

Embodied User Interfaces: Towards Invisible User Interfaces

Kenneth P. Fishkin, Thomas P. Moran & Beverly L. Harrison
Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA. {fishkin, moran, beverly}@parc.xerox.com

Abstract: There have been several recent examples of user interface techniques in which the user uses a computational device by physically manipulating the device. This paper proposes that these form an interesting new paradigm for user interface design, *Embodied User Interfaces*. This paper presents and defines this paradigm, and places it in the evolution of user interface paradigms leading towards the ideal of an invisible user interface. This paper outlines the space of design possibilities in this paradigm, presents a design framework for embodied user interface design, and articulates a set of design principles to guide design.

Key words: gestures, physical manipulation, user interfaces, tactile ui, tangible computing, kinaesthetics, design principles.

1. THE EVOLUTION OF INTERFACE PARADIGMS

There have been a number of recent explorations of emerging user interface paradigms in which the user interacts with a computing device by means of direct physical manipulation. In one paradigm, termed the *Tangible User Interface* by Ishii and Ullmer (1997), the user interacts with a physical object that then controls an "attached" virtual representation. Examples of this paradigm include the "bricks" of Fitzmaurice et al. (1995), the "doll's head" of Hinckley et al. (1994), the "meta-desk" and "phicons" of Ishii and Ullmer (1997), and the "marbles" of Bishop (in Smith, 1995). Alternatively, the user can interact with the computational device itself (usually a handheld device) by tilting it, translating it, rotating it, etc., in which case the manipulation and the virtual

representation are integrated within the same object, a paradigm which we term *Embodied User Interfaces*. Examples of this include scrolling a menu by tilting a device (Rekimoto, 1996), zooming text by pushing/pulling a device (Small and Ishii, 1997), navigating a list by tilting and squeezing a device (Harrison et al., 1998), and changing device behaviour by carrying it to another locale (Fitzmaurice, 1993; Want et al., 1995). These provocative user interface techniques present compelling examples of innovation that attempt to minimise the cognitive distance between a task goal and the human actions needed to accomplish that task. We believe these interaction paradigms are on an evolutionary path towards an ideal of the invisible user interface, as shown in Figure 1.

Keyboard UI —>
 Graphical UI (GUI) —>
 Gestural UI —>
 Tangible UI —>
 Embodied UI —>
 ... *invisible user interface*

Figure 1. Evolution of Computer User Interfaces

As user interface paradigms evolve, new features are added that make the interfaces more direct and more invisible. Figure 2 is a brief look at a few example techniques in the various user interface paradigms and the new features they exemplify.

<u>Paradigm</u>	<u>Example Techniques</u>	<u>New Features</u>
Keyboard	Unix shell	general command language
Graphical	mouse/window/desktop	“direct manipulation” on displayed objects; real-world metaphor
Handles	light pen on display bricks	direct contact with displayed objects multiple generic handles; bi-manual; more degrees of freedom in manipulation
Tangible	MIT dome phicon doll’s head	specific, tightly coupled input metaphor 3-D manipulations
Embodied	LCD lens over meta-desk page-turning tablet	feedback in device itself embodied task; tightly coupled input and output metaphors.

Figure 2. Perceptual-Motor Features of Evolving User Interface Paradigms.

From this figure, we see the emergence of metaphoric representations of real-world objects on displays and the ability to “directly manipulate” these displayed objects. Then, more general control “handles” are introduced providing greater manipulatory freedom of expression. Then “tangible” objects representing the objects of work are introduced; physically manipulating these produces effects on a display and, in one case, on the device itself. There tends to be a progression towards tighter embodiments, more

directness in manipulating the intended object, and more coincidence between input and output. We believe that this reflects a progression towards a more real-world interaction style, where there is no perceived mediation, i.e., an invisible user interface. However, computational systems must mediate (sense and interpret) the actions of the user. In order to provide interaction experiences that are more akin to real world experiences and that leverage everyday skills, we therefore attempt to design these mediators to become as transparent as possible.

In this paper we focus on the Embodied User Interface paradigm, the distinguishing features of which are:

- the *task is embodied* in a device .
- there is *coincidence of input and output* in the device.
- the device *provides highly specific and familiar affordances* for particular kinds of actions.

