

Gesture-Based Hybrid Approach for HCI in Ambient Intelligent Environments

Stefano Carrino, Elena Mugellini,
Omar Abou Khaled

Department of Telecommunications
and Computer Science
College of Engineering and Architecture of Fribourg
Fribourg, Switzerland
{Stefano.Carrino, Elena.Mugellini,
Omar.AbouKhaled}@hefr.ch

Rolf Ingold

Department of Informatics
University of Fribourg
Fribourg, Switzerland
Rolf.Ingold@unifr.ch

Abstract—In this paper we propose a novel interaction approach, based on gestures, aiming to recognize and enhance interactions between augmented human beings and augmented environments. We explain how this model, we called ARAMIS, allows interaction designers to fill the gap between real and virtual worlds augmenting the human itself. An enhanced interaction is achieved exploiting a hybrid approach. This approach is defined hybrid since it is the combination of several complementary techniques: wearable and pervasive computing paradigms, brute force, fuzzy and ML methods, virtual and real worlds, optical and non-optical sensing technologies. A framework implementing this concept has been developed. Finally, in order to validate our approach, we present a first prototype implemented according to the ARAMIS concept.

Keywords – HCI, Ambient Intelligence, gesture-based interaction, hybrid-approach

I. INTRODUCTION

People naturally communicate through the voice, gestures, expressions and movements. In HCI area several researches are addressing the design of natural user interfaces allowing the machine to understand the natural language of human beings [1]. In order to allow this kind of communication and bridge the gap between the human and the machine, natural interactions have to be understood by the system. With the same logic the system must provide a feedback as natural as possible. Although these are still opened research fields, at present Aml systems intrinsically have the potential to implement augmented natural-interaction¹. For instances many of the manipulations used on multi-touch screens, such as the pinch, have no natural counterpart in our interactions with everyday life (it is not possible to pinch a piece of paper to make it smaller, while preserving its aspect ratio). This concept can be highly extended taking into account possibilities provided (at present or in the next years) by novel interaction vectors as smart, mixed reality environments and the

approaching internet of things. In these cases the interaction experienced by the users can be extremely richer than the common interaction of every day [2].

The major contribution presented in this paper is the proposition of a hybrid interaction approach, called ARAMIS, aiming to enhance natural gesture-based interaction between humans and smart environments. Our approach is defined hybrid since it is based on four complementary approaches: environmental/wearable paradigms, brute force/fuzzy/machine learning methods, real/virtual worlds, optical/non-optical sensing technologies. ARAMIS uses this heterogeneity in order to exploit the strong points of individual approaches or technologies, and to overcome their weakness (such as light condition sensitivity in vision-based techniques or heavy training for neural network algorithms).

This paper is organized as follow: Section II presents some related work on the interaction between augmented human and smart environments. Section III briefly presents an envisioning scenario highlighting augmented interactions between the user and smart environments. Section IV describes the ARAMIS approach. Section V introduces how the ARAMIS framework handles the information. Section VI illustrates the implementation of the ARAMIS concept and a first prototype. Finally section VII concludes the paper.

II. RELATED WORKS

Studies on human augmentation through the technology have a long history in HCI domain. Since the works of Engelbart [3] the possibility to augment human capacities using the technology was taken into account. If the first studies focused on approaches to improve human abilities to deal with everyday issues, afterwards the problem changed with the years; technology evolved and from human-machine interaction scientific community starts to talk about human-machine cooperation [4], envisioning a kind of communication based on similitudes with the human-human interaction. Moreover, since the environment surrounding the human was progressively changing in something richer and at the same time more complex (in the interfaces, tools and objects), the

¹ Natural interaction is seen in this paper as the computers capability to exploit the perceptive abilities which so far have been a prerogative of human-human communication

need of an approach encompassing the wearable and pervasive dimension is becoming a central issue. Under this perspective Greenberg et al. [5] defines a novel interaction approach between users and pervasive devices based on their relative proximity (distance, orientation, identity, movement, and location). Pederson et al. [6] present an egocentric interaction model, in which wearable and pervasive paradigms are combined and interactions are classified based on human perceptual characteristic such as what can be observed and what can be manipulated. This model aimed at bridging the gap between real and mixed reality worlds. We differ from previous approaches by providing a perspective in which also the human has to be augmented in order to extend the interaction toward an augmented natural-interaction and, in addition, we focus on hybrid middleware able to take advantages from different techniques.

