Thinking out of the box – user experience as viewed from an egocentric interaction perspective

Thomas Pederson[1,2] **Rosa Lanzilotti**[2] **Francesca Montinaro**[2] **Maria Francesca Costabile**[2]

[1]Department of Computing Science
Umeå University
SE-90187 Umeå, Sweden
top@cs.umu.se

[2]Dipartimento di Informatica
Università di Bari
70125 Bari, Italy
{pederson, lanzilotti, montinaro, costabile}@di.uniba.it

ABSTRACT
This position paper focuses on our emerging conceptual framework for modelling and designing everyday activity support systems: *egocentric interaction*. Although we depart from a classical instrumental HCI design perspective, we identify various common interests with the emerging research field of User Experience. In particular we attempt to contribute to the workshop by providing a user-body centered model that bridges the gap between the physical and the virtual world, and by suggesting a personal computing platform inspired by wearable and ubiquitous computing research, which we expect to emerge within the coming 5-10 years. The paper is concluded with two example projects where the proposed egocentric interaction framework is applied in practice.

Author Keywords
Ubiquitous Computing, Mixed-Reality Systems, User Experience.

ACM Classification Keywords
D.3.2 [**Design Tools and Techniques**]: User interfaces; D.2.11 [**Software architectures**]: Domain-specific architectures; H.1.2 [**User/Machine Systems**]: Human factors; I.3.6 [**Methodology and Techniques]**: Interaction techniques; K.8 [**Personal Computing**]: Miscellaneous.

INTRODUCTION
As noted in the workshop call, "the conception of usability has been evolving, along with the emerging IT landscape and the everblurring boundary of the field of HCI".

In this position paper, we describe our design framework for modelling human activity in this "emerging IT landscape": *egocentric interaction*. Our aim is to grasp and reveal an individual human actor's perception of a specific mixed-reality environment by sensing the presence and state changes of physical (real) and virtual (digital) objects in the vicinity of his or her specific human body due to the performance of physical actions. By focusing the modelling and sensing of human activity around the specific human body itself, we believe we have a chance to acquire a better understanding both of the specific human actor's perception of the world (e.g. what can technically be experienced by the particular individual in a given mixed-reality context) as well as of how it is manipulated (e.g. where objects are moved and in what sequence). We believe that such an understanding, by being based on measurable attributes, could provide valuable input to a more holistic User Experience (UX) design approach, complementing the classical HCI models that tend to be device-oriented as well as the UX work focused on social aspects.

"Are user experience elements tractable, quantifiable, and measurable?" We would propose that the answer to this rhetorical question, expressed in the call for this workshop, is a partial "yes". A currently popular view within cognitive science is that much human activity is directly or indirectly affected, or even driven by, the surrounding environment. More precisely, the detailed execution of individual human actions do often depend on what the specific human actor perceives from whatever is around. We believe it is possible, for a computing system, to capture quantifiable parts of this human individual view of the world even though hard-to-grasp (subjective) cognition certainly has an impact on human perception[1]. One of the most interesting contributions from the field of UX to classical HCI is in

[1] One nice illustration of how cognition can impact (higher levels of) perception is the proverb that "if you are provided with a new hammer, things around you tend to look like nails".

fact the attempt to include subjective attributes like, for instance, aesthetic, emotional, and social aspects in a design space which has previously mainly concerned ease-of-use.

USER EXPERIENCE AND UBIQUITOUS COMPUTING

According to Hassenzahl and Tractinsky, UX is influenced by a) user's internal state (e.g. predispositions, expectations, needs, motivation, mood, etc.), b) the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.), and c) the context within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.) [2]. Just like UX expands the scope for HCI along cognitive and social dimensions, the field of Ubiquitous/Pervasive Computing (and related fields such as Wearable/Mobile Computing, Augmented/Mixed Reality, Context Awareness) expands the notion of what can be part of a computing environment. In particular, these new fields of technology-driven research have one thing in common: they attempt to include parts of the everyday real world in the computing environments to be designed and interacted with.