In essence, *all* aspects of the interaction are embodied in the device. At this point in the paper, we are only considering the perceptual-motor dimension of UIs, in which they are regarded as a means to manipulate objects. There is another dimension of UIs in which they are regarded as a means of communication between the user and the system. Even the most “invisible” UIs have a communication aspect to them, because systems must “know” (by sensing and interpreting) when and how the user is communicating with them. This aspect will arise later in this paper in the Design Principles section, where some communicative principles appropriate to Gestural, Tangible, and Embodied interfaces are presented.

Early user interface paradigms such as GUIs have been systematically and extensively studied, and design principles and evaluative criteria have been derived and documented (e.g. Preece et al., 1994). However, the recent paradigms, such as Tangible and Embodied User Interfaces, have thus far only been studied via isolated exploratory design probes into a potentially large and complex design space. Indeed, until very recently there has been no notion of Tangible and Embodied User Interfaces *existing* as coherent classes. This paper attempts to remedy this lack of analysis of the paradigm of Embodied User Interfaces by presenting, in Section 2, a framework for designing Embodied UIs and illustrating it with two examples of Embodied UIs. In Section 3, we show that the space of manipulations on Embedded UIs is enormous, which raises a multitude of design issues to be dealt with. In Section 4, we present a set of design principles for Embedded UIs, many of which are applicable to other UI paradigms. It is our hope that this framework and these principles will inspire a more systematic and principled investigation of novel and emerging user interface techniques and contribute to a more solid foundation for pursuing the invisible interfaces of the future.

2. A FRAMEWORK FOR DESIGNING EMBODIED USER INTERFACES

The design features presented above imply that the user's task environment should be *embodied* within a physical/computational device, and that this embodied task should be linked to an analogous *real-world* task. The stronger this coupling, the more natural and pervasive the metaphor(s) involved, the more naturalistic and transparent the interaction becomes. Ideally, these invisible interfaces reflect high degrees of embodiment, coincidence, and appropriate manipulations, and thus enhance the richness and intuitiveness of the interaction experience.

<u>Real-World Task</u>	<u>Device-Embodied Task</u>
Physical objects	Physical device
Real properties of objects	Represented objects & properties
Manipulations on objects	Manipulations on device <i>sensors on device</i> <i>input grammar and</i> <i>interpretation</i>
Feedback from objects	Feedback from device and representation

Figure 3. Design Framework for Embodied Uis.

In the real world, we wish to perform some task using a given physical object. That object has a set of real physical properties. These properties include the physical affordances (some manipulations are more natural to perform than others), the expected effects of certain manipulations, the atomicity of the object, and so forth. Some set of physical manipulations of the target object is performed to perform the task. We expect feedback, in a variety of forms, during and after the manipulation. Several items not required in real world interactions are necessary in the device interactions: sensors, a grammar, and an interpretation. These reflect the fact that a mediating agent of some sort must detect manipulations, form a command sequence based on these, and then interpret that command sequence.

We now illustrate this framework, and ground our presentation, by presenting two examples of Embodied interfaces from our own recent work (Harrison et al., 1998). We then proceed from the specific back to the general, analysing each component of the framework and presenting design principles appropriate to each.

Example 1. Navigation Within a Document or Book

In the first example, the Embodied Task consists of page-by-page navigation through a multi-page document, only one page of which can be displayed on the screen of a pen or tablet computer at a time. A book was chosen as the most appropriate physical object to map to this device object. While a tablet computer isn't a perfect match for a book, it does share a number of object properties: a flat surface for displaying text, a sequential organisation of displayed pages, obvious start/end points, and the user expectation that the device is often used for displaying and navigating text.

To turn to the next page in a book, users flick on the upper right corner from right to left (Figure 4a). To turn to the previous page, they flick on the upper left corner from left to right. To keep the interface embodied, coincident, and afforded by real-world expectations, our implementation therefore supported both of these manipulations (Figure 4b) on the device.



Figure 4. (a) Page-by-page navigation, real world and (b) virtual embodiment.

