Moving Towards Naturalistic Interactions With Invisible Interfaces: Can "Digital Simplicity" Be Achieved?

Beverly L. Harrison
Intel Research Seattle
1100 NE 45th Street, Seattle, WA 98105
beverly.harrison@intel.com
(206)545-2503

THREE CHALLENGES OF INVISIBLE INTERFACES

Past user interface design leveraged real world metaphors and the naturalistic associations of these to create what was hoped to be more easy to use technologies (e.g., file folders, the desktop, the classic trash can, now subtly renamed the more environmentally friendly "recycle bin"). Pictorial representations (icons) were intended to facilitate understanding, learning, and the direct manipulation of objects of interest. These waves of prior innovation enabled technologies to become accessible to a broader community of users in ways that are now considered quite pervasive.

Technology is now equally pervasive in mobile or handheld devices, household appliances, and is often embedded invisibly in the environment around us (e.g., sensors, cameras, wireless networking). Where formerly users had explicit interactions with particular technologies they deliberately selected (by virtue of using a particular computer or device, a particular input mechanism, and/or a particular application), they may now be interacting in implicit ways with imprecisely selected technology at almost any moment in time. These interactions can occur whether users are aware of them or not, and whether users intended them or not.

Early visions of the future presented by Wellner's DigitalDesk [38, 39] and by Weiser's ubiquitous computing [36, 37] have been extended upon and are reflected in substantial research over the last 10 years [e.g., 8, 9, 11, 17, 26, 28, 29, 30, 31, 33] (including my own work [e.g., 6, 7, 13, 20, 21, 35]). A goal of these emerging projects is to seamlessly blend the affordances and strengths of physically manipulatable objects with virtual environments, functionality or artifacts thereby leveraging the particular strengths of each. This seems like a natural progression towards making the next UI "metaphor" the real world itself: real objects having real properties that are linked to or embedded with the virtual artifacts that they control. This reflects a powerful, only slightly understood user interface paradigm of "invisible interfaces"; interfaces that are always on, embedded in our everyday objects or environment, and subtly invoked by our naturalistic interactions. However, as with any technology, getting the design right is crucial if our aspiration is for widespread accessibility, predictability, ease of use and ubiquity. A first grand challenge for invisible interfaces is better *articulating what the user interaction model is* and how the associated design principles need to evolve to adapt to this model.

While the explicit manipulation of an everyday object may influence or trigger embedded technologies, research has simultaneously extended to examine the broader context of this interaction. (This focus on context of use appears to be particularly emphasized in ubiquitous computing.) Deliberate user interactions may produce implicit and unintended consequences because of the contextual assumptions made in designing and embedding these technologies (e.g., picking up a medicine bottle to read the code on the label for a refill vs. picking it up to take the medication [34, 25]). Furthermore, some environmentally embedded technologies are activated simply by virtue of the user being in a particular location without any expressed interactions at all (e.g., in view of a camera system, in range of a particular wireless network, in range of a motion sensor [e.g., 3, 10]). The idea of "situated action" is not new [32] however, it seems that the importance of understanding context is of increasing importance in disambiguating user intent in this newer interaction space [e.g., 5, 15, 22, 25]. A second grand challenge for invisible interfaces is *understanding or correctly inferring user context* and how this impacts design.

As a direct consequence both of the changing nature of what it means to interact with technologies in this invisible interface paradigm and due to the increasingly important role context plays, there is a crucial transformation needed for evolving evaluation techniques and methodologies. These invisible interfaces are used in dynamic and often highly mobile contexts and locales. They often involve a complex mesh of infrastructure, devices and sensors all of which must work as a coherent whole and thus must be assessed as a system. Evaluation methods that might have worked well for single technologies or specific interaction methods do not seem well suited to this more challenging problem domain. Modified techniques are being adapted and tested to try and address these unique attributes and challenges [e.g., 4, 12, 16, 24]. However, research on the methodologies and evaluation tools is in the early stages with much more promising results still ahead. A third grand challenge for invisible interfaces is *creating or evolving evaluation techniques and methodologies*.

