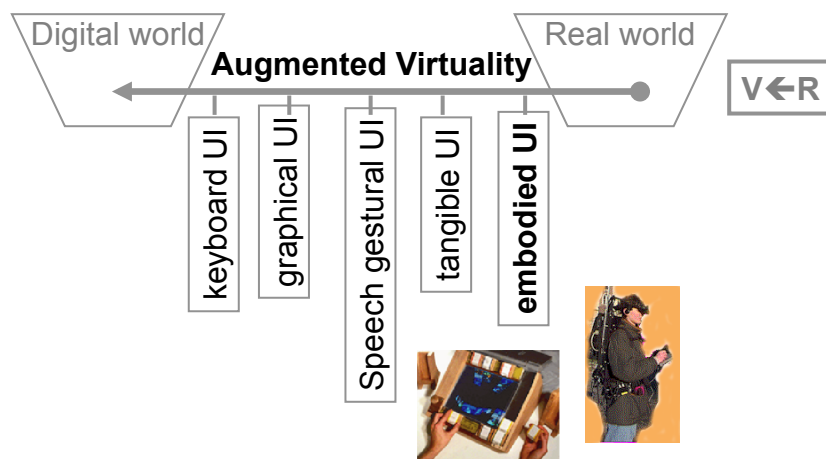
- Rapid Construction of physical interfaces
CMU

– Video: *PhysicalPrototype-UIST06.mov*



45

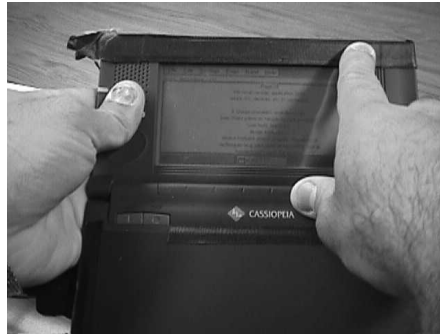
Augmented virtuality



46

Embodied User Interface

- Embodied UI: Rank Xerox



Embodied User Interface

- Embodied UI: Rank Xerox



Embodied User Interface

- Tilt and gesture based user input
 - Compaq project



Rock'n'Scroll

Video: *RocknScroll.mov*

49

Embodied User Interface



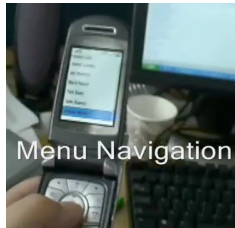
Peephole Displays

Ka-Ping Yee
Uni. of California, Berkeley
Video: *CHI03VIDEO/Yee.avi*

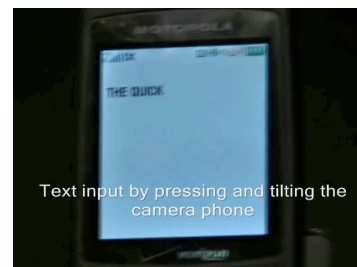
50

Embodied User Interface

- Camera phone based motion sensing



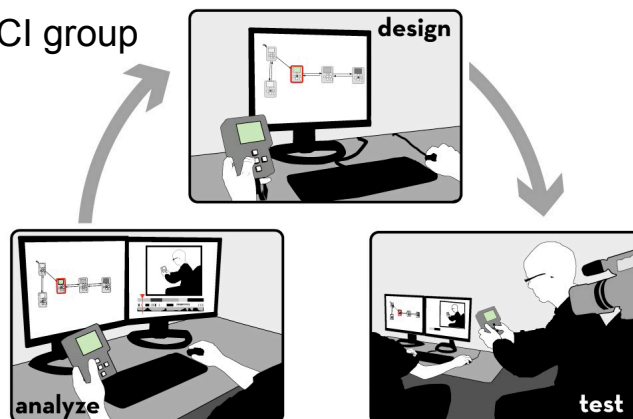
Uni. of California, Berkeley
& IBM
Video: Phone3DMvt-UIST06.mov



