4 Presence 2010: The Emergence of Ambient Intelligence

Giuseppe RIVA, Pierpaolo LORETI, Massimiliano LUNGHI, Francesco VATALARO, Fabrizio DAVIDE

Abstract. In the last five years we have seen significant advances in three promising technology areas: virtual environments, in which 3D displays and interaction devices immerse the user in a synthesized world, mobile communication and sensors, in which increasingly small and inexpensive terminals and wireless networking allow users to roam the real world without being tethered to stationary machines.

The merging of these areas allows the emergence of a new vision: the Ambient Intelligence (AmI), a pervasive and unobtrusive intelligence in the surrounding environment supporting the activities and interactions of the users. The most ambitious expression of AmI is the Mobile Mixed Reality: the enhancement of information of a mobile user about a real scene through the embedding of one or more objects (3D, images, videos, text, computer graphics, sound, etc) within his/her sensorial field.

Within this vision, the concept of presence evolves, too. The sense of “being there” covers only the simulation and immersive sides of the sense of presence. To be “present” in the augmented context offered by the AmI, the user has to be aware of its meaning and accept its hybridity. Only “making sense there”, the user experiences a full sense of presence.

Contents

4.1 Introduction ................................................................. 60
4.2 The emergence of Ambient Intelligence ...................... 63
4.3 Ambient Intelligence: technological requirements ........ 65
4.4 Mobile Mixed Reality: perspectives, challenges and open issues ....................................................... 70
4.5 “Making sense there”: a new vision of presence .......... 77
4.6 Conclusions ................................................................. 80
4.7 References ................................................................. 81
4.1 Introduction

Is presence just a “being there”? Imagine that you are walking through an European street, admiring the old buildings around you. At the end of the road you notice a big building, and you begin to move closer to it. As you get within a certain distance, you suddenly realize that it is an ancient stadium whose exterior is clothed in white marble.

The bottom three stories have many arches separated by different columns; the stories are separated by a thin architrave. The fourth story is a solid wall with thin pilasters. However, much of the stadium's outside walls are missing. You are “there”. In front of you there is an old ruin and around you different tourists are shooting at it. Suddenly, a woman holding a white umbrella start to speak:

“The Colosseum was originally called the Flavian Amphitheatre after its builders, the emperors Vespasian and Titus, both of the Flavian family. Construction began around 70 A.D. in a low lying area between the Palatine, Esquiline, and Caelian hills that had once formed the pond of Nero's Domus Aureus. The amphitheatre probably came to be called the ‘Colosseum’ because a colossal statue of Nero once stood near it. The Colosseum was completed in 80 A.D. and seated more than 50,000 spectators. Its opening was celebrated with 100 days of games in which thousands of animals and gladiators were killed. Occasionally the Colosseum was flooded in order to stage small naval battles. The emperor had his own entrance to the Colosseum, and from his private "box seat" he decided the fate of defeated gladiators.”

After these words, you are still “there”, but your presence in that place is totally different.

Now you are “aware” that this old stadium held 50,000 people and had all kinds of events, from wild animal hunts to gladiators to actual sea battles.

Turning your head you notice a virtual reality kiosk on your left. You approach it and take one of the special see-through glasses. After a few seconds a 3D animation overlaps the scene: the old stadium is replaced by a 3D model of the original Colosseum. Inside it, you can see two gladiators fighting with wild animals in the stadium crowded of shouting people. Even if you are still “there”, now you are more “there” than before. What is the difference? Before you were just “there”. The narratives and the VR experience make you “more” present in the context you were in.

This example clearly show that the existing definitions of presence based on the concepts of “immersion” or “to be in” are somewhat limited and are not able to explain emerging technologies such as augmented reality or mixed reality.

As already noted by Slater [1] the traditional description of presence as the “sense of being there” is no more satisfying:

“Presence is not simply a ‘sense of being there’ that might be assessed in a questionnaire, however long, complex and validated the questionnaire - it is the total response (italics in the original) to being in a place, and to being in a place with other people. The ‘sense of being there’ is just one of many signs of presence - and to use it as a definition or a starting point is a category error: somewhat like defining humor in terms of a smile” (p. 7).

What we propose in this chapter is a definition of presence as “making sense there”. According to this vision, researchers have to study presence by analyzing the user/s
interaction with and within the synthetic environment including all the different aspects that converge on it: the social relationships established, the physical and symbolic resources exploited and the cultural competence used (see also Gamberini and Spagnolli in this volume, Spagnolli and Gamberini [2]).

Moreover, we will also try to show how this vision may change the role of Information technology (IT), transforming the way people interact between themselves and with objects around them. A new role for the Information Technology (IT) is already visible, while technology's focus is gradually shifting away from the computer as such, to the user [3, 4]. This change of paradigm has the objective to make communication and computer systems simple, collaborative and immanent. Interacting with the environment where they work and live, people will naturally and intuitively select and use technology according to their own needs.