As noted repeatedly within the HCI community for the past 10 years or so, users might change and combine computing contexts in a way previously not technically possible. Users do not interact with one computer but with many, of different kind, and as an integrated part of everyday real-world activities. This is certainly a new source of experience for users, that needs to be deeply analysed and modeled in some way.

In the model we propose, technological details of the computer systems do not matter any more, turning the emphasis towards modeling physical (real) and virtual (digital) objects uniformly. The underlying assumption is that this resembles how domain experts mentally model physical and virtual objects of interest in mixed-reality environments. This model is thus the basis for building effective mixed-reality environments that will allow domain experts to concentrate on the daily tasks, reducing as much as possible any interference from the technological infrastructure.

EGOCENTRIC INTERACTION CORNER STONES

Our general aim is to design computer systems that facilitate everyday activities no matter if they take place in the physical world, the virtual world (mediated by some interactive computing device) or — which we believe will become increasingly common in the future — somewhere inbetween [4].

The egocentric view on interaction we propose differs from more classical HCI models by building on the following corner stones:

1. **A physical-virtual design perspective.** Objects of interest to a particular human agent are modelled uniformly no matter if manifested in physical space, virtual space, or physical-virtual space.
2. **One human, no user interface.** Focus is on interaction between a single human agent and objects of interest to that human agent, specifically ignoring infrastructure that mediates virtual objects, such as personal computers.
3. **Strengthening connections between physical and virtual artefacts.** Focus is on the role of potentially automatic interplay between a) physical and virtual objects [residing in] b) physical and virtual environments [making up] c) the physical and the virtual world.
4. **"Applications" emerging by doing.** Modelled and infrastructurally supported physical-virtual environments ("applications") are just as likely to emerge through everyday egocentric interaction by specific human agents, as being pre-defined by external physical-virtual environment designers.
5. **Support for living rather than just work activities.** The aim is to support personal everyday activity 24 hours, 7 days a week, without drawing the classical border between work and leasure, home and out.

The term 'egocentric' has been chosen to signal that it is the human body and mind of a specific human individual that (sometimes literally) acts as a centre of reference to which all interaction modelling and activity support is anchored.

Motivation for Corner Stones 1 & 2

What matters for human agents when manipulating objects for the purpose of performing well-known activities is the changes made to objects of interest (henceforth called domain objects), whether it is about calculating next years project budget in a spreadsheet application or knitting a pullower. What does not matter is the way the domain objects are accessed. E.g., expert PC users do not have to pay attention to how the physical mouse corresponds to the virtual arrow. Instead, they pay attention to the effects of their virtual activities as if they were inside the virtual environment themselves. Dragging a virtual document to the virtual trash can is as natural as throwing the corresponding paper document in the physical trash can below the desk. Based on observations such as these, the physical-virtual design perspective [4] suggests that physical and virtual domain objects should be as central for physical-virtual designers as they are for the users.

Motivation for 3

One dimension of the physical-virtual gap has to do with the lack of causality crossing the divide [4] . State changes of an object in the physical world typically does not have any effect on a corresponding object in the virtual world, or vice versa. Infrastructure keeping corresponding physical and virtual artefacts in synch would bring the two worlds closer to each other.

Motivation for 4 & 5

The space of possible activities is much larger in the physical (real) world compared to in the virtual world. It is harder as environment designer to have detailed control over human activity (and the way it is performed) in the physical world compared to the almost complete control offered when designing environments and behaviour in the virtual world.

The dynamics of "physical-virtual applications" caused by the extreme mobility (in physical space) of virtual environments, compared to the immobility of physical environments. The extreme mobility of virtual environments, paired with ever cheaper virtual environment providers (computing devices) makes it technically possible to provide virtual environments in almost any physical situation, not just the typical office setting. Thus, the design space for activity support more or less automatically expands from classical work settings to include just about any kind of human activity thinkable.