INTERACTION MODELS FOR "INVISIBLE INTERFACES"

My past work has investigated three fundamentally different mechanisms of interacting with technology in the ubiquitous computing domain: 1) physically manipulative interfaces through object handling and deformation via embedded sensors [e.g., 6, 7, 13]; 2) embedded "inferred" interfaces though object location/presence sensing via RFID and computer vision [e.g., 20, 21, 35]; and 3) interaction resulting from embedding technology in the environment such as cameras, sensors, or wireless networking. These may be deployed independently or in concert to create an overall user experience typically characterized by the absence of visible user interface controls. Independently each of these creates a different model of interaction and different design constraints and possibilities.

Physically Manipulative Interaction

These interfaces typically rely on pressure sensors, accelerometers, motion sensors, light sensors, sound input and output, motors, and wireless communication to enable users to physically press, push, tilt, or juxtapose handheld devices or objects in order to convey commands. (I would like to differentiate interactions where object "deformation" is used from those where merely picking up an object is the interaction and hence command, discussed next).

There are already some compelling prototype examples reported in the current HCI literature:

- Scrolling a menu on a PDA by tilting the PDA [26]
- Zooming text by pushing/pulling the device towards/away from the user[8]
- Rolling or turning knobs [2]
- Squeezing/physically interacting with plush toys [18, 19] or physically deforming object surfaces [23]
- Or in my own prior work [6, 7, 13], exploring user interface techniques for navigating and displaying documents in portable reading devices/ebooks by squeezing, stroking, flicking, and tilting.

We find that designing these interactions confronts us with a host of subtle design issues, and that there are no articulated design principles to guide us through them. We believe that the user interfaces above are the forerunners of a new paradigm for user interface design, user interfaces in which a computing device is directly and physically manipulated, which we have call *Manipulative User Interfaces* (MUIs).

The MUI paradigm extends the GUI paradigm in 2 fundamental ways: 1) GUIs are somewhat indirect, in that the mouse controls a remote cursor behaving as interaction intermediary, whereas in MUIs, users use their bodies to directly manipulate devices, objects and/or their displays to specify commands. 2) The MUI paradigm opens up a much richer space of interactions than just pointing. Consider, for example, the Tamagotchi [1] "virtual pet" toy, which you interact with by means of a number of buttons. The toy requires its owner to provide it with food (by pressing the feeding button), to give it attention (by pressing the petting button), etc. But from a MUI perspective you could interact with the Tamagotchi in much more affective ways. You could soothe the pet by stroking it, play with it by tossing it in the air, ease its hunger by rubbing its belly, and so forth.

These manipulations can be divided into three categories: manipulations that change the *spatial position* of the device (translation, rotation, orientation, tilting), manipulations that change the *structural properties* of the device (deformations, pressing, squeezing, flicking, shaking), and manipulations that change the *inter-device relationship* (juxtaposition, stacking, tiling, proximate association or dissociation). All of these can be temporary or permanent in effect, can be parameterized in various ways, and can be performed alone, simultaneously, or in sequence.

In MUIs the *extent of embodiment* can vary. In some cases there is a tight coupling between the manipulation, the perceived content this affects, and feedback (both tactile and visual) indicating the resultant effect. All are embodied within a single object to be manipulated. In other instances the manipulation frequently affects content/data via a controller intermediary and visual feedback of the effect are represented in an external system distinct from the controller. Examples of this include recent graspable-object interactions, such as "phicons" [11, 28], "bricks" [9], and "doll's head" [14] techniques, where the object being manipulated acts as a remote control for actions taking place outside the object. The boundary between these two approaches can be somewhat blurred.

Embedded "Inferred" Interaction

This class of interaction is one where the act of selecting an object or moving it triggers a technological response thereby communicating an implied or inferred "request" from the user. This is typically achieved by computer vision, RFID, or location detection where technological modifications to an object are invisible to the user (barcodes and glyphs being the exception) yet handling these instrumented objects has an effect. A number of prototypes have been built to demonstrate a variety of applications:

- Moving physical blocks representing buildings for urban planning [33]
- The pick-and-drop work that attaches virtual content to a physical transport mechanism or physically selecting a projections of objects to move items between surfaces or devices using gestures [26]
- And some of my own work [20, 21, 35] looking at augmenting books, staples, business cards, posters, and augmented in/out boards and whiteboards with phicons (Collaborage).