51

Embodied User Interface

- Prototyping tool for embodied UI
 - Stanford HCI group

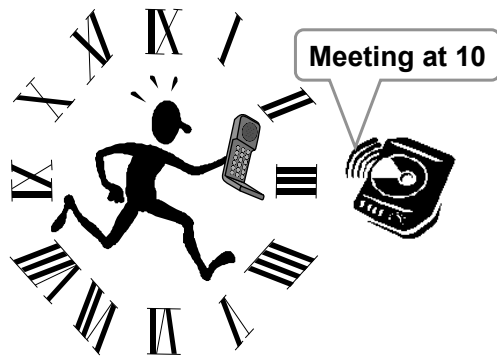


–Video: MapPDADesignplatform-UIST06.mov

52

3D sound

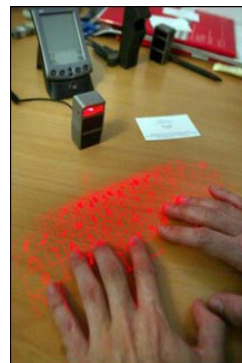
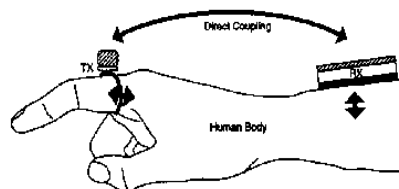
- Spatialized sound
- Mobile setting



Soundbeam
Neckset

53

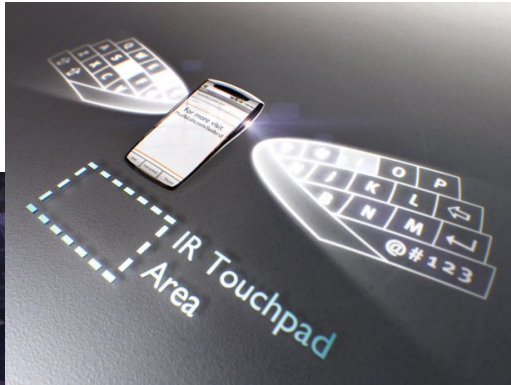
Keyboard



54

Keyboard

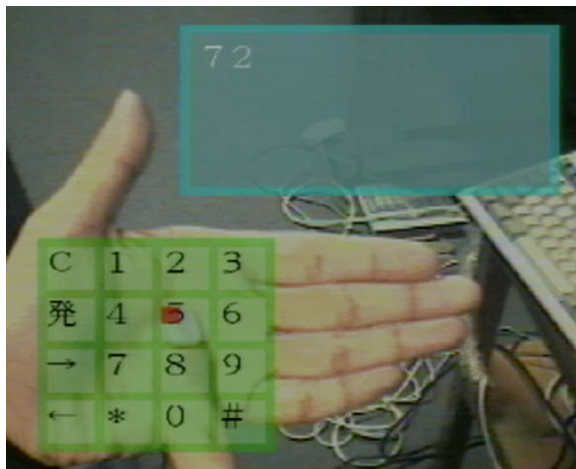
- 2010: result of an Open Web Concept Phone Project in which Mozilla collected suggestions and feedback from thousands of users



55

Keyboard

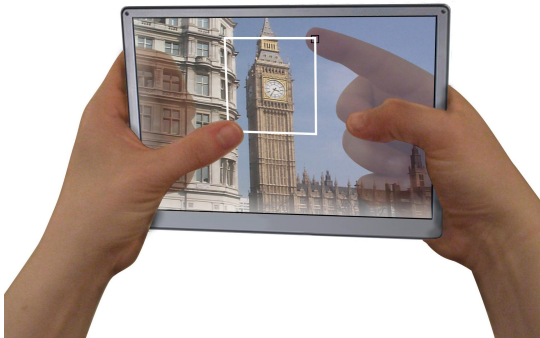
- Oulu University (Finland)
- Video: *Main.mpeg*



56

Interaction: LucidTouch:

- A See-Through Mobile Device
- LucidTouch.avi



57

Multi-surface interaction

- Pick&Drop

– Video: *Rekimoto_PickAndDrop.mov*

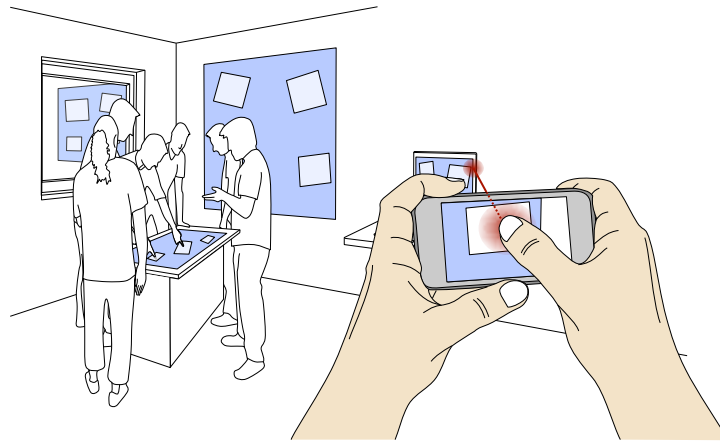


58

Multi-surface interaction

- TouchProject

– Video: TouchProjector.wmv



59

Direct manipulation and mobility

- Key issues:
Transparency/Usability of mobile devices

- Challenges for HCI

- Limited interactional resources



- Interaction in mobility



STAMPED
08

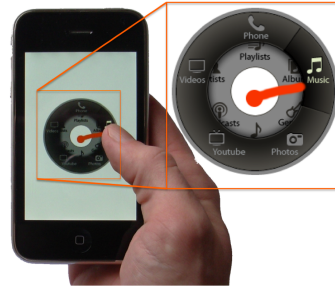
Direct manipulation and mobility

- Wavelet menu

- Video: MenuWavelet.video

- Problem space

- Space on screen
 - No keyboard for shortcuts
 - One-hand interaction
 - Eye-free interaction



61

MATCH

- Multimodal Access to City Help
- A Multimode Portable Device that accepts speech and pen gestures created by ATT&T



62

MATCH

- Part of a multi-million, multi-year contract from DARPA
- Enables users to interact using speech, pen, or synchronized combinations of speech and pen
- Essentially a testbed for designing portable multimodal applications

- *Video: DEMOMULTI/CityPlannerATT.mpeg*

63

MATCH Testing Statistics

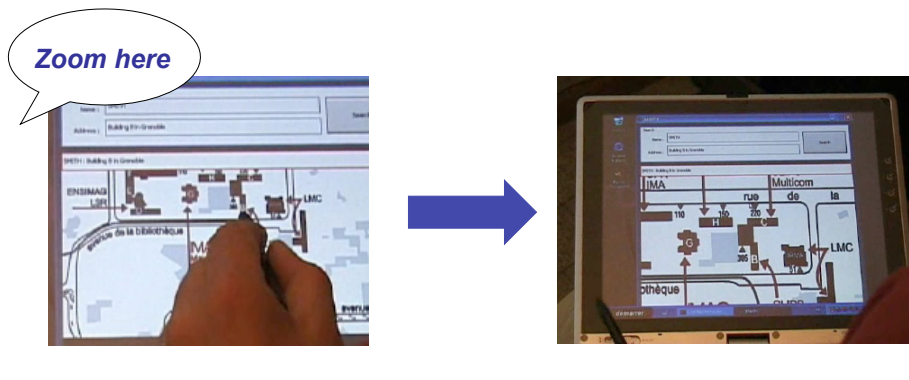
- Exchanges 338
 - Speech only 171 (51 %)
 - Multimodal 93 (28%)
 - Pen only 66 (19%)
 - GUI actions 8 (2%)



64

YellowPages

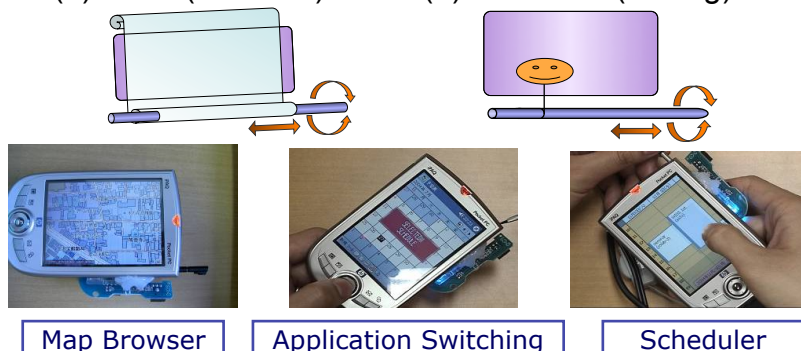
- YellowPages: LIG-IIHM
 - Voice command + Pointing using stylus
 - Video: *YellowPages-CLIPS.wmv*



65

RodDirect

- Two types of stylus movement in the stylus holder are associated with parameters
 - (1) Twist (Rotation)
 - (2) Push/Pull (Sliding)



