



HAL
open science

Multisensory vs multimodal interfaces. The central role of action. Some elements for a FP6 roadmap

Annie Luciani

► **To cite this version:**

Annie Luciani. Multisensory vs multimodal interfaces. The central role of action. Some elements for a FP6 roadmap. Information Society Technologies-Framework Programme 6, 2002, Luxembourg. pp.[8]. hal-00910571

HAL Id: hal-00910571

<https://hal.science/hal-00910571>

Submitted on 16 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IST – Interface Technologies in FP6 - Luxembourg - 13 & 14 May 2002

Multisensory vs Multimodal Interfaces

The central role of « action » Some elements for a FP6 Roadmap

Annie Luciani, ACROE-ICA
Computer Arts Laboratory

Institut National Polytechnique - Ministère de la Culture
Grenoble - France

I would like to propose a couple of elements for a roadmap in the domain of Interface Technologies. These elements are draft from observations related in the IPPA's documents.

The main conclusions of the five IPPA exercises are :

1. The scattering of the domain of Interfaces in several technologies and Key Action Lines. We can distinguish three circles :

the central circle, addressing the technologies of Multi-Modal & Multisensori Interfaces, of Language and speech technologies and of Visualisation, Virtual Environments & Image Processing

and the peripheral circles such as :

Micro and Opto Electronics and Micro – systems for the developments of new sensors
Signal and Data processing, for the matching of Multisensori Signals

Agent Technologies and Content Authoring tools, for the development and the use of VR and AR technologies

And Bio-informatics, neuro-informatics for the understand multisensoriality and the use of biological mimetics in artefacts

2. The KA line « Multimodal Interfaces » remain mainly in the field of the development of displays and camera, that signs the domination of the visual perceptive modality.
3. The core technologies for multisensoriality and cross-sensoriality, and the core understanding of the human sensory-motor characteristics are developed mainly in the FET.
4. There are a very few number of projects which deals with the developments of haptic devices and actuators, except new displays.

Even, if it is shocking to say, I think that we are always following a classical point of view, which derived from the historical concept of television and radiodiffusion in which the multimodality is understood :

1. as a domination of the perception modalities :
Visual channel,
Acoustical channel,
More recently, tentative to add the taste and smell perceptions
All this information are physical sensorial information received by the user.
2. the only information produced by the user was by mean of speech and it runs at symbolic level
3. The sensory information was understood as a collection of monomodal perceptions, in with low coupling between them
4. And in the same way, there is a low level of coupling between symbolic and physical information.

We can notice :

first the domination of perceptual modalities :

(1) visual channel, (2) the acoustical channel for verbal communication, (3) acoustical channel for non verbal communication

and second the Low level of coupling

(1) Between each physical information, and (2) Between physical and symbolic information

The beginning of interactivity introduces another emitting channel : the gesture as an input

That was the first physical information produced by the user. And we can notice the increasing of level of coupling particularly between input gestures and visual modality and input gestures and acoustical modality, which create a new but implicit coupling between visual and acoustic modality.

And the development of artificial intelligence introduces the face interpretation, which is a second kind of symbolic information produced by the user. Notice that it maintains the low level of coupling between symbolic and physical level.

More recently, the concept of force feedback appears, which create a new kind of perceptual signal. It plays at a physical level, but it implements in its principles a strong need of coupling between the gestural inputs and the gestural outputs.

We would like to insist that it is not only a supplementary way of interaction, but it is the way from which a qualitative and quantitative paradigm shift occurs :

- ◆ Qualitatively, with gestures (as input and output signals, in other words, received and produced), the question of multimodality is raised not as a collection of modalities but as sensori-motor loops : gesture to gesture, gesture to vision, gesture to audition.
- ◆ Quantitatively, the reactivity of the interaction changes deeply : the time response has to be less than 1 ms for haptic reactivity (that is fast) , about 5 ms for the other loops and only more than 20 ms for classical audio-visual superposition, that is very slow.

