Towards an Activity-Aware Wearable Computing Platform Based on an Egocentric Interaction Model

Thomas Pederson and Dipak Surie

Dept. of Computing Science, Umeå university, SE90187 Umeå, Sweden
{top,dipak}@cs.umu.se

Abstract. In this paper, we present our egocentric interaction model for recognizing and supporting everyday human activities. We explain how it allows designers of ubiquitous computing systems to view physical (real) and virtual (digital) objects as residing in one single space and how sets of objects in the vicinity of a specific human actor can be classified based on human perceptual characteristics such as what can be observed and what can be manipulated. We also propose a wearable computer architecture that is based on the egocentric interaction model which potentially could facilitate the development of Ubiquitous Computing applications by letting an operating system take care of maintaining communication with worn and instrumented sensors as well as computing devices. Finally, we present our first steps in implementing an activity-aware wearable support system for people suffering mild dementia based on the proposed model and architecture.

Keywords: Human-Computer Interaction, Ubiquitous Computing, Wearable Computing, Context Awareness, Virtual Reality.

1 Introduction

Our current user interfaces for personal computing make us less mobile than what state-of-the-art computing hardware would otherwise allow for. Miniaturization and power efficiency improvements of portable personal computers (PCs) as well as the diffusion of wireless networks has brought us to a state where all the computing power and connectivity most of us would need for a single day easily could fit into a small backpack. However, the way we have chosen to provide access to that computing power, through the WIMP[1] interaction paradigm, constraints the kind of individual human activities for which it can offer computational support. Limiting characteristics include

- the assumption that the human actor can dedicate all attention to the interaction with the virtual environment provided by the computer (e.g. does not bike or drive a car)
- the assumption that the real world environment in which the interaction takes place is always the same (quiet, in the shadow, etc.)

[1] WIMP — Windows, Icons, Menus, and Pointing device.

H. Ichikawa et al. (Eds.): UCS 2007, LNCS 4836, pp. 211–227, 2007.

- the assumption that input and output devices are few, and the same, at all times (i.e. screen, keyboard, and mouse)

As acknowledged widely by researchers in proactive HCI areas like Augmented/ Mixed Reality (e.g. Mackay et al. [11]); Ubiquitous/Pervasive Computing (e.g. Weiser [25]); Graspable/Tangible User Interfaces (e.g. Ishii & Ullmer [9]); Wearable Computing (e.g. Starner [21]); and Context Awareness (e.g. Dey et al. [5]), these assumptions do in general not hold, or severely limit interaction possibilities, in mobile and real-world dependent activity scenarios.

Thus, in an era where many real-world activities have become more or less PC-dependent, personal activities that for instance demand large amounts of attention to events in the real world and/or the ones better performed while on the move cannot typically make use of interactive personal computing power.[2] Wearable Computer research has tried to break free from some of the constraints in part by coming up with one-handed keyboards, head-up displays, and voice-based interaction but with little impact on the general public so far. Cellular phones/PDAs are becoming increasingly powerful but do not in any significant way extend the applicability of personal computing power due to the fact that the interaction paradigm by and large is inherited from the world of more full-grown PCs, apart from the improved portability of the hardware itself[3] To summarize: the currently ruling interaction paradigm limits the scope of applications for personal computing. It is obvious that new alternative or complementing interaction paradigms would significantly expand the application area for personal computing systems by providing support for new activities and therefore also previously neglected user groups.

In this paper we introduce our first steps towards an egocentric interaction paradigm and show how it has helped structure and inspire our efforts in developing future applications for personal computing involving activities in the real world. The term “egocentric” has been chosen to signal that it is the human body and mind of a specific human individual that (sometimes literally, as will be shown later) acts as centre of reference to which all interaction modeling and activity support is anchored. The model is based on two conceptual corner stones: 1) a physical-virtual design perspec-tive to enable a coherent view of domain objects residing in both physical and virtual environments, 2) a situative model of physical-virtual space for framing human object-centric activities in time and space.

The second half of the paper is devoted to the description of an envisioned wearable computer equipped with an operating system that offers modeling of basic activities in both the real and digital realms. The main motivation is that by embedding activity sensing and modeling into the operating system, developers of Ubiquitous Computing applications could be relieved from solving technological design problems that (today) frequently occur with every effort to design new task-specific UBICOMP systems. Finally, our own limited first steps towards such a wearable

---

[2] That is of course not to say that all personal activities would benefit from it.

[3] In fact, an often stressed selling-point for the most advanced cellular phones currently on the market is their ability to handle the same office-related data files as “real” personal computers.But, in other than exceptional cases, is it really an appropriate task for such small devices?

computer is presented in the light of developing a "cognitive prosthesis" for people suffering dementia, aiming at supporing patients with Activities of Daily Living.

2 Egocentric Interaction Model

2.1 A Physical-Virtual Design Perspective

Personal computing devices such as PCs, PDAs, and cellular phones play an increasingly important role in human activity. However, there are still things that most people would prefer to do "off-screen" in the physical (real) world, such as having parties, reading long text documents, or spending vacation. We argue that there exists a class of activities that are neither physical or virtual, but "physical-virtual" [15]. People frequently do parts of an activity in the physical world (e.g. proof-reading a text document under construction) and parts in the virtual world (e.g. adjusting paragraphs within "the same" document in a word processing environment). This behaviour is likely to become more common. Hence, future environments should be designed with such physical-virtual activities in mind.