JAIST: Japan Advanced Institute of Science and Technology, JAPAN
Video: *RodDirectDemo.mpg*

66

Multimodality on mobile phone

- Contact manager application on a mobile phone: LIG-IIHM
 - Devices : keyboard + microphone
 - Enabling forms filling using the stylus and speech commands



67

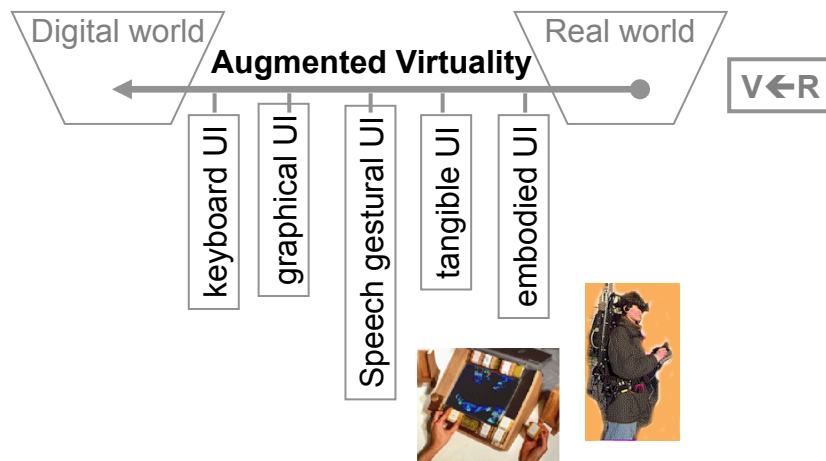
Commercial product

- www.kirusa.com



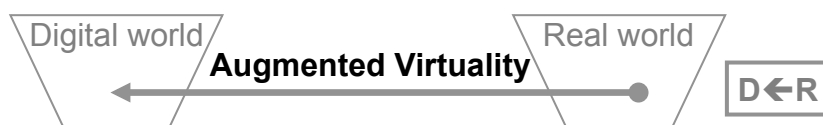
68

Augmented virtuality



69

Computer as tool Augmented Reality / Augmented Virtuality



Purpose of the task = computer

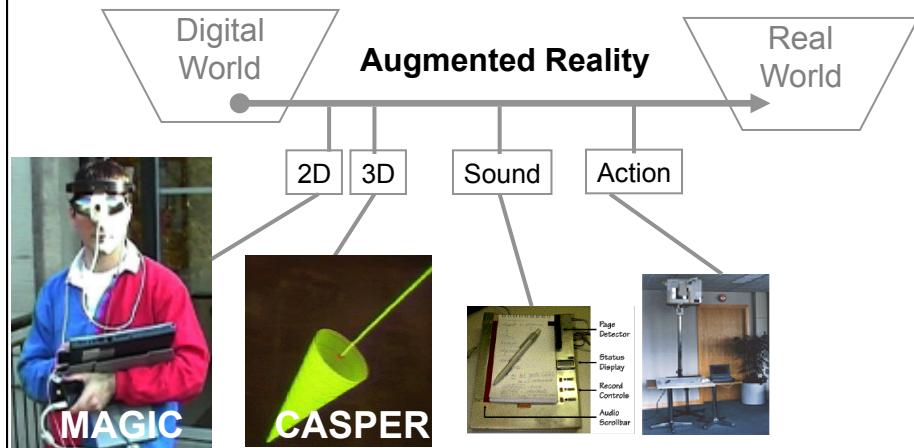


Purpose of the task = real world

70

Augmented Reality

- New interaction modalities

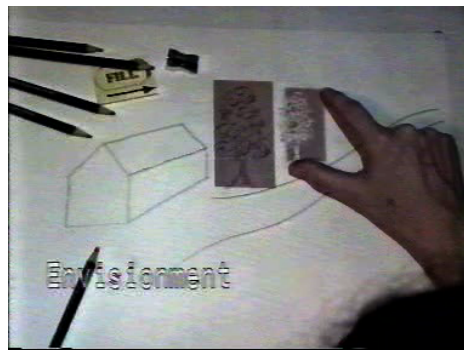


71

Augmented Reality

- First AR system
 - P. Wellner DigitalDesk

– Video: *DigitalSketching.mpg*



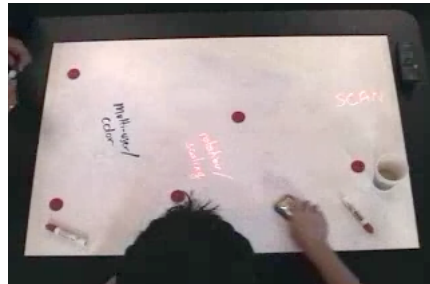
72

Augmented Reality

- MagicTable: LIG-IIHM

- Brainstorming

- Video: [MAGICBOARD/magicboard.mov](#)