This gives an idea of the reactivity of the computerized environments, whatever they are : communication networks wireless or not, Virtual reality systems, avatars, tangible media, augmented reality systems, etc... with which human beings have to interact.

So doing, the true nature of multimodality emerges, which could consider not each modality by itself, but in its strong correlation with others. The questions become more the reactivity of the multisensori-motor loops, designed as closed loops, than the co-operation of the media and of modalities at a phenomenological level. The terms of cross-modality or intersensoriality, reactivity, could replace the phenomenological link between each perceptual modality, opening new technological references :

- ◆ Sensors and actuators,
- ◆ micro and macro reactive multisensory - motor Computer Components and Architecture
- ◆ Regulation, control, servosystems, etc...

These observations point what I think to be the two major bottlenecks for the interface technologies, in the aim to define FP6 future developments :

- ◆ To Rise clearly, as a central question to be solved, and despite the current point of view, the question of multimodal communication in terms of reactivity of closed sensori-motor loops and not only in terms of a collection of perceptual channels.
- ◆ To Rise clearly, as a question to be solved, the question of the relationship between the all the multisensori-motor levels and the symbolic and cognitive levels.

First Bottleneck

Rising clearly, as a central question to be solved , the question of multimodal communication in terms of reactivity of closed sensori-motor loops and not only in terms of a collection of perceptual channels, causes a radical change in the technologies of interfaces (i-TV begin just a little to go in this direction) at several levels :

- a. at the level of the structure of the fonctionnal coupling between modalities of action and modalities of perception :
 - ◆ what kind of computer architecture (components and macrosystems) must to be developped ?
- b. at the level of the qualitative evaluation and needs in the different action-perception loops which play in the communication between human persons and artificial organisms (computerized worlds)
- c. at the level of the quantification and organisation of the distant communications :
 - ◆ what would be the quantitative capacities of that telecommunications networks to assure human multisensorial interactivity?
- d. at the conceptual level, we have to refer to the contemporary theories of action-perception-cognition (ecological and synergistic theories) and we have to assure the interdisciplinarity with the sciences of life (neurophysiology) and the sciences of human perception and cognition (psychology) to develop at a basic level a better undertanding of the human sensori-motor chains.

Second Bottleneck

Rising clearly as a question to be solved, the question of the relationship between the multisensory-motor level and the symbolic and cognitive level.

Today, these two levels are considered :

- ◆ either at the same level (for example the speech is considered as a modality at the same level of symbolic gesture)
- ◆ or completely disconnected.

This means that this question is NOT treated.

It appears in some applications which have to transform the signals level to symbolic level and vice-versa (animation, music, robotics, social organisations through sounds and images communications, emergent language).

We have to take care that this question does not be treated at a superficial level by buzz-words.

That is the case mainly today between the speech and the gesture for example, or between the speech and the image. The concordance between symbolic events extracted from speech and from digital video sequences are not trivial to be found and to be matched.

The general addressed question can be identified as :

The co-operation between multisensory closed – loops signals and symbolic events and languages through symbolic categorisation.

That is the condition to can perform efficient encoding between signals and cognitive patterns.

The emergence of new interfaces addresses 8 types of developments :

- Sensors and actuators
- Communication with computers and computer architecture
- Metaphors and models of interaction
- Data processing
- Data encoding (multisensory signals to and from symbolic events)
- Multisensory studies and human psychoperceptive evaluations
- Macro concepts (tasks, efficiency, skills, etc.)
- Societal implications and integrated applications

1.Sensors and actuators

Actuators may change fundamentally the structure of interfaces and of the underlying computer architecture and capabilities. They place at the centre of the technological challenge the question of the real-time and of the reactivity of the systems.

There are mainly four types of actuators

Mechanical actuators for the proprio-tactilo-kinesthetic perception (haptic modality)

Acoustical actuators

Optical actuators

Chemical actuators for the senses of taste and smell.

