Challenges in Communications Research Beyond the VICOM Project

F. Vatalaro[1], G. Cortese[2], F. Davide[2], A. Detti[1],
M. Leo[3], P. Loreti[3] , and G. Riva[4]

[1] Dipartimento di Ingegneria Elettronica - Università di Roma "Tor Vergata",
and CNIT,
Via del Politecnico 1, 00133 Roma, Italy
`info@vicom-project.it`
[2] Telecom Italia Learning Services S.p.A., Rome, Italy
[3] CNIT, Consorzio Nazionale Interuniversitario delle Telecomunicazioni, Italy
[4] Applied Technology for Neuro-Psychology Lab., Istituto Auxologico Italiano,
Milan, Italy

Abstract. The VICOM (Virtual Immersive COMmunications) project is a three-year project funded by the Italian Ministry of Instruction University and Research aiming at investigating innovative communication paradigms. The project represents a wide coordinated effort focused on integration of immersive and wireless technologies in view of the fourth generation of mobile communications. The main goal of the project consists of the design of a wideband system architecture for immersive services and of its validation through two distributed large test-beds. Starting from VICOM ongoing experiences some future challenges and objectives for the future situated and autonomic communications technologies are envisaged in the paper.

1   Introduction

The VICOM (Virtual Immersive COMmunications) project [1] is a project funded by the Italian Ministry of Instruction University and Research (MIUR) focused on investigating innovative communication paradigms. It is a three-year project started in November 2002.[1]

The main project goal is the design of a communication system architecture able to provide mobile immersive services. The architecture effectiveness will be demonstrated in two service test-beds identified in the project as Mobility in Immersive Environment (MIE) and Virtual Immersive Learning (VIL), respectively. The test-beds aim at being one first step towards new communication models focused to

---

[1] VICOM partners are: CNIT, a consortium of universities acting as the coordinating partner, and involving researchers from several universities, the Italian National Research Council (CNR), through their Bologna and Pisa units, the Polytechnic of Milan, the ISCTI (Istituto Superiore delle Comunicazioni e delle Tecnologie della Informazione) of the Ministry of Communications, and Telecom Italia Learning Services.

achieve a natural interaction with the communication media. VICOM objectives are integration of the immersive and virtual technologies with wireless access technologies, the development of ubiquitous interaction tools between humans and virtual and physical environments, the seamless integration of immersive interfaces and sensors with the environment to induce perception of a natural interaction, the development of non-invasive multimodal interfaces, wearable devices, "aural" networks and the evaluation of their ergonomics, the degree of user acceptance, as well as the psychological effects.

Starting from the description of the VICOM ongoing research the paper illustrates some future challenges and objectives for the future situated and autonomic communications technologies.

2   VICOM Research Framework

Multimedia mobile communications are starting to face the challenge of integrating audio, video and sensing interfaces, in order to realize new forms of services. Among these, mobile immersive services enriched with virtual contents will play a significant role: this is the technological framework for the VICOM project. In the future of telecommunications, beyond the present so called third generation (3G), virtual and immersive communication services will induce an augmented reality experience in the user through the integration of pervasive communication technologies and virtual multimedia contents. A unified architecture will seamlessly integrate all the needed system and service features. This is one first step towards the development of a auto-reconfigurable, scalable and technology independent communication system foreseen by the situated and autonomic paradigm for the 2020s.

To reach this purpose a system architecture has been designed to transparently support ambient intelligent services, to provide advanced context aware functionalities, to make available user content adaptation services and to offer self-organized network configurations. The system architecture, now under development, aims at providing a first benchmark to experiment virtual and immersive technology components in an exemplary distributed scenario. The selected service scenario is the so-called "VICOM campus" comprising multiple real campuses, augmented by virtual contextual information as dynamically required to service the user. Therefore, the campus will be a mixed reality distributed service area in which the user may benefit, on his/her personal devices and/or on ambient devices, of a predefined set of immersive services, based both on person-to-person and person-to-system communications.

Main research areas are communications, ambient and service intelligence, and virtual multimedia contents delivery. The specific studies undertaken are focused on novel communications paradigms [2],[3]. Nevertheless, the system architecture intends to be effective also with today's off-the-shelf virtual reality and wireless technologies. Moreover the system architecture aims to be "plug & play", seamlessly interfacing different technology components, which will live together in the experimental system. To this aim a set of standard interface are being defined.