III. INTERACTION SCENARIO

This section briefly presents a scenario envisioning the augmented interactions possibilities offered by augmented human in smart environments.

Context: Human-computer interactions (HCI) taken into account in this project have as context a smart environment, with real and virtual interactive devices available inside, and a set of wearable devices woven seamlessly on the user. Alice, the system user, coming back home, is wearing her glasses (actually a head-mounted display). The house is an environment enriched by real and digital elements. Alice through gestures can interact with the surrounding environment in a natural, augmented way.

Interaction with virtual objects (virtual screen): Alice seats in her sofa and with a simple gesture (recognized through sensors woven in her clothes and sensors in the environments) a virtual screen appears in front of her. Passing a hand on the virtual screen, the hand changes shape becoming a mouse pointer and allowing an easier interaction with the screen (without the problem of occlusion, since the virtual objects can be superposed to the real hand). To watch a movie from her sofa she chooses to enlarge the screen and put it on the ceiling: she grabs the virtual screen corners and stretches her arms until the ceiling tuning the screen position, size and orientation (showed in Figure 1). After some time, enjoying the comfort of the sofa, she starts reading a book that is a virtual representation of a real, book that is physically placed in another room.

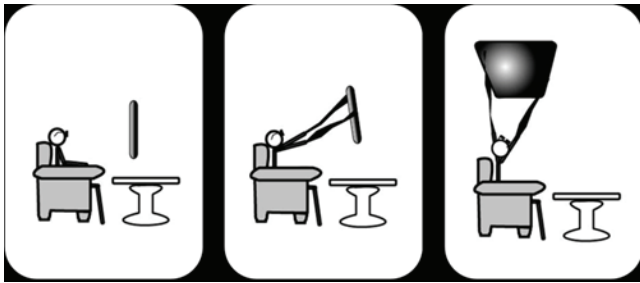


Figure 1. Interaction with a virtual screen

Virtual access to real objects: Alice decides to wear her precious earrings; hence, to open her strongbox, she has to digit the right password in a virtual display that is accessible only in the virtual dimension of the house. In this case the strongbox is a physical object in the real world whereas its command interface is in the digital world.

This scenario highlights the augmented interaction key points and the motivation behind the ARAMIS model. In the next sections it is explained how the presented augmented interactions are modeled in ARAMIS and how ARAMIS handles them.

IV. ARAMIS: BASIC CONCEPT AND BACKGROUND

Gestures are a natural way for a person to interact with the environment, tools and the objects around him [7]. For this reason ARAMIS is focused on gestures as main communication modality. However the ARAMIS approach can be used to model different type of modalities (e.g. voice, touch, etc.) and it can be easily extended toward multimodal approaches.

Since gestures lack in universality (gestures have not the same meaning everywhere) and in order to propose general principles, ARAMIS does not establish a gesture lexicon, but allows detecting different command gestures according to different application and usage contexts.

A. Why a Hybrid Approach

In this section the proposed hybrid approach is presented in its four components highlighting the advantages in the use of these complementary techniques.

Pervasive and Wearable paradigms

Coupling pervasive and wearable paradigms is possible to merge the advantages of both approaches. Table 1 presents the different features form the two approaches as presented in [8] and [9]).

Table 1. Features provided by the pervasive and the wearable approach

FEATURES	PERVASIVE	WEARABLE
Privacy		X
Personalization		X
Consistency		X
Localized information	X	
Localized control	X	
Resources	X	

Pervasive paradigm provides localized information, controls and resources; whereas wearable paradigm provides privacy, personalization and consistency. Using at the same time pervasive and wearable approaches some features, such as personalization and localized control can be combined to improve user interaction with the system. However the mutual improvement between these paradigms is further increased by our approach. For example information can be fitted and displayed for specified users according to their profiles and interaction context (information from the user and the environment); moreover exploiting virtual elements this

information is still private and personalized. Also the resource management will benefit of our hybrid approach. In ARAMIS data is elaborated in the wearable (mobile) system or/and in the environmental (fixed), according to the type of information to handle, the available devices and the quality of the connection. For instance, data that is troublesome to transfer, such as a high frequency data from the accelerometers, can be elaborated locally on the wearable machine and just the processed information (e.g. extracted features) is sent further. On the other hand information that is heavy to elaborate (e.g. images) can be elaborated at environmental side (more details can be found in [8] and [9]).

