Laurence Nigay

Interaction modality and multimodality

Joseph Fourier University
LIG Grenoble Informatics Laboratory
PERSYVAL-lab
Engineering for Human-Computer Interaction

JIRC Journées Informatique de la Région Centre

EHCI group

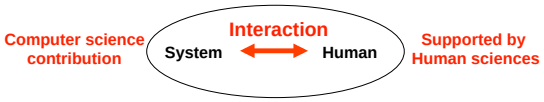
- "EHCI" for Engineering for Human-Computer Interaction
  - Created in 1990
  - 9 professors and assistant professors
  - 1 CNRS researcher
  - 1 CNRS research engineer part time
  - 14 PhD students
  - 5 project engineers
  - 1 visiting researcher
  - 10 master students

EHCI group

- EHCI group belongs to LIG (Computer Science Laboratory of Grenoble)
  - 188 academic researchers
  - 367 doctoral / post-doctoral students, visitors, engineers
  - 45 technical and administrative staff
- 23 autonomous research teams
- 4 principal themes
  - Infrastructure (networks and data)
  - Software (foundations and design models)
  - Interaction (perception, action and dialog)
  - Knowledge (learning, agent models and web-ontologies)

EHCI group: Research thematic

- Software Engineering for HCI



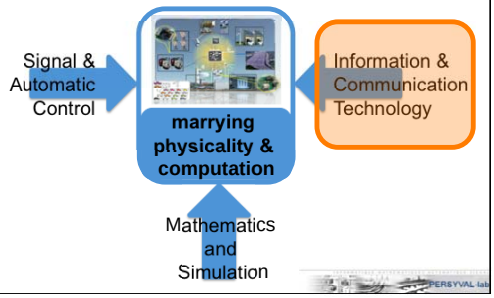
- Understanding the human-computer phenomena
- Establishing links between psychology-ergonomics and software engineering
- Designing, developing and evaluating interaction techniques
- Developing conceptual and technical tools based on HCI principles: Utility, Usability, Context

EHCI group: Research axes

- 5 complementary research axes
- Multimodal interaction
- Interaction with small handheld devices
- Mixed reality interaction
- User interface plasticity
- New interaction techniques

Laboratory of excellence Persyval-lab

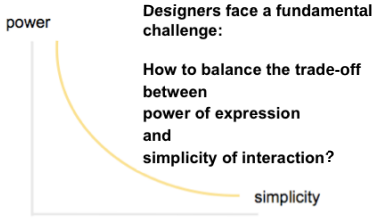
- Pervasive Systems & Algorithms at the convergence of the physical and digital worlds



Outline

- Research framework
  - Scientific themes
  - Research approach
  - HCI & Ubiquitous computing
  - Three research axes
- Foundations of my work
  - Interaction modality
  - Multimodality
- Combining the real and virtual worlds

Scientific themes: Design trade-offs

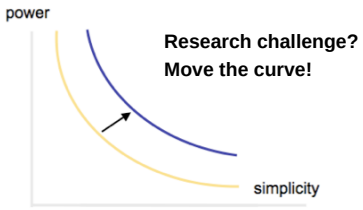


Designers face a fundamental challenge:

How to balance the trade-off between power of expression and simplicity of interaction?

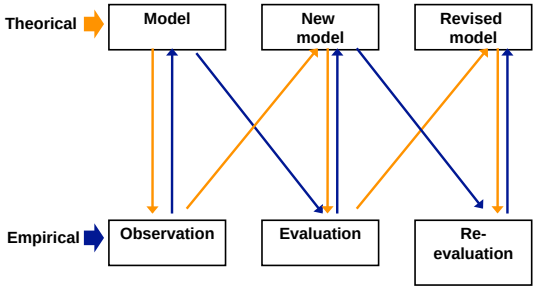
simple things should be simple
complex things should be possible

Scientific themes: Design trade-offs

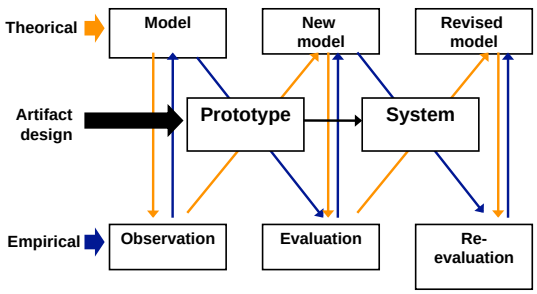


Research challenge?
Move the curve!

Scientific approach: deduction and induction



Scientific approach: deduction and induction



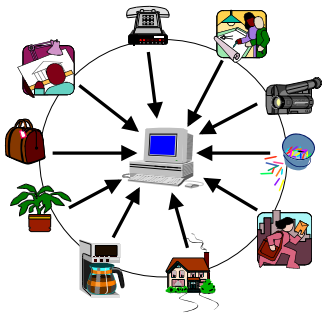
HCI & Ubiquitous Computing

- Ubiquitous computing 1991
  M. Weiser Scientific American
- Calm technology

- Invisible technology
- Technology available at any place
- Symbiosis of the real and digital worlds

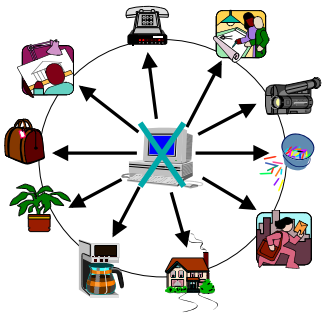
HCI & Ubiquitous Computing

- Virtual Reality



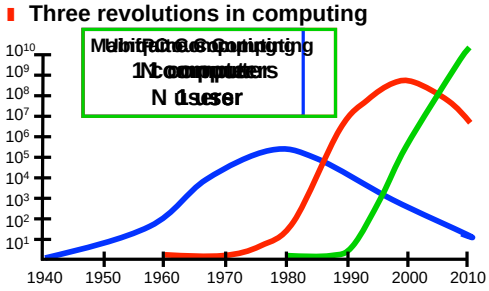
HCI & Ubiquitous Computing

- Embodied Virtuality



Ubiquitous Computing

- Three revolutions in computing



Machines sold per year
(M. Weiser, UIST 94)