A first sign of this change has been the creation of totally new interactive communication environments, such as Computer Mediated Communication (CMC) and Computer Supported Collaborative Work (CSCW) [5]. However, the final steps towards this vision will be allowed by three dominant trends:

- increase of richness and completeness of communications, through the development of multimedia technologies, towards "Immersive Virtual Telepresence" (IVT), including an increased attention to the aspects of human perception and of person-machine interaction;
- increasingly relevant role of mobility, through the development of mobile communications, moving from the Universal Mobile Telecommunications System (UMTS) "Beyond 3rd Generation" (B3G);
- pervasive diffusion of intelligence in the space around us, through the development of network technologies.

The merging of these trends, within the “being aware there” paradigm (see Figure 4.1), allows the emergence of a new vision [6]: the Ambient Intelligence (AmI), a pervasive and unobtrusive intelligence in the surrounding environment supporting the activities and interactions of the users.

Figure 4.1 Converging technologies in the “being aware there” paradigm
As noted by the ISTAG group [6]: one or more objects (3D, images, videos, text, computer graphics, sound, etc) within his/her sensorial information [7].

“Such an environment is sensitive to the presence of living creatures (persons, groups of persons and maybe even animals) in it, and supports their activities. It ‘remembers and anticipates’ in its behavior. The humans and physical entities - or their cyber representatives - together with services share this new space, which encompasses the physical and virtual world” (p. 6).

The AmI can be seen as the integration of functions at the local level across the various environments, enabling the direct natural and intuitive interaction, and also dialogue, of the user with applications and services spanning collections of environments - including the cyberspace level - enabling knowledge, content organization and processing.

The most ambitious expression of AmI is Mixed Reality (MR). Using Mixed Reality it is possible to seamlessly integrate computer interfaces into the real environment, so that the user can interact with other individuals and with the environment itself in the most natural and intuitive way.

Within MR, a key role will be played by Mobile Mixed Reality (see Figure 4.2): the enhancement of information of a mobile user about a real scene through the embedding of one or more objects (3D, images, videos, text, computer graphics, sound, etc) within his/her sensorial information [7].

As indicated by the AmI framework, the embedded information is based on factors like location and direction of view, user situation/context aware (day of the time, holidays of business related, etc), user preferences (i.e. preference in terms of content and interests), terminal capabilities and network capabilities.

![Figure 4.2 An early prototype of Mobile Mixed Reality (From [7])](image-url)
However, this is not an easy task. In fact, the development of effective AmI and MR tools requires the concurrent efforts of different disciplines ranging from engineering to ergonomics, from communications to psychology.

4.2 The emergence of Ambient Intelligence

What is Ambient Intelligence? Ambient Intelligence (AmI) refers to a new paradigm in information technology, in which people are empowered through a digital environment that is aware of their presence and context, and is sensitive, adaptive, and responsive to their needs, habits, gestures and emotions.

According to the vision of AmI provided by the Information Society Technologies Advisory Group (ISTAG) to the European Commission, all the environment around us, homes and offices, cars and cities, will collectively develop a pervasive network of intelligent devices that will cooperatively gather, process and transport information [8]. Therefore, in this sense AmI is the direct extension of today’s concept of ubiquitous computing, i.e. the integration of microprocessors into everyday objects. However, AmI will also be much more than this, as the AmI system should adapt to the user’s needs and behavior.

In fact, as underlined by the AMBIENCE Project, AmI can be defined as the merger of two important visions and trends: "ubiquitous computing" and "social user interfaces":

“It builds on advanced networking technologies, which allow robust, ad-hoc networks to be formed by a broad range of mobile devices and other objects (ubiquitous- or pervasive computing). By adding adaptive user-system interaction methods, based on new insights in the way people like to interact with computing devices (social user interfaces), digital environments can be created which improve the quality of life of people by acting on their behalf. These context aware systems combine ubiquitous information, communication, and entertainment with enhanced personalization, natural interaction and intelligence”.

(http://www.itea-office.org/projects/facts_sheets/ambience_fact_sheet.htm)

Figure 4.3 The AmI Space (adapted from [9])
Further developing these points, ISTAG introduced the concept of *Aml Space*. The Aml Space is composed of *networked* (using a changing collection of heterogeneous network) *embedded systems* hosting *services* which are dynamically configured distributed components (see Figure 4.3).

The *Aml Space* can be seen as the integration of functions at the local level across the various environments, and enables the direct natural and intuitive dialogue of the user with applications and services spanning collections of environments - as well as at the cyberspace level - allowing knowledge and content organization and processing [9].

In particular hey should offer capabilities to:

- *Model the environment and sensors available to perceive it*, to take care of the world model. This deals with the list of authorized users, available devices, active devices, state of the system, et cetera.
- *Model the user behavior* to keep track of all the relevant information concerning a user, automatically builds the user preferences from his past interactions and eventually abstracts the user profile to more general community profiles.
- *Interact with the user* by taking into account the user preferences. Natural interaction with the user replaces the keyboard and windows interface with a more natural interface like speech, touch or gestures.
- *Control security aspects* to ensure the privacy and security of the transferred personal data and deal with authorization, key and rights management.
- *Ensure the quality of services* as perceived by the user.

Within this frame, Immersive Virtual Telepresence (IVT) and wireless technologies will play a fundamental role in helping the Aml vision to cope with the need for natural user interfaces and for ubiquitous communication.

