A Generalized Peephole Metaphor for Augmented Reality and Instrumented Environments

Andreas Butz, Antonio Krüger,
Saarland University, Computer Science, 66123 Saarbrücken, Germany
email:(butz,krueger)@cs.uni-sb.de

***Abstract*— We present the generalized peephole metaphor, a model of interaction in augmented reality and instrumented environments. After briefly introducing previous research, which led to the model, we describe a specific scenario in which we introduce the core idea and some implications. The model nicely matches research in human perception, which is also discussed briefly. A short outlook describes our current research, of which the generalized peephole metaphor has become an integral aspect.**

I. Introduction

A. Peepholes, Toolglasses and Magic Lenses

Historically, screen space and screen resolution have always been scarce resources, and this has remained true until today. In order to make better use of existing screens, and enhance the (logical, not physical) focus of a graphical display, the interface metaphor of the magic lens[1] was developed. The magic lens is the electronic equivalent of a magnifying glass and allows the display of more detail or alternative types of information in screen areas, over which it is (metaphorically) held. To allow interaction and input through the lens, it is combined with the toolglass, a filter modifying the meaning of the mouse cursor acting through it. Bier et al. give a systematic overview of tools and types of interaction with these see-through interface widgets in [2].

While the magic lens metaphor was designed for 2D screens, there is an equivalent within physical 3D environments. In [3], George Fitzmaurice devised the vision of creating situated information spaces by making palmtop computers situation aware. Making Information situated or localized in space provides a very natural access to it. By mapping virtual information to the physical world, many types of information can be provided in places, where they would belong, were they physical entities. This transforms the question *how* to access a given information into the question, *where* to access it and thus makes use of the very well developed spatial memory humans possess. A spatially aware palmtop device is just held to the physical location in question, and the associated information appears. Fitzmaurice's prototype was called the chameleon and used a portable TV monitor as the palmtop device.

Palmtop computers have become commonplace today in the form of electronic organizers. They are powerful computing devices and provide access to various types of information. Therefore, a logical step was to combine them with 3D tracking to further work on Fitzmaurice's vision and create a large personal information space accessible with a small PDA by moving it in space. A corresponding prototype is described in [4], introducing the concept of *peephole displays*. The prototype described can act within a coordinate system fixed in space, e.g., over a table, or in a coordinate system fixed relative to the user of the device. In the first case, it can be used just like a magnifying glass or a torch in the dark, to reveal information bound to positions in space. in the second case, it uses the strong sense that humans have about their own body coordinate system, to position certain information or applications relative to the user's body.

B. Instrumented Environments and Augmented Reality

Instrumented environments provide a sandbox in which ubiquitous computing scenarios can be investigated. Instruments and devices embedded in the environment give things and places the power to display or sense information. They provide the connecting points between the physical world, in which they are located, and the virtual world to which they are connected. In this sense, the instrumentation of an environment augments it without the use of classical AR tools, such as projection or head-worn displays (see also [5]).

While in the vision of Ubiquitous Computing as devised in [6], objects or spaces have the ability to make themselves known, display information, or sense input, augmented reality research mostly looks at the external augmentation of objects or spaces. For a single user, the result is ideally the same, no matter whether an object contains a display, information is projected onto it, or overlaid to its visual appearance in a video-based or see-through head-worn display: The user sees the information on the object. For multiple users, care has to be taken to provide a common perception of the virtual overlay, if this is created with head-worn displays [7], but their experience will be shared automatically in the case of projection or embedded displays. Issues arising from shared experiences, such as privacy, synchronization and common reference points are increasingly addressed in the multi-user VR and AR literature, such as for example [8].

If we want to create environments, in which multiple users can effortlessly experience the same augmentation, instrumentation and projection provide the ideal means for this. Mark Weiser and Daniel Russel [9] make clear, that creating a harmonized view on multiple devices and modalities is a prerequisite for the successful exploitation of ubiquitous computing and AR/MR in the future. In the following sections we will describe a specific instrumented environment and then propose such a harmonized and generalized view on instrumented environments, which we call the generalized peephole metaphor.