73

Context-aware interactive system

- Sensing modalities



74

Context-aware interactive system

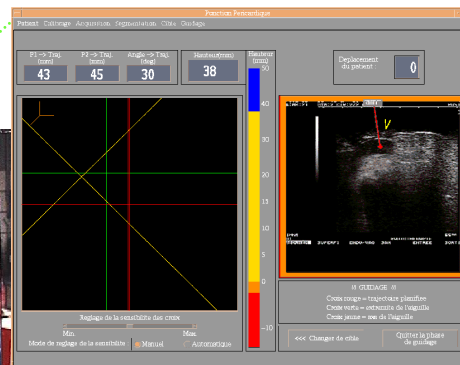
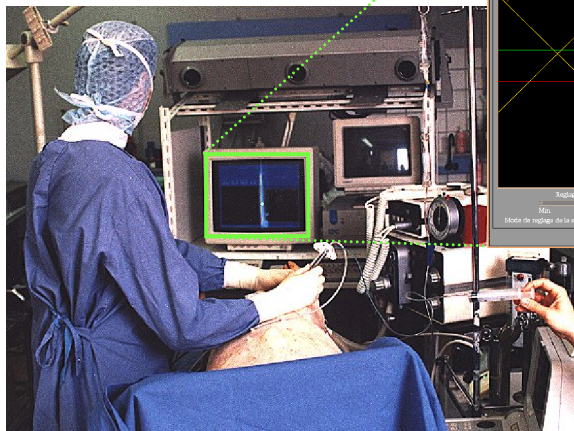
- Input modalities
 - Sensing modalities
 - 3D Location
 - 3D Orientation
- Output modality
 - Textual information displayed on HMD
 - Integration of virtual information and actions in the real world of the user through modalities