Fig. 1 summarizes the research topics providing foundations to the two VICOM immersive technology test beds (MIE and VIL) to be developed.
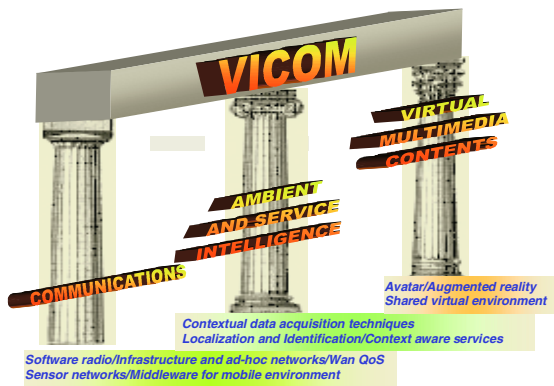

Fig. 1. VICOM research pillars

The MIE test-bed conjugates immersive and mobility aspects, based on wideband wireless techniques, in order to offer "light" virtual multimedia contents to the user on-the-move, both pedestrian (indoor, outdoor) and vehicular. On the other hand, the VIL test-bed is mainly focused on "heavy" virtual contents distribution within a fixed high-speed content delivery network. In both the test beds service intelligence is achieved by means of a context-aware service approach, while the ambient intelligence components are off-the-shelf hardware components (microphones, cameras, etc.) enriched with software applications specially developed or adapted on purpose.

3   Architecture and Services for the Test-Bed

3.1   Architecture

The system architecture is designed to support the main system functionalities: shared context function and content adaptation. The shared context function is in charge of disseminating context data among the users. The content adaptation function changes and/or scales the services presentation accounting for the device capabilities [5]. These functions will generate a data flow composed of natural and synthetic flows, such as avatars or contextual iconic representations that enable several forms of presentation. In fact the data collected in the test-bed environment, gathered by heterogeneous sensors (i.e. cameras, microphones, wireless sensors, etc.), and low level context data (i.e. number of persons in a room, available communications ports, etc.) will be fused together in order to provide the application layer high level context information (i.e. in the room there is a meeting and which communication means are usable).

Moreover a middleware layer will provide both context data dissemination and session-oriented communications; the network functionalities will provide the needed communication capabilities integrating fixed and wireless network elements. It in-

cludes ad-hoc networking features offering the best QoS management. Fig. 2 illustrates the solution for the architecture of the software platform.
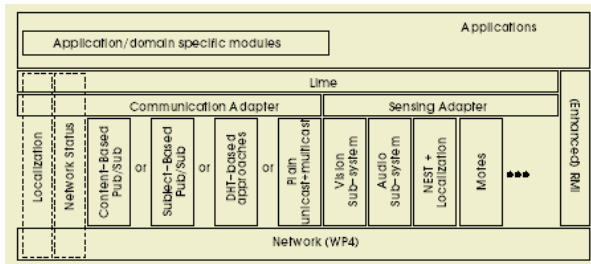

Fig. 2. The VICOM architecture

The analysis of requirements of the application scenarios has made clear a number of requirements for the middleware useful to support the development of the various functionalities that require support for:

- Proactive and reactive operations. Applications must be able to proactively request for information, as well as be asynchronously notified of relevant events.
- Content-based information access. Applications must be able to select the information they need to access, using filtering criteria that are not necessarily fixed a priori and that are based on the information content itself.
- Data sharing. User applications need to share with others their data, both of applicative and contextual nature, and need to do so regardless of possible network reconfigurations.
- Decentralized, peer-to-peer programming model. If the fundamental requirement of mobility is to be addressed radically, the programming model cannot explicitly rely on a server. Even if mobility is not part of the picture, the large-scale characteristics of VICOM scenarios demand for a high degree of decentralization.

To support these functionalities the middleware layer, used initially in static environments and specifically redesigned for the mobile environment, is adopted as the main interface towards the applications. It introduces rules for the management of data in the shared spaces. This programming model encompasses all the four aforementioned requirements.

The software architecture foresees that the middleware layer provides the overall programming interface for the user applications and that the Communication Adapter, below the middleware, decouples it from the specific transport layer. This way allows that the middleware supports flexible and cooperative interactions among user applications, both in ad-hoc and in infrastructure mode, which are specific of the MIE scenario, as well as for the VIL test-bed.