The former will enable the citizen to interact with the Aml and to control it in a natural and personalized way through voice and gestures. The latter will provide the underlying network and will also enable electronic devices to communicate with each other and with the user.

However, the Aml requirements are not just technological. ISTAG identified a series of necessary characteristics that will permit the eventual societal acceptance of Aml [8].

Aml should:

- facilitate human contact;
- be orientated towards community and cultural enhancement;
- help to build knowledge and skills for work, better quality of work, citizenship and consumer choice;
- inspire trust and confidence;
- be consistent with long term sustainability both at personal, societal and environmental levels;
- be controllable by ordinary people - i.e. the ’off-switch’ should be within reach (these technologies could very easily acquire an aspect of ’them-controlling-us’!).

Moreover, to be successful, the various Aml markets will require specific contents.

In particular there is a need for content-oriented tools and services to support multi-cultural content generation, its engineering and management.
4.3 Ambient Intelligence: technological requirements

4.3.1 Mobile radio technologies

Although there is a tendency to refer to the future "beyond 3G" (B3G) mobile radio technologies as 4G or even 5G, there is a significant discontinuity in aims and modalities in the expected development of technology [6]. In fact, 1G, 2G and 3G have been based on the approach of defining and standardizing "monolithic" systems having a certain degree of internal rigidity. On the contrary, B3G is being conceived to be characterized by flexible, reconfigurable and scalable system elements and by the seamless integration of different systems, both mobile and fixed systems.

Flexibility in network interconnection and evolution in network system features will have a limited extent within 3G. Starting from 3GPP (3G Partnership Program) Rel. 99 issuing of specifications, a possible evolution path will include at least two additional steps [10]:

- 3GPP Rel.04, still with separate transport and signaling control paths, but with unified "Media Gateways" to jointly process circuit switched and packet switched traffic;
- 3GPP Rel.05, with full network integration of transport and control signals, merging into a single IP paradigm.

This evolutionary approach is suggested by a combination of business and technical reasons. One of the main concepts in UMTS is that the network establishing connections is separated as much as possible from the network parts creating and maintaining services.

To implement this concept 3GPP Rel. introduces network components, known as service capabilities. In 3GPP Rel. service capabilities are still proprietary and can be vendor-dependent, and there is no common service creation environment.

This is overcome with 3GPP Rel.04 and Rel.05 with the Open Service Architecture (OSA). Figure 4.4 collects some service capabilities already defined in 3GPP Rel. (they will be expanded within Rel.04/05):

- WAP (Wireless Application Protocol), to offer to the end user a web browser tailored to the wireless environment;
- LCS (Location Communication Services), to provide other service capabilities with the user position information;
- MExE (Mobile station application Execution Environment), to provide other service capabilities with information about the ability of the terminal to handle information;
- USAT (UMTS SIM Application Toolkit), to offer the tools required for SIM card handling.

A further service capability of special importance is CAMEL (Customized Applications for Mobile network Enhanced Logic). Within 3G the concept of Intelligent Network (IN) will evolve.

The concept of IN is derived directly from the PSTN/ISDN network environment, and thus has some deficiencies when adapted to a mobile network environment.

Originally, IN allows transfer of user service information within the user’s own home network.
However this is not adequate when he/she can roam between different networks. CAMEL is therefore the service capability necessary to transfer service information between networks: it acts as a "service interconnection point" and serves other service capabilities.

The objective is to allow completion of transactions involving different networks transparently to the user. One example of CAMEL implementation is through the concept of VHE (Virtual Home Environment): thanks to VHE the user is provided with the same personalized set of services, no matter through which UMTS network gets connected. This concept is depicted in Figure 4.5.

A key problem to be faced is how to break through the barrier between the individual and the network, so realizing efficient person-machine interfaces. Achieving ever high degrees of integration between users and the AmI implies the existence of connections between electronic and telecommunications devices the user wears and the surrounding environment. An advanced network to allow the interconnection of a range of different wearable devices would constitute a Body Area Network (BAN) [11, 12]. A BAN is a heterogeneous network, which meets at least the following requirements:

- is capable of connecting both complete devices and individual device components in an easy and reconfigurable way;
- supports a range of different connection technologies - using both wired and wireless techniques including radio, infrared, and optical;
- is easy to use and to configure;
- supports a range of different classes of data (real time audio and video, Internet, etc.);
- allows the user to connect to the outside world;
- ensures security with respect to connections with the outside world.
As the BAN has to support interoperability between devices from different manufacturers it is necessary to define a common communications protocol. In this setting a BAN device could be either a complete piece of equipment (for example a cell-phone) or a single component (such as an earphone or a keyboard). BANs do not usually have to communicate over distances greater than 1-2 metres.

Personal Area Networks (PANs) are very similar to BANs. A PAN is defined as a short range (about 10 m) wireless network allowing the creation of ad-hoc, configurable connections to/from the outside world.

While BANs provide local management of devices worn by the user PANs create a network around the user.

Typical applications for a PAN include a range of personal computational services provided via laptops, PDAs and mobile phones. A PAN, might, for instance, provide access to a printer located in the vicinity of the user or Internet access via a local port.

