COLLABORATION IN VIRTUAL REALITY OF INTELLIGENT AGENTS

Ádám VÁMOS*, István FÜLÖP*, Barna RESKÓ*,**, Péter BARANYI*
*Cognitive Informatics Research Group, Computer and Automation Research Institute,
Hungarian Academy of Sciences, Budapest, Hungary
**Óbudai University, Alba Regia University Center, Budapest, Hungary

ABSTRACT

This paper presents a system for high level collaboration between humans and intelligent agents in a virtual reality. The proposed system is capable of creating virtual worlds with the simulation of Newtonian physics and their visualization as interface for humans. Furthermore, with voice recognition and synthesization the human interface provides a user friendly way for conversations with intelligent agents. The proposed system is able to connect the different communication modalities allowing the human user to control the intelligens agents using the brain-in-the-loop control method. The system has a standard modular build-up in order to connect different virtual worlds to each other. Finally the capabilities of the system are demonstrated on two examples.

Keywords: *Virtual Collaboration, 3D Internet, robot telemanipulation*

1. INTRODUCTION

The continuous evolution of computers creating virtual realities like in movies are not science fiction any more.

In the everyday life there is a strong demand on virtual reality applications. For example in the popular game called The Sims users may design their own houses, meet new virtual friends, go to work or just watch TV - in their virtual world.

With the spread of Internet into almost every household on-line communities became very common. Many people use instant messengers instead of meeting a person in real life. On community sites like Facebook users can share information about their lives with each other.

The game called Second Life went even further with integrating an on-line community with a virtual reality engine. In Second Life residents may register their own avatars with the ability to change their appearance like clothes or hair style as they wish. It is also possible to create new objects which can be used or sold later. Users can control their avatars wandering around in the virtual world interacting with objects or they can communicate with other people through their avatars. Second Life even has its own economic system and currency, which avatars can use in the virtual world.

High level communication between two humans is called cognitive communication. This communication process does not restrict on verbal communication channels, but it also includes non-verbal and para-verbal channels, for example tone, facial gestures, etc. These communication channels use all the human input modalities, which allow very dense information flow. However, low level communication between two computers is called infocommunication. When high level input modalities and communication channels are used for communicating with computers, we are talking about cognitive infocommunication.

Communication with virtual realities may also require special devices to raise the level of this communication to the level of cognitive infocommunication. Vision and speech are the two most natural modalities of communication between humans, which should be used in the case of virtual realities as well.

Until recently the widely available computer displays supported only a 2D projection of the 3D virtual world. However 3D vision is also one of the most natural and obvious perceptual modalities of species with stereoscopic visual abilities. This implies that a 3D display can play the role of a cognitive infocommunication device that links the user to the virtual world. This is why the consumer market also sees a big opportunity in the virtual reality and in 3D in general. 2010 was named to be "the year of 3D"[1]. Manufacturers are introducing better and better equipments displaying 3D content on 3D projectors or 3D LCD panels. 3D input devices like stereo cameras or camcorders also appeared on the market.

As all computers, robots can perform tasks of low or medium level of intelligence fast and precisely but they don't have the ability to make higher level decisions, such as deciding strategies or reacting on unforeseen problems. On the other hand humans are able to do such complex tasks, but they do it relatively slow and loosely compared to robots. With the so called *brain-in-the-loop* control method, humans give commands to robots that they can autonomously execute, using their own artificial intelligence [1] [2]. Humans oversee the working process while robots execute the commands fast.

A future application shall integrate the power of unlimited possibilities of virtual reality, the ergonomy of cognitive infocommunication including 3D visualization and the high level intelligence achievable by the brain-in-the-loop control.

Our goal in this paper is to introduce a system that tries to integrate the above three concepts, and to show its applicability in robot telemanipulation and control.

The rest of this paper is organized as follows: in Section 2 the structure of the proposed system is introduced including the communication channels (2.1 and 2.2), the simulation of the virtual world (2.3), and the component-based architecture (2.4). In Section 3 the implementation issues

[1]NVIDIA Corporation on 2010 INTERNATIONAL CONSUMER ELECTRONICS SHOW, Las Vegas, Jan. 6, 2010

ISSN 1335-8243 © 2010 FEI TUKE

of the base component is discussed, while in Section 4 two test cases demonstrate the operation of the proposed system (4.1 and 4.2). Finally, Section 5 concludes the paper.