The counterpart of the Communication Adapter is the Sensing Adapter, which effectively decouples the middleware from the specific of the sensing techniques employed to extract contextual information from the environment. The middleware essentially takes care of disseminating context information across the system. The design context information management is central for the effective development of immersive systems architecture.

Fig. 3 shows the basic logical components required for context information processing, i.e. the components for generation of context information from sensors and for distribution of these data to applications.
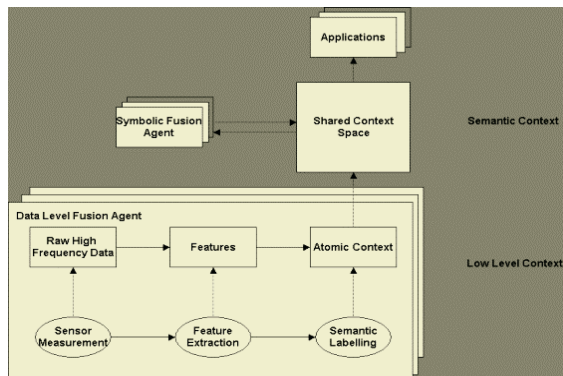

Fig. 3. Context Data Acquisition, Fusion, and Distribution

Low level sensor data are filtered, acquiring raw data and performing some analysis on them, into meaningful 'semantic context' information.

At level of semantic layer the context data are published in a shared data space; applications and symbolic fusion agent components access and possibly subscribe to this information. In this last case, they receive automatic notification of new information. Moreover symbolic fusion agents generate additional context information at a higher level of abstraction.

The user applications access the context space, either in a proactive (query) or publish subscribe style. Applications must have a way to discover sources of context information (i.e. to join the context space) as well as the formats and semantics of data in the space.

- The Shared Context space pattern implements a common, distributed "blackboard" offering publish-subscribe style coordination to a number of peers.
- Providers of information to this space include:
- sensors in the physical environment (cameras, environmental sensors, etc.);
- sensors in the user's Personal Area Network (GPS, motion sensors, cameras, speakers etc);
- inference components for high level;

In addition to basic publish-subscribe mechanisms the context space should provide more sophisticated features such as:

- the ability for applications to obtain time series in addition to single observations;
- aging mechanisms (allowing context information to 'expire' after some predefined time);
- history mechanisms to ease storage of observations, and queries on these observations;
- policy-based control over the quantity of data stored and refresh intervals for sensor observations;
- extensibility and ontology discovery: data in the context space should belong to well-defined ontologies; but flexibility is needed to allow sensor and inference components to implement new ontologies. Middleware should provide for (provider) publication and (client) discovery of context ontologies.

Fig. 3 shows the basic logical components required for context information processing, i.e. the components for generation of context information from sensors and for distribution of these data to applications.

3.2   Test-Beds

For the MIE test-bed, we plan to extensively use ad-hoc networking during the experiment as an extension of the infrastructure wireless coverage. The selected person-to-system trial application is the virtual guide service that augments real world viewed by the user with guide information, in order to create the immersion of the service in the natural user environment. The main functionalities that will be validated with this test-bed are:

- prove the sensing, storing and distribution of the context data (location and identity) both in ad-hoc and in infrastructure wireless environment;
- use of spatial model of the ambience within the application;
- use of augmented reality for the presentation of virtual contents.

In the VIL test-bed, the trial service is the immersive teaching in which each student is reported on a virtual room where real classrooms or home stand alone users are represented by video-stream or by "avatar" depending on their communication capabilities. The teacher is provided with mean to enhance his "student feedback" with respect to actual e-learning approach. Moreover, new interfaces based on virtual reality representation enable the students in accessing remote laboratory instruments in a shared virtual environment fashion. In this test-bed, we resort to fixed network with QoS management, due to the large bandwidth request. The main functionalities that will be validated with this test-bed are:

- prove the use of novel technologies of coding, transport and presentation of natural, synthetic, and mixed multimedia flows enriched with contextual "icons";
- verify the psychological impact of this form of novel e-learning on users.

3.3   User Applications

The virtual immersive guide application will guide the test-bed user towards a specific target. As soon as the user will enter in the test-bed area he will be identified and located. The technologies for the identification and for the localization will be of several different kinds, nevertheless a data fusion module will provide to use the data to extract a single information on the user to disseminate in the shared environment. After the incoming phase the user will choose his target and he will be guided by means of instructions displayed on his device. The instructions and the messages exchanged with the environment and the other devices will be different as function of the user device; they can be provided by a speaking avatar, or by an arrow superimposed on the user view, or by simple textual messages. When the user will reach the target he will be able to get some more information by viewing some information tags that he meets also during his path (Fig. 4).