HCI & Ubiquitous Computing

- HCI in the context of "Ubiquitous computing"
  - a seamless interactive environment
  - unobtrusive, everywhere
  - often invisible and yet in our consciousness

HCI & Ubiquitous Computing

- HCI in the context of "Ubiquitous computing"
  - "Our surrounding is the interface"
    to a universe of integrated services. This will enable citizens to access IST services wherever they are, whenever they want, and in the form that is most "natural" for them

HCI & Ubiquitous Computing

- HCI in the context of "Ubiquitous computing"

- Unprecedented challenges for interaction design
  - Combining the real and virtual worlds
  - Multiple interaction devices/modalities
    - Small and large interaction surfaces
  - Dynamic contexts of use

HCI & Ubiquitous Computing

- HCI in the context of "Ubiquitous computing"

- Unprecedented challenges for interaction design
- … while defining an opportunity for the users' acceptance of innovative interaction techniques going beyond the desktop paradigm

The Xerox Star has reached its limits

HCI & Ubiquitous Computing

- My scientific goals:

1. Design and realisation of interaction techniques for large information spaces (*interactive visualisation*)
2. Design and realisation of interaction techniques on mobile devices (*mobile techniques*)
3. Design and realisation of interaction techniques for combining the real and virtual worlds (*combining real and virtual worlds*)

HCI & Ubiquitous Computing

- My scientific goals:

1. Design and realisation of interaction techniques for large information spaces (*visualisation*)
2. Design and realisation of interaction techniques on mobile devices (*mobile techniques*)
3. Design and realisation of interaction techniques for combining the real and virtual worlds (*combining real and virtual worlds*)

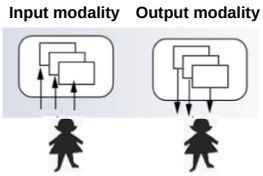
From the point of view of **Interaction modality and Multimodality**

Outline

- Research framework
  - Scientific themes
  - Research approach
  - Ubiquitous computing
  - Three research axes
- Foundations of my work
  - Interaction modality
  - Multimodality
- Combining the real and virtual worlds

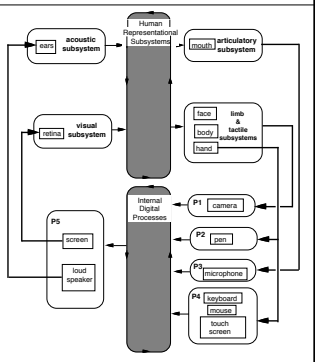
Interaction modality

- Modality = (device, interaction language)
  - A set of sensors (input devices) or effectors (output devices)
  - A processing facility based on a language



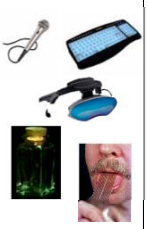
Interaction modality

- Theory ICS
  - APU Cambridge
- ICS as predicting cognitive resources involved in using and choosing modalities

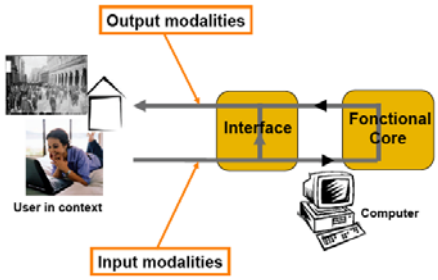


Interaction modality

- Modality = (device, interaction language)
- Recent interaction paradigms such as perceptual UI tangible UI and embodied UI open a vast world of possibilities
  - M1 = (microphone, natural language)
  - M2 = (keyboard, command language)
  - M3 = (mouse, direct manipulation)
  - M4 = (PDA, 3D gesture) *embodied UI*
  - M5 = (HMD, 3D graphics) *AR*
  - M6 = (bottle-sensor, 3D gesture) *tangible UI*
  - M7 = (GPS, localization) *perceptual UI*
  - M8 = (Tongue display, 2D shape)



Interaction modality



Interaction modality

- INPUT Modality = <d, I>

Modality = < , natural language>

MATIS project

Interaction modality

- INPUT Modality = <d, I>

M = <camera-finger, gesture>

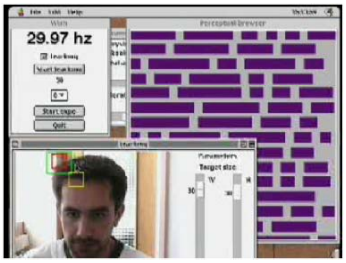
M = <micro, NL>

M = <PDA, gesture>
Embodied UI

M = <stylus, direct manipulation>

Interaction modality

- Input M = <camera-head, gesture>

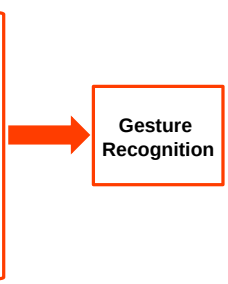


Perceptual browser project

Interaction modality

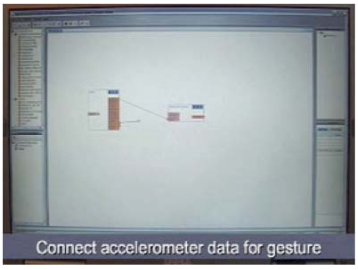
- Input M = <devices, gesture>
  - Devices: wiimote, iPhone, trackIR, Shake …

SHAKE
triple axis accelerometer, two capacitive sensors and a vibrotactile actuator