2. STRUCTURE OF THE PROPOSED SYSTEM

The proposed system is built up of several components, which can be connected to each other according to the users' needs. The main component of the system is the 3D Visualization with the Virtual Reality itself. The Virtual Reality as a collaboration arena consists of a physical space which describes the physical attributes - e.g. location, speed, look, etc. - of the entities and a communicative space which connects the intelligent entities. Humans and other intelligent agents can join the virtual reality with their avatars as visual representations.

The 3D Visualization is the bridge between the user and the Virtual Reality. It allows the user to interact with the Virtual Reality and its participants that can be other humans, robots or other intelligent agents. The user has its 3D Visualization which is connected to the Virtual Reality and controlled by the user, but more virtual realities can be connected together in order to connect participians of different virtual realities into a virtual reality system. It extends the capabilities of a separate virtual reality and provides a standard for *3D Internet* as a network of virtual realities.

Non-human participants of the virtual world are called *cyber devices*, which can be physically existing or purely software-based components. When a cyber device is connected to the virtual world its visual representation (avatar) appears in the 3D Visualization. The cyber devices connected to the virtual world can then operate according to their own internal algorithms. They can communicate with other cyber devices or humans, or manipulate existing objects in the Virtual Reality.

In common virtual reality systems the virtual world is independent from the real one. In our system cyber devices and humans are the buckles connecting the real world to the virtual one, so they can have effect on each other. Cyber devices can gather information from the virtual world and can act upon them in the real world and can execute commands, so the real world is influenced, as well as they can update the physical attributes of their representation and humans can directly manipulate existing objects, so the virtual world is influenced. If a physically existing cyber device can change its physical attributes - e.g. location, orientation - they can be updated either by the cyber device itself - internal update - or by another entity - external update - which is connected to the virtual reality.

Modalities of humans and other intelligent agents are very different, that's why the system provides input and output interfaces for both. For example a computer program can easily understand that object1 was collided with object2 at location x,y,z, which is hardly understandable for humans. On the other hand the 3D graphical output accompanied with a short noise is an efficient cognitive communication channel to inform the human users about the collision. Such 3D output is even more spectacular and efficient if the display of the user supports 3D stereoscopic visualization, linking the user into the virtual reality even more.

In order to accommodate the information needs of the cyber devices about the virtual world, the system has an output where high level information is provided about the virtual reality. This output can be used by the cyber devices to react on events and situations in the virtual reality.

Emitting sounds or synthesizing voice is also supported by the system. This feature makes the cyber devices able to speak to the users instead of just some console outputs.

2.1. User input channels

Users can connect different types of input devices to the system. The main manipulation device is the "magic wand", a 3D pointer inside the Virtual Reality. The pointer acts as the virtual forearm of the user. It can be used to select and move objects or interact with the system or with cyber devices.

The 3D Visualization component supports several interfaces through which the Virtual Reality can be manipulated. The interfaces make it possible to connect various input devices to the system. Joysticks, data suits, glyph tracking systems [3] [4] or accelerometer-based devices like Nintendo's Wii or Apple's iPhone helps the user to move around in the 3D world and use the magic wand.

The camera is the window through which the user can see the virtual world. Manipulating the camera allows the user to adjust the position and looking direction of this window. This can be useful to move around in the virtual room, and also to supply this input from a head tracker device. Head position and looking direction information can be used by the 3D Visualization component to compensate for an asymmetrical viewing frustum. The camera and the magic wand can be controlled with the mouse and the keyboard by default.

If a stereoscopic visualization is available, it allows the user to be fully involved in the virtual reality. The user sees the virtual world like real and can control its avatar using its own body by motions and gestures. If the virtual world is shared between several users, it is possible for example to play a football match with motion capture suit and stereoscopic screens on the walls of the room.

It is also possible to use the built-in speech recognizer and send words or sentences understandable by the virtual reality component, which in turn performs accordingly.

2.2. Communication with cyber devices

Cyber devices are also informed about the speech messages therefore they are able to react accordingly. The Virtual Reality component sends high level messages to the cyber devices about various events occurred in the virtual reality. Such events may be collisions, speech messages from users, new entities, etc.