Unlike environmentally sensed interactions, in embedded inferred interactions the participation of both a user and an embodied object are necessary. While the technologies deployed in both situations have similarities (wireless communication, detecting changes in location across wireless networks, detecting presence/absence of an object or person of interest), embodied inferred interactions are instantiated in particular devices or objects that users hold, carry and interact with. Whereas in environmentally sensed interactions, the environment infers interactions from the user him/herself regardless of the presence of absence of any particular object (e.g., the user's presence triggers a motion sensor, floor pressure pad, camera system).

As with most taxonomies or categorizations, there are situations where the distinction between an object-based inferred interaction and an environmentally aware application is somewhat blurred, for instance, in cases where an object has a tight affiliation with a user's identity (e.g., a cell phone the user habitually carries even when not in use) but the object's presence is virtually forgotten. The object (or device or sensor) is used to transmit information to environmental sensors as a back channel (e.g., cell phone transmits GPS location data) while the user is not deliberately interacting with this object and may be unaware of this active background transmission. One could argue that there is an interaction mechanism embodied in an object co-present with a user and thus this scenario is an embedded inferred interaction. However, the user has not participated actively (or perhaps even knowingly) in this interaction, the object's role is passive (from the user's perspective) and the application is crucially more dependent upon the environmental infrastructure sensing the user's presence (rather than the object's presence) thus it could equally easily be considered an environmentally sensed interaction. Perhaps the value of categorizing scenarios and applications will best be determined as we evolve design guidelines for these categories.

Environmentally Sensed Interaction

The proliferation and availability of largely invisible communication, camera and sensor-based technologies have created new possibilities for environmentally sensed interactions: global cellular networks, city-wide and building-wide wireless networks, city-wide and building-wide camera coverage, room or house scale sensors, etc. It is becoming more commonplace to see applications that utilize this infrastructure to sense and infer things about individuals, groups, behavior patterns, and trends (including divergences from norms). As describe above, individuals may be participants (knowingly or unknowingly) by virtue of habitually carrying devices that afford sensing or by themselves occupying or moving through a space that is instrumented.

There have been a number of implementations of "smart home" and "smart kitchen" prototypes that used sensors to detect and even identify individuals and their activities [e.g., 3, 25]. One could argue that home alarm systems are a simple but early instance of these. City wide camera-based technologies are deployed in the UK and in some US cities for security and traffic monitoring. Any number of location sensing applications have recently been built to track people, the places they frequent, the routes they take, and the activities they are doing [e.g., 3, 10, 15, 16, 25].

While all three of the above interaction categories have implications for inadvertent use, uninformed use, and privacy, environmentally sensed interaction is perhaps the most problematic and challenging [e.g., 10 presents an excellent overview of issues and studies]. There are no "objects" channeling users' intentions and express interactions. Nor can we leverage such objects to inform users of system activity. Deriving interaction models and design guidelines for this class of application must incorporate these considerations.

Conscious, semi-conscious, and unconscious interactions

We characterize the interaction models as being *conscious* (fully aware of interactions with an object), *unconscious* (fully unaware of interactions by virtue of having no explicit instantiation), and *semi-conscious* (sporadic awareness and forgetfulness of objects that afford interactions whether or not the user is aware).

We deliberately wish to use these as metrics to categorize 5 dimensions of interaction: awareness, intent, ubiquity, ease of use/design and skill. In fact, one might argue that as interaction mechanisms and technologies become more pervasive and ubiquitous (if well designed), they may migrate from being consciously invoked, manipulated, or monitored to semi-consciously or sporadically manipulated/monitored to unconsciously embedded in habitual routines without much explicit planning or thought about use. This raises interesting and difficult questions. How does the acquisition of skill and expertise relate to technology consciousness? How is design related to this and does bad/good design support or hinder consciousness? Is the migration of technology from consciously aware usage to semi- or unconscious use a positive progression?