As stated above, the merging of these paradigms brings the combination of their related strong points. However to make this possible a lot of issues have to be faced (e.g. heterogeneous devices working asynchronously and with different data rates must be able to rapidly exchange information in order to effectively take advantage one from the other).

Mixed Reality

Gestures are performed in a mixed (real and virtual) world. As presented in the scenario, the presence of digital components is the enabling factor to enhance the interaction. It allows avoid physical constraints (object can be suspended in the air, people can stretch their arms, etc.) and defining new ways of interaction. In ARAMIS two entities are augmented by the technology: the human and the environment. A user avatar is used to represent a human being, and especially the body parts involved in the interaction, in a mixed world (e.g. in a system designed for hands and arms interaction only these parts will be augmented). Since the human representation is digital, the avatar is not tied to physical laws and not necessarily something of visible (e.g. the avatar is not an anthropomorphic digital representation of the user, an avatar can be a simple pointer representing the user hand). Hence, the user avatar can change considerably from an application to another since it depends on the gesture vocabulary, the context and the kind of mixed reality implemented for the interaction and the feedback. Finally the avatar has the goal of facilitating and enhancing the interaction.

Human augmentation becomes valuable in the case the surrounding environment provides rich interfaces of interaction and feedback modalities: we called such ambience augmented (or smart) environment. Our approach defines augmented environments as composed by two kinds of elements: objects and places. 1) Objects can be pure real, pure virtual or enhanced. With the adjective enhanced we classify objects that have the double nature: real and virtual. It means that it is possible to interact with them as normal or virtual objects. For example a book, that can be read directly (physically) or displayed on a screen (virtual representation), belongs to this category. Moreover the simultaneous existence of a real object, its digital counterpart and a suitable feedback interface helps the user in the interaction with tools solving part of the problems tied to the use of a ubiquitous, invisible computer [10]. 2) Places are mixed reality entities integrated in the ARAMIS system that describe interactive objects and their state (id, position, activation state, etc.).

Classification techniques

ARAMIS hybrid approach combines different classification techniques in order to implement decision tasks: brute force, machine learning, fuzzy logic, or neuro-fuzzy [11]. According to the interaction to be implemented and the context of use one or a combination of these techniques can be used.

With a brute force approach is possible to interpret simple data in order to have less computation and, therefore, improve performances [12]. Training is not needed and algorithms are generally simple. This improves speed in recognition but brings limitations in the gestures complexity that can be handled; additional complexity must be introduced in order to make the system user adaptive. With gesture of higher complexity, statistical model and machine learning approaches could be used. Analyzing complex gestures is often not possible or too complicated using brute force algorithms. Machine learning (ML) algorithms can be a solution (especially Hidden Markov Model (HMM) and Artificial Neural Networks (ANN) are used in the gesture recognition field) [13]. However ML solutions require often a large quantity of data to train and to learn the system (for example a neural network can be initialized with random values but it will need training sessions to update these values). Moreover to achieve good results, an algorithm structure may require several calibration sessions and user knowledge of the system can speed up this practice. However tuning such systems can take hours or days of computer time because networks may converge on an inappropriate solution or rule or may fail to converge at all [14]. Neural researchers have proposed hundreds of schemes to alleviate this problem, but none has removed it. Even after a lengthy tuning session, the final rules may not be much better than the first set. Moreover “an impediment to a more widespread acceptance of ANN’s is the absence of a capability to explain to the user, in a human-comprehensible form, how the network arrives at a particular decision. Neither can one say something about the knowledge encoded within the black-box” [11].

Facing these issues and aiming to improve the velocity of the computation is possible using a fuzzy logic (FL). Fuzzy sets are considered to be advantageous in handling higher order processing. Although generally faster than ML often FL systems require experience (of the system developer) in order to configure properly the system. In fact the goal of FL is to simulate the behavior of an expert addressing complex problems, instead of analyzing a problem in detail [15].

Rather than relying on an expert to supply a training set of data and correct a network in the process of learning, unsupervised neural networks learn simply by observing an expert's decisions. In this way, an adaptive fuzzy system can learn to spot rule patterns in the incoming data. In other words adaptive systems can help fuzzy systems learning rules and the merits of both neural and fuzzy systems can be integrated in a neuro-fuzzy approach (NFA). In fact NFA are used to realize a fuzzy system through the connectionist networks, with the advantage that is it possible to extract rules without the need of an expert [16].