BANs and PANs will be able to follow the user wherever he/she goes. Strong interactions between BAN/PAN and the environment will allow AmI to get accurate perception of the user. Strong and advanced forms of interaction between the PAN and the BAN will allow the user to receive richer feedback from the surrounding AmI. The underpinning technologies required to construct the AmI landscape cover a broad range of ICT technologies. In the following we highlight some of the required technologies.

### 4.3.2 Enabling Technologies

The realization of the AmI vision requires the development of a number of enabling technologies [6]: some of them are today under study, and a few are sketched in the following. We only concentrate below on interface technologies and mobile communications technologies.
The former are focal to the development of IVT technologies. The latter aim to cope with the requirements of flexibility, scalability and efficient spectrum usage.

- **Vision technologies.** Human vision is based on perception\(^1\), a complex hypothesis testing mechanism grounded on the two optical images provided by the eyes. Today’s 3-D vision technologies are two: stereoscopy and holography. The former are more developed and are based on the parallax effect, exploited by the human vision system to perceive depth. Brain attempts to overlap the two different retinal images as much as possible by rotating the eyeballs towards a single focal point. The IMAX-3D movie vision system is based on stereoscopy. The scene is filmed with two orthogonally polarized cameras located 6.5 cm apart (approximately the eyeball distance). The two images are then projected on one single movie theater screen. The spectator wears the IMAX-3D visor with polarized lenses matched to the two incoming waves. Holography is a means to create an image without lenses. Differently from photography, in addition to the intensity of the light wave, holography memorizes its phase. Today it is still difficult to provide holographic moving images, so this technology is considered suitable for applications in a perspective beyond a ten year vision. A number of other vision technologies are under study: this is one technology area in which some main breakthroughs will certainly happen: 3-D computer displays [14], miniaturized cameras for direct injection of images into the eye, etc. Another related technology is eyeball tracking.

- **Smart dust.** “Smart dust” [15] is a cloud of tiny speckles, each one of millimeter dimension, of active silicon: the prototype under development at Berkeley University is shown in Figure 4.6. Mote senses, communicates and is self-powered. Each mote converts sunlight into electricity, locally elaborates information, localizes itself, both in absolute and relative to other particles, communicates with other ones within a few meters; furthermore, they jointly act as a coordinated array of sensors, a distributed computer, and as a communications network. A similar technology is “smart painting”, a random network of wall painted computers studied at the MIT. Some features of smart dust are: free-space communication at optical frequencies, bit rates in the order of kilobits per second, power consumption of a few milliwatts, adoption of power management strategies, directional transmission within a narrow beam.

- **Radio reconfigurability.** Radio reconfigurability (also known as “software radio”) [16] is a set of techniques and rules aiming at allowing the radio system reconfiguration through software downloading, so that the system can readapt to different standards. The ideal software radio system is a single wireless platform configured in software to support any wireless standards for any use, in any place, at any time. Expected impacts of software radio on system design are widespread: from base stations to terminals, and to network infrastructures, from standards definitions to spectrum regulation issues, etc. This will be a fundamental technology, to reach the great objective of flexibility at

\(^1\) “Perception is not determined simply by the stimulus patterns; rather it is a dynamic searching for the best interpretation of the available data. ... The senses do not give us a picture of the world directly; rather they provide evidence for checking hypotheses about what lies before us. Indeed, we may say that a perceived object is a hypothesis, suggested and tested by sensory data.” [13] D. Drascic and P. Milgram, Perceptual Issues in Augmented Reality, in: S.S. Fisher, J.O. Merritt and M.T. Bolas (Eds.), *Stereoscopic Displays and Virtual Reality Systems III*, SPIE Press, 1996, pp. 123-134.
equipment level, the objective of scalability at system level, and that of optimum use of frequency spectrum. **Radio reconfigurability.** Radio reconfigurability (also known as "software radio") [16] is a set of techniques and rules aiming at allowing the radio system reconfiguration through software downloading, so that the system can readapt to different standards. The ideal software radio system is a single wireless platform configured in software to support any wireless standards for any use, in any place, at any time. Expected impacts of software radio on system design are widespread: from base stations to terminals, and to network infrastructures, from standards definitions to spectrum regulation issues, etc. This will be a fundamental technology to reach the objective of flexibility at equipment level, the objective of scalability at system level, and that of optimum use of frequency spectrum.

- **Smart antennas.** Smart antennas aim at dynamically controlling interference produced into the system and external interference captured from it. In transmission a smart antenna ideally radiates power only towards the desired user. In reception it ideally captures power only from the desired user [17]. These antennas are now already considered in cellular system base station designs for inclusion as an option within 3G systems. In the future, their use will be extremely widespread, as spectrum efficiency will be one major target in system design.

