Multimodal User Interfaces and Agent Systems

Rami Saarinen
University of Tampere

In this presentation:

- Agent Systems
- Multimodality and multimodal user interfaces
- Current work & future options
- Conclusion

Agent Systems

Agent System is very ambiguous term. Loose definition: agent system is a distributed system that is composed of autonomous applications called agents. These agents have a method to find other agents and to communicate with each other.

Agent Systems

The development and research of Agent Systems was hot from mid 90s' till year 2000. After that there has been a decline, but that may very well be caused by the discovery of other “new” methodologies such as SOA & ubiquitous computing.

Agents

Also the term Agent is very ambiguous with many different conflicting definitions.

Shoham [1997]:

- By definition: a representative who acts on behalf of other entities.
- An entity whose state is viewed as consisting of mental component such as beliefs, capabilities, choices and commitments.
- Common ground: an autonomous entity that functions continuously in a environment in which other processes take place and other agents exist.

Agents

Bradshaw [1997] lists common properties for agents:

- Autonomy
- Reactivity
- Persistence
- Personality
- Communication
- Cooperation
- Adaptability
- Mobility
- inferential capability

Agents

Agent in general:

- Can communicate with other agents
- Has ways to observe and manipulate its immediate environment.
- Can be purely reactive or proactive (predicting / goal oriented), or more often somewhere in between.

Common aspects of agent systems

- Agents are usually contained in some shared isolated environment that provides the means for agent management and communication.
- The system usually contains a central server that provides the “yellow-pages” service for the agents.
- Communication is done in well-defined communication protocol.

Standardization

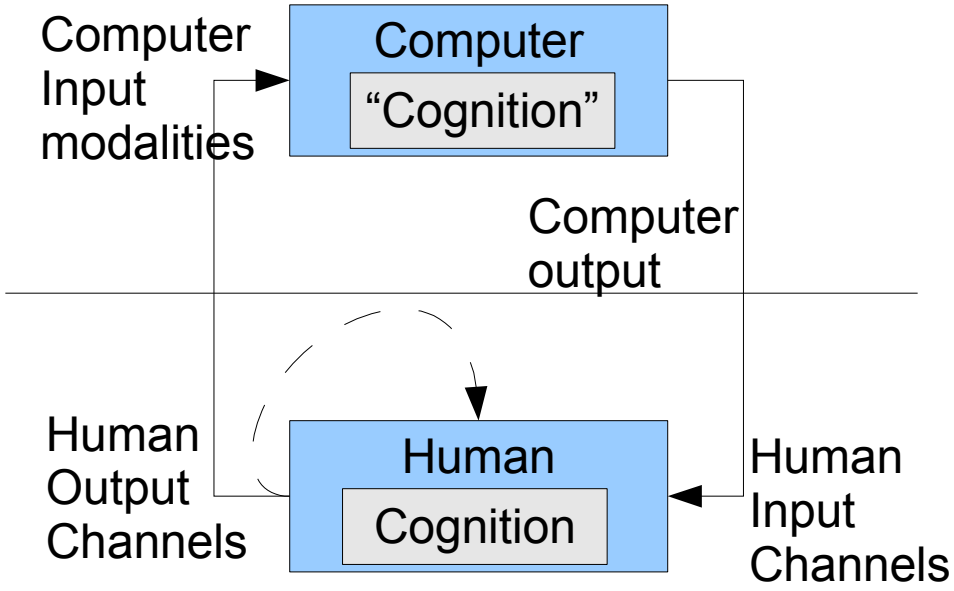
- CORBA (OMG)
  - Common Object Request Broker Architecture
  - Client – Server architecture
  - Communication done in IIOP (Internet Inter-ORB Protocol)
- KQML & KIF
  - Communication language: KIF, Knowledge Information Format.
  - Communication protocol: KQML, Knowledge Query and Manipulation Language.
- FIPA
  - Foundation of Physical Intelligent Agents
  - Abstract “model” architecture.
  - Communication protocols
  - ACL: Agent communication language

Well known agent systems

- JADE (Java Agent DEvelopment Framework)
  - Java
  - FIPA compliant
- IBM ABLE (Agent Building and Learning Environment)
  - JavaBean based
  - FIPA compliant
- OAA (Open Agent Architecture)
  - C++ & others
  - no conformance

Multimodality

- The taxonomy of multimodal interaction: communication between the human and the computer using several information sharing channels through a common interface [Shomaker *et. al.* 1995].