The following diagram (in Figure 2) schematizes the strong points and the weakness of these four approaches. The parameters taken into account are: speed of the classification,

need of a large training dataset, capability to recognize complex gestures and need of experience to “tune” the algorithm.

The main idea if ARAMIS concept is to combine these methods flexibly, aiming to take advantage of the more suitable solution according the information to analyze.

Optical and Non-Optical sensing technologies

Combining more acquisition technologies brings several advantages. Different sensors, working with different technologies make possible to have a more robust and accurate dataset. So far, several technologies (vision-based, inertial and acoustic) have been considered. Researches highlighted advantages and drawbacks for any technology. Therefore an approach is usually adapted to specific use cases and contexts [17]. For example optical technologies give very useful information but are expensive in term of processing power if more than one sensor is considered. In addition, sensibility to lighting changes is still an open issue in computer vision domain. Otherwise inertial sensors are more resistant to environmental changes, can be easily embedded in small devices, but suffer for systematic and incremental errors, necessitating a recalibration phase for long measures.

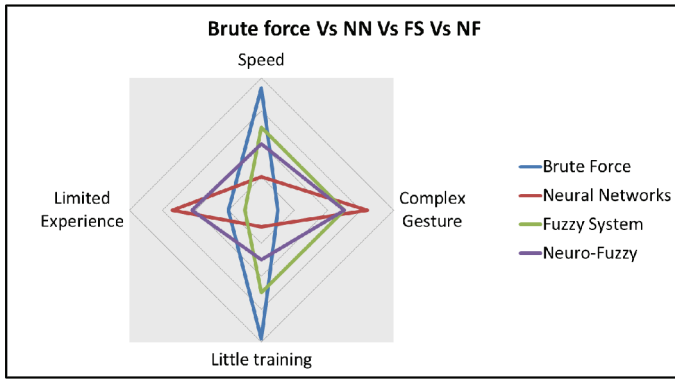


Figure 2. Algorithms comparison

B. Working Configurations

Thanks to the proposed hybrid approach ARAMIS is suitable for three working configurations from totally wearable (augmented human) to totally pervasive (augmented environment):

- Augmented human in augmented environment. This is the standard case of use of our system. Our model provides the maximum of augmentation in the interactions. Information provided at the two levels (pervasive and wearable) are merged and exploited to recognize the interaction and provide the feedback.

- Human (without wearable devices) in augmented environment. A reduced interaction lexicon will be recognized at the environment level (e.g. through fixed cameras), and the feedback is presented back to the user in different ways such as acoustic, or visual output.

- Augmented human in normal environment. Also in this case a reduced vocabulary is recognized and the interaction with the environment takes the form of contextual information

provided by the wearable devices (e.g. approach used in smart-phones for some augmented reality applications).

The major issue in a combined wearable-environmental approach is the data transmission and fusion between the different entities. Data are heterogeneous and asynchronous. Hence, the elaboration can be complex and, above all, too slow to achieve real-time interaction. Therefore, in order to demonstrate the feasibility of our concept a concrete framework has been realized.

V. ARAMIS: THE FRAMEWORK

Figure 3 shows an overview of the ARAMIS framework and how it handles the information and the decision mechanisms. This framework is service-oriented; it means that each component exposes some method as service that can be flexibly used by other components running on the same machine or exploiting a network.

The framework representation in Figure 3 reflects the hybrid approach presented above separating the environmental and the wearable modules.

A. Information handling

Usually, environmental modules have more processing capabilities than wearable modules but the transfer of the information between the different devices can be quite expensive in term of time and bandwidth occupation. In order to optimize the elaboration task the ARAMIS framework allows the information from any sensors to be handled according to three different paths. Raw data that are expensive to be processed (e.g. data from accelerometers needing the extraction of several features) are transmitted to the environmental side and then elaborated (path 1). At the contrary, data that is heavy to transfer (e.g. high quality frames from the camera) has the main features extracted in the module in which is sensed and then sent ahead for further elaboration (such as features fusion according to the path 2). Whereas the information from the worn devices is enough to be analyzed autonomously, gesture can be recognized separately at environmental and wearable side and only the final decision