II. Elements of a Specific Instrumented Environment

The instrumented environment which we are currently setting up, is called the SUPIE (Saarland University Pervasive Instrumented Environment). It currently consists of a single (office) room containing a desk and a chair. The desk is augmented by a projector in order to create visual elements on its surface, and by a camera providing both a steady video stream and high resolution pictures upon request. On the wall is a large touch screen, corresponding to an instrumented blackboard. A small touch screen display in the shape of a picture frame can display a decorative image by default, but also provide information or serve as an area of interaction when needed. In addition, The whole room can be augmented

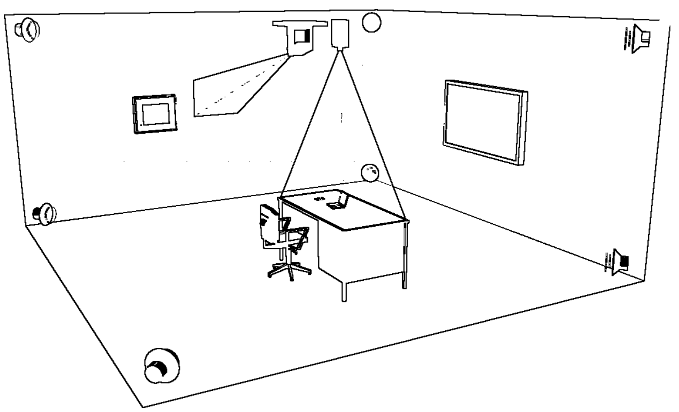

Fig. 1. A conceptual sketch of SUPIE

by projection from a ceiling-mounted movable projector, serving the same purpose as IBM's Everywhere Display [10]. This steerable projector effectively turns the whole room into a large continuous display, since it can display onto any surface with direct line of sight to its position. Spatial audio is provided by speakers distributed in the SUPIE environment. This allows the generation of spatialized sound in any location of the room.

Furthermore, the room contains regular contemporary computers of various sizes, such as desktop PCs, laptops, tablet PCs and PDAs. These provide display space in more or less well known positions, as well as input facilities of sorts (keyboards, mice, pens, touch screens). A common software layer called the environment manager will connect these devices and take over management of their input and output modalities, similar to a window manager on conventional 2D screens. While it is too early to discuss technical details of this infrastructure, its design will be guided by the model of interaction described below.

III. The Generalized Peephole Metaphor

A. Peephole Output

The idea of looking at AR and instrumented environments in a way consistent with peephole displays [4] first became obvious, when we thought about how to control the ceiling mounted projector. In accordance with [10], a 3D model of the room is used to eliminate spatial distortion in the projected images. Technically, images are texture-mapped onto surfaces in an otherwise dark 3D space, and then the virtual camera (mounted on the virtual ceiling) is aligned with the physical projector. If we point the physical projector to a certain position in space, it will display whatever information is there at the same position in the virtual 3D model. The projector thus makes visible the information assigned to a given position in space, only when it is pointed at this position. Information in other positions is still there in the virtual model, but not visible in the physical room. The projector effectively provides a peephole into the virtual space. In theory, with fast enough head and gaze tracking and projector control, this could be used to constantly align the projector with the user's gaze, which would make all information visible to the user by just looking at it. In practice, however, this becomes somewhat less compelling (to say the least) because of the lag between user tracking and the mechanical motion of the projector.

B. Peephole Input

Symmetrically to the peephole view on output modalities, input can be looked at through a peephole. Just as IBM's ED projector is equipped with a camera, so is SUPIE's steerable projector. Similarly to the desk, a camera is mounted aligned with the projector and provides a permanent low resolution video stream as well as high resolution images on demand. This camera provides marker recognition via the AR Toolkit [11] library and thus serves as an input modality by recognizing tagged objects and tracking them in space. This input modality is, of course, only available in the area at which the steerable projector is currently pointed, which conveniently creates an interactive area in the room by simultaneously providing input and output there.

C. Multiple Peepholes

An ideal display continuum would provide display capabilities at any time on any given surface in the room. An electronic wallpaper would be a good step in this direction. In reality, the steerable projector can not point at all surfaces simultaneously. It has to share display availability between different places. In addition to the peephole provided by the projector, all available screens can serve as additional peepholes into the same information space. They just become islands of high resolution and permanent availability within an otherwise time-shared and low resolution display continuum. Similarly, all available input channels and devices in the room can be seen as additional islands of permanently available input. Touch screens, keyboards, mice, or interaction objects such as the Tuister [12] provide additional peepholes for input to the virtual layer of the environment. This implies, that basically all devices in the environment can be used as peepholes of sorts into the virtual realm. It also means, that by peeping into the same virtual layer, they provide a consistent view on it and thus a shared experience for multiple users.

D. Privacy Implications

The peephole metaphor for instrumented environments also allows a consistent treatment of privacy issues. Public displays, such as large wall-mounted displays provide shared peepholes, while small and private displays, such as PDAs or head-worn displays provide private views into the virtual. The issue of access control to information in shared environments leads to the need to formulate a set of rules, which displays will become blind for which types of information. The various peepholes can simply contain filters, which work similarly to magic lenses to make sure that only information appropriate to that specific display is displayed. This is also a generalization of the vampire mirror and lens concept introduced in [7]. The filters can then be switched or adjusted in order to make, for example, public displays show content, which would normally be considered private, if this is wanted by the content's owner.