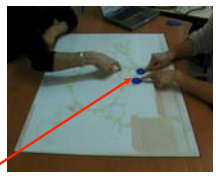
Interaction modality

- **Input M = <devices, gesture>**
  - **Examples: wiimote and trackIR**



Definition of a modality

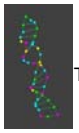
- **Input M = <camera-token, direct manipulation>**



Tokens

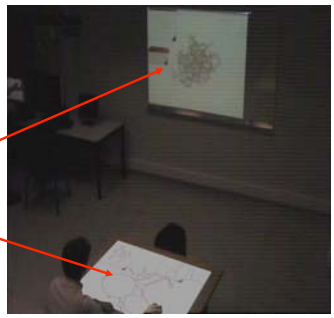
The position of tokens are tracked by a vision based mechanism.
A tracking video camera is fixed on the top of the table.
NAVRNA: A system to visualize, explore and edit RNA.

Definition of a modality



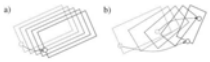
Tertiary structure

Secondary structure



Definition of a modality

- **Input M = <camera-token, direct manipulation>**



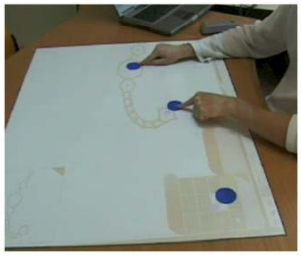
a) One token:
panning action

b) Two tokens:
zoom + rotate

Definition of a modality

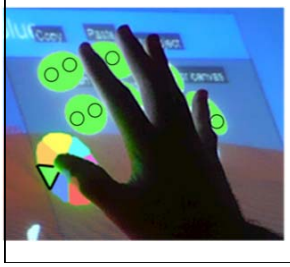
- **Input M = <camera-token, direct manipulation>**

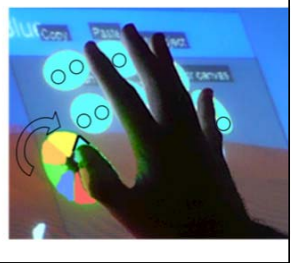
Two-handed interaction: Bending one branch of RNA



Definition of a modality

- **Input M = <multi-touch table, direct manipulation>**
- **Multitouch menu: MTM**

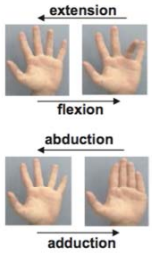




Definition of a modality

- Input M = <multi-touch table, direct manipulation>
- Multitouch menu: MTM

15 degrees of freedom
Thumb: 5 degrees of freedom



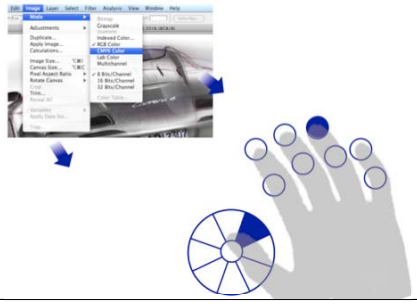
Bulchholz, B., Armstrong, T. J., (1992) A Kinematic model of the human hand to evaluate its prehensile capabilities, J. Biomecanics.

Definition of a modality

- Input M = <multi-touch table, direct manipulation>
- Multitouch menu: MTM

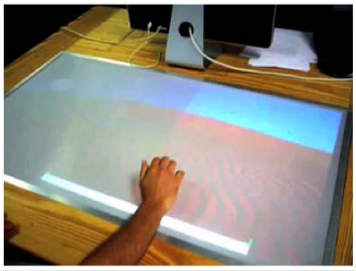
Definition of a modality

- Input M = <multi-touch table, direct manipulation>
- Multitouch menu: MTM



Definition of a modality

- Input M = <multi-touch table, direct manipulation>
- Multitouch menu: MTM



Definition of a modality

- Input M = <multi-touch surfaces, interaction langage>
- Device: cubtile
  - Immersion *www.immersion.fr*
  - First 3D multitouch interface
  - 5 multi-touch surfaces
  - Rear Diffused Illumination
  - Multiuser

Definition of a modality

- Input M = <multi-touch surfaces, interaction langage>
- Device: cubtile
- One interaction langage per surface
- Two-handed interaction

Definition of a modality

- **Input M = <multi-touch surfaces, interaction langage>**
- **Two handed-interaction**
  - **Psychological Theory -Kinematic chain – Y. Guiard**
    - Right-to-left reference: The right hand performs its motion relative to the frame of reference set by the left hand
    - Asymmetric scales: Different temporal-spatial scales of motion
    - Left hand precedence: The left hand precedes the right: for example, the left hand first positions the paper, then the right hand begins to write
    - Right hand preference: Is the one finishing the action, touching the world

Definition of a modality

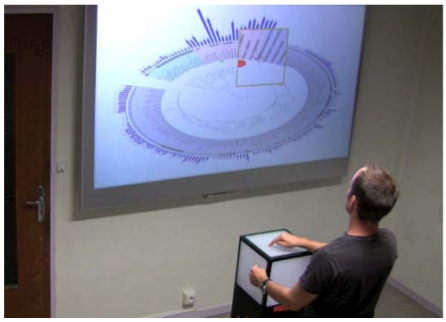
- **Input M = <multi-touch surfaces, interaction langage>**
- **Two handed-interaction**

Action within the lens

Moving the Lens

Definition of a modality

- **Two handed-interaction**



Definition of a modality

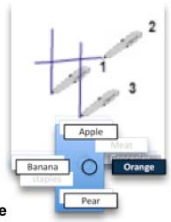
- **Input M = <small interactive surface, direct manipulation>**
- **Wavelet menu on iPhone**
  - **Space on screen**
  - **No keyboard for shortcuts (novice mode)**
  - **The best way to interact is to use only one-hand**
  - **Eye-free interaction**