- **Stratospheric platforms.** Stratospheric platforms, also known as High Altitude Local Orbiters (HALOs), are among possible new system developments to provide direct and immediate WAN access to telecommunications resources [18]. They are quasi-geostationary platforms located at an altitude around 25 km above ground operating at
millimeter wave frequencies. HALOs can provide cellular coverage in line-of-sight. Their very large elevation angles (more than 60 deg) promise visibility much better than that achievable with geostationary satellite systems; the usage of millimeter wave frequencies can allow use of very large bandwidths, which can result in system capacity much larger than achievable with cellular systems. Main expected applications are last mile interactive multimedia services distribution to low cost rooftop terminals. Each stratospheric platform acts as the highest tower in town, providing high density, high capacity, high speed services, with low power requirements, and no latency, mainly to metropolitan and suburban areas. On-board switching is adopted for direct connection between subscribers within the same platform coverage area. Traffic for subscribers outside the stratospheric platform coverage area will be routed through ground stations and/or by means of inter-platform links.

- **Ultra wide band communications.** Ultra wide band (UWB) communication (also known as "Impulse radio") [19] is a technique to spread the signal bandwidth over an exceptionally large bandwidth (well beyond those of 3G CDMA systems), starting from frequencies close to baseband and with cancellation of the continuous wave. Through the use of extremely narrow pulses (duration of a few nanoseconds) the r.f. bandwidth is in excess of 1 GHz. UWB allows large capacity multiple access, very large values of individual bit rates (e.g. 100 Mbit/s), and very low power spectral densities. Furthermore, the intrinsic diversity provided by the ultra wide bandwidth, combined with the use of RAKE receiver structures, allows multipath fading immunity and better and stable transmission quality. Finally, the usage of low frequencies provides improved penetration of radio waves through walls. UWB will allow spectrum coexistence of different signals for different services overcoming the present problem of inefficient use of large portions of spectrum, permanently and exclusively allocated and poorly used. Main expected applications are indoor interactive multimedia communications, localization and distance measurements. UWB is therefore one possible technology candidate to wirelessly distribute multimedia services provided through last mile technologies, such as e.g. HALOs. Presently, the Federal Communications Commission in the U.S. started the standardization procedure to allow commercial use of UWB [20].

### 4.4 Mobile Mixed Reality: perspectives, challenges and open issues

#### 4.4.1 Perspectives

As we have seen in the Introduction, the most ambitious expression of AmI is *Mobile Mixed Reality* (MMR). MMR is based on the embedding of one or more objects (3D, images, videos, text, computer graphics, sound, etc) within the sensorial field of a mobile user [7].

Moreover, following the AmI paradigm, any embedded object is context aware and based on factors like location, direction of view, user situation, user preferences, terminal capabilities and network capabilities.

The possibilities offered by MMR are huge. By integrating within a common interface a wireless network connection, wearable computer and head mounted display, MMR enhances users’ experience by providing information for virtually 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 presented three-dimensionally and is integrated into the real world. Recently Christopoulos [21] 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 surrounding environment; interaction can be through voice or gestures, and the virtual guide can be an animated guide and provide assistance in different scenarios based on location and context information.
- **Augmented Reality or Virtual Reality combined with conversational multimedia (or “virtual immersive cooperative environments”)**: Conversational multimedia can be added to a VR or AR scenario, as shown in Figure 4.7, 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 the future the terminal will be able to sense the presence of a user and calculate his/her current situation. Throughout the environment, bio-sensing will be used to enhance person-to-person and person-to-device communications. Biometrics technology will be used to enhance security by combining static (facial recognition) and dynamic information (voice and lip movement, uncontrolled user gestures), as well as user’s habits, which the network will be able to acquire and maintain.

![Figure 4.7 MMR combined with conversational multimedia (from [21])](image)
4.4.2 Challenges and open issues

According to the Book of Visions 2001 elaborated by the WWRF (Wireless World Research Forum) “During its first century of existence, the development of communication systems has experienced a serious degradation of end user’s ease of use.” [22]. In fact, the increasing complexity of technology heavily impacts on user friendliness of telecommunications networks.

Conversely, the fast penetration of wireless communications has put into evidence the user’s need to get easily connected anywhere and anytime at an affordable price. On the one hand, wireless communications clearly proved that the most a technology provides simple access means, added to freedom of movement and increased security, the most the user is willing to accept it.

On the other hand, the most a technology is complex and costly, the less the user is prone to accept it, in spite of possibly large potential advantages, which are generally not reachable by the average user not interested in spending time and energies in acquiring the underlying technology fundamentals. As a consequence, the successful systems of the future will adhere to the paradigm of “disappearing technologies”, both valid for communications and computing, and will provide improved ease-of use at the expense of an increased, but invisible to the user, complexity of the underlying systems and networks necessary to transport and process the information in the different multimedia forms and usage contexts.

As the user must be firmly put at the “center of the universe of technologies”, it is clear that the elaboration of a purely technical vision of the future of telecommunications and computing is not only insufficient but even dangerous. Rather, any viable technical solutions must be put into a much wider perspective.

According to [22] we mainly need:

- a solid user-centered approach, looking carefully at the new ways users will interact with the wireless systems;
- innovative services and applications made possible with the new technologies;
- new business models and value chain structures overcoming the traditional "user - network provider - service provider" chain.

It is expected that the major innovative thrust will come from new ways of interaction with the wireless system and among systems. The emerging need is bridging the user’s real world and virtual world and to continuously put them in contact.

The challenge for technology is therefore seamless communication amongst individuals, things, and cymans (our synthetic counterparts in the cyberworld - a sort of autonomous avatars) to provide the citizen of the future with a natural enhanced living environment.