- The use of multiple (possibly the same) modalities for data input and output.

Multimodal Fusion

- The integration of commands received via different input channels [Nigay & Coutaz, 1993]:
  - Using modalities sequentially: fusion combines the inputs
  - Using modalities sequentially without fusion
  - Using modalities concurrently with fusion combining the inputs
  - Using modalities concurrently without fusion
- Use of modalities to support command recognition

Multimodal Fission

Produce multimodal output:

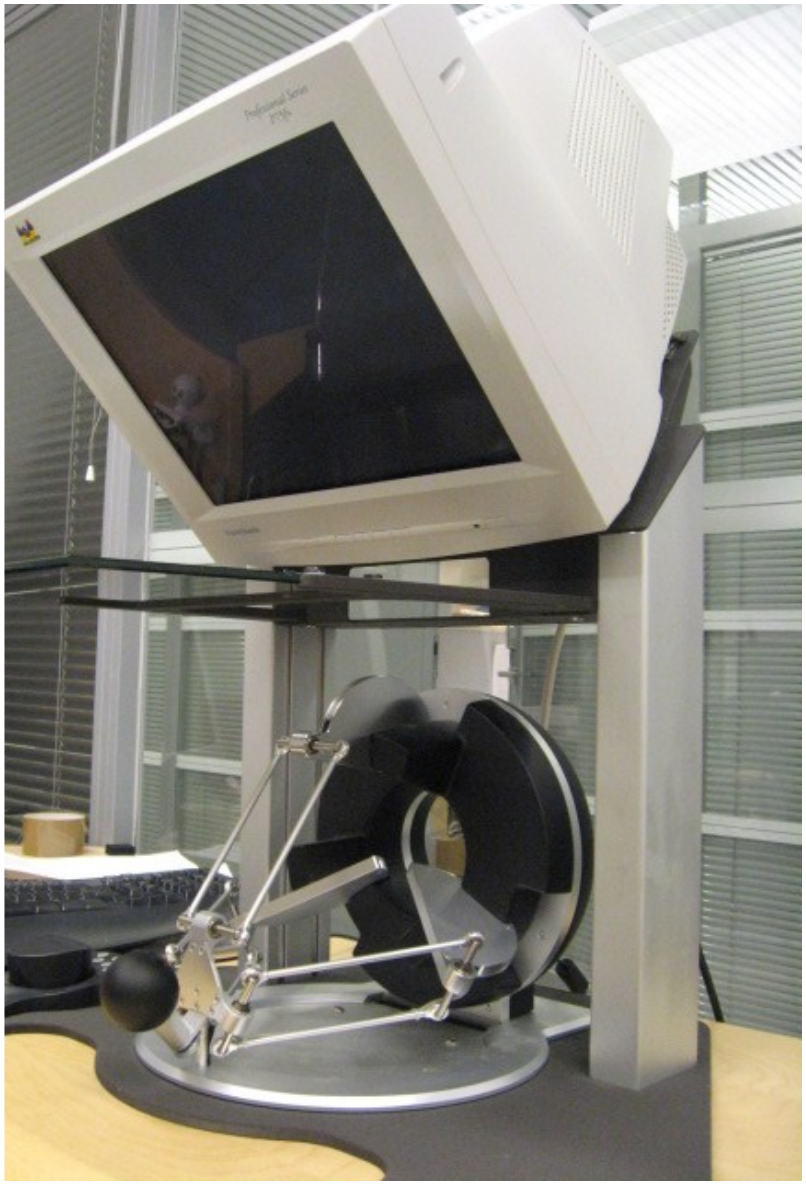
- Modality re-mapping
- Replace modality with another
- Support the recognition of output with multimodal feedback
- Intelligent and context-aware use of modalities

Haptics

- Feedback through the sense of touch
  - Force feedback
  - Vibration
  - Tactile feedback
- Requires a special device
- Sometimes coupled with 3D virtual environment
- It's all about forces:
  - motion based (spring, damper, friction, inertia)
  - time based (constant, periodic, impulses).