The Physical World, Virtual World, and Human Activities. We have chosen to rely on the following definitions in order to model simple human activity across the physical-virtual gap:

- The physical world is the world built of and containing matter directly perceptible to humans, and whose state is defined by arrangements of such matter in places, constrained by and modified according to laws of nature, within a geometrical three-dimensional space, at any time instant partially perceptible by humans through their senses. [15]
- The virtual world is the world built of and containing digital matter (bits) that after transformation into physical phenomena becomes perceptible to humans, and whose state is defined by arrangements of such phenomena in places, constrained by and modified according to (human-designed) laws of logic, within a topological multidimensional space, at any time instant partially perceptible by humans through displays (possibly multi-modal and audio-visually up to three-dimensional) built into computational devices residing in the physical world. [15]
- Human activities are described using three levels of abstraction: activity, action and operation inspired by activity theory [12]. An activity has an objective and is comprised by a set of actions that have well-defined goals and are accomplished by largely unconscious operations.

From Input and Output Devices to Plain Direct Manipulation. The proposed physical-virtual perspective is a way to deal with the gap between the physical and the virtual world, and to facilitate the exploration of designing infrastructure for helping human actors bridging it. The assumption is that a reduced physical-virtual gap means less "friction" for physical-virtual activities, i.e. activities that involve frequent switching between physical actions (Fig. 1) and virtual actions (Fig. 2) [15].

The physical-virtual design perspective differs from more classical HCI models by explicitly ignoring input and output devices of interactive computers such as PCs,

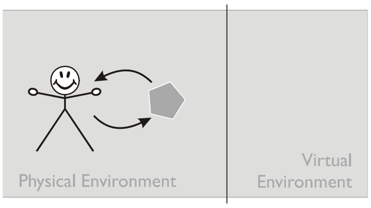

Fig. 1. Physical action: Human actor interacting directly with an object

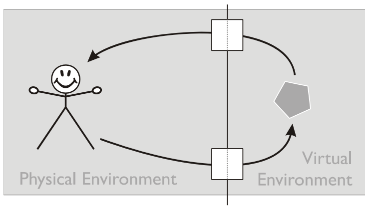

Fig. 2. Virtual action: Human actor interacting indirectly with an object through input and output devices (the white squares) creating the illusion of direct manipulation [20]

PDAs and cellular phones, seeing them as completely transparent mediators for accessing virtual objects. In practice, this means to make no difference between the situation shown in Fig. 1 and the situation shown in Fig. 2. Taking such a stance permits the modeling of real-world and digital entities as if they were situated in the same Euclidean space, which we believe is advantageous when modeling mobile everyday computing applications where the interaction complexity wastly surpasses what can be sufficently described using a classical human-computer interaction dialogue model. Furthermore, such a view aligns well with findings in psychology indicating that expert users of tools (whether being a tennis racket or a computing device) tend to focus on the domain object they are working on (whether it is a tennis ball or an email) and forget about the details of the tool (such as the handle of the tennisracket or the workings of particular input device) itself.

2.2 A Situative Model of Physical-Virtual Space

The situative model of physical-virtual space is intended to capture what a specific human actor can see and not see, reach and not reach at any given moment in time (Fig. 3). The model is based on the physical-virtual design perspective briefly outlined in the previous section, meaning that physical and virtual domain objects are treated as being located in the same space. The model is for the emerging egocentric interaction paradigm what the virtual desktop is for the PC/WIMP interaction paradigm: more or less everything of interest to a specific human actor is assumed to, and supposed to, happen here. Although spatial and/or topological relationships between objects within a particular space (e.g. the manipulable physical-virtual space) certainly could be of interest, we have in our work so far only cared for whether an object is present in a space or not. Applying the model in this simple way generates a set of objects for each space at any given time instant.[4]

[4] The content of the innermost ellipsis in Fig. 3, "object manipulation", is actually not just a representation of spatial properties between the human body and a set of objects. Instead, it also provides information about the state changes to objects currently manipulated by the human actor. For reasons of simplicity we are nevertheless treating this information entity as a space unless noted otherwise.

An Example of Applying the Model to an Actor's Situation. If a glass of juice is in the right hand of a specific human actor, and an email just brought forward on the cellular phone held up in front of the same human actor's face by the left hand (see Fig. 5 left for such a situation captured in a Virtual Reality environment), both objects would be considered to reside in the object manipulation space in Fig. 3. A paper newspaper on the table just in front, and the keys in the same persons pocket would instead be modeled as inside the manipulable space. A painting on the opposite side of the table (but not the one behind the actor's back) would be in the observable space. Finally, all technically perceivable objects in the physical-virtual world which at least for the moment not happen to be perceivable by the specific human actor are regarded as situated in the world space, outside the spaces mentioned earlier.