E. The Information Space

The generalized notion of peepholes and peephole interfaces allows a consistent integration of all available devices in a room. Every information has a position in space and a set of access restrictions assigned to it. Displays with known positions and level of privacy can display corresponding information. Movable displays can be moved to where information is available and be used to display it. Conversely, information objects can make themselves known by opening a peephole which shows them. A virtual alarm clock attached to an object can, for example, request the steerable projector to display it when it goes off and simultaneously play back an audio signal from its position. Searching a specific object then just means to ask it to actively open a peephole to make itself known. The spatial audio channel itself is a second display continuum with perfect simultaneous availability in the whole environment and can be used to guide the user's attention to given locations. Both the visual and the audio channel can be used separately to open a peephole. In the example of the alarm clock or search function it might also be sufficient to ring an alarm bell at the position of the object and then display it only on explicit request of the user. This may avoid interruptions, if the projector is already used to display other information at the same time.

IV. Implications for HCI in Instrumented Environments

A. Cognitive Peepholes

While traditional models of visual attention assume that everything we see is accumulated in a visual buffer and that all subsequent cognitive processing relies on this buffer, newer research seems to disprove the existence of a homogeneously detailed buffer. Experiments have shown that visual disturbances in an image, which attract the visual attention of observers, often make them fail to notice big changes in the image while looking at the image at the same time. This effect, known as *change blindness*, provides strong arguments against the existence of a visual buffer for highly detailed information everywhere. The model described by Rensink [13] explains the role of attention to establish the visual coherence between objects. By guiding our attention to a certain object, we retrieve highly detailed visual information which is lost when our attention is guided to other parts of the environment. A much coarser representation of the object remains in memory which allows us, for example, to remember the object's position in space.

This view on the cognitive representation for visual processes has some striking similarities to the generalized peephole metaphor. Instead of displaying every visual information permanently, we rely on guided attention to (re-)establish spatio-visual consistency between objects in the environment. According to Rensink's model only the objects needed for the visual task at hand are represented in high detail. The argumentation for our peephole metaphor is very similar: Only the real objects needed for the user's actual task are augmented with virtual information. In the same way the brain is saving processing power and memory, the instrumented environment can save similar resources when presenting and obtaining information to and from users. Rensink's model implies, that even if only a small part of the environment is visible in high detail, users will still be able to interact with the whole environment effectively.

Another related cognitive model is the concept of *active vision*. If a user's attention is directed at a given object, it already perceives much more detail than in the visual periphery, because of an area of higher resolution in the center of the retina, called fovea. On top of this, humans actively looking at an object become even more sensitive to changes by building up hypotheses of expected events and actively adapting their vision, looking for these events. If we extend this concept to other sensory channels, we can speak of *active sensing*, if we also include output, we might have to speak of *active communication*. The generalized peephole metaphor can strongly support this cognitive model by providing detailed information only in areas of the user's attention.

B. Active vs. Passive Peeping

Two styles of interaction can be distinguished in an instrumented environment adopting the peephole metaphor: In *active peeping* the user accesses information directly at a certain location in space whenever he or she wants. Users can actively open a peephole by controlling the steerable projector directly via gestures or a 3D-input device. This allows the user to refresh or look up information which is expected at a certain position in space.

In *passive peeping* the attention of the user is guided by means of a peephole to support the user's actual task. Attention can be guided from one place to another, e.g., by rendering a virtual sound source at a certain position in the real environment or by using the projector to follow a virtual object that moves from one location to another. Passive peeping involves knowledge about the user's current tasks and goals in the instrumented environment and, of course, information on all accessible peepholes (e.g. high resolution displays). The instrumented environment needs to rely on sensor information received through peephole input channels, as well as a sophisticated model containing information on all devices available in the environment.

C. Saving Cognitive and Technical Resources

As mentioned before, electronic wallpapers would be an excellent, but currently technically still unfeasible solution for visually presenting information in an instrumented environment. However, in the light of the discussed cognitive models, even if we assume availability of simultaneous ubiquitous display capability, maintaining the peephole metaphor will have a number of advantages over displaying all information simultaneously. One example are permanently and rapidly changing environments. As known from theories of preattentive vision [14], small changes in the field of view are easily attracting our attention (known as *pop-out effect*). Too many of those changes will lead to a chaotic environment distracting users, thus counteracting the vision of calm computing.

As described in [15], this problem can be overcome by using the change blindness effect again, if visual information is only updated while the user's gaze is absent. Unfortunately this requires high fidelity tracking of the user's gaze which currently is only possible for well defined settings in limited spaces (e.g., in front of a desktop computer). For scenarios adopting the peephole metaphor, only a small part of the environment is augmented with information and even without gaze tracking most of the environment remains calm most of the time.