One possible setup

- 3D virtual environment is projected on a semi-transparent mirror.
- The user uses 3D glasses and looks at the virtual environment from the mirror.
- The haptic device is located just under the mirror providing a co-centric use of the device as well as integration of the real physical world (the device, hands) to the virtual world (with the graphical representation of the device).



Direct manipulation

Basic properties of direct manipulation [Shneiderman, 1997]:

- Continuous representation of the objects and actions of interest.
- Physical actions or presses of labeled buttons instead of complex syntax.
- Rapid incremental reversible operations whose effect on the object of interest is immediately visible.

Multimodal Agent Systems

- Dire need to distribute the computational intensive tasks to different computers
- Haptics itself require a lot of computing power to display both the virtual environment and haptic feedback
- Redundancy and improved command recognition
  - Several competing recognizers
  - Command prediction
  - Adaptive / learning recognizers

But that was “hot” in 60's!

- Richard A. Bolt *et. al.* Presented the legendary “Put-that-there” multimodal application in 1980.
  - Douglas Engelbart, 1963!
- For example OAA had several multimodal applications built on top of it, such as: Automated Office, Multimodal Maps, CommandTalk etc. etc.

So what's new?

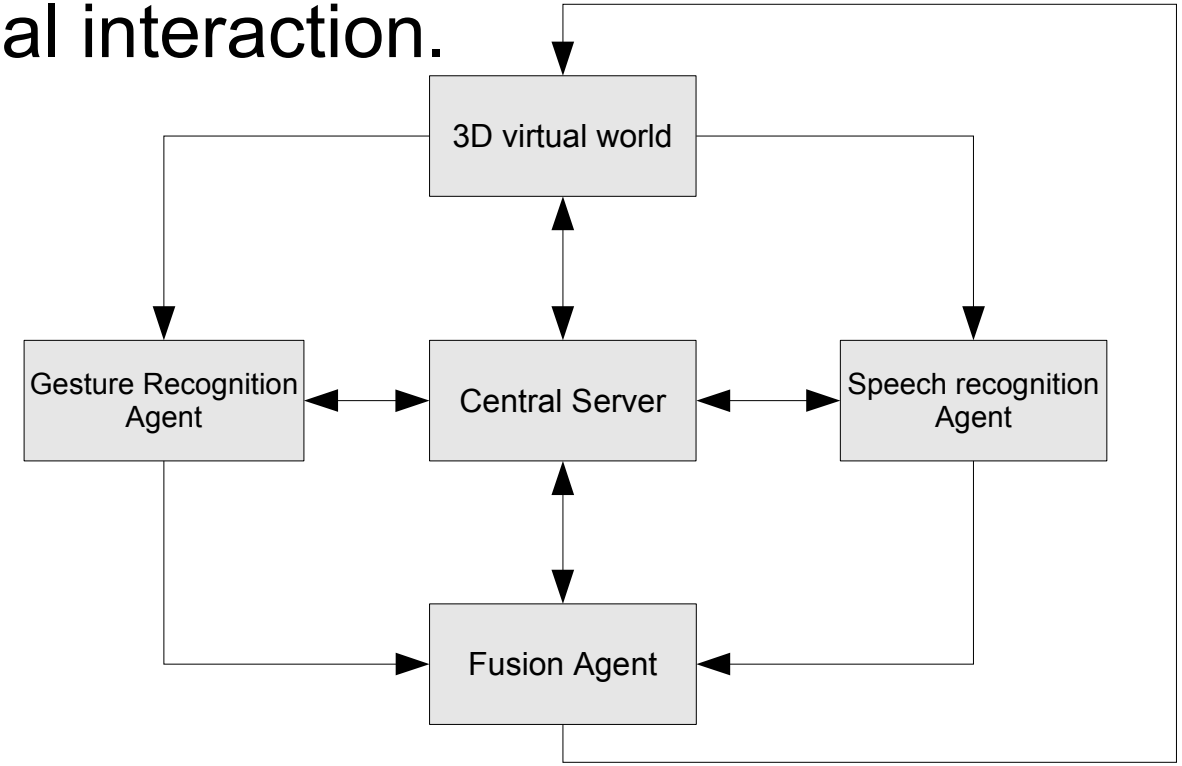
- A lot of prior work focus on applications not the underlying properties of a functional distributed agent system.
- Emerging technologies:
  - Haptics
  - “brain hat”
  - biofeedback
  - Ubicomp