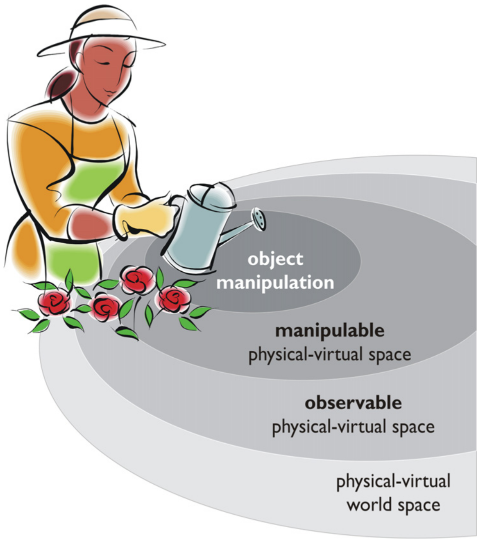

Fig. 3. A situative physical virtual space model adapted from [15]

In the physical world, the border between the observable and manipulable spaces can be approximated and described in Euclidean terms: manipulable things are typically closer than things only observable and the border is somewhere in-between. This spatial relationship is reflected in Fig. 3. Determining a corresponding border in the virtual world is somewhat more complex and depends on situative access to input and output devices. Due to the nature of the application area towards which our current system development efforts is targeted (see next section of this paper), we have chosen to temporarily suspend the work on investigating how object manipulation and navigation should be best modeled in virtual environments (e.g. WIMP-kind) to fit into the situative space model shown in Fig. 3. However, experiences from a first attempt [15] has convinced us that it should be possible.

Using the Model for Guiding Ubiquitous Computing Application Design. We consider the borders of the observable space to define the set of objects that can possibly be part of a physical-virtual "application"[5]at any given time-instant for a

[5] Clarification: With "application" we mean in this context a predefined environment for supporting specific tasks like the kind of interactive software programs executed and used by end-users in virtual environments within the PC/WIMP paradigm, with the significant addition that "applications" in the egocentric interaction paradigm are imagined to potentially span the physical-virtual gap, encompassing objects residing in both the physical and the virtual world.

specific human actor. If a computing system displays information outside the observable space, it will not be noticed. If access to a desired virtual object is provided through a computing device currently outside of the manipulable space, the human actor is forced to change physical location. As it happens, this view aligns well with the WIMP/direct manipulation paradigm for virtual-world [20] interaction where successful application design as well as use very much depends on keeping the right objects "on screen" at the right time. Drawing from experience in using that interaction paradigm, and the human everyday strategy to arrange objects in the real world into places, it is probably safe to generalise and recommend to designers of Ubiquitous Computing systems to make sure that their applications do not enforce or provoke frequent fundamental changes to the observable and manipulable spaces.

3 An Activity-Aware Wearable Computing Platform

The computer platform we envision has at least two unique characteristics: 1) it complements rather than attempts to replace the existing flora of computing devices such as PCs, cellulars or PDAs, by supporting the higher-level activity in which the interaction with those other devices takes place, 2) it incorporates real-world objects into the pool of resources which it monitors and manages.

3.1 System Architecture

Being a relatively abstract perspective on future Human-Computer Interaction, or rather Human-World Interaction, the egocentric interaction model does not assume any particular implementation approach. For instance, computing and sensing technology for tracking physical activity of a specific human actor could be imagined to be either worn by the human actor herself, or be situated in the surrounding physical environment. The same goes for any computing device that provides access to the virtual world which could be both worn by their owner or ubiquitously distributed throughout the physical environment like in Mark Weiser's vision [25].

For reasons of privacy, efficiency, design complexity, feasibility, and cost, we have found an implementation approach based on wearable sensing and computing power most attractive. The basic idea is to make the wearable system as self-sufficient as possible, reducing the Pervasive Computing problem of "uneven conditioning" [18]. Information about activities performed using devices external to the wearable computing system (e.g. in the case when the human actor is using a desktop PC) need to be transmitted through some standard wireless communication protocol to the wearable computer for analysis. Complemented with the real-world object manipulation information from wearable sensors (although, depending on the application the wearable sensors probably need to be complemented with sensors instrumented in the environment), the egocentric interaction system would (at least in theory) be able to gracefully model activities across the physical-virtual gap.

The system architecture (see Fig. 4) is made up of a wearable computer offering computing power and storage space for activity-related data generated by an egocentric interaction sensor pool monitoring object-centric phenomena within the observable physical-virtual space of a specific human actor. Furthermore, the wearable computer runs an activity-aware operating system hosting both advanced physical-virtual applications developed by software developers as well as simpler programs designed by the user her/himself. Applications can incorporate the manipulation of both physical objects (e.g. an interactive item in a theme park) and virtual objects (e.g. a web page describing the same item). Explicit interaction with the activity-aware operating system (and the applications running on top of it) can be performed through user interface clients running on general-purpose devices like PCs or cellular phones. Inspired by Intel's Personal Server concept [24], the wearable computer does not provide any means for interaction itself. Implicit interaction [19] with the activity-aware operating system emerges whenever the user interacts with a physical or virtual object inside the manipulable physical-virtual space (see Fig. 5) monitored by the egocentric interaction sensor pool. We present the architecture in more detail below.

The Activity-Aware Operating System

A software system running on the wearable computer which receives and preprocesses detailed data about ongoing human activity from the egocentric interaction sensor pool, and makes the information available to high-level software applications executing on top of the operating system. The activity aware operating system also provides support for these applications whenever they need to interact with the human actor in a more explicit manner by taking the responsibilty of identifying possible input and output channels.