75

Augmented surgery: CASPER

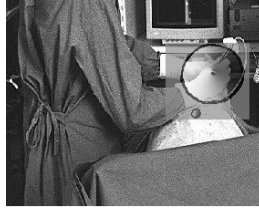
- LIG / TIMC



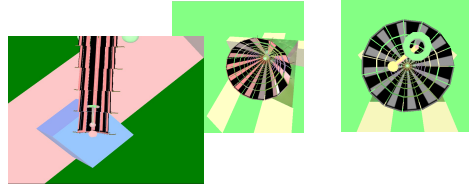
76

Augmented surgery: CASPER

- Output modality
 - Perceptual continuity



- Cognitive continuity



77

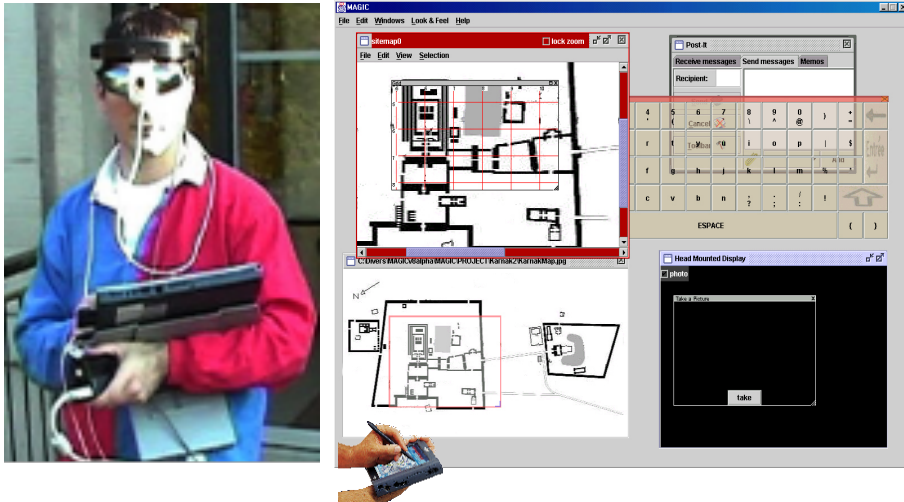
MAGIC hardware

- LIG-IIHM



78

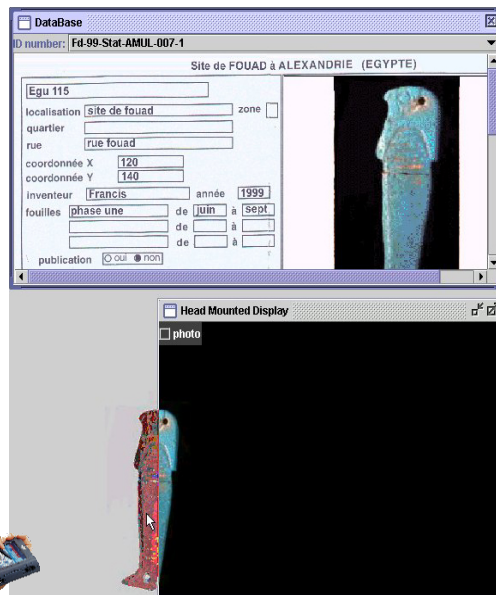
MAGIC platform



79

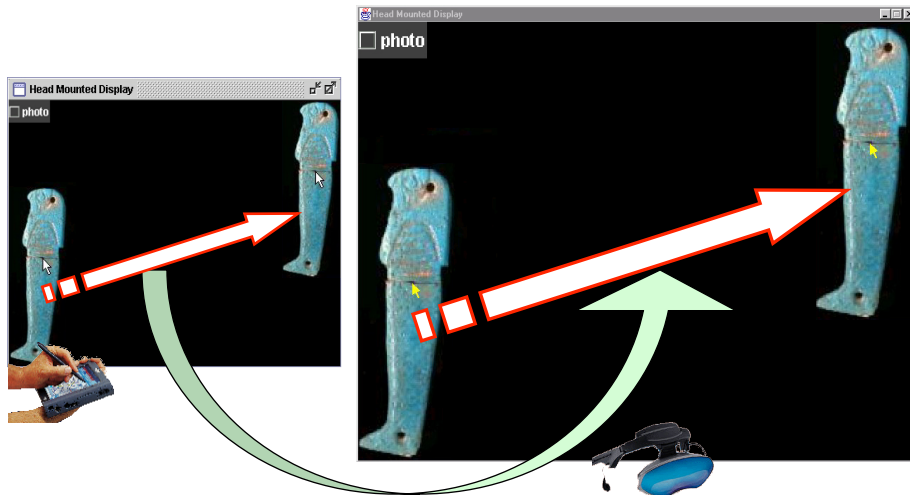
Combining Real And Virtual Entities: Comparing Both Kinds Of Entities

- Getting a picture from database
- Dragging it to the gateway
- and then positioning it according to real objects