As a page turns in a book, feedback is provided visually (pages move, the new destination page shows, a new relative location shows), aurally (the sound of pages turning), and kinaesthetically (tactile pressure on finger or thumb, tactile feedback of pages moving or sliding). In our implementation, therefore, visual feedback mimicked real world feedback; pages changed with animation, and after a page turning manipulation, both the page number and the contents changed to reflect the new page. A “page-turning” audio cue, similar to the real-world sound, was also provided. However, no kinaesthetic feedback was provided.

The framework therefore guides the linkage between the real world task and an embodied task, modified by the constraints imposed from particular device choices, manipulations, and feedback from the user interface. Implementation constraints and sensor constraints are introduced as a result of having to use a mediating agent. This embodied manipulation requires that the left and right upper corners detect a finger press, the direction of a stroke, and a release of pressure. This was implemented by attaching pressure sensors to the surface of the device. This allowed us to “retro-fit” pressure-sensing technology onto a normally pressure-insensitive device, while also

maintaining the “book” metaphor – the book stroke gesture is typically made on a corner, hence also in our implementation.

The interface was therefore specific to the task at hand – each manipulation corresponded to a specific command within one specific application.

Example 2. Navigation Through Sequential Lists

In our second example, the Embodied Task was to scroll through a sequential list displayed on a small handheld computer, a Palm Pilot™. A Rolodex was chosen as the physical target object which would best map to this target object. In this case, then, the real world task representation assumes physically manipulable items or cards, a circular sequential organisation, and a knob that controls the Rolodex (see Figure 5a). While it was impractical for us to glue a knob onto the side of the computer, our implementation did support circular sequential organisation, and displayed virtually manipulable cards of roughly the same size and shape as in the physical Rolodex.

In this real-world task, physical manipulations include turning the knob (with a rate parameter), and stopping at a desired location. The direction of turn determines the direction of list traversal. Our embodied task manipulations were similar. We used card items with visual tabs arranged in a sequence (Figure 5b). Turning the circular list towards the user would begin flipping from A towards Z (assuming an alphabetised list) and vice-versa. On a physical Rolodex, users turn the knob rotationally (at some rate of speed) (Figure 5a), tilting the wrist. Similarly, on the Pilot™, the user manipulation was a wrist tilt movement away from a neutral resting position. Extent or degree of tilt corresponded to rate or speed of turning. Turning harder moves faster through the list, similar to (Rekimoto, 1996). To stop at a particular item, the user either ceases to tilt (i.e., maintains the list container in a neutral position), or squeezes the device, (roughly akin to grasping the Rolodex card).



Figure 5. (a) list navigation, real-world, (b) virtual embodiment.

In the real-world task, visual feedback includes the flipping of cards, and the display of the new card. Auditory feedback is the sound of the cards flipping. Kinaesthetic cues include finger pressure, extent of rotational movement, and direction of rotation. Due to time constraints, our implementation contained only some visual feedback – users often commented that this lack of feedback affected the “naturalness” of the interface, which we regard as more evidence in favour of this as a design component.

The mediating agent used to implement the tilting manipulation was a tilt sensor, attached to the top of the Pilot, parallel to the plane of the display. As the device is angled towards or away from either plate, the amount of fluid in contact with the plate varies and impedance varies accordingly. We were therefore able to use the sensor readings as a crude measure of the computer’s orientation relative to gravity.

In the real world, the rotation manipulation is sufficiently gross (a significant wrist action), and the penalty for a “false positive” is sufficiently small, that no “safety” or “clutch” is necessary to protect against flipping a Rolodex past the target. However, in our embodied task, this was not the case — the manipulation was rather small, and the penalty for scrolling when the user did not wish to scroll was significant. Accordingly, the implementation required a second manipulation, signalling an enabling/disabling of the scrolling feature. In our case, we decided to use an initial squeeze of the device to indicate the desire to navigate through the list, followed by a second squeeze to “grasp” the desired item, thereby ending the navigation task. To avoid muscle stress, users did not have to maintain the squeezing pressure during navigation. The device was padded with foam to further suggest squeezing capability. By mismatching the kinaesthetic

motion scales, therefore, we had to backtrack from invisibility in other aspects of the design.