THE PHYSICAL-VIRTUAL DESIGN PERSPECTIVE

Computers, embedded in the "background" as well as more obtrusive artefacts (e.g. PCs, PDAs, 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" [4]. 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 Classical User Interface Can Be Ignored

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 information technology for helping human agents bridging it. The assumption is that a reduced physical-virtual gap means less "friction" for physical-virtual activities. Physical and virtual space is modelled together, and automatic mechanisms for synchronising related phenomena in both worlds are imagined to be offered by systems that have been developed with the physical-virtual perspective in mind. By viewing the physical and virtual worlds as one, we believe the chance to make them one increases.

Adopting the physical-virtual perspective involves abstracting away the classical HCI concepts of input and output devices, giving them a background role as Inter-World Event Mediators (IWEMs). Fig. 1 and 2 illustrate the basic cases of physical and virtual human action (object manipulation). IWEMs are shown as white squares.

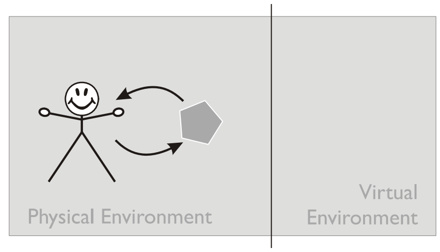

Fig. 1. Physical action [4]: Human actor interacting directly with an object

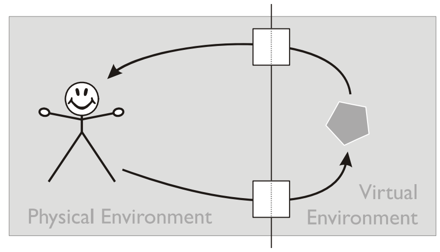

Fig. 2. Virtual action [4]: Human actor interacting indirectly with an object through IWEMs creating a direct experience

In order to arrive at a definition of physical-virtual activity it is useful to define human border-bridging activity on a lower level of abstraction first:

Definition 1: A physical-virtual action pair consists of two actions belonging to the same activity and often time-wise adjacent, where the first action is constrained (by lack of action support in the current environment) or chosen (e.g. based on individual preferences) to be performed in the physical world and the other action is constrained/chosen to be performed in the virtual world, or vice versa [4].

Physical-Virtual Artefacts

Among physical-virtual action pairs we can sometimes identify one or several information-mediating objects that are subject to indirect or direct human manipulation in both actions, objects that transcend the physical-virtual border by being present in both worlds. Such objects are referred to as Physical-Virtual Artefacts (PVAs) and for denoting the presentations of them in the two different worlds, the term PVA manifestation is used. A text document presented in both the physical (e.g. printed on paper) and the virtual world (e.g. within a word processing environment) would serve as a good example of a PVA, where each manifestation affords different kinds of manipulation.

Definition 2: A physical-virtual action is an action on a PVA where both the physical and virtual manifestations are directly controlled and/or monitored by the agent [4].

Fig. 3 and 4 illustrate two possible kinds of physical-virtual actions. Finally, the concept of physical-virtual activity is defined as follows:

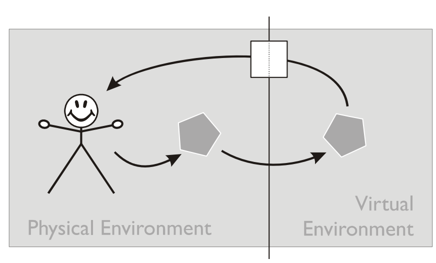

Fig. 3. Physical →virtual action [4]: Human actor observing the virtual-world effect of manipulating a physical object

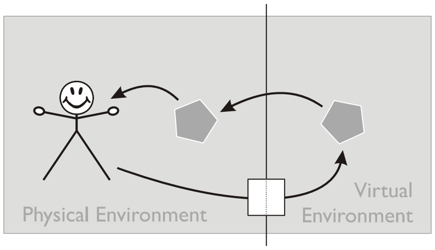

Fig. 4. Virtual →physical action [4]: Human actor observing the physical-world effect of manipulating a virtual object

Definition 3: A physical-virtual activity is an activity consisting of a sequence of actions containing a) at least one physical-virtual action pair or b) at least one physical-virtual action. [4]

One Space, One Magnifying Glass

By viewing the physical and the virtual worlds as equally important for human activity, the proposed physical-virtual perspective makes terms tied to an implicit virtual-world bias such as "context awareness" obsolete. It also expands the meaning of "location tracking" (currently having an implicit physical-world bias) to include also space and place in the virtual world. It invites the viewing of the relationship between physical and virtual environments from unconventional angles. For instance, why should not the current (local) state of the virtual world influence how activity in the physical world is interpreted? Could it not be of use for physical-world "applications" to be aware of their virtual-world context? And why is virtual-world location tracking (e.g. web pages browsed by a human agent) not considered when designing interactive environments?