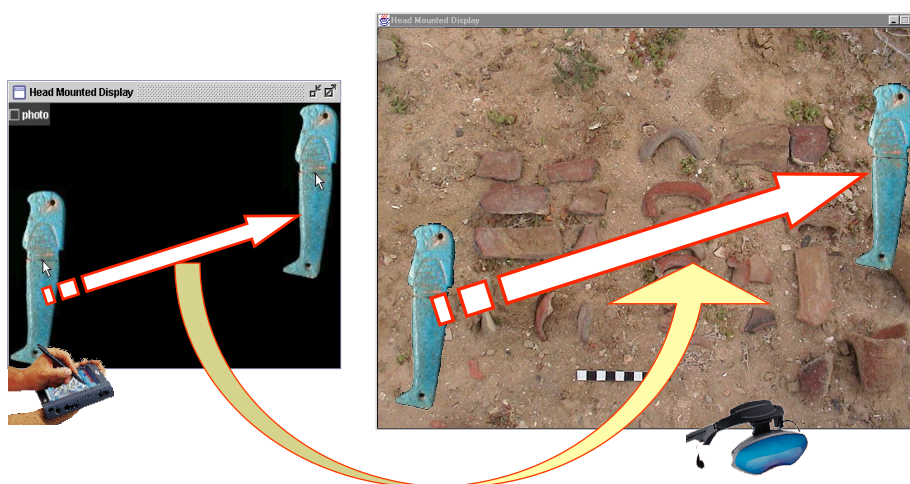
80

Combining Real And Virtual Entities: Comparing Both Kinds Of Entities



81

Combining Real And Virtual Entities: Comparing Both Kinds Of Entities



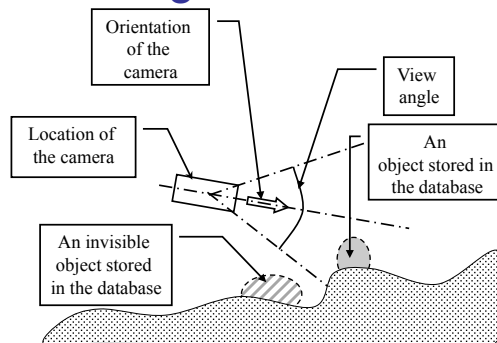
82

Combining Real And Virtual Entities: Comparing Both Kinds Of Entities



83

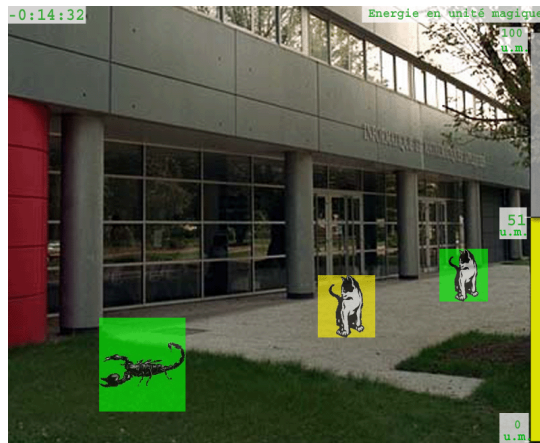
Combining Real And Virtual Entities: Augmented Stroll



- Collaborative (asynchronous): sharing of finds
- Mobile: walk on the excavation ground
- Augmented Reality: augmentation of the evolving excavation ground by saved information