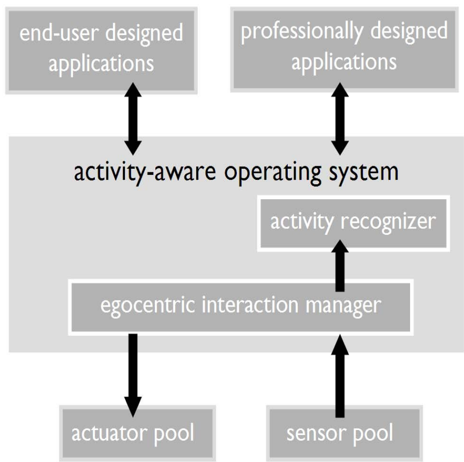

Fig. 4. Conceptual architual of the activity-aware computing platform

In contrast to how end-user interaction in virtual environments offered by current PC operating systems tend to be centered around (virtual) file management and tools for manipulating these files at various levels of abstraction, we imagine the proposed operating system to be centered around everyday objects residing anywhere in the physical-virtual continuum: From purely physical (such as a coffee cup) over physical-virtual (e.g. a specific newspaper manifested on paper as well as online) to purely virtual (e.g. an email). The operating system would represent all known objects as Physical-Virtual Artefacts (PVA)[6] no matter where in the continuum they happen to be situated. We believe this generalisation from "file" to PVA to paradoxally both relax and give structure to the design space for UBICOMP application developers. The relaxation comes from the relatively easy incorporation of both physical and virtual events into the applications

[6] A Physical-Virtual Artefact is, in short, an object having at least one representation in the physical world and at least one representation in the virtual world [15].

under development, the structure comes from a standardized view on a range of everyday physical and virtual events that by necessity has to be defined and agreed on among object and device manufacturers in order for the activity-aware operating system to handle the information.

Why include a real-world object model as low down as on the OS level? Because the world is not made up of “files” and “folders”. If application designers are to design Ubiquitous Computing applications for human everyday activity support, it is best to embed mechanisms for linking physical phenomena to virtual as deep down as possible in the environment in which applications are developed. In this paper we are focusing on two components that we feel to have general applicability: a situation modelling component (the egocentric interaction manager) and an activity recognition component (the activity recognizer). Other more specific components might be needed depending on the specific application area UBICOMP developers choose as target.

The Egocentric Interaction Sensor Pool. The main purpose of this system component is to capture objects’ state changes within the situative space model (Fig. 3) and transmit information about them to the egocentric interaction manager. The sensor pool consists of a configuration of both sensors for physical-world phenomena and sensors for virtual-world phenomena. The physical-world sensors are made up of a fix body-worn sensor configuration potentially complemented with a more dynamically available sensor set instrumented in the environment, changing as the human actor changes physical location. The virtual-world sensors are in fact software clients assumed to be installed on interactive computing devices and which wirelessly informs the egocentric interaction manager about object manipulation performed by the human actor on virtual objects which the specific device is making available. For example, information about text editing operations performed by the specific individual on a PC are expected to be transmitted by that PC to the activity-aware OS running on the wearable computer worn by the same individual.

The Egocentric Interaction Actuator Pool. The purpose of this component is to enable the applications running on top of the activity-aware operating system to affect the state of objects in the world space. They do this through the egocentric interaction manager, ensuring that the changes are perceived by the human actor. In analogy with the sensor pool, the actuator pool consists of both physical actuators (e.g. digital display, a light switch, or a door lock) and virtual actuators. Just like physical actuators change the state of physical objects, virtual actuators change the state of virtual objects and both kinds are treated in the same way by the egocentric interaction manager (discussed in the next paragraph). Virtual actuators are provided by client deamon processes running on external interactive computing devices, able to pop up dialogue boxes on the device’s display, but also to perform more detailed actuation if application software running on the devices allow it. Furthermore, just as in the case of the sensor pool, the actuator pool is imagined to consist of a combination of both wearable and instrumented actuators.

The Egocentric Interaction Manager. This component channels and directs the major part of the communication between the human actor, the activity-aware operating system, and the surrounding physical-virtual environment. By maintaining an up-to-date high-level representation of object manipulation, the manipulable space, and the observable space (Fig. 3), it simplifies the work for other operating system components (such as the activity recognizer) as well as for applications. For instance, if an application needs to urgently inform the user about something, it can rely on the egocentric interaction manager to ensure that the information will be technically perceivable by the human actor. In this situation, the egocentric interaction manager will attempt to determine the orientation of 2D displays (belonging to the current actuator pool) in relation to the human actor's visual perception. Virtual objects shown on such a display should only be regarded as being within the observable space if the viewing angle of the display permits the objects to be perceived by the human actor. This is a special case of the tricky problem of detecting visual occlusion where objects obstruct other objects. We are convinced that the orientation of 2D displays in relation to a specific human actor's gaze is measurable using existing sensing technology although such technology is not yet widely available and used on the market today.

The Activity Recognizer. Automatic activity recognition is receiving a growing interest in Ubiquitous Computing and related research communities. Its wide applicability has motivated us to include such a mechanism in the proposed operating system. As a specific human actor changes physical and/or virtual location, objects come into and leave the observable physical-virtual space in a dynamic fashion. Inspired by situated action theory [22] and the proximity principle: "Things that matter are close. Things that are close matter." [10], we believe that the content of the observable and manipulable spaces of the model indicates the specific human actor's intent with the currently ongoing activity. The purpose of the activity recognizer is to continuously determine the ongoing activity, to store details about this activity, and to provide this information to other operating system components as well as applications on request.