Input modalities on small devices

- **Marking menus**
- Advantages
  - Circular design
  - Fluid transition
  - Scale independance
- Limitations
  - Screen space requirement
  - Number of commands
    - Error rate in expert mode
    - Ambiguous marks in expert mode

Input modalities on small devices

- **Marking menus**
- **Multi-stroke marking menus**
- **Instead of considering a spatial compound stroke,**
- **Multi-Stroke menus introduce a serie of simple strokes**
- **Require less physical input space in novice & expert modes**
  - **A submenu is displayed on top of its parent menu**
  - **Overlapped marks**
- **Increase accuracy in expert mode**
- **Increase the number of items**
  - **No ambiguous gestures in expert mode**

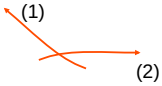


Input modalities on small devices

Wave menus

- **Few physical input space**
- **A submenu is displayed at the center of the menu system. In order to remain visible, its parent menu is then enlarged to surround this submenu**
- **In expert mode, Wave Menus work exactly the same way as Multi-Stroke Menus**

Novice Mode



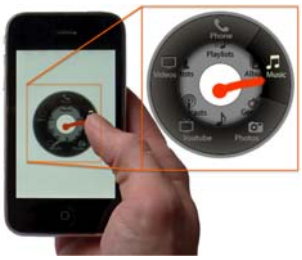
Expert Mode

Input modalities on small devices

Wave menus

Input modalities on small devices

Wavelet menus: extension on Iphone



Input modalities on small devices

Wavelet menus: extension on Iphone

- The Wavelet menu appears centered around the contact point.
- By drawing a stroke towards the desired item, the first level is enlarged permitting progressive appearance of the submenu.
- A second stroke selects an item in the submenu.

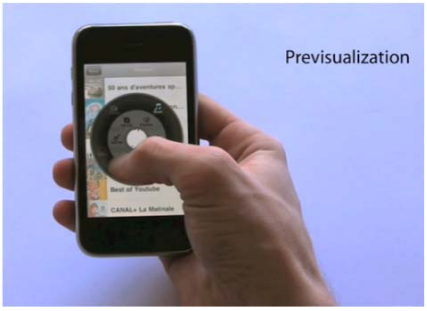
Input modalities on small devices

Wavelet menus: extension on Iphone

Long lists management: the linear list appears in the center of the Wavelet menu and is surrounded by its parent menus.

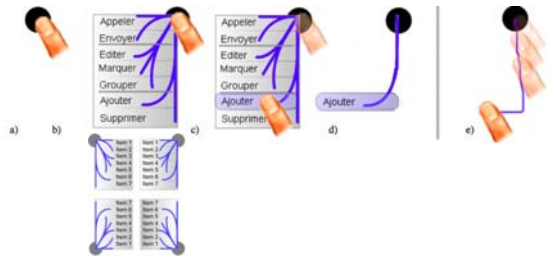
Input modalities on small devices

Wavelet menus:



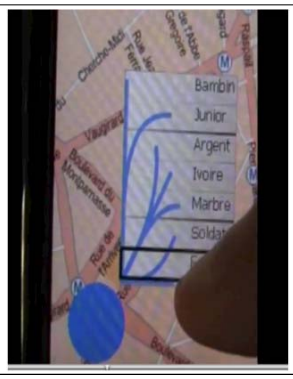
Input modalities on small devices

- LeafMenu: extension of linear menus



Input modalities on small devices

- LeafMenu: extension of linear menus



Definition of a modality

- Input M = <bottle-sensor, gesture>

Ambient room
Media Lab MIT

- Input M = <Object-sensor, manipulation>

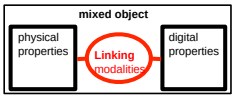
Movements Pressure Torsion

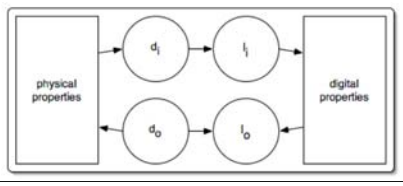
Definition of a modality

- Input M = <Object-sensor, manipulation>
- Object-Sensor = Mixed objects
  - MIM: Model of Mixed Interaction
    A conceptual model for designing augmented objects
  - OP (Object Prototyping): A toolkit based on the conceptual model MIM for prototyping mixed objects

Definition of a modality

- Input M = <Object-sensor, manipulation>
  - MIM: Model of Mixed Interaction

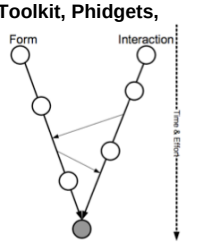




Definition of a modality

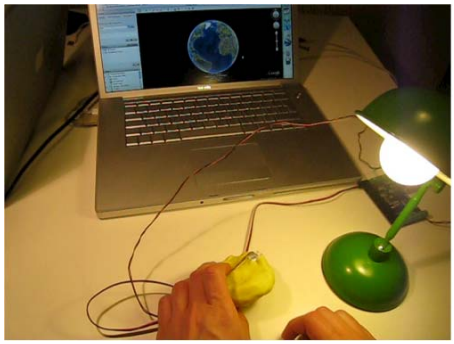
- Input M = <Object-sensor, manipulation>
  - OP: Object prototyping
    - Toolkit developed in Qt
    - Toolkit that includes ARToolkit, Phidgets, Interface-Z sensors

OP - Prototyping mixed objects: Combined development of **form and interaction**.