Upon connecting to the Virtual Reality component, cyber devices can register their visual representation and also they can register a menu containing commands specific to the given cyber device. For example a robot dog may have a command "Fetch!", but this command does not exist in a KUKA industrial robot. Because menus are completely available using voice commands only, and cyber devices can synthesize speech through the system, even complicated human-robot collaboration tasks can be achieved purely with speech.

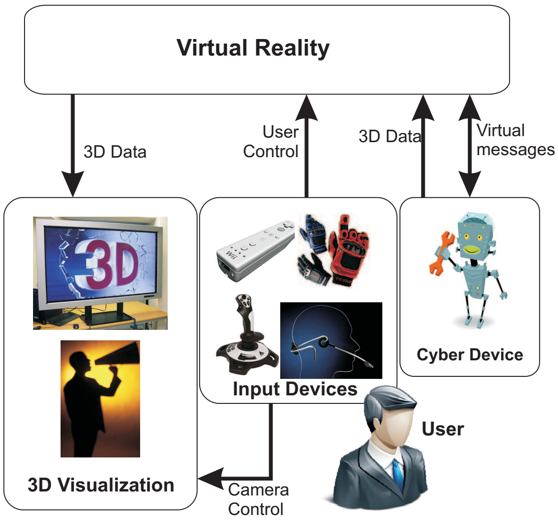


Fig. 1 The communication interfaces and connectable devices of the 3D Visualization component

2.3. Simulating the physical world

The Virtual Reality also includes a physical engine responsible for the physical simulation of the objects in the virtual world. There are several types of objects distinguished by the physical simulation. Environmental objects are the most basic objects as their physical attributes are not changing. They represent the parts of the static working environment like walls or tables. Cyber devices, as external entities update their avatars by sending low level position and orientation data to the Virtual Reality component. Virtual objects exist only in the virtual world and Newtonian laws of physics are applied on them. They can collide with other virtual or environmental objects or with avatars. Menu objects and the magic wand are special objects as they don't have a physical body only a graphical representation.

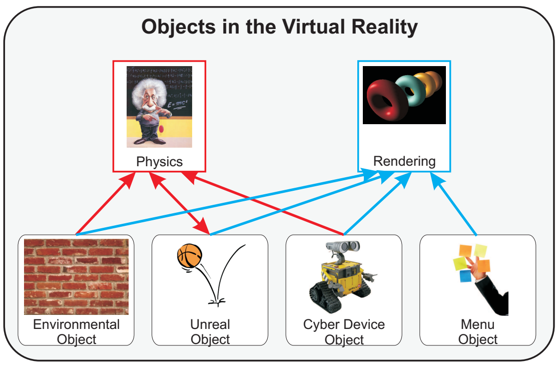


Fig. 2 Different types of the physical objects

2.4. Modular configuration

As mentioned above, the system is built up using components. Components register themselves to a system-wide known name server and publish input and output interfaces through which they can communicate. Above the existing interface types unique types can also be defined. Components with the same type of input and output interface pairs can be connected to each other adding more functionality to the existing ones. Cyber devices and user input devices connect directly to the Virtual Reality component as shown in Fig. 1. For example, a glyph tracking device, which produces 3D coordinates, can be used as position information for a cyber device or as user input, since both inputs require the same type of data. Connections can be added to or removed from the system during runtime without interfering the communication on other interfaces. This feature makes the configuration of the participants of the virtual reality fast and convenient.

The proposed system can be expanded with new modules. The interfaces used for communication and management are free open-source softwares and that's why there is no restriction in the details of implementation of the modules, they can be written in any popular programming language like C++, Java, C#, etc. Connected components don't need to be running on the same computer as they can communicate over the Internet. The well-defined interfaces make it possible to completely hide secret or sensitive details of the implementation of a given component and make only the communication interfaces public making even reverse engineering impossible.

Not only new modules can be added to the system, but different systems can be connected in order to join the virtual worlds as shown in Fig. 3. If two virtual worlds are joined, they can be accessed from the other forming a system of virtual realities which can provide the basis of *3D Internet* of the future as a world-wide virtual reality system. This feature multiplies the capabilities of the system providing infinite possibilities of collaboration.