Current status

- Currently working on two different setups that both use the self-made agent architecture.
- The base architecture is almost ready except there is still work to do with the modality fusion agent.
- The agent communication is done with a text based messages (currently plain text, XML in near future).
- Used technologies: C++, Qt, H3D and Python.

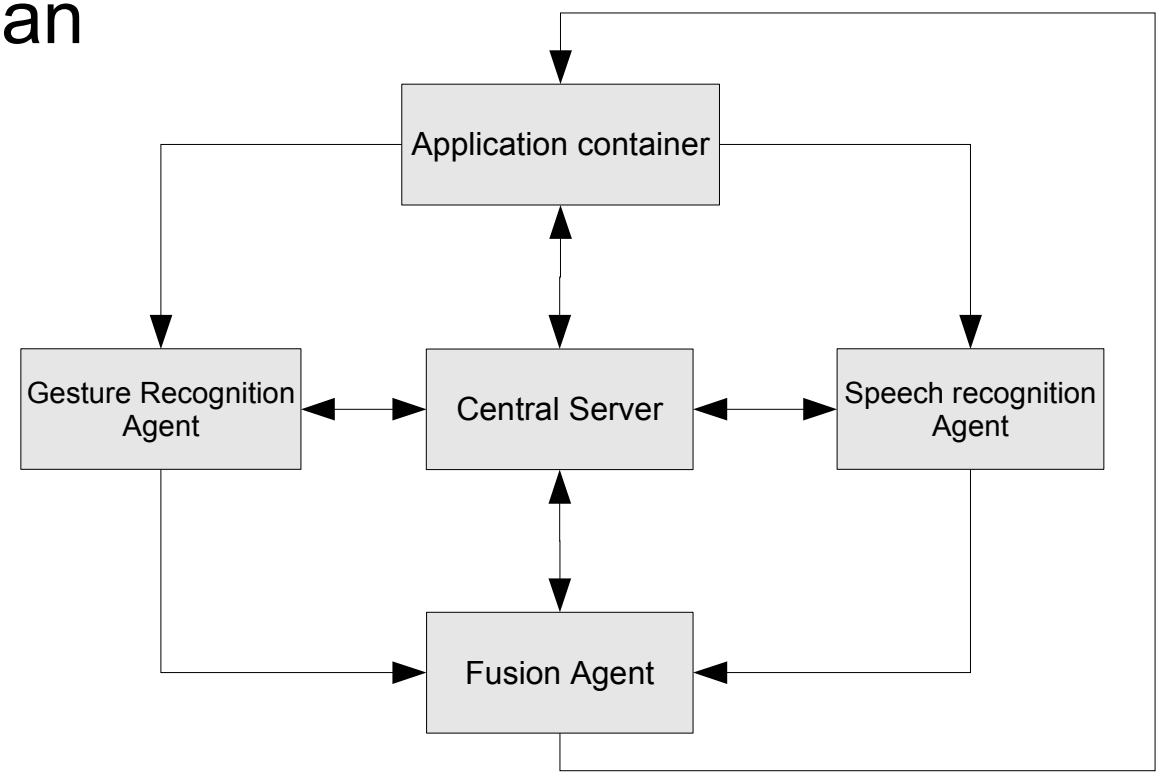
Research Question One

To build a 3D haptic application that uses gestures (the movement of haptic device) and speech for command input. The current focus is the multimodal interaction.