Mechanical actuators

According to the generalization of the electrical (rather than mechanical) commands in all the daily objects and instruments, we may assume that all these objects will be equipped with a large amount of sensors of every types (cars, home and professional instruments, etc.. Conversely, the maintaining of the human skill in action, require the restitution of basic physical feeling which sustain the human skills and performance. That is the main reason that associated actuators will be implemented in all these equipments. For example, in the car driving action, that is a fact that electrical command have often to be assisted by force feedback feeling. That is the same evolution in some professional issues such as technical manipulation with high risks (surgery), manual assistance, training and learning, etc.

Conversely, the sense of touch is one of the more sense used by human to perform instrumental tasks which cannot be artificially performed. It is also probably the sense which convey preferentially the appreciation of the reality.

The question raised by this type of actuator constitutes today the best technological difficulty probably at the same level of complexity that the emergence of raster visual displays in the 1970 : (1) they have to convey forces and torques, (2) they have to be associated with adapted sensors in a high demanding real-time closed loops (reactivity less than 1 ms), (3) high quality of the mechanical part (no friction), (4) they have to be flexible in their morphological aspects, (5) small size, (6) they have to be generic and modular to serve a large amount of applications (embedded actuators, desktop manipulators, etc). A chart giving all the technological requirements to be solved for this category of actuators : Number of degrees of freedom, types of variables, resolution in positions and forces, morphological appearance, etc...

Acoustical actuators

They are actuators running at a less macroscopic scale of energy than mechanical actuators for the haptic sense. This modality can be function without action : in human sense, it is an a vision exteroceptive sense. When it is necessary that it function with action the time response of the reactivity system is less than in the haptic modality (about 500 Hz). The development can address the possibility to have a lot of small sheared acoustical transducers.

Optical actuators

In this categories, there are all the developements in display (flat, wearable, etc...) developed in micro and opto-electronics

Chimical actuators for the senses of taste and smelt

These actuators adress the domain of micro-valves for the restitution of molecules and micro-fluidics for combining flows. They can be implemented on diary objects or placed in an ergonomic way on human beings.

We can assume that the number (by degrees of freedom) of actuators would be at the same order than the sensors.

To realize the multisensory-motricity in the communication between human beings and computerized environements, the space of these paires of sensors and actuators have to implement:

- ◆ Multidimensionality ($n \times 3D$ for images, $n \times 1D$ for sounds, $n \times xD$ for the gestures)
- ◆ Multiscale in space : acoustical, mechanical, optical, chimical, micromechanical)
- ◆ multi-reactivity in time : gestural action / haptic perception, gestural action/3EP, EP/EP)
- ◆ that are hard scientific and technological questions.

(EP : Exteroreceptive perceptions : vision, smelt, audition, etc.)

3.Communication with computers and computer architecture

After their « sensorialization », aiming the fact that all the instruments that human beings manipulate will be more or less « computerized », displacing the purpose of the interfaces from the monomodality and perception to the multimodality and action-perception loops, provokes deep changing in the design of the communication between human beings and computerized sensori-motor environements. Two requirements would support this shift:

- ◆ the reactivity of the action-perception loops
- ◆ the quality of the restitution for each modality and for their combination

The main stream has developped the quality of the restitution for each modality.

Implementing high degree of « reactivity », would have deep technological consequences on :

- ◆ The basic components for calculation
- ◆ The organisation of those components
- ◆ The Operating systems

We can assume that :

- ◆ we have work for the modification of OS by including generic the sensor-actuator real-time drivers at the low – level of operating systems
- ◆ we have to develop others types of computer architectures with other types of distributed components for example.

4. Metaphors and models of interaction

To illustrate this point, let us to give some examples :

- ◆ in the sense of presence, what are the optimal combination of our sensori-motor modalities and with which level of quality in terms of time reactivity and of perceptual rendering ?
- ◆ To achieve a complex and risked tasks such as tele-surgery, what are the optimal choices we have to perform between all our senses : is it the quality of the image rendering ? is it by adding Virtual Images by introducing optical actuators (pairs of video sensing – video actuator) for mixed reality ? is it by adding force feedback ?
- ◆ Is the interaction model the same with the « helicopter view metaphor » or « tink fish metaphor » to explore unknown worlds (virtual, distant, abstract), « potter or diamantary metaphor » for manipulating accurate objects, etc.
- ◆ How kind of metaphors and understanding conveys the association of sounds and images without any real or cognitive reference to the actions ?
- ◆ What about the lack of the spatiality and locality feelings in exploring abstract worlds such as in WWW, net, or konwledge databases ?