Professionally Designed Applications and End-User Applications. Just like any other operating system, the activity-aware operating system will not provide any added value to a person's activities without applications running on top of it. We can only speculate in what kind of applications that could be developed using the activity-aware operating system as base. As explained in more detail in the next section, we are currently developing an everyday activity support application for people suffering mild dementia. Furthermore, just like the UNIX operating system allows skilled users (system administrators) to easily automate tasks within the domain of file processing by offering automation mechanisms at the right level of abstraction, we believe that an operating system monitoring human everyday activities on a high level could offer good support for end-users (the human actors using the system) to automate their everyday routine tasks, or at least the actions within an activity which the egocentric interaction actuator pool can affect. The paradigm of 'programming by demonstration' seems suitable due to its low learning threshold.

4 Experiences from the Development of a Cognitive Prosthesis for People Suffering Mild Dementia

We have applied the concepts discussed earlier when building a wearable "cognitive prosthesis" for supporting Activities of Daily Living. Such activities include getting dressed, preparing breakfast and activities related to personal hygiene. Dementia patients have problems in remembering to perform their everyday activities, or actions within those activities. Other problems include not being able to get started in the first place; not being able to continue after having been interrupted; or missing some operations that are mandatory for the completion of an activity. A system that

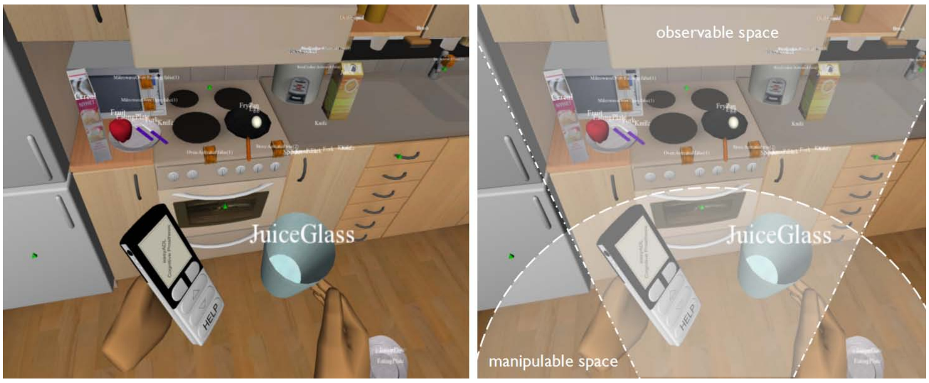

Fig. 5. A snapshot of the Virtual Reality home environment as seen by the user (left) and as seen by the egocentric interaction manager (right)

could help overcome the above mentioned problems would enable patients to stay in their home for a longer period of time, have a normal independent life, and also reduce the burden on family members and caregivers. Our long-term goal is to build a dementia tolerant home environment using wearable and ubiquitous computing technologies. By favourising the wearable approach, we hope to also open for the possibility of a future version of the system which also works outside the home. We are using a Virtual Reality environment as a test-bed in developing the cognitive prosthesis in order to 1) speed up the design (and re-design) process and 2) to compensate for the limitations with the currently available sensing technologies. A Virtual Reality model using the Colosseum3D real-time physics platform [2] is used to simulate a physical home environment augmented with an egocentric interaction sensor pool mainly consisting of wearable sensors. Fig. 5 (left) shows a snapshot of the environment. Based on the performance of the system in the VR environment presented in more detail elsewhere [23], we have also started to develop a hardware prototype in parallel with continuing the development of the cognitive prosthesis in the VR environment.

4.1 Operationalization of the Situative Space Model in Fig. 3

Observable Space (OS). It is the set of objects within a cone in front of the human actor's eyes with this cone following the head movements as shown in Fig. 6 (left). The height of the cone is limited by the walls of the indoor environment and visual occlusion is considered in determining the set of objects within this space.

Manipulable Space (MS). It is the set of objects within a hemisphere in front of the human actor's chest as shown in Fig. 6 (right). Such a shape is motivated by the fact that humans have two hands and the assumption that they manipulate objects within reach of their hands. The hemisphere follows the human actor's chest movements and the origin point represents the centre of the manipulable space hemisphere. The radius of this hemisphere is equal to the maximum distance between the origin point and a hand.

Object Manipulation (OM). When objects are manipulated by an actor, two events can be generated: grabbed event or released event. Both events include information about the object manipulated by the actor. The actor can manipulate objects with both hands. We do currently not make any distinction between the right hand and the left hand since our objective so far has been to only know what object the actor is currently manipulating and its state changes.

4.2 Implementation of the Activity-Aware Platform

To implement the complete activity-aware computing platform as shown in Fig. 4 including the activity-aware operating system is beyond our current development resources. So far, using the VR environment described earlier, we have chosen to focus our attention to components and parts of components that are of most immediate use for our aim in developing the cognitive prosthesis: the egocentric interaction sensor pool, the egocentric interaction manager, the activity recognizer, and a "professionally designed application": an ADL support application.