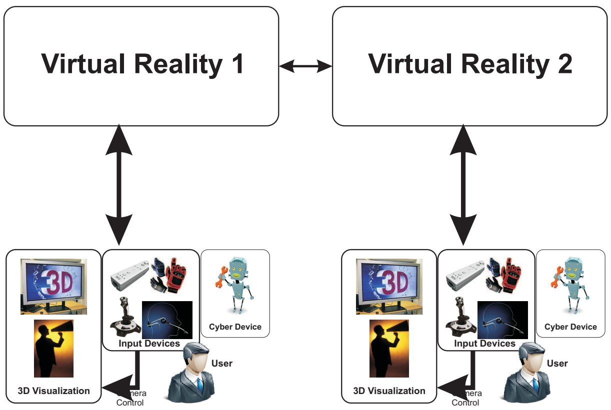


Fig. 3 Connecting virtual realities into a virtual reality system

Components can be connected and activated with a simple GUI manager as can be seen on Fig. 4. The components are displayed as rectangles. Blue color means an activated, white an inactive component, while erroneous ones are marked with red. Interfaces are filled spots or semicircles depending on if they are for input or output. Components which are registered in the name server can be added to the virtual reality using only a few clicks.

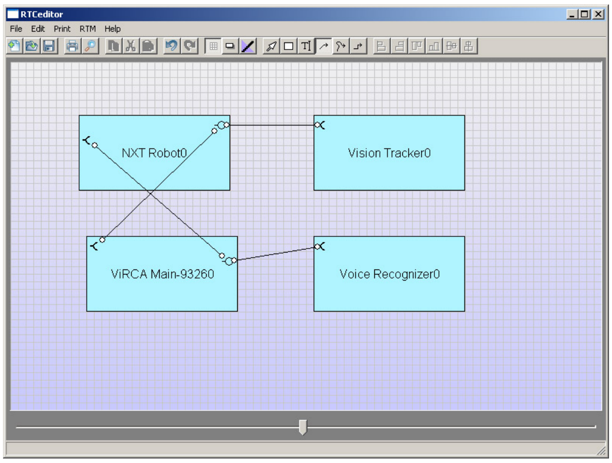


Fig. 4 An example of the GUI of the Component Manager application

3. IMPLEMENTATION ISSUES

The management of the components is implemented using the *Robot Technology Middleware* or *RTM* framework. This open-source free framework uses *CORBA* technology for communication. RTM is used in Japan by various projects and companies efficiently. If a robot component is RTM-compatible it is guaranteed that it will be compatible with any other RTM-compatible system and can be used without any glitch.

CORBA, the *Common Object Request Broker Architecture* is the industrial de-facto standard in distributed applications for its robustness. However development of distributed applications, such as components in the proposed system, can be difficult to a developer without preliminary knowledge in CORBA-based systems. Upon the overall system design easy expansibility with externally developed cyber devices was an important aspect. That's why the *Internet Communication Engine* or *ICE* framework is used for communication between the modules. ICE was named to be designed with the experience of CORBA but easy usability was also an important design aspect [5]. However development with ICE is quite easy, design was tended to be made even more easier with hiding the implementation of the communication interfaces into pre-written classes as much as possible. In practice a developer who is not familiar with the concept of distributed programming can implement the communication interfaces needed for a cyber device to communicate with the Virtual Reality in about 2-3 hours. This time can later be brought down to flexible about half an hour.

The Virtual Reality component has a modular architecture, while the internal communication is event-based. Modules are separated by the kind of task they do. Internal events are managed by the Event Engine. This module acts as a mailbox to other components with filtering and postponing events to them. The most important modules are the Rendering and Physics Engine as these two modules are responsible for bringing the Virtual Reality to life. For rendering, the OGRE Graphic Engine [6] was chosen because it's open-source and free and still a very advanced system. For the physical simulation the Bullet Physics Library is used. The Menu Controller - as the name suggests - pops up menus in the virtual world and upon selection, executes the assigned commands. The rest of the modules implement the various interfaces of the Virtual Reality component. The Repository Controller completed with the Object Synchronizer receives data from cyber devices like registering new avatar or updating position information. Human input actions can be received on three interfaces. These interfaces are responsible for the following:

- Camera movements and rotation
- Magic wand movement, rotation and actions
- Text-based commands

As shown in the above list, human speech recognition is not strictly part of the 3D Visualization component. However, by connecting an existing speech recognition and the 3D Visualization component to each other in the Component Manager shown in Fig. 4, the resultant system will be capable of reacting on verbal commands. The speech synthesization module listens to text-based messages received from other modules of the Virtual Reality component or from cyber devices, and synthesize speech accordingly. Keyboard and mouse can also be used in the virtual reality as manipulators of the camera and the magic wand. Inputs received from these basic devices are always active and they can be used as fall back or auxiliary inputs. In Fig. 5 an overview can be seen about the internal structure of the Virtual Reality component.