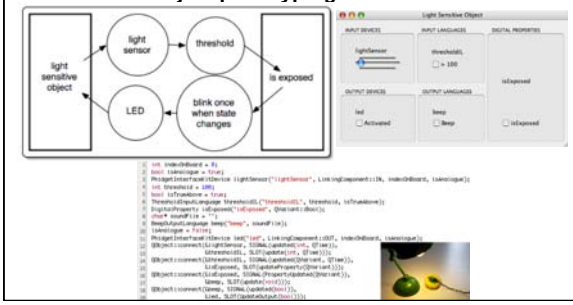
Scott E. Hudson HCI CMU

Definition of a modality



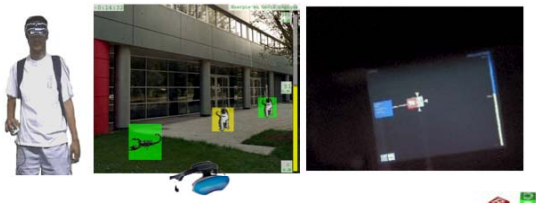
Definition of a modality

- Input M = <Object-sensor, manipulation>
  - OP: Object prototyping



Definition of a modality

- Input Modalities (*sensing modalities*)
  - M1 = <GPS, localization>
  - M2= <magnetometer, orientation>



TROC project

Definition of a modality

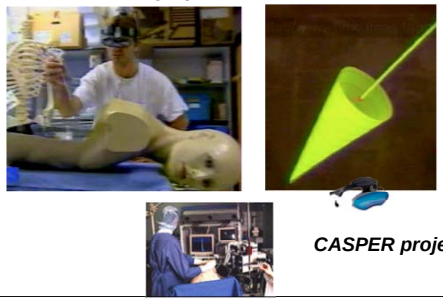
- Input Modalities

Get

  - M1 = (micro, command)
  - M2 = (trackpad, 2D gesture)
  - M3 = (camera, 3D gesture)

Definition of a modality

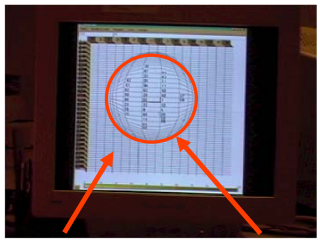
- OUTPUT Modality = <d, I>
  - M = <HMD, 3D graphics>



CASPER project

Definition of a modality

- OUTPUT Modality = <d, I>



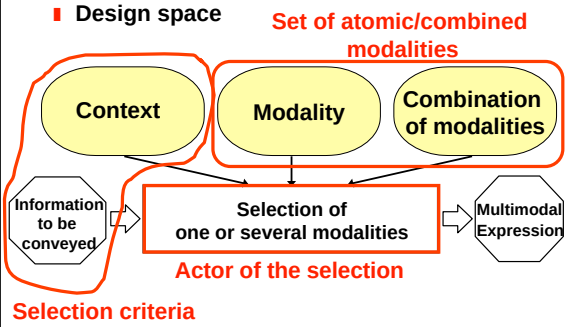
M = <screen, table> M = <screen, deformed table>

Outline

- Research framework
  - Scientific themes
  - Research approach
  - Ubiquitous computing
  - Three research axes
- **Foundations of my work**
  - Interaction modality
  - **Multimodality**
- Augmented Reality

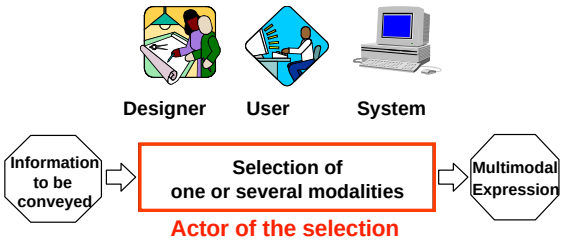
Multimodality

- **Design space**



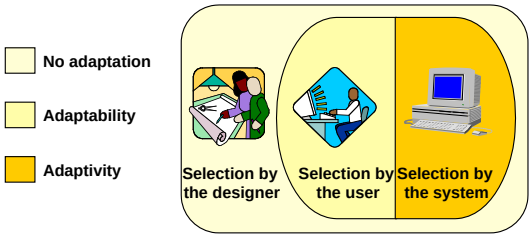
Multimodality

- **Who is performing the selection**



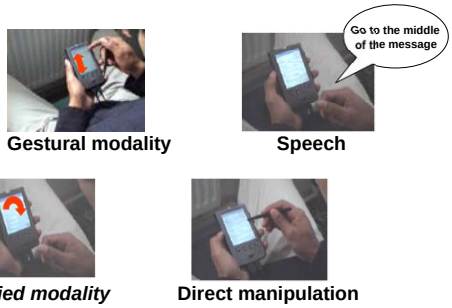
Multimodality

- **Who is performing the selection**



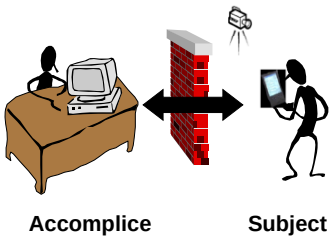
Multimodality Adaptability

- **Selection of the modalities by the user**



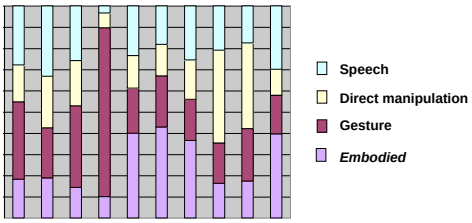
Multimodality Adaptability

- **Wizard of oz**



Multimodality Adaptability

- Use of the modalities
- All sessions / All subjects



- The subjects used all of the modalities
- Individual preferences leading in some cases to specialization

Multimodality Adaptability

- Selection of the modalities by the user


- M1 = (trackpad, 2D gesture)

Tactile

- M2 = (camera, 3D gesture)

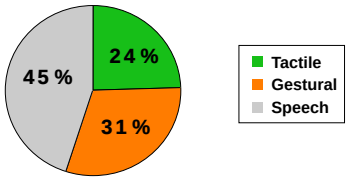
Gestural

- M3 = (micro, command)