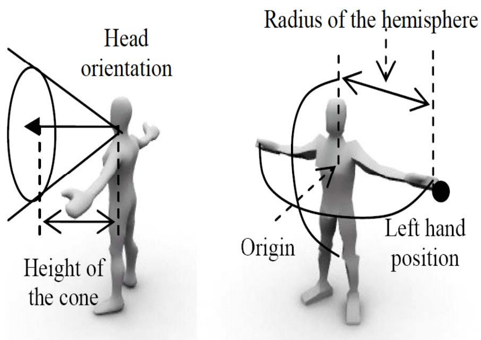

Fig. 6. Observable space (left) and manipulable space (right) as defined in the VR environment

The Egocentric Interaction Sensor Pool. The currently implemented sensor pool captures the identity of physical objects in OS, MS, and events generated by OM. It also captures simple button-pressing on a cellular phone although not in accordance with the physical-virtual design perspective since these operations are treated as being physical and the simulated cellular phone is not transmitting the manipulation of the virtual objects it shows on the display to the egocentric interaction manager. We are planning to make the sensing of interaction with the cellular phone work in accordance with the physical-virtual design perspective (e.g. that objects shown on its display would be treated similarly to real-world objects) and also to model the cellular phone display as part of an egocentric interaction actuator pool. We are also currently expanding the wearable sensor pool with a selection of instrumented physical sensors in order to improve the possibilities for providing detailed activity support.

The Egocentric Interaction Manager. The current egocentric interaction manager feeds the sensor data it receives from the sensor pool directly to the activity recognizer. We plan to modularize these components better and create a minimalistic activity-aware operating system at the same time. Currently, the egocentric interaction manager handles only input from the sensor pool. As soon as the display of the cellular phone is modeled as an actuator, the egocentric interaction manager will be enhanced with actuation control, allowing higher-level applications such as our cognitive prosthesis to display visual messages without bothering about what display device that could be the most suitable in a given situation – a decision then made by the egocentric interaction manager on the basis of the situative space model (Fig. 3) and the availability of visual actuators (display devices) in the observable space for the specific human actor.

The Activity Recognizer. Our current activity recognizer is based on a combination of Growing Neural Gas [7] and Hidden-Markov Models [17] for each activity directly mapped onto the situative space model (Fig. 3), monitoring object manipulation, the manipulable space, and the observable space. The presence patterns of objects in the different spaces are compared to previously trained patterns for the specific individual, resulting in an average recognition precision of 89% at 98% recall ability on the action level (e.g. "setting the table for lunch"; "preparing coffee"; "frying vegetables"; "doing the dishes") among 10 trained activities within a VR-simulated kitchen environment. The precision and recall values on the action level (higher-level of detail, lower level of abstraction) which basically tries to determine what action within the same set of activities the human actor is performing (e.g. actions such as "to place the forks on the table"; "to fill the coffee maker with water"; "to turn on the stove"; "to place the plates in the sink") gives a precision of 76% and recall of 96%. The lower figures for high-detail activity recognition aligns well with findings in the literature. One problem is the high amount of general-purpose objects in kitchen environments, making object-use-based activity recognition particularly hard. We have several ideas for how the numbers can be improved and are also convinced that other environments containing a higher proportion of specialized tools and everyday objects would provide better recognition results. For more information on the activity recognizer and the evaluation of it we refer to [23].

The ADL Support Application. In order to learn about our user group, we have arranged over 10 brainstorming and interview sessions with professional dementia caregivers at the Geriatric Medicine Department at the University Hospital of

Northern Norrland in Sweden. One outcome is the fact that dementia affects patients in different ways, implying the need for individually tuned support. We have addressed this issue in part by training the activity recognizer on an individual basis (meaning in the end that the cognitive prosthesis becomes a very personal kind of tool), in part by identifying the kind of support which a majority of the patients could benefit from. The ADL support application we are currently developing intervenes when the human actor 1) has omitted to change the state of an object considered mandatory for the activity currently performed, or 2) has omitted to perform a mandatory action. An object's or action's level of "mandatoriness" for a given activity is determined based on the training data stored in the activity recognizer, resulting in a probability value. This probability value is generated automatically using unsupervised learning algorithms.

The overall design philosophy is that the ADL support application should remain completely silent as long as activities are performed as during the training phase with a certain tolerance for changes in sequence and omission of non-mandatory events. The interface modality of the intervention is decided by the ADL support application based on the current configuration of the actuator pool (e.g. what displays that are present in the observable space in Fig. 4) and based on the particual individual's abilities to perceive information in a certain modality.

Assessing the Model and the VR Design Approach

We have also conducted ethnographical studies in five households to further investigate details of everyday object manipulation and to learn about real world everyday activity. Based on a wearable video camera system capturing the environment in the direction of the subject's head (see Fig. 7), we have been able to study the dynamics of everyday object use withing the observable and manipulable spaces in the real world.

One result from this study is the fact that real world activities and activities performed in our VR environment differ with respect to the detail in which activities are performed [3] helping us to determine to what extent the results in the VR world is transferable to the real world.