Research Question Two

To work out the bootstrapping problem by making the agent network configuration available via HTTP protocol from the central server. Already built-in is also the ability to distribute the actual client application with the configuration data (initial client is an "empty shell").



Current Issues

- Is there a possibility to support direct manipulation?
  - That is, what are the limitations for a constant feedback to the user in the distributed environment?
  - Network Lag?
  - Multiple users?
  - Would a multi-level fusion or somehow predictive command recognition work?
- The user is somewhat unimportant in the description of several agent systems. Multimodal interfaces are centered around the human agent. That is, how to design and build a **user interface** that is realized as a agent system.

Immediate plans

- Paper focusing on multimodal fusion, how it's been done in related work and what I have done in my research. SoSE paper writing workshop.
- Possibly another paper later this year with focus on more technical aspects of multimodal systems.

Future possibilities

- Redundancy / competing recognizers.
- predictive / adaptive input recognizers and fusion.

More?!

- Fission
- Multiple users
- Mobile devices

Publications

Rami Saarinen, Janne Järvi, Roope Raisamo, Jouni Salo, Pentti Hietala, Arto Hippula, Kari Peltola, Eva Tuominen, and Marjatta Kangassalo, Proactive Agent Environment for Children's Exploratory Learning. NordiCHI 2004 workshop "HCI Issues in Proactive Computing". Tampere, Finland, October 24, 2004. http://www.hiit.fi/uerg/nordichi04-proactive/, 8 pages.

Marjatta Kangassalo, Roope Raisamo, Pentti Hietala, Janne Järvi, Kari Peltola, Rami Saarinen, Eva Tuominen, Arto Hippula. Proactive Agents That Support Children's Exploratory Learning. Julkaisussa: Yasushi Kiyoki, Bengt Wangler, Hannu Jaakkola, Hannu Kangassalo (Eds.), Information Modelling and Knowledge Bases XVI, IOS Press, 2005, 123-133.

Rami Saarinen, Janne Järvi, Roope Raisamo, & Jouni Salo. Agent-based architecture for implementing multimodal learning environments for visually impaired children. In Proceedings of the 7th International Conference on Multimodal Interfaces (ICMI'05), pages 309316. ACM Press, 2005.

Rami Saarinen, Janne Järvi, Roope Raisamo, Eva Tuominen, Marjatta Kangassalo, Kari Peltola, & Jouni Salo. Supporting visually impaired children with software agents in a multimodal learning environment. Virtual Reality, 9(2-3):SpringerVerlag, London Ltd. 108117, 2006.

Pietrzak, T., Martin, B., Pecci, I., Saarinen, R., Raisamo, R., and Järvi, J. 2007. The micole architecture: multimodal support for inclusion of visually impaired children. In Proceedings of the 9th international Conference on Multimodal interfaces (Nagoya, Aichi, Japan, November 12 - 15, 2007). ICMI '07. ACM, New York, NY, 193-200.

Conclusion

What

- Agent Systems
- Multimodal interaction
- Haptics

Why

- Need to distribute the processor heavy operations to different computers
- Improve input recognition and fusion
- Better support for multiple users
- The trend of ubiquitous / smart home technology

References

[Bradshaw, 1997] Jeffrey M. Bradshaw. An introduction to software agents. AAAI Press / MIT Press, 1997.

[Nigay & Coutaz, 1993] Laurence Nigay & Jóelle Coutaz. A design space for multimodal systems: Concurrent processing and data fusion. In In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 172 - 178. ACM Press, 1993.

[Shoham, 1997] Yoav Shoham. An Overview of Agent-Oriented Programming. AAAI Press / MIT Press, 1997.

Thank You!

Feel free to contact me:
Rami Saarinen
rs@cs.uta.fi