Speech

Multimodality Adaptability

- Use of the modalities
- All sessions / All subjects



- Individual preferences leading in some cases to specialization
  - Subject 4: Tactile modality for rotating a puzzle piece
  - Subject 9: Speech for rotating and collecting a puzzle piece

Multimodality Adaptivity

- Selection of the modalities by the system
- Context-aware systems

Ring

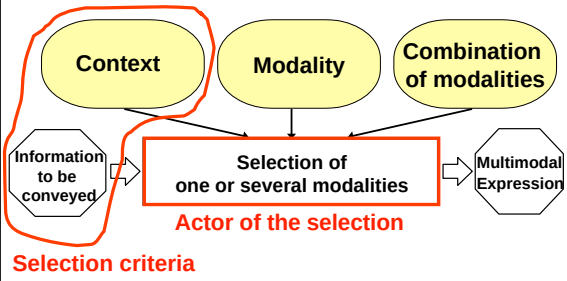
Vibration

Multimodality Adaptivity

- Selection of the modalities by the system
- Context-aware systems
- Grand challenges
  - To what extent can modality selection & configuration be automated?
    - Proactiveness (take the initiative), Predictability, Transparency
  - Where full automation is not possible or desirable, how can human involvement be supported?
    - End-user programming of context-sensitive environments

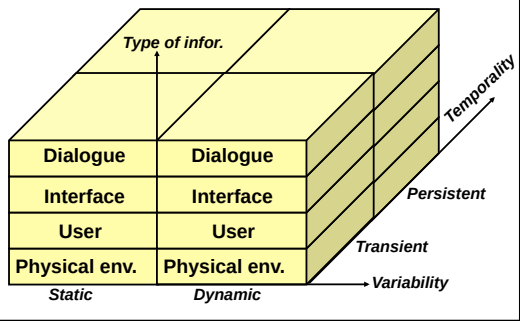
Multimodality

- Design space



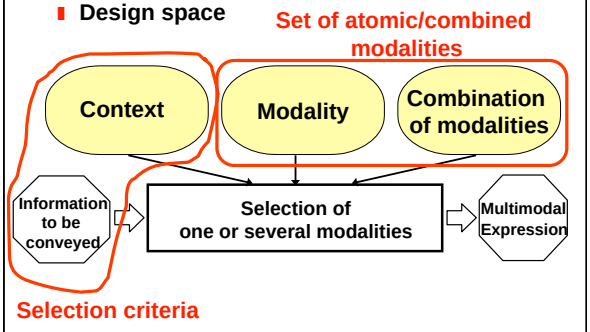
Multimodality

- Selection criteria: Context



Multimodality

- Design space



Atomic and combined modalities

- Interaction modelling at the modality level
  - Rich enough to express differences

  - Abstract enough to enable reasoning
    - among modalities (vast world of modalities)
      - any physical object can be involved in interaction
    - between modality and task/system/service/ context issues

Atomic and combined modalities

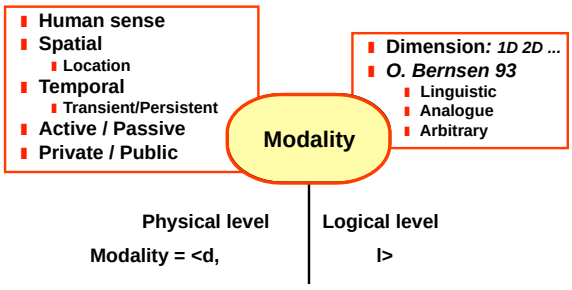
- A vast world of atomic and combined modalities
  - any physical object can be involved in interaction as a device
- We can no longer expect to model each input and output modality in all their diversity at the concrete level
- We need to reason about modalities at a higher level of abstraction

Multimodality

- Characterisation of a modality
- Active modalities
  - For inputs, active modalities are used by the user to issue a command to the computer such as a pedal to move a laparoscope in a CAS system
- Passive - Implicit modalities
  - 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 (PUI). For example tracking position

Multimodality

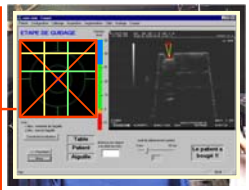
- Characterisation of a modality



Multimodality

- Characterisation of a modality

- Physical level
  - Human sense: Sight
  - Spatial: Location = screen
  - Temporal: Persistent
  - Public
- Logical level
  - 2D
  - Non Analogue
  - Arbitrary



Multimodality

- Characterisation of a modality

- Physical level
  - Human sense: Sight
  - Spatial: Location = operating field
  - Temporal: Persistent
  - Private
- Logical level
  - 3D
  - Analogue
  - Non arbitrary

Multimodality

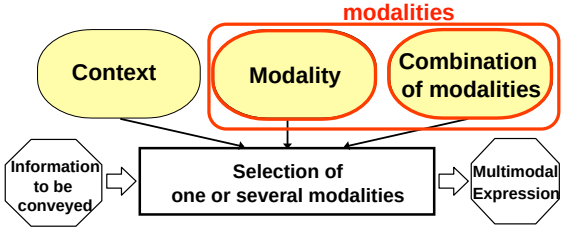
- Characterisation of a modality

- Physical level
  - Human sense: Sight
  - Spatial: Location = screen
  - Temporal: Persistent
  - Public
- Logical level
  - 3D
  - Analogue
  - Non arbitrary