It is generally considered feasible that this vision may enter into effect around 2010 [22, 23]. In the novel user-centric approach being developed for B3G there are three main issues to be covered:

- integration of mobile technologies;
- user modeling;
- interaction interfaces.
Integration of mobile technologies. It is useful to rely upon a reference model for the interaction between the user and surrounding technologies. In Figure 4.8 we show a conceptual layering of the individual space, where we recognize the following main elements:

- Body area network (BAN), for the friendly management of sensors and interfaces in contact with the body and around it;
- Personal area network (PAN), for downloading of information on personal and shared peripherals;
- Local area network (LAN), for the nomadic access to fixed and mobile networks, and to the Internet;
- Wide area network (WAN), for the access and routing with full mobility and the desired degree of QoS (includes mobile and fixed networks);
- "The Cyberworld", where users avatars and cymans seamlessly interact.

All the layers must be integrated to cooperate in service provision to the user. The outmost layer represents the cyberworld, where avatars of the users can stay in touch with other synthetic agents, so to get services and perform transactions.

Wireless technologies are the key elements for allowing this seamless interaction inside the real world and between it and the cyberworld. WWRF is elaborating the so-called MultiSphere Reference Model, with similar structure as the model referred above [22].

![Figure 4.8 The individual space layers](image-url)
User modeling. Most of the system and application design today is technology driven only because we do not have yet the tools to incorporate user behavior as a parameter in system design and product development. A strong effort must be made in the direction of user modeling to achieve in user understanding the same level of confidence we have reached in modeling technology.

The main workhorse has been identified in the development of user scenarios. This is an activity very marginally involving engineering skills: rather, the focal competencies can be provided by a wide set of professional categories (including psychologists, movie directors, science fiction writers, etc.). The logical flow for assessment of viable scenarios is depicted in Figure 4.9. Scenarios must be designed to encompass societal, economic as well as technology developments and form a logical framework in which use cases can be fitted.

The European Commission and research organizations, such as the WWRF, encourage scenario-based approaches for pushing the research in the right way [30]. Experts have to analyze the scenarios drawing consequences and future research topics. The main output of these modeling efforts will consist in the "pieces of technology" needed to provide the functionalities envisaged within the reference scenarios.

One main result of user modeling efforts is the user-centric context definition. Much more than simple "user profiling" and "service profiling", abstract user context definition will provide the description of different observable dimensions (i.e. attributes) characterizing a situation in which the user is immersed. The context oriented systems will be able to answer questions, such as: Where are you? Whom are you with? What resources are close to you? What is your mood today? Defining abstract context types, the “context metadata” (such as home, office, travel, distress, etc.), will help assisting the user in different environments. Dynamical management of context metadata will be based on creation of a database of common abstract context types.

Figure 4.9 Logical flow for scenarios modeling. (From [22])
The European Commission and research organizations, such as the WWRF, encourage scenario-based approaches for pushing the research in the right way [30]. Experts have to analyze the scenarios drawing consequences and future research topics. The main output of these modeling efforts will consist in the "pieces of technology" needed to provide the functionalities envisaged within the reference scenarios.

One main result of user modeling efforts is the user-centric context definition. Much more than simple "user profiling" and "service profiling“, abstract user context definition will provide the description of different observable dimensions (i.e. attributes) characterizing a situation in which the user is immersed. The context oriented systems will be able to answer questions, such as: Where are you? Whom are you with? What resources are close to you? What is your mood today? Defining abstract context types, the ”context metadata” (such as home, office, travel, distress, etc.), will help assisting the user in different environments. Dynamical management of context metadata will be based on creation of a database of common abstract context types.

Needed context information will be also a function of local culture and on society organization.

This approach has recently received the status of international standard, through the International Organization for Standardization's ISO 13407 "Human centered design for interactive systems".

According to the ISO 13407 standard [24], human-centered design requires:

- the active involvement of users;
- clear understanding of use and task requirements;
- appropriate allocation of function;
- the iteration of design solutions;
- a multi-disciplinary design team.

Moreover, it is based around the following processes:

- understand and specify the context of use;
- specify the user and organizational requirements;
- produce designs and prototypes;
- carry out user-based assessment.

A sample of VE developed using the ISO 13407 guidelines is the IERAPSI surgical training system [25, 26].

Interaction interfaces. Recent advances in wearable computing and intelligent fabrics promise to realize truly ubiquitous communication and computing. While technology used in everyday life is increasingly complex, the basic human capabilities evolve very slowly.

In developing new interaction technologies limitations in sensory and motion ability, in short and long term memory, as well as in brain processing power must be taken into account. In order to fully exploit the technology offer we need to extend the interaction with more senses (touch, smell, and taste) and at the same time make better use of the senses used today (hearing and vision) by also exploring peripheral vision and ambient listening. With even a longer perspective we may consider combining perception through multiple senses and "amplifying” them with new technology prostheses. In the short term the increase of efficient use in human interaction interfaces will be the first step, essentially
extending to communication and computing the most intuitive and well assessed means of use of senses and of elementary devices of common use. In the following we present two examples of new interaction interfaces. The first is the phone glove of the Ericsson Research Usability and Interaction Lab [22, 27], which provides an intuitive extension of the touch sense in communication (see Figure 4.10).