In some situations, users should be notified about changes in the environment occurring outside of all currently open peepholes. How this can be done depends on the importance of the information. About rather important changes, the user should be actively notified, for example by directing his or her attention to the position of change itself (as described above). Less important changes may be summarized when the user peeps at a location again, for example either by small animations or temporal diagrams (similar to stock charts) that visualize the changes since the last visit.

V. Conclusions and Ongoing Work

We have introduced the generalized peephole metaphor, providing a consistent way of describing input and output in instrumented environments and augmented reality environments with partial augmentation. The metaphor provides a consistent way to describe some limitations of current AR technologies, namely the limited availability and capabilities of display and input devices. It is also consistent with current models of visual perception and allows the user to build up a spatial model of the virtual layer which is overlaid to the physical environment. Simultaneously, it makes the environment calm by eliminating unwanted distraction, and thus eases understanding and interacting with localized information. The peephole metaphor also supports concepts of privacy and shared experience in multi user augmented realities.

While most of the hardware of our instrumented environment SUPIE is set up at the time of writing, we're currently designing a software environment for the management of all elements of instrumentation. This software will provide consistent management of input and output resources in the environment, similar to the device driver layer of an operating system. On top of this layer, an environment manager will be built, serving the same purpose as a window manager on 2D screens. Together, this will provide a sandbox in which a variety of interaction techniques and usage metaphors can be implemented, evaluated, and refined.

A fundamental guideline in building this environment will be the peephole model of interaction. While this paper provides mainly the formal framework, supported by a number of evidences from research on visual perception, the generalized peephole metaphor remains to be tested and verified in user studies to prove its scientific validity. We would like to discuss it at the STARS workshop in order to get additional opinions, experiences, and feedback.

References

[1] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose, "Toolglass and magic lenses: The see-through interface," *Computer Graphics*, vol. 27, no. Annual Conference Series, pp. 73–80, 1993.

[2] E. A. Bier, M. C. Stone, K. Fishkin, W. Buxton, and T. Baudel, "A taxonomy of see-through tools," in *Proc. ACM CHI '94*. ACM Press, 1994, pp. 358–364.

[3] G. W. Fitzmaurice, "Situated information spaces and spatially aware palmtop computers," *CACM*, vol. 36, no. 7, pp. 38–49, July 1993.

[4] K.-P. Yee, "Peephole displays: Pen interaction on spatially aware handheld computers," in *Proceedings of CHI 2003*. ACM Press, 2003.

[5] "Augmenting buildings with infrared information," in *Proceedings of the International Symposium on Augmented Reality ISAR 2000*. IEEE Computer Society Press, 2000.

[6] M. Weiser, "The computer for the 21st century," *Scientific American*, vol. 3, no. 265, pp. 94–104, 1991.

[7] "Enveloping users and computers in a collaborative 3d augmented reality," in *Proceedings of the International Workshop on Augmented Reality IWAR '99*. IEEE Computer Society Press, 1999.

[8] D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalavari, L. M. Encarnacao, M. Gervautz, and W. Purgathofer, "The studierstube augmented reality project," *Presence*, vol. 11, no. 1, pp. 33–54, February 2002.

[9] D. M. Russell and M. Weiser, "The future of integrated design of ubiquitous computing in combined real & virtual worlds [late-breaking results]," in *CHI '98*. ACM Press, 1998, pp. 275–276.

[10] C. Pinhanez, "The everywhere displays projector: A device to create ubiquitous graphical interfaces," in *Proc. of Ubiquitous Computing 2001 (Ubicomp'01), Atlanta, Georgia*, 2001.

[11] M. Billinghurst, S. Weghorst, and T. Furness, "Shared space: An augmented reality approach for computer supported collaborative work," *Virtual Reality*, vol. 3, no. 1, pp. 25–36, 1998.

[12] A. Butz, M. Groß, and A. Krüger, "Tuister: a tangible ui for hierarchical structures," in *Proceedings of PI03: Workshop on Real World User Interfaces, held at Mobile HCI 2003, Udine, Italy*, 2003.

[13] R. Rensink, "Internal vs external information in visual representation," in *Proceedings of Smart Graphics 2002*. ACM Press, 2002, pp. 63–70.

[14] A. Treisman, "Preattentive processing in vision," *CVGIP*, vol. 31, pp. 156–177, 1985.

[15] S. Intille, "Cange blindness information displays for ubiquitous computing environments," in *Proceedings of UbiComp 02*. Springer, 2002, pp. 91–106.