Fig. 7. A snapshot of the observable space for a human actor doing the dishes in the real world, as approximated by the area covered by a video camera lens worn on the head [3]

4.3 Current System Architecture

Although most of our system development efforts have been directed towards getting an activity-aware platform to work in the VR environment, we have accounted for porting the same system to a real-world platform. Fig. 8 shows our current architecture that in principle allows us to change between running the activity support

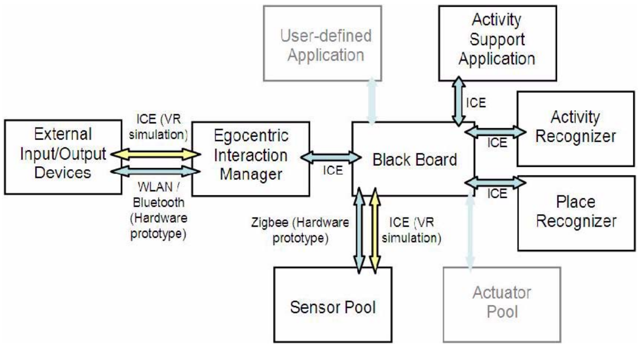

Fig. 8. The current architecture of the wearable activity-aware cognitive prosthesis. “User-defined applications” and the “actuator pool” are left for future work. The “external input and output devices” (including various kinds of visual and auditory display clients; keyboard clients, etc.) is under development.

system in the VR environment and running it in the real world by simultaneously exchanging a) the VR sensor pool with a real-world sensor pool, and b) the input and output devices simulated in the VR environment with real-world input and output devices.

5 Related Work

The idea of using computers for assisting individual human actors in everyday life is not new but has gotten increased relevance in the last 10 years because of increased capacity of mobile and wearable devices. One example is the research performed at Georgia Tech investigating the possibilities in creating an always present, context-aware “digital assistant” [21]. The egocentric view differs from their and most other similar “intelligent assistant” approaches, by focusing on detecting presence of physical (and virtual) objects rather than places or persons, for detecting and contextualizing human activity. The approaches are, of course, complementary. However, as mentioned earlier, by taking a world-neutral physical-virtual design stance, the egocentric view on interaction differs from most existing HCI modelling approaches by not seeing the state of the real world as merely context to virtual activity but an inseparable part of it. Aura [8] and Christensen and Bardram [4] have explored an activity-centered computing paradigm. Our work defers from theirs by attempting to provide support for physical-virtual activities. Philipose et al. [16] and Patterson et al. [13] explore activity recognition based on RFID-tracked object manipulation. Using the situative space model, we extend this approach by also considering the changes to the set of objects seen and to the set of objects reachable by an actor. Dey [5] proposes a framework for developing Ubiquitous Computing applications inspired by the PC/WIMP interaction paradigm and location-based services. We propose an alternative approach, centered around the human body and based on human perceptional characteristics.

6 Conclusions

We have presented an interaction model which _ealizati some of the drawbacks of the PC/WIMP interaction paradigm such as the demand for the human actor to devote complete attention to a specific computing device and the lack of computing support for activities involving everyday object manipulation. The proposed solution involves an "egocentric interaction manager" which continously monitors to what extent input and output devices are observable and manipulable for a specific user. We have also proposed a wearable activity-aware platform based on the model, as well as our first steps in implementing it for the purpose of developing a "cognitive prosthesis" for people suffering mild dementia. The main advantage of the wearable approach in designing activity-aware support is the potential independence on instrumented sensors, enabling computing and activity support literally everywhere. The proposed interaction model has proven to be successful for activity recognition in a Virtual Reality environment, motivating us to proceed with implementation in the real world. We are fully aware of the fact that the realisation of the components in the proposed activity-aware computer system faces numerous technological challenges of which we ourselves up until now have already addressed a few. However, many of these challenges are already actively addressed in isolation within areas related to Ubiquitous Computing, for instance in the shape of solutions for ad-hoc sensor networking (e.g. the Global Sensor Network middleware [1]). We see the main contribution of the work presented in this paper to be our proposal for how advances in these fields can be combined to an integrated system that has the potential of opening for new application areas for digital computing, and as a consequence, introduce the power of personal computing to new human communities beyond the classical office worker. We have also demonstrated the utility of Virtual Reality simulation as a tool for early assessment of Ubiquitous Computing systems.

Future Work. As stated in section 4, there are still several parts of the proposed activity-aware platform that needs to be developed in order to arrive a fully working cognitive prosthesis based on a sensor pool simulated in Virtual Reality. We are currently working on the implementation of these components. In parallel, we are experimenting with an RFID-based [6] sensor pool (a reader on the chest with antennas on the hands and on the chest; tags on the objects) connected to an ultra-portable PC for tracking object manipulation and the manipulable space. Such information has been reliably identified in many applications including [14], [16] and [13]. Presence of physical objects in the observable space is probably the hardest challenge but our tests in a Virtual Reality environment indicate that at least for activity recognition purposes, the manipulable space can to some extent substitute the observable space (with a reduction of 4% in activity recognition precision) [23]. With regards to sensing virtual objects in the spaces we imagine this being done by communication with the devices presenting them. The protocol and a suitable abstraction level at which the events in virtual environments (provided to the human actor by computing devices such as portable and stationary PCs, cellular phones, etc.)

are to be described and transmitted to the wearable activity aware computer is currently also an open issue.