EGOCENTRIC INTERACTION

The egocentric interaction perspective is based on a situative model of what a specific human agent can see and not see, reach and not reach at any given moment in time (Fig. 5).

The model is based on the physical-virtual design perspective briefly outlined in the previous section. Thus, physical and virtual domain objects are treated as being located in the same space. As a specific human agent changes physical and/or virtual location, objects come into and leave the observable physical-virtual subspace in a dynamic fashion. Somewhat simplified, one can say that it is the borders of the observable subspace which defines the set of objects that can possibly be part of a physical-virtual "application" at any given time-instant for the specific human agent.

The idea of using computers for assisting individual human agents in everyday life is not new but has gotten increased relevance in the last 5-10 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" [8]. 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 in this respect. 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.

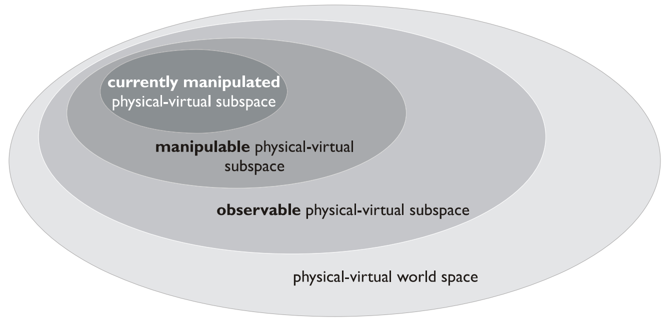

Fig. 5. A situative physical-virtual space model [4]

Computing Infrastructure for Egocentric Interaction

The egocentric approach follows the current HCI trend, breaking with classical Task Analysis that assume human agents to perform all actions based on rational decisions for reaching well-defined goals most efficiently. Egocentric computing systems do not necessarily have to actually know what the modeled activity is about but rather what the human agent seems to need (in time and space) in order to perform it, mainly based on historical data of object use. Thus, we foresee that emerging individualised physical-virtual "applications" rather than traditional pre-defined general-purpose ditto designed by application designers.

Being a relatively abstract perspective on future HCI, the egocentric view does not assume the use of any particular kind of technology for supporting interaction between the human agent and the physical-virtual world. For instance, computing and sensing technology for tracking physical activity of a specific human agent could be imagined to be either worn by the human agent herself, or be situated in the surrounding physical environment. The same goes for virtual environment providers (computing devices providing access to the virtual world) which could be both worn by their user or ubiquitously distributed throughout the physical environment like in Mark Weiser's vision [9].

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 egocentric computing system as self-sufficient as possible, reducing the problem of "uneven conditioning" [6]. Information about activities performed using devices external to the wearable egocentric interaction system (e.g. in the case when the human agent is using a desktop PC) need to be transmitted through some standard wireless communication protocol to the wearable egocentric interaction device for analysis. Complemented with real-world object manipulation information from sensors, the egocentric interaction system would (at least in theory) be able to gracefully model activities across the physical-virtual gap.

The conceptual system architecture (illustrated in Fig. 6) is based on a wearable computing/sensing hardware configuration consisting of a *private black box* offering computing power and storage space for data generated by an *egocentric interaction sensor pool* monitoring object-centric phenomena within the observable physical-virtual subspace of a specific human actor.