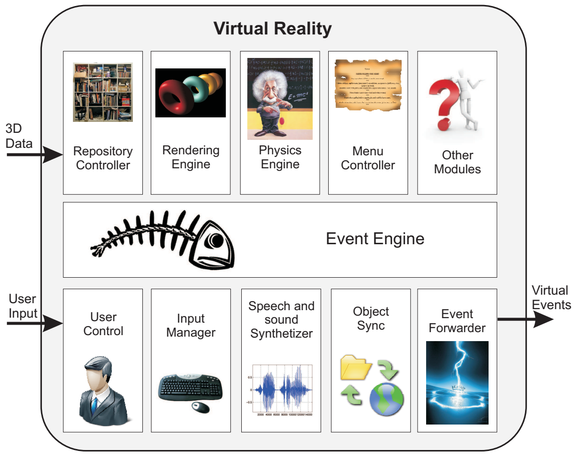


Fig. 5 Internal structure of the 3D Visualization component

4. TESTS AND DEMONSTRATION

In this section the operation of the Virtual Collaboration Arena will be demonstrated by showing how users can collaborate with two example robots using high level communication channels only.

4.1. LEGO Mindstorms NXT robot

LEGO Mindstorms NXT robots incorporate the diversity of the various LEGO bricks with the easy programmability in a C-like language. [7] The hull of the robot can be built around the central processing unit of the NXT.

The batteries, USB and Bluetooth communication interfaces and ports to the sensor and servo motors can be found in this unit. NXT robots can be extended with various types of sensor units:

- Touch sensor
- Light sensor
- Object and movement sensor
- Gyroscope
- Temperature sensor
- Compass sensor
- 3-axis accelerometer
- Color sensor
- IR communicator

As it can be seen in the above list, inexpensive robots can easily be assembled for various types of tasks. In Fig. 6 two example configurations can be seen. On the top of both robots a red and a cyan marker is mounted for vision based position tracking.

Fig. 6 Example NXT robot configurations

Testing the balancing robot in the virtual reality

The balancing robot is a good example to demonstrate how the Virtual Collaboration Arena could help in cooperating with robots. Let's take the task of moving around on the floor and pushing some objects to a given place. The balancing robot can be moved by controlling the power of its two servo motors. Without using any controller it is an impossible task to manually balance the robot. Furthermore, even a simple task as moving straight is hard because the power of the two servo motors always differ from each other a bit. In turn these tasks can easily be done by using the NXT sensors and an on-board program. In practice, the balancing of the robot is done by using an accelerometer while the straight line moves and turns are controlled by a gyroscope.

In order to work with the balancing robot in the Virtual Collaboration Arena, a cyber device component is needed to be implemented. The task of this component is to exchange data between the real robot and the virtual world by updating the avatar of the robot in the virtual world and postponing commands to the real robot. Because the robot can't determine its position in the real world, coordinates are gained from a visual tracker system, which is also a component in the Virtual Collaboration Arena. For easier operations during the tests, the voice recognizer component is connected to the Virtual Reality component. The connection of the components can be seen on Fig. 4. Upon starting the NXT component, a name has to be chosen for the avatar of the robot. In the demo *"Peter"* will be used as name.

Peter can be selected in the virtual world by either clicking on it with the magic wand, or calling the name "Peter". The 3D Visualization component then informs the user by the speech synthesizer module that the robot was selected (Fig. 7a). Peter can be navigated with the magic wand or with speech as well, but the speech reaction of the component can be achieved only by speech control. Using the speech modules in both direction, the user can actually take a conversation with Peter, and manage the work that way. After selecting him, the "Show functions" voice command browses the available functions of Peter (Fig. 7b). Without speech recognition the menu of Peter can be recalled by right clicking on it. The appearance of the function list is different in the two approaches, but the content is the same.