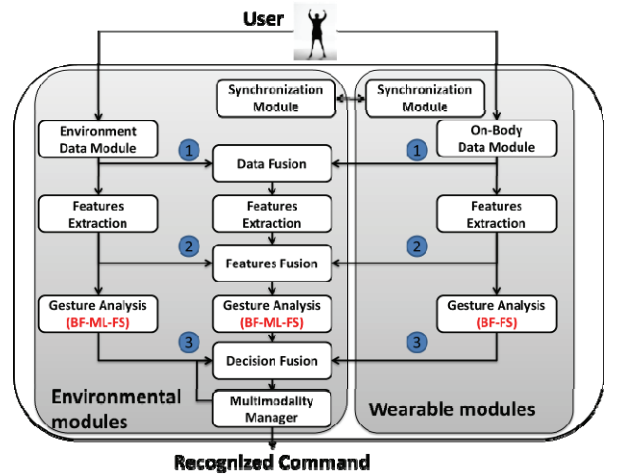


Figure 3. ARAMIS framework – Overview

(path 3) will be pondered and merged (more information about the framework can be bound in the paper [9]).

B. Decision algorithms handling

The gesture analysis modules analyze features in order to classify the input provided by the user. Once again because of the different capabilities between the wearable and pervasive modules it becomes interesting to differentiate and tune the algorithms to be adopted according the type of input and the used devices. At wearable side it will be better to use algorithms fast and requiring small training dataset such as brute force for simple gesture and fuzzy systems to more complex ones. At environmental side all possibilities are opened and also more complex algorithms could be furnished by the framework (Simple Vector Machine, Hidden Markov Model, Neural Networks, Decision Trees, etc.).

VI. ARAMIS: THE IMPLEMENTATION

The previous sections have introduced the ARAMIS approach and how it models the interaction between augmented human beings and the surrounding augmented environments. In this section is presented a first implementation of a system based on the ARAMIS model developed in order to demonstrate its feasibility and benefits. The prototype with the used algorithms, the set-up, and the test of use cases are presented in this section.

A. Gesture lexicon

The gesture lexicon we implemented is composed by several dynamic gestures performed with one arm: goUp, goDown, goLeft, goRight, shake left-right, shake up-down, shake back-forward, rollClockwise, rollCounter-clockwise, rollClockwise-&Return. The rolls gestures have to be performed by the arm wearing the gyroscope in order to be recognized. Some more gesture representing numbers drawn in air by one hand are recognized in special context (see section D). In addition a pointing gesture has been introduced in order to indicate and select the device that is the second term of the interaction. These gestures are been chosen for the sake of simplicity and the minimum effort (cognitive and physic to realize the gesture).

B. Algorithms

Several algorithms have been used in order to face the implementation of a first prototype demonstrating the feasibility of our approach. The complexity of the problem can be resumed in the following main challenges: gesture analysis (features extraction, gesture segmentation, and gestures classification), user recognition and tracking.

Table 2. Algorithms in wearable components

<i>Algorithm</i>	<i>Goal</i>
PTAMM	Pointing
PTAMM	World Creation and user tracking
Optical Flow	Dominant arm gesture recognition
Brute Force	Non-dominant arm gesture recognition

Table 3. Algorithms in environmental components

<i>Algorithm</i>	<i>Goal</i>
Haar cascade	Face recognition -> hand color extraction
Kalman/ CAMShift	Hand tracking
Speed threshold (or OCR-like)	Gesture Classification
Synchronization	Inspired from [18]

The choice of a simple gesture vocabulary allowed us to use simple classifiers based on brute force to perform recognition at wearable and environmental side. However, as stated before, the system can allow more complex classifiers; for example, at present, is equally possible to use an OCR-like classifier in order to recognize gestures performed facing the camera in the environment and that have reference to alphabet characters. In addition we plan to offer a Hidden Markov Model (HMM) classifier in order to analyze the data captured by inertial sensors to analyze more complex gestures. However, our simple gesture vocabulary allows us to use simpler and faster brute force classification algorithms without significant changes in the recognition task. A simple vocabulary is possible since a gesture is interpreted differently according to the target of the interaction; this solution allows us to limit to five the number of gestures to be classified in parallel (reducing considerably the possibility of errors). The presence of the user on the in face of the environmental camera is detected by a face detection algorithm (also used to determine in which part of the image is possible to find skin colored elements, as shown in Figure 7).