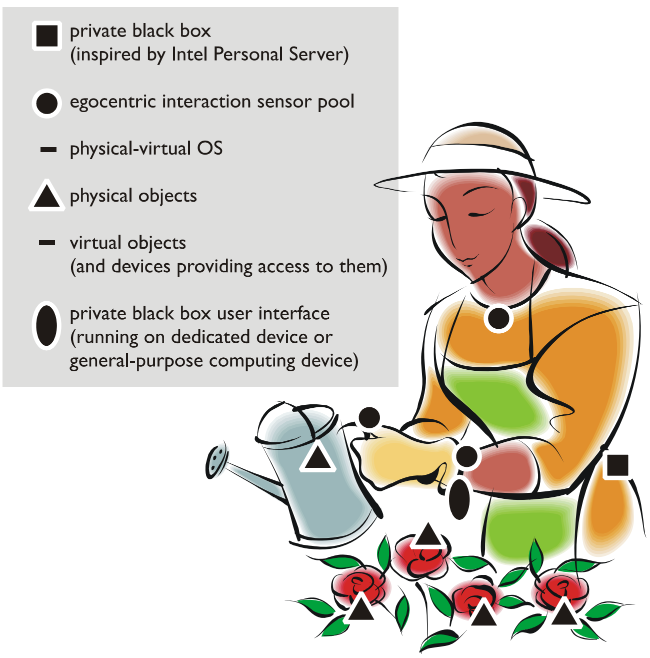

Fig. 6. General components of an egocentric interaction system (virtual objects not pictured)

Furthermore, the private black box runs a *physical-virtual operating system* hosting both advanced physical-virtual applications developed by software developers as well as simpler programs designed by the user her/himself. Such applications can incorporate the manipulation of both physical objects (e.g. a sculpture at a museum) and virtual objects (e.g. a web page describing the same sculpture). Explicit interaction with the physical-virtual operating system is performed through a *private black box user interface*, either fitted onto the private black box itself, or running on a general-purpose device like a PC. Implicit interaction [7] with the physical-virtual operating system emerges whenever the user interacts with a physical or virtual object inside the manipulable physical-virtual subspace (see Fig. 5) monitored by the private black box. The local computing of the private black box can optionally be enhanced by communication with publicly and ubiquitously accessible *shared object knowledge repositories*, distributing anonymous data about objects and their everyday use.

As an example of a useful application, wearable egocentric interaction technology has the potential of ensuring that human agents always have the necessary physical and virtual objects at hand for successfully performing the activities they like to. Such systems can act in the conscious "foreground", reminding the wearer to bring this or that physical object along when changing physical location, or in the "background" by setting up physical-virtual environments prior to the human agent's arrival, making the emerging physical-virtual world a slightly smoother place to be in.

APPLYING THE FRAMEWORK

As it is common in the explorative branches of HCI, the interaction theory presented in this paper has been and will continue to be developed hand in hand with prototype systems. The aspects of the egocentric interaction framework related to the physical-virtual design perspective have inspired and been inspired by the development of the Magic Touch system [5; 3] which among other things provided limited support for end-user design of Physical-Virtual Artefacts based on wearable position and object identification sensors.

easyADL

A more extensive application of the egocentric interaction framework is currently undertaken within the easyADL project (www.cs.umu.se/research/easyadl) where the focus lies on recognition and support of everyday activities based on the situative physical-virtual space model (Fig. 5). easyADL is a two-year research project started in June 2005 investigating new forms of computer technology for facilitating everyday life of individuals suffering dementia disease. The goal is to develop a wearable computer-based "cognitive prosthesis" which to some extent could compensate for mild cognitive impairments when performing everyday tasks such as getting dressed, preparing dinner, and going to the bathroom.

In order to speed up the design process, and in order to compensate for limitations with currently available sensing technologies, Virtual Reality models are used to simulate wearable real-world sensors, feeding "clean" data to the activity modelling algorithms under investigation. The method also facilitates the experimentation and comparison between different wearable sensor configurations for successful recognition of activities based on a set of objects nearby and/or manipulated by the human agent within a given time span. The intention is to substitute the VR environment with real-world environments when "backbone" egocentric interaction components (e.g. activity recognition) work satisfactory.

CHAT

In the upcoming project CHAT, "Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies", we intend to develop a software infrastructure that allows services accessed through thin