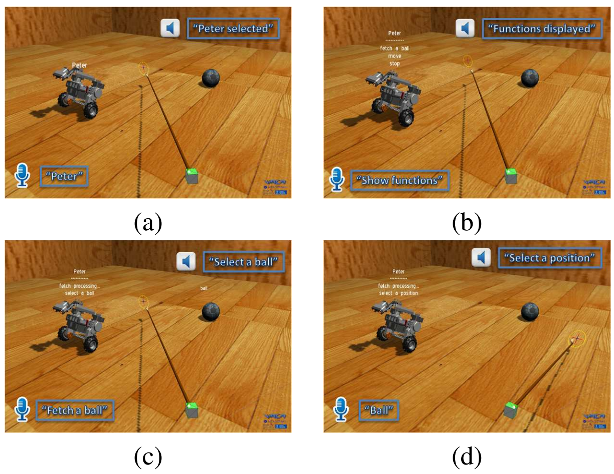


Fig. 7 Giving orders to the balancing robot using voice recognition and the magic wand

For using the displayed functions, the speech commands can be used henceforward. By speaking out the name of the function, the 3D Visualization component instructs about the required further steps of the function. In this example the "Fetch a ball" function was called. Hereupon Peter informs the user that he needs a ball that is to be pushed (Fig. 7c). After the desired ball has been chosen by pronouncing its name or by clicking on it, Peter asks for a location where to push it (Fig. 7d). The desired position can also be selected by the speech command "Click", which acts as a click with the magic wand. At this step however, the magic wand is indispensable because it has to be maneuvered over the desired position (Fig. 7d). After every parameter is given, Peter goes to the ball, and pushes it to the selected position on its own. In the meantime, the actions of Peter can be tracked in both the 3D Visualization component and, if it's possible, in the reality (Fig. 8).

(a)

(b)

(c)

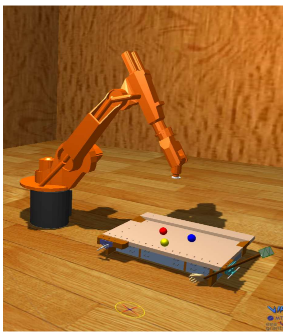
Fig. 8 After the necessary details are given the robot is processing the command on its own

4.2. KUKA industrial robot

The KUKA Robot Group supplies industrial robot solutions for countless working environments including the ones for as low payload as 6 kg up to the extreme heavy duty as even a ton. KUKA robots can be equipped with special manipulators designed exactly for the tasks of the factory production line. Among other things, KUKA robots can be used for:

- Assembling small or large parts
- Sewing leather seat covers
- Grinding and polishing
- Loading and unloading
- Spot welding
- Painting

(a)



(b)

Fig. 9 KUKA industrial robot in the real world (a), and in the Virtual Collaboration Arena (b)

Manipulators are usually placed on high accuracy, 6-degree-of-freedom (6DOF) arms, which supply with accurate information about the positions of the elements. An example can be seen in Fig. 9a[2]. The robot can turn around its base, expand or contract using the two joints of the middle section, rotate and bend the head and finally rotate the manipulator itself. The avatar in the Virtual Reality component can be seen in Fig. 9b.

Testing the KUKA robot in the virtual reality

In this test application a light KUKA loader/unloader robot will be taught to place balls from a given position to an other, using high level modality communication channels. In the demo *"KUKA"* will be used as name for the robot.

As an industrial robot KUKA has few capabilities but it can exploit them fast and precisely. KUKA is programmed to place a small plastic ball from one position to an other. The real robot is capable of moving the head over a given position. When grabbing the ball, KUKA is able to compensate the small deviations of the position of the ball. Still the selection of the ball and the desired location is a hard task for a human but easy with the Virtual Collaboration Arena. Similar to the method introduced in section 4.1, KUKA can also be easily connected to the Virtual Reality component. After this, the Virtual Reality component translates the high level command from the human user to 3D coordinates for KUKA, similar to the method introduced in section 4.1.

5. CONCLUSION

A method for human-robot and robot-robot collaboration has been proposed in this paper. The proposed system implements numerous interfaces needed for high level modality communication channels including:

- Stereoscopic 3D visualization of the virtual world
- Speech recognition and synthesization
- 3D pointer and manipulation devices such as data suit or Wii controller
- Glyph tracking

On the other hand the system also implements interfaces which are easily usable by computers.