Acknowledgoments. We would like to thank Fabien Lagriffoul for developing large parts of the activity recognition component; Daniel Sjölie for developing the VR sensing infrastructure and the blackboard communication mechanism; Lars-Erik Janlert for many helpful suggestions on the physical-virtual design approach and the egocentric interaction model; Robert Bhatt for the valuable ethnographical study. We would also like to thank Gösta Bucht and Björn Sondell from the Dept. of Geriatric Medicine, as well as Kenneth Bodin, Anders Backman and Marcus Maxhall at VRlab. Final thanks go to Michel Banâtre and Paul Coderc at IRISA/INRIA, Rennes, France for supporting the initial development of the idea of a physical-virtual operating system. This work has been partially funded by the EC Target 1 structural fund program for Northern Norrland.

References

1. Aberer, K., Hauswirth, M., Salehi, A.: The Global Sensor Networks middleware for efficient and flexible deployment and interconnection of sensor networks. Technical report LSIR-REPORT-2006-006
2. Backman, A.: Colosseum3D - Authoring Framework for Virtual Environments. In: Proceedings of EUROGRAPHICS Workshop IPT & EGVE Workshop, pp. 225–226 (2005)
3. Bhatt, R.: Comparing the Performance of ADLs in “Virtual” and “Real Life” Environments. Dept. of Computing Science, Umeå university, report UMINF-06.40 (2006)
4. Christensen, H.B., Bardram, J.: Supporting Human Activities - Exploring Activity-Centered Computing. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, pp. 107–116. Springer, Heidelberg (2002)
5. Dey, A.K.: Providing Architectural Support for Building Context-Aware Applications. Ph. D. Thesis Dissertation, College of Computing, Georgia Tech. (2000)
6. Finkenzeller, K.: RFID Handbook, 2nd edn. John Wiley and Sons, New York (2003)
7. Fritzke, B.A: Growing neural gas network learns topologies. In: Tesauro, G., Touretzky, D.S., Leen, T.K. (eds.) Advances in Neural Information Processing Systems, vol. 7, pp. 625–632. MIT Press, Cambridge, MA (1995)
8. Garlan, D., Siewiorek, D., Smailagic, A., Steenkiste, P.: Project Aura: Towards Distraction-Free Pervasive Computing. IEEE Pervasive Computing, special issue on Integrated Pervasive Computing Environments 21(2), 22–31 (2002)
9. Ishii, H., Ullmer, B.: Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In: Proceedings of CHI 1997, pp. 234–241. ACM Press, New York (1997)
10. Janlert, L.-E.: Putting Pictures in Context. In: Proceedings of AVI 2006, pp. 463–466. ACM Press, New York (2006)
11. Mackay, W., Fayard, A.-L., Frobert, L., Médini, L.: Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control. In: Proceedings of ACM CHI 1998, Los Angeles, pp. 558–565. ACM Press, New York (1998)
12. Nardi, B. (ed.): Context and Consciousness: Activity Theory and Human-Computer Interaction. MIT Press, Cambridge (1995)

13. Patterson, D., Fox, D., Kautz, H., Philipose, M.: Fine-Grained Activity Recognition by Aggregating Abstract Object Usage. In: Ninth IEEE International Symposium on Wearable Computers (2005)
14. Pederson, T.: Magic Touch: A Simple Object Location Tracking System Enabling the Development of Physical-Virtual Artefacts in Office Environments. In: Journal of Personal and Ubiquitous Computing, vol. 5, pp. 54–57. Springer, Heidelberg (2001)
15. Pederson, T.: From Conceptual Links to Causal Relations — Physical-Virtual Artefacts in Mixed-Reality Space. PhD thesis, Dept. of Computing Science, Umeå university, report UMINF-03.14 (2003) ISBN 91-7305-556-5
16. Philipose, M., Fishkin, K., Perkowitz, M., Patterson, D., Fox, D., Kautz, H., Hähnel, D.: Inferring Activities from Interactions with Objects. IEEE Pervasive Computing, 50–57 (October 2004)
17. Rabiner, L.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. In: Proceedings of the IEEE, vol. 77(2), IEEE Computer Society Press, Los Alamitos (February 1989)
18. Satyanarayanan, M.: Pervasive computing: vision and challenges. IEEE Personal Communications 8(4), 10–17 (2001)
19. Schmidt, A.: Implicit Human Computer Interaction Through Context. Personal Technologies 4(2&3), 191–199 (2000)
20. Shneiderman, B.: The future of interactive systems and the emergence of direct manipulation. Behaviour and Information Technology 1, 237–256 (1982)
21. Starner, T.: The Challenges of Wearable Computing: Part 1 & 2. IEEE Micro 21(4), 44–52, 54–67 (2001)
22. Suchman, L.: Plans and situated actions: the problem of human machine interaction. Cambridge University Press, Cambridge (1987)
23. Surie, D., Pederson, T., Lagriffoul, F., Janlert, L., Sjölie, D.: Activity Recognition using an Egocentric Perspective of Everyday Objects. In: UIC 2007. Proceedings of IFIP 2007 International Conference on Ubiquitous and Intelligent Computing, Hong Kong, Springer, Heidelberg (July 11-13, 2007)
24. Want, R., Pering, T., Danneels, G., Kumar, M., Sundar, M., Light, J.: The Personal Server: Changing the Way We Think about Ubiquitous Computing. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, Springer, Heidelberg (2002)
25. Weiser, M.: The Computer for the 21st Century. Scientific American 265(3), 66–75 (1991)