The pointing gesture requires special effort since the position of the user and the target must be known. A customized version of PTAMM algorithm (presented by Castel et Al. [19] [20]) proved to be suitable for this goal. Moreover it offered the basis to build virtual maps on a real environment realizing the augmented world we need. In fact we use this method to initialize new environments and introducing the interactive objects in the system. A summary of the used algorithms is shown in Table 2 and Table 3 (classified according their use in the wearable or pervasive paradigm).

C. Set-up

Wearable Set-Up

In the implementation of our system we used the wearable devices as illustrated in Figure 4.

The user body is divided in two main parts: the dominant arm and the non-dominant arm. The first is the arm used for most pointing gestures, whereas the second is used for symbolic gestures and as feedback (haptic stimuli) [21]. As segmentation gesture we chose a double rotation of the non-dominant arm (rollClockwise&Return) that has the feature to be easily recognized by gyroscopes and accelerometers (99% of recognition rate with trained users, with negligible false po-

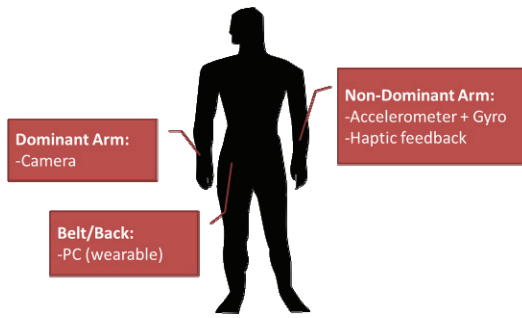


Figure 4. Wearable components set-up

sitives). In the dominant hand (on the upper side of the wrist) is placed a camera. At present the wearable processing part are done on a PC on the user; the use of the only PDA as input, output and processing device is taken into account for the next evolution of the system.

Environmental Set-Up

The augmented environment is realized inside a smart living room. The set-up consists in several pervasive components. A large display (projection on a wall) is used for a visual feedback for the interaction with virtual objects. Using a projection instead of a head mounted display decrease the burden on the user, simplifying the implementation of the scenario. A fixed camera is disposed in front of a couch; this turns the couch into an interactive area in which sensors on the user and in the room can be used at the same time. The enhanced object are a media center PC to play multimedia content, wireless-controlled switches to control some lamps in the room, and processing units hidden in the room performing data elaboration and message routing between elements in the system.

D. Prototype

According the working configuration (section IV.B) we developed three interaction scenarios and we present interaction supported by ARAMIS according to the three working configurations.

Case 1 (on Figure 5): augmented human in augmented environment. The user can point and select the media center and start interacting with the menu displayed on a small LCD (such as select the content of a DVD and control the sound).

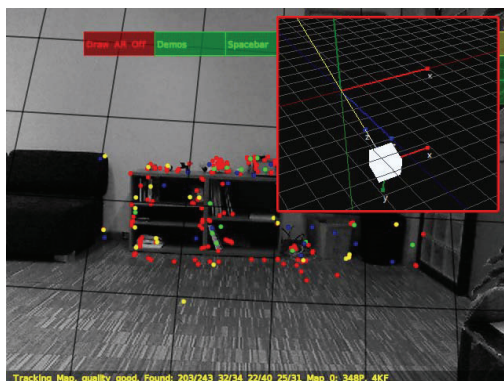


Figure 5. PTAMM information during the word creations. Points of interest (left). Camera orientation (up-right corner).

In addition, complex interactions are allowed, user can grasp a virtual representation of the LCD screen and move it on the wall by pointing at it, now the media center interface will be shown on the wall. With the same gestures the augmented user can point at remote switches and turn on and off the room lighting system. The same gesture is interpreted differently by the system according to the target of the interaction. Whereas the user sits on the couch, a vision-based system for gesture recognition is turned on and the user hands are tracked. Inputs from worn devices and the environment are merged enabling more gesture commands.

Case 2: augmented human in standard environment. The user enters in the environment for a first time. He can include it in the system (creating a new world, as shown in Figure 6) and add virtual objects, which are coupled with the environment only for the position in the space. Due to the lack of feedback (only sound or vibrations), the user should choose conveniently the position of the virtual objects (e.g. a virtual media center on the top of a real media center).