By giving orders to robots that they can autonomously execute, human users can expand the possibilities of the robots according to the brain-in-the-loop control method.

The high level communication channels used in the communication between the human users and robots make the humans more comfortable in the virtual reality and in this way they can concentrate more on their real tasks.

ACKNOWLEDGEMENT

The research was supported by HUNOROB project (HU0045, 0045/NA/2006-2/ÖP-9), a grant from Iceland,

[2]source: http://www.foundrymag.com/Content/Site018/articles%5C06_01_2007%5C68507KukaTitang_00000044521.gif

Liechtenstein and Norway through the EEA Financial Mechanism and the Hungarian National Development Agency.

REFERENCES

[1] SOROS, G. – RESKO, B. – SOLVANG, B. – BARANYI, P.: “A cognitive robot supervision system,” in *Applied Machine Intelligence and Informatics, SAMI 2009*, 2009.

[2] KORONDI, P. – SOLVANG, B. – BARANYI, P.: “Cognitive robotics and telemanipulation.” in *15th International Conference on Electrical Drives and Power Electronics (EDPE 2009)*, 2009.

[3] WOODS, E. – MASON, P. – BILLINGHURST, M.: “Magicmouse:an inexpensive 6-degree-of-freedom mouse.” in *Proceedings of Graphite 2003*, 2003.

[4] KATO, H. – BILLINGHURST, M. – POUPYREV, I. – IMAMOTO, K. – TACHIBANA, K.: “Virtual object manipulation on a table-top ar environment.” in *In proceedings of the International Symposium on Augmented Reality*, 2000.

[5] HENNING, M.: “A new approach to object-oriented middleware,” *IEEE Internet Computing*, vol. 4, 2004.

[6] JUNKER, G.: *Pro OGRE 3D Programming*. Apress, 2006.

[7] GASPERI, M.: *LabVIEW for LEGO Mindstorms NXT*. National Technology & Science Press, 2009.

Received January 10, 2010, accepted April 26, 2010

BIOGRAPHIES

Ádám Vámos, 2009: M.Sc. Degree in Technical Informatics, Budapest University of Technology and Economics
Ádám Vámos has published 1 international conference paper and 1 journal paper about VIRCA, a Cognitive informatics framework system, and currently working at SZTAKI on the VIRCA project.

István Fülöp, 2008: M.Sc. degree in Technical Informatics, Budapest University of Technology and Economics (Integrated Intelligent Systems Specialization)
István Fülöp is member of the Cognitive Informatics Research Group of the Computer and Automation Research Institute (MTA-SZTAKI). He is involved in the development of VirCA, the Virtual Collaboration Arena developed by the research group.

Barna Reskó,
- 2004: M.Sc. degree in Technical Informatics, Budapest University of Technology and Economics
- 2009: Ph.D. degree in Intelligent Systems, Budapest University of Technology and Economics
- 2004: Research student, Norwegian University of Science and Technology
- 2005 - 2007: Research student, University of Tokyo

Barna Reskó has published 38 international conference papers and 9 journal papers about the Cognitive informatics modeling of the primary visual system and its applications.
Memberships: IEEE Hungarian section, John V. Neumann Computer Science Society, Hungarian Fuzzy Association

Péter Baranyi, Brief Biographical History:
1994. M.Sc. degree in electrical engineering, Budapest University of Technology and Economics
1995. M.Sc. degree in education of engineering sciences, Budapest University of Technology and Economics
1999. Ph.D. degree, Budapest University of Technology and Economics
2006. D.Sc degree, The Hungarian Academy of Sciences
1996. and 1998. The Chinese University of Hong Kong
1997. The University of New South Wales, Australia
1996. CNRS LAAS Institute, Toulouse, France
2000 - 2001. Gifu Research Institute, Japan
2002 - 2003. University of Hull, U.K.
2005. Leader of the Cognitive Informatics Research Group at the Computer and Automation Research Institute of the Hungarian Academy of Sciences.
Membership in Learned Societies:
- Founding member of the Integrated Intelligent Systems Japanese - Hungarian Laboratory
- Founding member of the Information Technologies in Manufacturing Norwegian-Hungarian Joint Laboratory
- Member of IEEE