Fig. 4. Virtual guide examples

With the application for the interaction with the environment the user will be able to interact with the surrounding ambient. Superimposed to the scene he/she is viewing and will able to display some information tags. These tags will be placed both on objects such as printers or computers and on personal devices. For each tag the user will be able to make different actions using a pointing tool. For example by pointing on a printer he will display a menu containing several choices that will be able to print a document command or execute a file transfer with another user (Fig. 5).

Fig. 5. Environmental interaction example

The virtual cooperation in ad-hoc environment application will allow to interconnect many users in an ad-hoc environment. The test-bed scenario will be composed by several independent ad-hoc islands (Mobile Ad-Hoc Network - MANET) interconnected among them. In this environment each user will be able to communicate both directly with the users inside his MANET or the remote users located in external MANET. As shown Fig. 6 a user can join a discussion room talking with users located in one or more different MANETs.
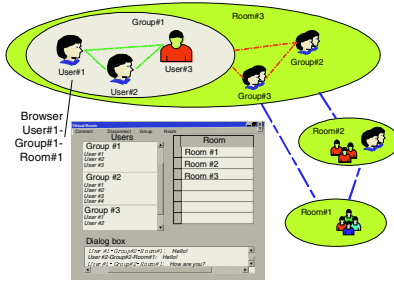

Fig. 6. Cooperation in ad-hoc environment example

In the test-bed VIL the immersive learning application will provide the user of a multimedia board in which natural and synthetic data flows will merge to offer an interactive interface. Using context-data the synthetic objects will superimpose to natural data flows to improve the interaction during one e-learning session. Moreover immersive technologies will increase the sense of being there so the user will have the experience to be in a virtual classroom. The context data will play an important role in this test-bed because the environment data (i.e. temperature, number of people, etc...) will be used to increase the information content of data flows. The virtual classrooms will be equipped with environmental sensors to provide for the identification and the localization services. They will be interconnected among them and the middleware will share the context data detected. The context management engine will be able to recognize human gesture and other conventional signs useful to have a natural interaction among all the users. Main targets will be the capability to perform the identification and to precisely locate the users that take part to the lesson. Another issue of the test-bed will be the capability for the system to adapt the data flows to the connection availability (Fig.7).
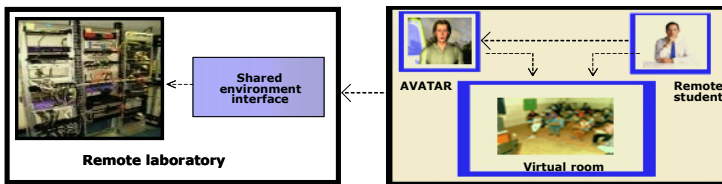

Fig. 7. Immersive lesson example

4   Research Challenges Beyond VICOM

According to the recent "ISTAG SCENARIOS FOR AMBIENT INTELLIGENCE 2010" [7] the evolutionary scenarios will be rooted within three dominant trends[8]:

- pervasive diffusion of intelligence in the space around us, through the development of network technologies and intelligent sensors;
- increasingly relevant role of mobility, through the development of mobile communications, moving from the Universal Mobile Telecommunications System (UMTS) "Beyond 3rd Generation" (B3G);
- increase of the range, accessibility and comprehensiveness of communications, through the development of multi-channel multimedia technologies.

The convergence of biosensors, 4G mobile communication and multi channel multimedia technologies manifests itself as the next frontier of ICT (Information and Communication Technology). An important role will be played by intelligent environments in which complex multimedia contents integrate and enrich the real space.

Within this process two trends are expected to shape the future challenges beyond VICOM: Ambient Intelligence (AmI) and Immersive Virtual Telepresence (IVT).

AmI is an emerging interface paradigm in which the computer intelligence is embedded in a digital environment that is aware of the presence of the users and is sensitive, adaptive, and responsive to their needs, habits, gestures and emotions.