UNDERSTANDING AND INFERRING USER CONTEXT

Many interactions with invisible interfaces rely on sensors to help make educated guesses about the user's context. If these sensors and interaction mechanisms are embodied within a device or object, parameters scoping the user intention may sometimes be inferred. In fact, stronger assumptions might be possible when considering sequences of object interactions [e.g., 25], for instance, activity inferencing that characterizes high level activities like "making a cup of tea" based on seeing interaction with the water tap, kettle, cup, tea box, and milk. One significant research problem is reliably and accurately abstracting these lower level actions into higher level activities, when users vary, lower level actions vary, sensors vary, and the data is noisy and ambiguous. However, surprisingly good results have come from targeting particular activity types that seem more amenable to prediction, especially if combined with supervised learning [e.g., 5, 16, 22, 25]. For instance, it is easy to detect someone in motion versus someone who is stationary or a change in location for a particular object. It is more difficult (and potentially intrusive) to detect contexts that have few physical cues (e.g., changes in mental state or switches in cognitive tasks). Determining task attributes that make certain activities most amenable to inferencing, evaluating training and inferencing algorithms, and assessing how reliable the inferencing needs to be in order to be useful are important areas of ongoing and further research. There is a significant body of work going on in machine learning, computer vision, speech and signal processing, and generally in context-aware computing to address many of these issues.

Knowing something about a user's context can greatly enhance the ability of invisible interfaces to behave in predictable or anticipated ways by providing data to help disambiguate user intentions. A combination of sensor-based data for activity inferencing and user supplied training data (used to establish ground truth) are proving to be interesting and useful techniques. This combination of system log data and user supplied field data are also a crucial component of more general evaluation strategies.

EVALUATION TECHNIQUES AND METHODOLOGIES

Ubiquitous computing and invisible interfaces pose particular challenges for evaluating whether or not technologies and applications are usable or successful. Traditional empirical studies can assess whether or not individual technologies work reliably and predictably in controlled situations. However ubiquitous computing typically requires a collection of technologies to work in concert and thus isolating, modeling, and evaluating individual components does not provide good indicators for whether or not the more complex system will be usable. Ubicomp technologies are used in dynamic contexts and in changing locations, where demands on the users' visual and mental attention varies dramatically and somewhat unpredictably. These complexities are compounded by the lack of design guidelines and interaction models to guide in developing usable and predictable "invisible interfaces" (if there is no interface per se, how do you "design" it?).

Evaluation methodologies have evolved to combine controlled laboratory studies and Wizard of Oz prototypes with in-situ field methods. In addition to ethnographic studies (observed user behavior) and diary or journal studies (self-reported user behavior), new techniques and metrics are being tested and applied [e.g., 4, 15, 16]. Ideally, these capture data that are in-situ, involve multiple participants, take place over (sometimes extended) periods of time, and are quantitative and qualitative in nature. Most notably, experience sampling methods (ESM) are eliciting user responses from the field by using sensors, inferred events, or contextual inferences to time prompting for users to answer questions delivered over mobile devices. These questions can thus be tuned to fit the nature of the inferred situation and to increase the likelihood of responses (because they are timed for less disruptive moments).

We are further interested in exploring methods for quantifying and measuring not only usability/predictability but also effort. While still in its early stages, we are investigating the application of subjective and objective measures of mental and physical workload such as those typically applied in traditional Human Factors Engineering, most notably the NASA Task Load Index, SWAT (Subjective Workload Assessment Technique), and SWORD (Subjective Workload Dominance). We are hoping that the metrics captured by these techniques may usefully quantify key factors in ubicomp applications and technologies that enable us to compare, contrast and systematically assess new approaches in design.

SUMMARY

I wish to summarize by highlighting some questions and issues that I believe the research community needs to address. Does categorizing the extent to which an interface is embodied in an object help us in formulating design principles, interaction models and evaluation metrics? How do we define and go about designing "invisible interfaces" if interaction mechanisms aren't visible? What is the interaction model? What is the role of user's context and how do we best make use of that? How accurately do we need to infer context? What do we communicate to the user about what we are inferring, when we infer, where we infer, and whether it is correct? What kinds of evaluation methodologies will most help us in assessing new usage patterns, new technologies, and invisible interactions? Will this evolution of technology result in outcomes that are "digitally simplistic"? Should this be an aspiration? If not, what are the measures for success?