3. DESIGN ISSUES IN THE SPACE OF MANIPULATIONS

These two sample application developments illustrate the essential features of an embodied user interface: embodiment, coincidence of input and output, and the use of specific metaphorically appropriate manipulations. We believe they also illustrate the need for a more disciplined design process and the creation of a set of articulated design principles appropriate for this space. The design space is enormous, and confronts us with a host of subtle design issues. To illustrate the size of the design space, consider the number of design parameters that may vary for each manipulation:

- Type of manipulation. Manipulations may be spatial (the device is translated, rotated, tossed, spun, shaken, flipped, etc.), structural (the device is squeezed, folded, curled, etc.), or environmental (the device is heated, lit, etc.). Many manipulations exist, each with their own affordances, user expectations, and kinaesthetic profiles.
- Portion of device affected. The manipulation may affect the entire device, or can be restricted to any arbitrarily small portion or portions of the device. For example, conventional keyboards, by looking at the portion of the device affected (which keyboard button is spatially depressed), map each button to a different semantic effect.
- Body part(s) employed. While we envision most manipulations as being performed by one or both hands, many other possibilities exist. Not only may other body parts (elbow, arm, feet) be employed, but portions of said body parts (e.g. fingers) could also be employed, and semantics can differ depending on which portions of which body part(s) are employed. For example, Harrison et al. (1998) shows a handwriting device that alters its display depending on which hand is used. Furthermore, implements other than the body, such as pens, rods, etc can perform the manipulation.
- Magnitude. The manipulations may be gross in nature (e.g. carrying a device across a room), but in general the magnitude of the manipulation can be arbitrarily small, a barely perceptible nudge or a tiny increase in pressure.
- Simultaneity. A manipulation can take place either singly (a device is tilted), in temporally simultaneous groups (a device is tilted while squeezing), or sequentially (a device is shaken, then tilted, then brought to eye level).
- Direction. Each manipulation has other parameters, corresponding to the physical nature of the manipulation: the vector of a translation, the angle of a rotation, the displacement vector of a depression, and so forth. All of these parameters can be mapped to varying semantic effects.

In general, how are designers to choose which subset of this great space of manipulations should be employed? How does one choose the real world task upon which the embodied task is based? When is specificity useful (versus generality)? If a device is to support a range of operations, and a range of manipulations, what inter-manipulation principles apply? Tangible and Embodied designs have thus far answered these questions by inspiration and intuition on a case-by-case basis. To attempt to redress this, we now discuss specific design principles that address some of these issues.

4. DETAILED DESIGN PRINCIPLES

Returning to the framework of Figure 3, we now analyse each component for a more detailed set of design principles to guide the design and assessment of interfaces that strive towards invisibility. While these design principles are targeted at Embodied interfaces, some are also applicable to neighbouring paradigms from the spectrum presented in Figure 1.

4.1 Embodiment Principle

The notion of mapping has often been used in user interface design theory, with principles presented to guide the mapping from real world tasks to internal representations of the tasks within the system (Moran, 1983). In physically manipulable interfaces, the mappings are more specific, relying on the physicality of the device and the user's handling of it. In most real-world situations, manipulation of an object produces changes in the same object. Therefore, for the most consistent mapping, the elements of the embodied task should be represented by parts of the device's hardware or on the device's displays (which can be visual, auditory, kinaesthetic, etc.).

The various states of the task need to be embodied in the device, so the user can perceive the states of the task and, further, can think of the task as physical transformations of the parts of the device representing elements of the task. The user can then affect the physical transformations by directly physically manipulating all or part of that same device. The adherence to the Embodiment principle is precisely what distinguishes Embodied from Tangible interfaces.

4.2 Physical Effects Principle

Given a particular embodied task, the designer has a choice of many manipulations for the user to effect the task. HCI theory has argued for a consistent mapping from task to action (e.g., Moran, 1981; Payne and Greene, 1986) for user interfaces. Studies on human gestural systems (Skelly, 1979) stress the importance of direct physical pantomime in making these mappings natural to use and remember. As the Tangible and Embodied interface paradigms employ such physical manipulations, we propose that the

best manipulations can be conceptualised as physically transforming the device in a way that accomplishes the physical transformation of the embodied task. The expected physical effects of the manipulation account for the task accomplished. We term this the **Physical Effects Principle**.