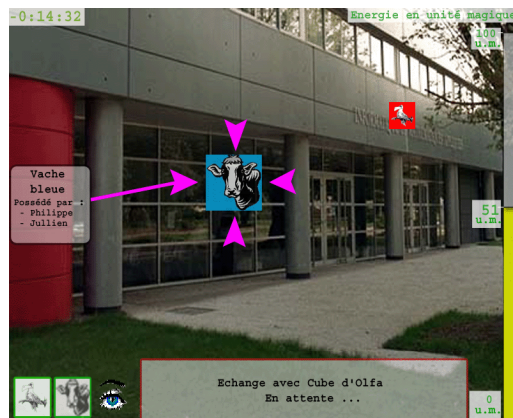
84

TROC



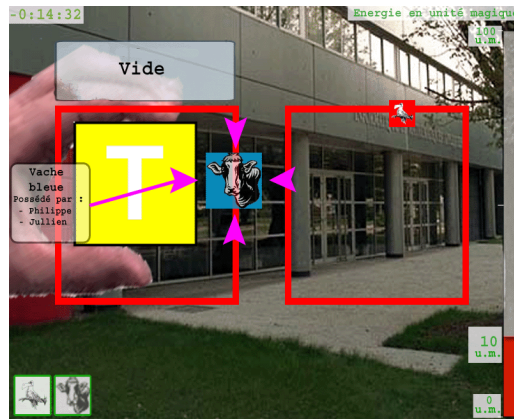
85

TROC



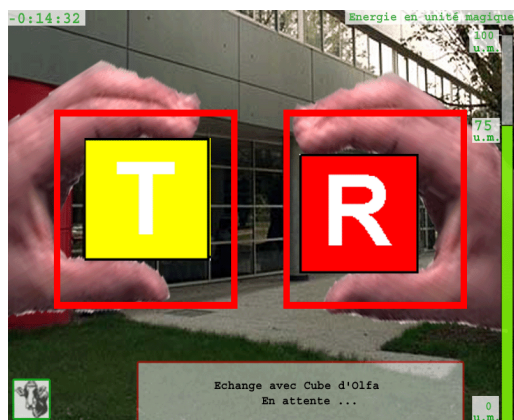
86

TROC



87

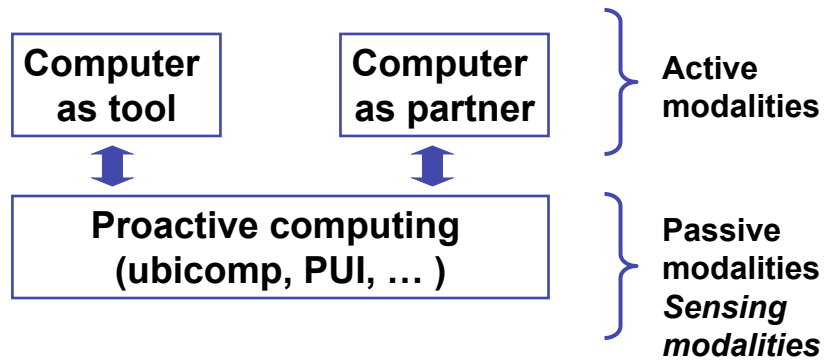
TROC



88

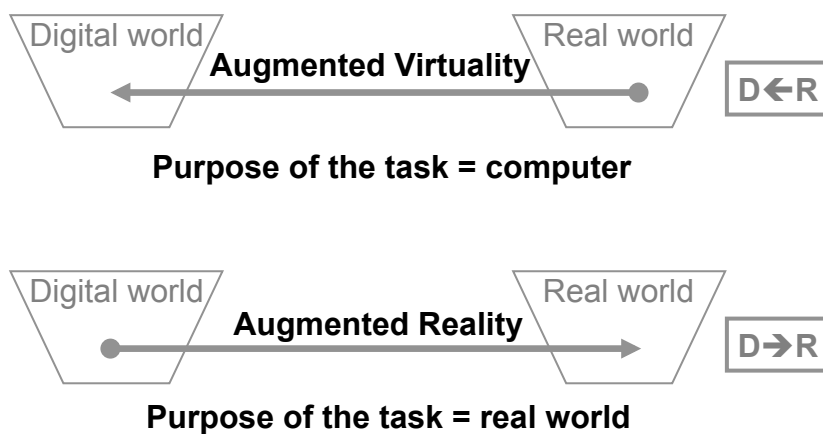
Conclusion

- Three paradigms for multimodality



89

Conclusion: Computer as tool



90

Conclusion

- Modality and multimodality:
- **A VAST space of possibilities to be explored**
- Multimodality is an integrating vector for several recent interaction paradigms that include:
 - perceptual user interfaces
 - tangible interfaces
 - Visualisation:
 - Modality = Output graphical interaction technique
 - Augmented Virtuality / Augmented Reality:
 - Modality based on physical objects
 - Multimodality:
 - Real world (Action/Perception)
 - Digital world (Action/Perception)

91

Readings

- Multimodal Communication for the Blind
<http://www.limsi.fr/Individu/bellik/francais/meditor.htm>
- Non-Visual Interfaces for Wearable Computers
http://www.dcs.gla.ac.uk/~stephen/papers/IEE_wearables_00.pdf
- MATCH: An Architecture for Multimodal Dialogue Systems
<http://www2.research.att.com/~johnston/matchacl02.pdf>
- MIPAD: A Multimodal Interaction Prototype
<http://research.microsoft.com/pubs/75373/2001-huang-icassp.pdf>
- RodDirect: <http://ist.mns.kyutech.ac.jp/miura/roddirect/index.php>
- DigitalDesk: Communications of the ACM Volume 36 , Issue 7 (July 1993) Special issue on computer augmented environments: back to the real world

92

Readings

- Berkeley Institute of Design - Uni. of California at Berkeley
<http://bid.berkeley.edu/>
- Stanford HCI group
<http://hci.stanford.edu/research/>
- Georgia Tech <http://www.gvu.gatech.edu>
- MIT Media Lab <http://www.media.mit.edu/research/>
- Carnegie Mellon Uni. - HCI Institute
<http://www.hcii.cmu.edu/>
- Uni. of Glasgow - Multimodal Interaction Group
<http://www.dcs.gla.ac.uk/~stephen/>
- Microsoft Research <http://research.microsoft.com/>
- Hasso-Plattner-Institut
http://www.hpi.uni-potsdam.de/forschung/fachgebiete/human_computer_interaction.html

93