The second example refers to extension of functionalities through objects of common use. The pen in Figure 4.11, at first glance is a conventional pen to write on a piece of paper; however, it contains a video camera that picks up and transmits the written text.

It also contains a set of sensors able to feel how much the nib is being pressed, and the way the pen is handled. If used on a special paper containing invisible identification it is able to associate what is written with the location of writing [28].
4.5 “Making sense there”: a new vision of presence

4.5.1 From “being there” to “making sense there”

In the previous paragraph we introduced the concept of virtual reality. But what is virtual reality?

VR is sometimes described as a particular collection of technological hardware. In particular, many people identify VR with a collection of devices: a computer capable of 3D real-time animation, a head-mounted display, data gloves equipped with one or more position trackers, etc. However, this exclusive focus on technology is somewhat disappointing [29]. As noted by Steuer [30] this approach:

“Fails to provide any insight into the processes or effects of using these systems, fails to provide a conceptual framework from which to make regulatory decisions and fails to provide an aesthetic from which to create media products” (p. 73).

If VR cannot be simply reduced to a collection of hardware, where should we look to identify its "essence"? Virtual Reality (VR) can be considered as the leading edge of a general evolution of present communication interfaces like telephone, television and computer [31-33].

The ultimate objective of this evolution is the full immersion of the human sensorimotor channels into a vivid and global communication experience [31]. Following this approach, VR can be defined in terms of human experience [30] as "a real or simulated environment in which a perceiver experiences telepresence," where telepresence can be described as the "experience of presence in an environment by means of a communication medium" (pp. 78-80).

As in traditional communication media and in narrative, the possibility of experiencing a sense of presence is strongly related to the possibility of defining a context. For instance, Polany [34] underlined that for a discourse to be a narrative at all, it must provide contextualizing orientative information as well as an account of the events that occurred. In particular, the events must be embedded within an appropriate spatialtemporal context. A good stand-alone narrative must provide its own context: it must orient the listener to where and when the events occurred, who the participants in the events were, what objects were involved, and so on.

This issue is well known by the visitors of theme parks. A recent research analyzed a sample of Walt Disney World’s Epcot visitors to identify the elements able to produce a more realistic ride experience [35]. The research showed that aspects of immersive interfaces, including displays, graphics and control device quality were not as important to the users as the “physics fidelity” (e.g. motion) of the rides, their background stories and goals.

This research underlines a key point for the development of successful VEs: VR is an hybrid technology with two faces: symbolic communication system and simulation tool. Within this vision, the main characteristic of the sense of presence is its hybridity [2] through a careful balance of simulation and symbolic communication. As noted by Fencott and Isdale [36]:

“We can observe that some VEs, such as flight simulators, tank simulators, architectural and urban planning VEs etc, are largely simulations in that, for
very good reasons they seek to emulate some aspect of reality as closely as possible. Other VEs, virtual artworks and some video games, function very much as symbolic systems where simulation is less important that the connotation of meaning. Examples of these are virtual training environments and many videogames.” (p. 1).

If we follow these authors, it’s clear that the sense of “being there” covers only the simulation side of the sense of presence. To be “present” in the context offered by a symbolic system, the user has to be aware of its meaning. Only “making sense there”, the user really experiences a full sense of presence. This point clarifies why in the scenario proposed in the Introduction the narratives and the VR experience make the subject more “present”.

But how we can make sense within a VE? Spagnolli and Gamberini [2] identified three key issues that shape the hybrid relationship existing between the user/s and the VE (pp. 429-431):

- **Expanded setting**: the environment the user is inhabiting while navigating expands beyond the boundaries of the immersion and include elements from different environments in the same setting.
- **Multiple action**: people are able to keep multiple setting running at the same time, acting on each of them simultaneously or in close alternation.
- **Uncertainty**: sometimes an element does not fit, or different scenarios compete to take over. This experience produces a meaning breakdown that requires a sense making process to turn the oddity in something familiar and manageable.

Extending this position, any VR experience, including the Virtual Immersive Cooperative Environment (VICEs) discussed above, can be considered as a process by which an actor or a group of social actors, in a given symbolic system, negotiates the meaning of the various situations that arise around and between them. In this definition we use the term “negotiations” because the sense making process is not automatic but requires a process of progressive discovering and sense making. In this sense, the most effective way of clarifying the meaning of the situation is to relate it to a shared context of meaning [37].

4.5.2 “Being aware there” in Virtual Immersive Cooperative Environments

In most VR simulations and in AmI the role of awareness is usually balanced both by the focus on the specific activity and by the previous knowledge of the user. However, the situation is different when virtual environments or AmI are used as a medium for collaboration between remote participants, as usually happens in VICEs.

As Slater and colleagues [38] found in their experiment, exists a positive relationship between presence of being in a place and copresence - the sense of being with the other people. So, in a VICE, the users require awareness of both the other users and the symbolic system surrounding them.