For example, consider an Embodied Task of telling a computer to compress a file. In the physical world, one way things are compressed is by squeezing them. In this task, the file is embodied on the display, so the Physical Effects Principle suggests that the device's case should be squeezed to signal the act "compress". The logic of the design is that the casing surrounds the display, and squeezing the casing causes the device, and its contents, to be compressed.

The Physical Effects Principle applies to tasks that extend beyond the boundaries of the device. For example, consider the task of sending email. The email message is embodied on the display of a wireless device. In the physical world one way things are sent off a surface to which they are attached is by flicking. Therefore, since the message is embodied as an entity on the surface of the display, "send" could be signalled by flicking the device, thereby flicking the embodied object into the air. If there are various nearby devices that could serve as the destination of the message, then the direction of the flick could indicate which device is to receive the message.

4.3 Metaphor

There are times when obvious appropriate direct physical motions don't apply. In this case, user interfaces have traditionally relied on the power of metaphor. For example, in GUI interfaces, the user manipulates windows, but they are only "metaphorical" windows. They aren't made of glass; they can be dynamically resized, and so forth. However, they still appeal to an underlying physical real-world metaphor, that of "windows" into different virtual worlds. We therefore extend the Physical Effects Principle: when there are no direct physical interpretations for manipulations to do a task, then the next best manipulations are those that appeal to a metaphor.

For example, consider a computer that can compile C++ files. There is no obvious direct real-world physical manipulation that causes things to be compiled. However, there are metaphors to encompass this. For example, we can think of compiling as a process in which "raw material" is transformed into a more "digestible" state, like a meat grinder with a rotary handle on its side. By tightly orbiting the device around some external point, "turning the crank", the embodied task is signalled metaphorically.

There will be times where no natural physical or metaphorical analogy obtains. For example, suppose our compiler can compile from a number of languages: C++ is only one of many. The manipulatory sequence must now include a manipulation that symbolises "C++". There is no obvious physical or metaphorical analogy to this word: some arbitrary manipulation must be assigned. However, we believe that in practice this will rarely be the case. For example, sign languages routinely use both physical and metaphorical analogies in their mappings and rarely resort to truly abstract mappings. Furthermore, this worst case is no worse than the average case of existing user

interfaces. For example, the interface can have two buttons, one marked “C” and another marked “+” (others are possible...), and the manipulation for “C++” can be depression of a portion of the device (the “C” button), followed by two sequential depressions of another portion of the device (the “+” button).

Sets of related tasks should be considered together during design, so that metaphors can be shared across tasks to avoid confusion between them. For example, consider the task of randomising a list of items. The items are embodied on the device display as a vertical list. A manipulation for randomising the list might be to shake the device back and forth, the logic being that the physical transformation is to mix the items on the display. Now consider the additional task of sorting the list in ascending or descending order. A natural manipulation for this task might also be a shaking action. However, ambiguity then exists between the “randomise” and “sort” operations. A single shared metaphor can be achieved by decomposing the tasks into two parts: moving the items, and arranging them in a certain order. A shaking manipulation is appropriate to indicate the first part. Tilting can specify the ordering: an upward tilt for an ascending sort, a downward tilt for a descending sort, and no tilt for randomisation. The metaphor is that the “heavier” items will sink to the bottom.

4.4 Kinaesthetic Manipulation Principles

Embodied interfaces, like all interfaces, signal device operations via a set of kinaesthetic motions on the part of the user. Various aspects of good kinaesthetic design have been well studied. Their lessons are particularly germane to Gestural, Tangible, and Embodied interfaces, as the possible range of kinaesthetic motions and muscles used is so great. Some principles from these studies that apply here are:

Comfort: The best manipulations are the ones that can be comfortably performed. The range of comfortable physical movements is determined by established guidelines in anthropometry and physiology (e.g. avoid movements that stress joints such as putting both hands behind the back). Furthermore, the sequence of manipulations should be designed such that they comfortably flow into each other. Sturman (1992) discusses this for the specific case of kinaesthetic sequences involving the hand.