5. Processing of multisensory-motor data

Coding, Representating, Processing, Composing, Communication, Storing multisensory-motor data

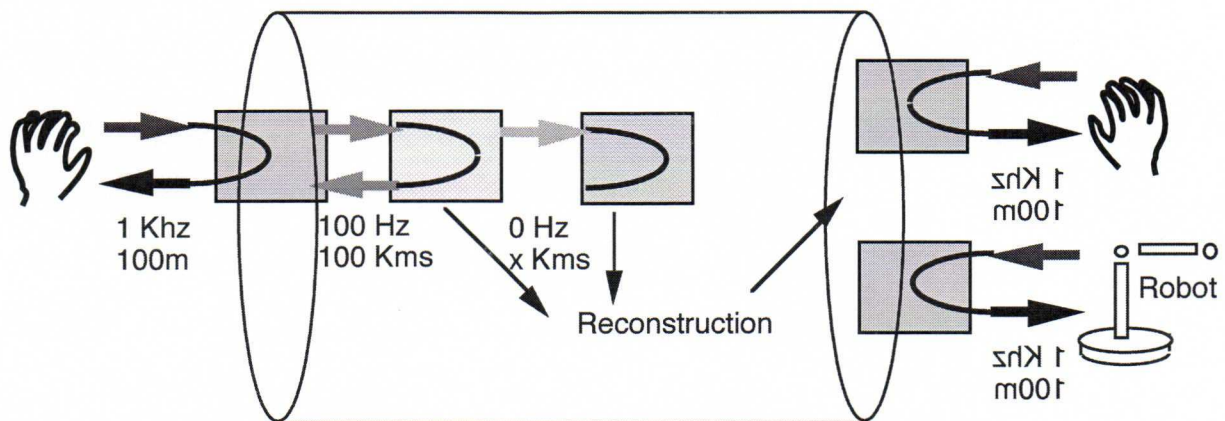
a. Multisensory-motor data and telecommunications

Standards

b. Fine Gesture and telecommunications :

Hardware : bandwith and types of communications means, physical structure of the networks

Software : multi-layered real-time to quite real-time servers (it adresses also the symbolic coding of multisensory-motor data)



c. Multisensory Atlas and databases

d. Interfaces to process, compose and create multisensory-motor data

6. Data encoding (multisensorial signals to and from symbolic events)

Generic processes for :

- Encoding multisensory-motor signals in symbolic events
- Decoding multisensory-motor signals in symbolic events to reconstitute the senses.

Applications :

- for symbolic manipulation to create new multimedia content (music, sound, animation)
- for efficient telecommunication
- to link with sense speech and language to actions and perceptions.

7. Multisensorial studies and human psychoperceptive evaluations

Multidisciplinary researches : neurophysiology, psychophysics, human behaviors, technological issues

Each modality :

- develop the knowledge in haptic modality (very low to day) including action and perception : for this use technological developments (VR, haptic devices, etc.)
- Quantitative (bandwidth, resolutions, amplitudes, etc.)
- Qualitative evaluations (, rôles of the forces, rôles of the forces human beings emit and receive, etc)

All the bi- modalities which include action

- Gestural action / haptic perception
- Gestural action / Visual perception
- Gestural action / acoustical perception
- All the bi- modalities which not include action

Visual perception / acoustical perception, id with smell and taste

All the tri (or more)- modalities which include action

- Gestural action / haptic perception/ Visual perception
- Gestural action / haptic perception/ acoustical perception

Multisensorial modalities studies

- Gestural action / haptic perception/ Visual perception / acoustical perception ...

Etc.

Grenoble, le 13 mai 2002

Annie LUCIANI