IVT is a new hybrid platform including shared virtual reality environments, wireless multimedia facilities - real-time video and audio – and advanced input devices – tracking sensors, biosensors, brain-computer interfaces. For its features IVT can be considered an innovative communication interface based on interactive 3D visualization, able to collect and integrate different inputs and data sets in a single real-like experience.

A typical first generation IVT system is virtual reality [9](Fig. 8). In VR, using visual and auditory output devices, the user can experience the environment as if it were a part of the world. Further, because input devices sense the operator's reactions and motions, the operator can modify the synthetic environment, creating the illusion of interacting with and thus being immersed within the environment.

IVT, however, is not only a hardware system. According to different authors the essence of IVT is the inclusive relationship between the participant and the synthetic environment, where direct experience of the immersive environment constitutes communication [10]. In this sense, IVT can be considered as the leading edge of a general evolution of present communication interfaces like television, computer and telephone. Main characteristic of this evolution is the full immersion of the human sensorimotor channels into a vivid and global communication experience: IVT provides a new methodology for interacting with information.

For this reason, next generation IVT systems will have an improved focus on the communication capabilities. A possible future IVT application is Mobile Mixed Reality (MMR) [11]. This application foresees the enhancement of information of a mobile user about a real scene through the embedding of one or more information objects

within his/her sensorial field. These objects may be part of a wider virtual space – the AmI Space - whose contents can be accessed in different ways and using different media (cellular phones, tablet PCs, PDAs, Internet, etc.).
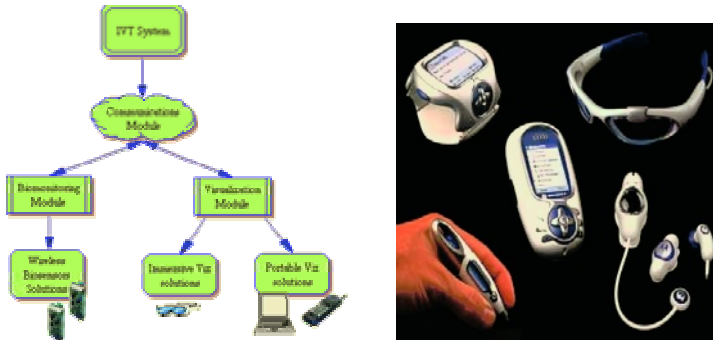

Fig. 8. The IVT system functional architecture and a prototype under development by Motorola

The possibilities offered by MMR are huge. By integrating within a common inter-face a wireless network connection, wearable computer and head mounted display, MMR virtually enhances users' experience by providing information for any object surrounding them. They can manipulate and examine real objects and simultaneously receive additional information about them or the task at hand.

Moreover, using Augmented or Mixed Reality technologies, the information is pre-sented three-dimensionally and is integrated into the real world. Recently, Chris-topoulos [12] identified the following applications of MMR:

- Smart signs added to the real world: Smart signs overlaid on user real world may provide information assistance and advertisement based on user preferences.
- Information assistant (or "virtual guide"): The virtual guide knows where the user is, his/her heading, as well as the properties of the sur-rounding environment; interaction can be through voice or gestures, and the virtual guide can be an animated guide and provides assis-tance in different scenarios based on location and context informa-tion.
- Augmented Reality or Virtual Reality combined with conversational multimedia (or "virtual immersive cooperative environments"): Conver-sational multimedia can be also added to a VR or an augmented reality scenario, where a user can see the avatar of another user coming into the scene and a 3D video conference is carried on. If we use VR, given the position and orientation information of the first user in the world, the second user can put the first one (or his/her avatar) in a 3D synthetic world.

In general, the IVT perspective is reached through:

- the induction of a sense of "presence" or "telepresence" through multi-modal human/machine communication in the dimensions of sound, vision, touch-and-feel (haptics);
- the widening of the input channel through the use of biosensors (brain-computer interface, psycho-physiological measurements, etc.) and advanced tracking systems (wide body tracking, gaze analysis, etc.).