In fact, a VICE offers its users a digital environment, which can be populated with users replicas (cymans) and objects. Within the digital environment they are free to navigate, meet other users and manipulate data usually represented in the form of three dimensional objects. At the same time they can use audio and video channels to interact with other users using both verbal and non-verbal forms of communication.
The main feature which distinguishes a VICE from a normal Computer Supported Cooperative Work (CSCW) is the use of a *Shared Space* which represents both the users of the environment and the data they are manipulating [39]. However, to be effective, the use of VICE calls for conceptual mechanisms with which groups can be built and vehicles through which groups can express themselves [40].

Benford and colleagues [41] discussed extensively the social significance of space as a resource for activity and interaction in VEs. These authors describe how the use of space - or, rather, the avatars in a meaningful spatial configuration - allows the support or indeed emergence of social mechanisms for control of scarce resources. In particular, they suggest that *continual awareness* of others allows people to flexibly modify their own behavior in social situations (e.g., a user heading across the room towards another probably indicates an interest in beginning an interaction). This result was confirmed also by successive studies [42-44]. In particular, Slater and colleagues [38] found that socially conditioned responses, such as embarrassment, were generated in a VICE, even though the individuals were presented to one another by very simple avatars.

In general the designer of a VICE aims at building tools that make it possible for users - individually or in small groups - to work autonomously. At the same time the tools should monitor the effectiveness of the interaction. To reach this goals developers usually have:

- a *user model* where the information is evaluated and provided according to user needs, knowledge, and preferences;
- and a *usage model* where the information about the user's environment and client hardware and software is held up-to-date.

This implies, for example, that the virtual environment has to allow changes in the way in which the user is represented and monitoring what is going on in the environment [42, 44].

Moreover, as we have just seen, the user internal organization of the environment depends on a shared interpretation of its "meaning". These issues have different key implications for the development [3, 5]:

- *Any model of a VICE has to allow actors to develop shared interpretations of its meaning* Decision making depends on a negotiated reduction in the gap between different actors' individual frames of reference.
- *Any model of cooperation within a VICE is tightly dependent on the specific application area for which the system has been developed*. The designer has to create a model in such a way as to reflect the characteristics of the situation and the goals of the actors.
- *Any model of cooperation will influence the way in which users interact with the system*: When VICE systems introduce new cooperative technologies they inevitably modify interactive processes. As we have already seen on several occasions new technologies are never transparent to users. The designer of a VICE has to realize that the first thing users will do is to try to give a sense to the artifacts they find there. In general terms the users of an VICE face two problems: on the one hand they have to perform their work within the group; on the other hand they have to grip the new technology and learn how to use it appropriately.
- *Any model of cooperation has to make users aware of situations where the model is no longer adequate to lead the interaction*: In many VICEs individual users receive only a limited degree of feedback. This implies that they are not always capable of
evaluating their own performance and the way in which this performance contributes to their general goals. For this reason it is essential that the cooperation model informs users about the way that the system is responding adequately to the situation.

- **Any model of cooperation has to be able to predict the phases in which individual users will organize their work:** A number of studies have shown how collaborative work in daily life involves alternating phases of group work, individual activity and work in restricted sub-groups. Given that users do not always share the same activities, skills and interests, VICE designers have to provide users with tools to approach group activities in a modular fashion. It is particularly important to create tools to allow individual users or small groups of users to carry out their activities autonomously.

4.6 Conclusions

In the last five years we have seen significant advances in three promising technology areas: *virtual environments*, in which 3D displays and interaction devices immerse the user in a synthesized world, *mobile communication* and *mobile computing*, in which increasingly small and inexpensive terminals and wireless networking allow users to roam the real world without being tethered to stationary machines.

The merging of these areas allows the emergence of a new vision: the *Ambient Intelligence* (AmI), a pervasive and unobtrusive intelligence in the surrounding environment supporting the activities and interactions of the users.

The most ambitious expression of AmI is the *Mobile Mixed Reality*: the enhancement of information of a mobile user about a real scene through the embedding of one or more objects (3D, images, videos, text, computer graphics, sound, etc) within his/her sensorial field.

In the chapter we detailed the mobile radio technologies and the different enabling technologies - vision technologies, smart dust, radio reconfigurability, smart antennas, stratospheric platforms, ultra wide band communications – required for the full development of this vision.

Within this vision, also the concept of presence evolves. The sense of “being there” covers only the simulation side of the sense of presence. To be “present” in the augmented context offered by the AmI, the user has to be aware of its meaning. Only “making sense there”, the user experiences a full sense of presence.

Following this position, any AmI experience, including the Virtual Immersive Cooperative Environment (VICEs) discussed above, can be considered as a process by which an actor or a group of social actors, in a given symbolic system, negotiates the meaning of the various situations that arise around and between them. In this definition we used the term “negotiates” because the awareness process is not automatic but requires a process of progressive discovering and sense making.

However, the development of full AmI and MMR paradigms it is not an easy task. As noted by Ducatel and colleagues [8], “Change is fast and it is up to us as entrepreneurs and technologists to engage in constructing the future: these things won’t happen automatically. A focused effort is needed starting now in order to give a shape to these new technologies.” (p. 1). Presence researchers will play a key role in this process.
4.7 References