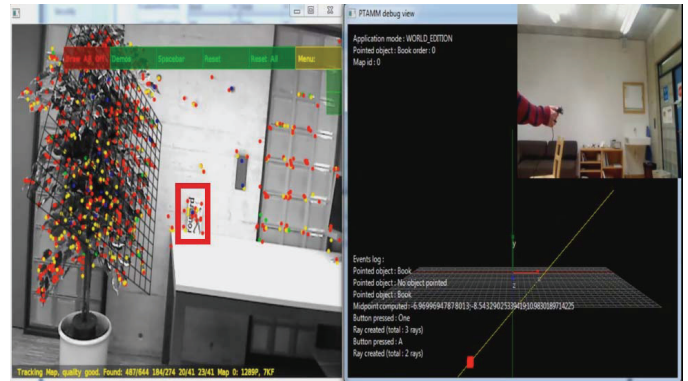
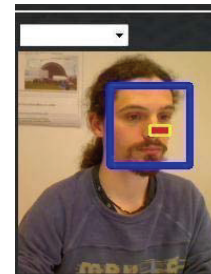


Figure 6. User creating a map. Left: user pointing a book (inside the red rectangle) to add it to the world. Right: ARAMIS perspective, the red cube is the book pointed by the camera.

Case 3 (Figure 7): normal human in augmented environment. In this case the system offers one augmented interaction in a special spot of the room: the couch facing the camera. The interaction is similar to the case 1 with just the gestures from the environment and a simpler gesture vocabulary. Using only a camera an augmented interaction is still possible. Interaction output: only the environment.



a)

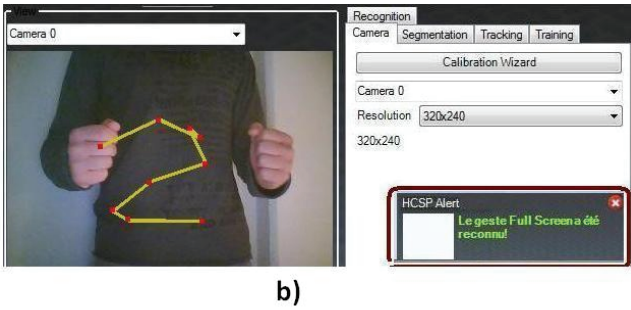


Figure 7. A user facing the camera. a) Haar based face detection and skin color extraction. b) Hand tracking by the environmental camera and with Kalman tracker (working configuration 3)

E. Discussion

The realized scenario implements a part of the hybrid approach described in the section IV. Case 1 (the only having wearable and environmental modules working at the same time), after some speed tests in different configurations, has been implemented according the following principles:

- Wrist worn camera: images are elaborate at wearable side from the features extraction to the gesture analysis (path 3 in Figure 3).
- Worn accelerometers and gyroscopes: raw data is sent at environmental side for features extractions and elaboration (path 1)
- Environmental camera: features are extracted and the gesture analyzed at environmental side.

Data is finally merged at decision level.

It is interesting to highlight that the system is user-location aware: gestures performed in front of the camera are recognized exploiting both the accelerometers and the camera, whereas gestures performed elsewhere (with a reduced lexicon) are recognized only by the worn sensors. For this prototype we have designed a simple gesture vocabulary of ten commands that are interpreted in different ways according to the position of the user in the room and the target of the interaction. The gesture recognition task is made easier by the gesture protocol we used, with five classifiers running in parallel in the worst case (two-three in average).

We have concentrated our efforts to demonstrate the feasibility of the proposed approach. From a technical point of views we evaluate the systems paying attention to real-time issues. We achieved gesture recognition (from features extraction to gesture classification) in less than 250ms. The most of the time is spent by the system sending data (serialization, transfer, and deserialization) between devices on the user and in the environment. It was perceived as acceptable during the tests and it motivates us to proceed with usability tests. However, at this moment, also the proposed gesture language needs more assessments to be deeply evaluated independently of the worn devices (still too cumbersome).

VII. CONCLUSION

In order to achieve novel augmented interactions between human and smart environments, we propose to adopt a hybrid gesture-based approach. Different interaction paradigms and techniques are used at the same time:

- Pervasive and wearable computing paradigms;
- Mixed reality interactive environments;
- Brute force, neuronal, fuzzy or neuro-fuzzy classification approaches;
- Optical and non-optical sensing techniques.

To validate our concept of hybrid approach we realized a framework allowing the system to flexibly handle the information. Data from each sensor can be elaborated in the more suitable way at wearable or environmental side.

We have implemented key elements of the ARAMIS approach demonstrating how to augment the interactions of a user in three different working configurations. The prototype, used to perform feasibility evaluations has demonstrated the system can work in real-time modeling augmented interactions.