ACKNOWLEDGMENTS

Much of my prior work was done at Xerox PARC in collaboration with Ken Fishkin, Roy Want, Anuj Gujar, and Tom Moran. Some of my work on activity centered design and mobile devices was done at IBM Almaden Research Center in collaboration with Tom Moran, Shumin Zhai, and John Barton. Most recently I have been working with researchers at Intel Research Seattle, including James Landay, Anthony Lamarca, Sunny Consolvo, Ian Smith and Matthai Philipose. The efforts and influences of these collaborations are gratefully acknowledged.

REFERENCES

1. Bandai Co. Ltd, the Tamagotchi Page. http:// www.urban.or.jp/home/jun/tama/
2. Brave, S. Ishii, H. and Dahley, A. Tangible Interfaces for Remote Collaboration and Communication. *Proc. of CSCW'98*, ACM Press, 1998, pp. 169-178.
3. Brooks, R. A., "The Intelligent Room Project," *Proc. of the 2nd International Cognitive Technology Conference*, IEEE CS Press, 1997, pp.217-8.
4. Consolvo, S. and Walker, M. Using the Experience Sampling Method to Evaluate Ubicomp Applications, *IEEE Pervasive Computing*, April-June 2003, pp. 24 – 31.
5. Dey, A. K., Kortuem, G., Morse, D. R., Schmidt, A. (Eds) Special Issue on Situated Interaction and Context Aware Computing, *Personal and Ubiquitous Computing*, 2001, Vol 5, Springer-Verlag.
6. Fishkin, K. P., Gujar, A., Harrison, B. L., and Want, R. Embodied User Interfaces for Really Direct Manipulation. *ACM*, Sept. 2000, Vol.43 No.9, pp. 75-80.

7. Fishkin, K. P., Moran, T., and Harrison, B. L. Design Principles for Manipulative User Interfaces. *Proc.s of Engineering for Human-Computer Interaction*, Heraklion, Crete, 1998.

8. Fitzmaurice, G. Situated Information Spaces and Spatially Aware Palmtop Computers, *CACM*, 36(7), July 1993, pp.38-49.

9. Fitzmaurice, G., Ishii, H., and Buxton, W. A. S. Bricks: Laying the Foundations for Graspable User Interfaces. *Proc.s of CHI'95*, ACM Press, 1995, pp. 422-449.

10. Friedman, B., Kahn, P. H., Hagman, J., and Severson, R. L., The Watcher and the Watched: Social Judgements about Privacy in a Public Place, In press. 2005. To appear in the *Human Computer Interaction*. 2005.

11. Gorbet, M. G., Orth, M., and Ishii, H. Triangles: Tangible Interface for Manipulation and Exploration of Digital Information Topography. *Proc. of CHI'98*, ACM Press, 1998, pp. 49-56.

12. Harrison, B. L., Cozzi, A., Moran, T. P. Roles and Relationships for Unified Activity Management, *Proc. of Group 2005*, ACM Press, 2005, pp. 236-245.

13. Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C., and Want, R. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces, *Proc. of CHI'98*, ACM Press, 1998, pp. 17-24.

14. Hinckley, K., Pausch, R., Goble, J. and Kassel, N. Passive Real-World Interface Props for Neurosurgical Visualization, *Proc of CHI'94*, ACM Press, 1994, pp. 452-458.

15. Intille, S. S., Bao, L., Tapia, E. M., Rondoni, J. Acquiring In Situ Training Data for Context Aware Ubiquitous Computing Applications, *Proc of CHI 2004*, ACM Press, 2004, pp. 1-8.

16. Intille, S. S., Rondoni, J., Kukla, C., Ancona, I., and Bao, L. A Context-Aware Experience Sampling Tool. Posters in *Proc. of CHI 2003*, ACM Press, 2003, .pp. 972-973

17. Ishii, H. and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms. *Proceedings of CHI'97*, ACM Press, 1997, pp. 234-241.

18. Kaminsky, , Dourish, P., Edwards, K, Lamarca, A., Salisbury, M., Smith, I. SWEETPEA: Software tools for programmable embodied agents. *Proc. of CHI'99*, ACM Press, 1999, pp. 144-151.

19. Marti, S. and Schmandt, C. Physical embodiments for mible communication agents. *Proc. of UIST 2005*, ACM Press, 2005, pp. 231-240.