Typically, the sense of presence is achieved through multisensorial stimula such that actual reality is either hidden or substituted via a synthetic scenario, i.e. made virtual through audio and 3-D video analysis and modelling procedures. In high end IVT systems, multimedia data-streams, such as live stereo-video and audio, are transmitted and integrated into the virtual space of another participant at a remote system, allowing geographically separated groups to meet in a common virtual space, while maintaining eye-contact, gaze awareness and body language. Presence with other people who may be at distant sites is achieved through avatar representations with data about body movement streamed over a high-speed network. Following these premises, a general system functional architecture for a high-end IVT systems should includes three main modules:

- The Visualization Module will use virtual environments and augmented reality to provide totally new users services and interfaces. The research will focus on the characteristics and components of wearable personal virtual reality systems with augmented reality display systems, tracking systems, wireless communications and wearable computing. Wireless communication is needed between components of the system and also between personal augmented reality system and networks services, such as world models and other users or avatars.
- The Biomonitoring Module will give the access to a wide range of bio-metrics data to support highly individual services. Biosensors are a neural interface technology that detect nerve and muscle activity. Currently, biosensors exist that measure physiological activity, muscle electrical activity, brain electrical activity, and eye movement.
- The Core Module within the system manages the information flows both internally within the software and externally within the environment to allow remote access and interrogation. This model requires unique messaging services that make the IVT database accessible to external authenticated users.

Transforming this vision in reality is not an easy task: the most a technology is complex and costly, the less the user is prone to accept it. Significant efforts are still required to move AmI and IVT into commercial success and therefore routine use. Possible future scenarios will involve multi-disciplinary teams of engineers, computer programmers, and users working in concert. Information on advances in IVT and AmI technology must be made available to the research community in a format that is easy-to-understand and invites participation. Future potential applications of these tools are really only limited by the imaginations of talented individuals.

5   Conclusions

The main target of the VICOM project is the development of two large test-beds in which mobile communication and virtual environment will be merged to offer to the user an immersive experience in a mobile environment. This is only the first step in the achievement of new communication modes pervasive and immersive. The knowledge developed through the VICOM project is carrying out new skills and competences to invest in innovative proposal to the development of full AmI and MMR paradigms.

References

1. VICOM Website: http://www.vicom-project.it
2. F.Vatalaro, "Dalle Telecomunicazioni alla Telepresenza Immersiva", Notiziario Tecnico Telecom Italia, Anno 11, n.3, Dic. 2002, pp. 111-116 (in Italian)
3. G. Riva, P. Loreti, M. Lunghi, F. Vatalaro, F. Davide, "Presence 2010: The emergence of Ambient Intelligence", from "Being There: Concepts, effects and measurement of user presence in synthetic environments", IOS Press 2003, section 4, pp. 59-82, available at: http://www.vepsy.com/communication/volume4/4Riva.pdf
4. Amy L. Murphy, Gian Pietro Picco, and Gruia-Catalin Roman. Lime: A Middleware for Physical and Logical Mobility. In Proceedings of the 21 st International Conference on Distributed Computing Systems (ICDCS-21), May 2001
5. G.P. Picco, S. Cicero, G. Cortese, D. Frey, A. L. Murphy, E. Trevisani, A. Vitaletti - "Software Architecture and Middleware for VICOM: Concepts, API, and guidelines for Application Development"- VICOM project internal document
6. F. Davide, A. Detti, E. Gregori, F. Vatalaro, "Wireless Networking for Virtual Immersive COMmunications: the VICOM Project" in Proc. Personal Wireless Communications (PWC), Venezia, September 2003
7. K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J. C. Burgelma, "Scenarios for ambient intelligence in 2010 (ISTAG 2001 Final Report)," IPTS, Seville 2000
8. F. Davide, P. Loreti, M. Lunghi, G. Riva, and F. Vatalaro, "Communications through virtual technologies," in Advances Lectures on Networking, E. Gregori, G. Anastasi, and S. Basagni, Eds. Berlin: Springer-Verlag, 2002, pp. 124-154
9. G. Riva and B. K. Wiederhold, "Introduction to the special issue on virtual reality environments in behavioral sciences," IEEE Transactions on Information Technology in Biomedicine, vol. 6, pp. 193-7, 2002
10. G. Riva and F. Davide, "Communications through Virtual Technologies: Identity, Community and Technology in the Communication Age," in Emerging Communication: Studies on New Technologies and Practices in Communication, G. Riva and F. Davide, Eds. Amsterdam: Ios Press. Online: http://www.emergingcommunication.com/volume1.html, 2001
11. L. Rosenblum, "Virtual and Augmented Reality 2020," IEEE Computer Graphics and Applications, vol. 20, pp. 38-39, 2000
12. C. Christopoulos, "Mobile Augmented Reality (MAR) and Virtual Reality," Wireless World Research Forum, Stockolm September 17-18 2001