In the first prototype we implemented brute force algorithms for simple gesture recognitions, as next step we plan to extend our prototype including a fuzzy approach in order to recognize more complex gestures in real-time. Finally a quantitative evaluation comparing the different solutions (brute force, machine learning and fuzzy approaches) will be done to complete the formalization of the presented interaction.

REFERENCES

- [1] OpenNi, <http://www.openni.org/>.
- [2] Dünser, A., Grasset, R., Seichter, H., Billinghurst, M. (2007) "Applying HCI principles to AR systems design" Charlotte, NC, USA: Mixed Reality User Interfaces: Specification, Authoring, Adaptation (MRUI'07: 2nd International Workshop at the IEEE Virtual Reality 2007 Conference), 11 Mar 2007. Mixed Reality User Interfaces: Specification, Authoring, Adaptation (MRUI'07) Workshop Proceedings, 37-42.
- [3] D. C. Engelbart, "Augmenting Human Intellect: A conceptual Framework," Summary Report, Contract AF 49(638)-1024, SRI Project 357B, Stanford Research Institute, Menlo Park, California (October 1962), AD 289565.
- [4] Hoc, J.-M. "From human-machine interaction to human-machine cooperation". *Ergonomics*, 43(7), 2000, pp. 833-843.
- [5] Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M. "Proxemic interactions: the new ubicomp?" *Interactions* 18, 1 (January 2011), 42-50.
- [6] Pederson, T. "Towards an activity-aware wearable computing platform based on an egocentric interaction model". UCS 07 Proceedings of the 4th international conference on Ubiquitous computing systems. 211-227 (2007).
- [7] Cook, D.J., Das, S.K. "How smart are our environments? An updated look at the state of the art". *Pervasive and Mobile Computing*. 3, 53-73 (2007).
- [8] Rhodes, B.J.; Minar, N.; Weaver, J.; "Wearable computing meets ubiquitous computing: reaping the best of both worlds", *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*, vol., no., pp.141-149, 1999
- [9] Carrino, S., Mugellini, E., Abou Khaled, O., Ingold, R. "ARAMIS: Toward a Hybrid Approach for Human-Environment Interaction" 14th International Conference on Human-Computer Interaction, HCI International 2011, Orlando, FL, USA, 09-14 July 11.

- [10] Rehman, K., Stajano, F., Coulouris, G. "Interfacing with the invisible computer". Proceedings of the second Nordic conference on Human-computer interaction - NordiCHI '02. 213 (2002).
- [11] Mitra, S., Acharya, T. "Gesture Recognition: A Survey". IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews). 37, 311-324 (2007).
- [12] Wobbrock, J.O., Wilson, A.D., Li, Y. "Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes". Proceedings of the 20th annual ACM symposium on User interface software and technology - UIST '07. 159 (2007).
- [13] Elmezain, M., Al-hamadi, A. "A Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition". Engineering and Technology. 31, 394-401 (2008).
- [14] Yuan, X., Wang, Y. "SVM Based Adaptive Inverse Controller for Excitation Control". Advances in Neural Networks-ISNN 2007. 469-478 (2007).
- [15] Nauck, D., Klawonn, F., Kruse, R. "Foundations of neuro-fuzzy systems". John Wiley & Sons, Inc. New York, NY, USA (1997).
- [16] Mitra, S., Hayashi, Y. "Neuro-fuzzy rule generation: survey in soft computing framework". IEEE transactions on neural networks a publication of the IEEE Neural Networks Council. 11, 748-68 (2000).
- [17] Crevoisier, A., Polotti, P. "Tangible acoustic interfaces and their applications for the design of new musical instruments". Proceedings of the 2005 conference on New interfaces for musical expression. p. 97-100 National University of Singapore (2005).
- [18] Castle, R.O., Klein, G., Murray, D.W. "Combining monoSLAM with object recognition for scene augmentation using a wearable camera". Image and Vision Computing. 28, 1548-1556 (2010).
- [19] Castle, R.O., Murray, D.W. "Object recognition and localization while tracking and mapping". 2009 8th IEEE International Symposium on Mixed and Augmented Reality. 179-180 (2009).
- [20] Römer, K. "Time synchronization in ad hoc networks". In Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing (MobiHoc '01). ACM, New York, NY, USA, 173-182.
- [21] Karam, M. "A framework for research and design of gesture-based human computer interactions". Thesis. Southampton: University of Southampton, (2006)