Multimodality

- Design space

Set of atomic/combined modalities



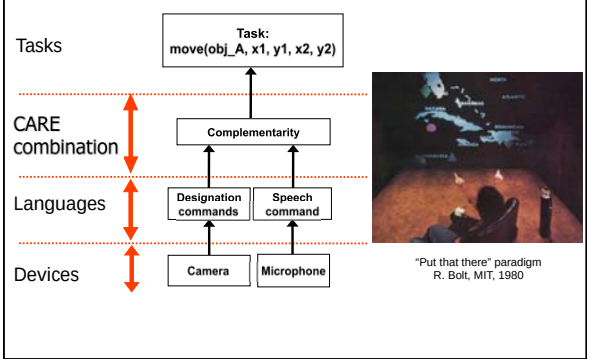
Multimodality

- Combination of modalities

- Several studies
  - UOM 94 / TYCOON 95 / CARE 95

- CARE properties
  - Relationships between Devices, Interaction languages and Tasks
    - C : Complementarity
    - A : Assignment
    - R : Redundancy
    - E : Equivalence

Multimodality

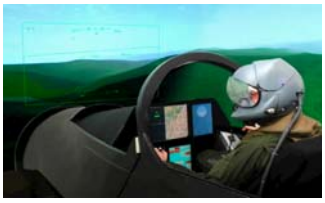


"Put that there" paradigm
R. Bolt, MIT, 1980

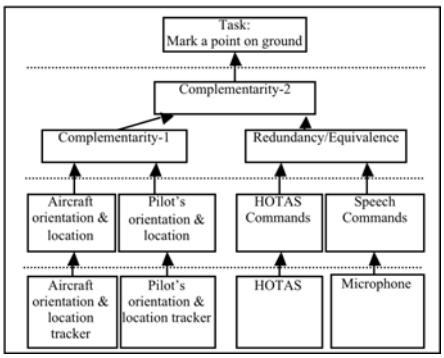
Multimodality

- **Flight simulator of a military fighter, used for studying future interaction techniques in the cockpit**

Modalities:
M1: Aircraft location
M2: Pilot's orientation
M3: HOTAS commands
M4: Speech commands



Multimodality



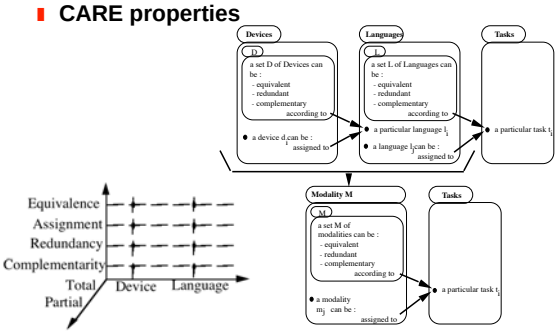
Multimodality

- **TROC: a game based on the technique of barter**
- **M1 = (Magnetometer, orientation)**
- **M2 = (GPS, location)**
- **Complementarity of M1 and M2 for selecting an object**

Multimodality

Multimodality

- **CARE properties**



Multimodality

- **Combination of modalities**
- **CARE properties**
- **New combination space**
  - **Different schemas and aspects of combinations**
  - **5 aspects: temporal, spatial, articulatory syntactic and semantic**
  - **5 schemas: [Allen 83]**

Multimodality

- Combination of modalities

Combination schemas

Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Articulatory	Independence	Fission	Fission Duplication	Partial Duplication	Total Duplication
Syntactic	Difference	Completion	Divergence	Extension	Twin
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

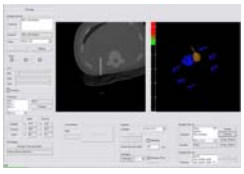
Multimodality

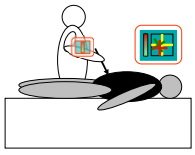
- AR Puzzle

Turn 48° to the left

Multimodality

- PERM system: a CAS system for kidney puncture





Multimodality

- Puzzle

M1 = <screen, 2D image>

M2 = <screen, color>

M3 = <mini-screen, crosses>

Multimodality

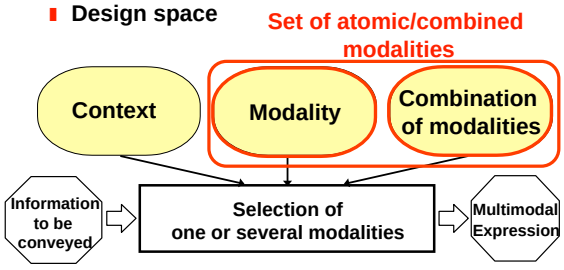
- Combination of M2 = <screen, color> and M3 = <mini-screen, crosses>

Combination schemas

Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Articulatory	Independence	Fission	Fission Duplication	Partial Duplication	Total Duplication
Syntactic	Difference	Completion	Divergence	Extension	Twin
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

Multimodality

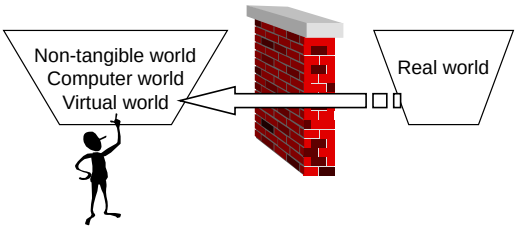
- Design space



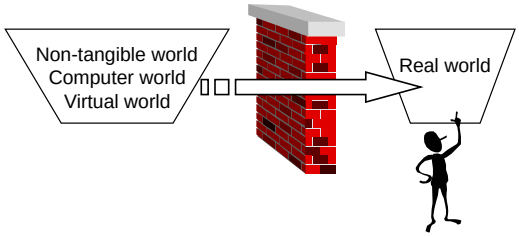
Outline

- Research framework
  - Scientific themes
  - Research approach
  - Ubiquitous computing
  - Three research axes
- Foundations of my work
  - Interaction modality
  - Multimodality
- **Combining the real and virtual worlds**