20. Moran, T. P., Saund, E., van Melle, W., Bryll, R., Gujar, A., Fishkin, K., Harrison, B. L. The Ins and Outs of Collaborative Walls: Demonstrating the Collaborage Concept. Short paper in *Proc.s of CHI'99*. ACM Press. 1999, pp. 192-193.

21. Moran, T. P., Saund, E., van Melle, W., Gujar, A., Fishkin, K., Harrison, B. L. Design and technology for collaborage: collaborative collages of information on physical walls. *Proc. of UIST'99*. ACM Press, 1999, pp. 197-206.

22. Moran, T.P., and Dourish, P. (Eds). Special Issue on Context Aware Computing, *Human Computer Interaction*, 16, 2-4 (2001). ACM Press. NY, NY.

23. Murakami, T., Hayashi, K., Oikawa, K., and Nakajima, N., DO-IT: Deformable Objects as Input Tools, *SIGCHI '95 Conference Companion*, ACM Press, 1995, pp. 87-88.

24. Mynatt, E. D., Ugarashi, T., Edward, K., and Lamarca, A. Flatland: New dimensions in office whiteboards. *Proc of CHI99*. ACM Press, 1999, pp. 346-353.

25. Philipose, M., Fishkin, K. P., Perkowitz, M., Patterson, D., Fox, D, Kautz, H., and Hahnel, D. Inferring Activities from Interactions with Objects. *IEEE Pervasive Computing*, Issue: Oct-Dec 2004, pp. 10-17.

26. Rekimoto, J. and Saitoh, M. Augmented Surfaces: A Spatially Continuous Workspace for Hybrid Computing Environments. *Proc of CHI99*, ACM Press, 1999. pp.378-385.

27. Rekimoto, J., Tilting Operations for Small Screen Interfaces. *Proc. of UIST '96*, ACM Press, 1996, pp.167-168.

28. Small, D., and Ishii, H. Design of Spatially Aware Graspable Displays. *Extended Abstracts of CHI'97*, ACM Press, 1997, pp. 367-368.

29. Streitz, N. A. and Russell, D. M. Basics of Integrated Information and Physical Spaces: The State of the Art. *Proc. of CHI'98*, ACM Press, 1998, pp. 273-274.

30. Streitz, N. A. Integrated Design of Real Architectural Spaces and Virtual Information Spaces. *Proc. of CHI'98*. ACM Press, 1998, pp. 263-264.

31. Streitz, N. A., Konomi, S., and Burkhardt, H.-J. Cooperative Buildings: Integrating Information, Organization, and Structure. *Proc. from the 1st International Workshop CoBuild'98*, Springer-Verlag. 1998.

32. Suchman, L. A. *Plans and Situated Action: The Problem of Human Machine Communication*. Cambridge University Press. 1987.

33. Underkoffler, J. and Ishii, H. URP: A Luminous Tangible Workbench for Urban Planning and Design. *Proc.s of CHI'99*, ACM Press, 1999, pp. 386-393.

34. Wan, D. Magic Medicine Cabinet: A Situated Portal for Consumer Healthcare, *Proc. Of 1st Int'l Symp. Handhel and Ubiquitous Computing (HUC9*9), Springer-Verlag, pp. 352-355.

35. Want, R., Fishkin K., Gujar, A., Harrison, B. L., Bridging Physical and Virtual Worlds with Electronic Tags. *Proc. of CHI '9*9, ACM Press, 1999, pp. 370-377.

36. Want, R., Schilit, B. N., Adams, N. I., Gold, R., Petersen, K., Goldberg, D., Ellis, J. R., and Weiser, M. An Overview of the ParcTab Ubiquitous Computing Experiment. *IEEE Personal Communications*, December 1995, pp. 28-43.

37. Weiser, M. The Computer for the 21st Century. *Scientific America*, *265*(3), 1991, pp. 94-104.

38. Wellner, P. Interacting with paper on the Digital Desk, *CACM*, *36*(7), July 1993, pp. 86-96.

39. Wellner, P. tactile manipulation on the DigitalDesk. Video in CHI'92 Special Video Program, *ACM SIGGRAPH Video Review 79*, 1992.