Combining the real and virtual worlds



Combining the real and virtual worlds



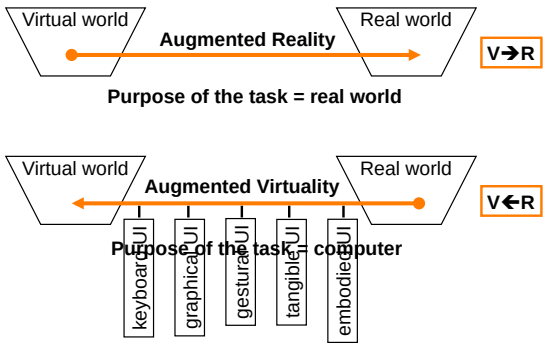
Combining the real and virtual worlds

- **Profusion of terms**
  - **Virtual reality**
  - **Bit / Atom**
  - **Computer Augmented Environment**
  - **Augmented Video**
  - **Augmented Interaction**
  - **Augmented Virtuality**
  - **Augmented Reality**
  - **…**

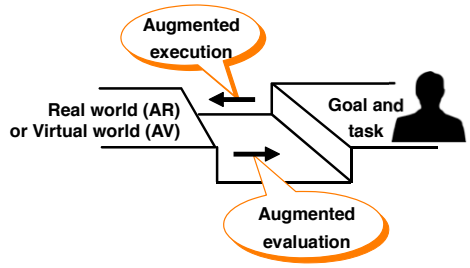
Combining the real and virtual worlds

- **Common objective**

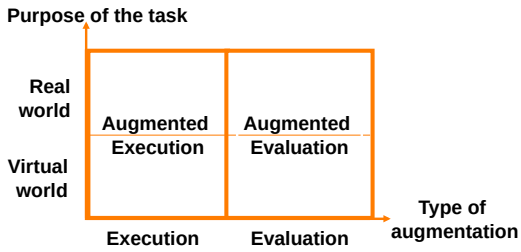
Augmented Reality Augmented Virtuality



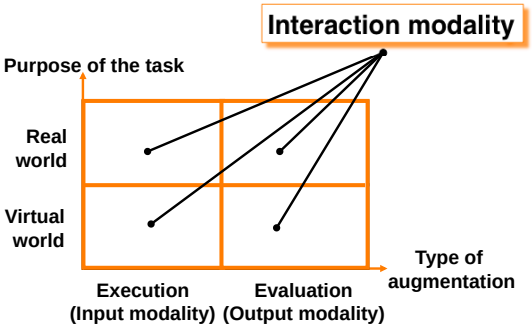
Augmented Reality Augmented Virtuality



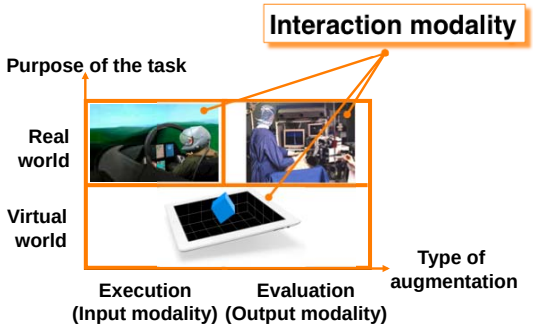
Augmented Reality Augmented Virtuality



Augmented Reality Augmented Virtuality



Augmented Reality Augmented Virtuality



Augmented Virtuality

- New input modality on mobile devices
  - intuitive
  - Inspired by natural movements

  - based on head tracking with the front facing camera

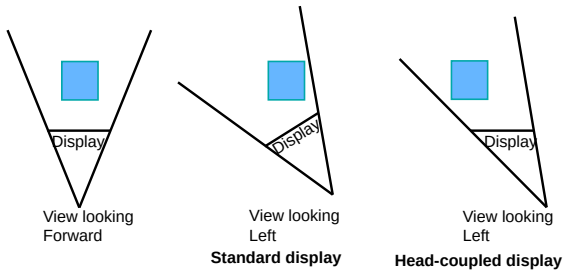
Augmented Virtuality

- Head tracking
  improves both input and output capabilities
- Input: Head movements to navigate
- Output: 3D effect

Augmented Virtuality

Augmented Virtuality

- Head-coupled perspective for 3D display
- Off-axis projection



Augmented Virtuality

- Off-axis projection
- Well adapted for 3D UI

Augmented Virtuality

Augmented Virtuality

- Video: viewed more than 2 million times – various articles (Wired 2011, MacStories 2011, …)
- Application on AppStore:
  i3D: Glasses-free monocular 3D
  - Downloaded 1,5 million times

Outline

- Research framework
  - Scientific themes
  - Research approach
  - Ubiquitous computing
  - Three research axes
- Foundations of my work
  - Interaction modality
  - Multimodality
- Combining the real and virtual worlds
- **Conclusion**

Conclusion

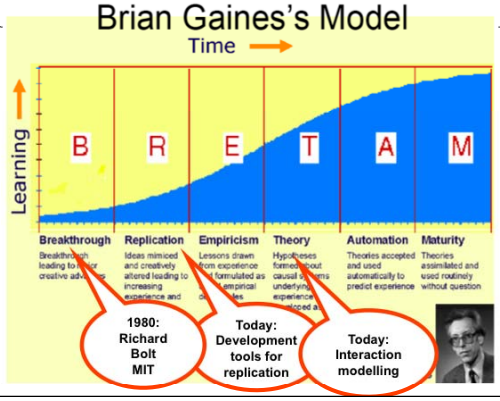


- **Modality and multimodality:**
  A VAST space of possibilities to be explored

- **Augmented Virtuality / Augmented Reality:**
  - **Multimodality:**
    - **Real world (Action/Perception)**
    - **Digital world (Action/Perception)**

- **Multimodality is an integrating vector for several recent interaction paradigms that include:**
  - **augmented reality**
  - **augmented virtuality**
  - **tangible interfaces**

In the 80's, Brian Gaines introduced a model on how science technology develops over time



Thank you