Comfort applies to a single manipulation, a manipulation sequence, and also to the entire sequence performed throughout the course of a day – carpal tunnel, tendonitis, and other woes of the modern era must be avoided. By restricting the device areas manipulated, and the magnitude of said manipulations, we can hopefully create user interaction sequences no worse than, and potentially better than (due to the increased degrees of freedom) existing user work practices.

Appropriate Modifiers. These manipulations can be parameterised in many ways: the speed, strength, force, and area of their affect, for example. These modifiers may have varying precision, degrees-of-freedom, and temporal duration. The corresponding modifications to the kinaesthetic motion should be a good match. For example, a modifier with a fine precision (e.g. the amount by which to resize an image) shouldn’t be bound to a coarse motion such as wrist motion (Sturman, 1992). The list-navigation

interaction described earlier in this paper suffered from lack of attention to this design principle: a modifier with a coarse precision (move to next element in list) was bound to a very fine motion (a slight wrist tilt).

Roles of the Hands. Humans use both hands working in parallel to express a single concept at a time. This is done either by having both hands perform the same action, or with the non-dominant hand setting the context for the dominant hand, as described by Guiard and Ferrand (1995) and applied by Buxton and Myers (1986). While interfaces could be designed in which the two hands operated in parallel to express multiple concepts (one per hand), such use simply does not seem to be natural for human users, whether in explicit communicative systems or in spontaneous gestures accompanying speech (McNeill, 1992). Therefore, an interface should not be designed to support, let alone require, this type of operation.

Socio-Cultural Factors. Given the context for our application domain, i.e., the design and use of portable computational devices, we must be aware of the limitations of the physical environment. For example, such devices may be used in locales where expansive gestures or manipulations that might be intrusive or embarrassing must be eliminated. The expansiveness of a gesture can be limited to the range of space that limits one's personal zone, as determined by proxemics (Hall, 1966). Potentially embarrassing or culturally problematic gestures or manipulations can be based on current cultural anthropology norms (Morris, 1977). These are more difficult to address since the cultural meaning of a gesture varies with location.

4.5 Sensing Principles

The best-designed manipulation is worthless if it can't be properly sensed. There are a number of constraints introduced by the sensing technology and the user's conception of that sensing technology. In an ideal setting, the sensing technology will exist completely transparently, such that users are unaware of how and when sensors activate. However, realistically this is not generally the case, for it takes a great deal of care to implement natural-feeling gestures.

4.6 Communication Principles

The design principles presented thus far have focused on the Perceptual-Motor features of user interface design. There is another axis of analysis, however, in which user interfaces are regarded as mediating agents for communicative dialogues. As Card et al. (1991) put it, "the design of human-machine dialogues is, at least in part, the design of artificial languages for this communication." We now consider manipulations on devices as a communicative language designed to impart a command sequence to a device, and we seek to apply some principles from communications theory and linguistics to further aid interface design.

This gestural command language, as sensed by the computer, is a sentence consisting of gestural units, each unit representing a single word (noun, verb, adverb, etc.) in a

command sequence, each unit derived from a set of manipulations. As discussed earlier, we may assume that such words don't temporally overlap (manipulations overlap, gestures don't). The sentence may start with a 'START' symbol, and end with a 'STOP' symbol.

While design principles for user interfaces conveying information via text or speech have been well studied, we are unaware of literature analysing interfaces such as these in which information is conveyed via manipulation. However, there is a great body of literature from other disciplines that analyse gestural systems created to communicate information kinaesthetically. While none of these systems are exactly equivalent to manipulative systems, there may be lessons we can learn from them by extracting principles that hold in common across various kinaesthetic systems. Such systems include:

Spoken Language-Based Sign Languages. Gestural languages have been developed in which arbitrary messages may be formed. The syntax and grammar of these languages is borrowed, in part or whole, from a spoken language known to the gesturers. Examples of these languages are American Sign Language (ASL) (Butterworth and Flodin, 1991), and the sign languages of Australian Aborigines (Umiker-Sebeok and Sebeok, 1978), and Cistercian monks (Barakat, 1975).

Home Systems. When deaf or deaf-mute children are not taught a sign language, they develop their own gestural systems, called home systems. These systems are totally free from any spoken language, as the children know none. While each child develops their language in isolation, studies have shown (Feldman, 1975; Goldin-Meadow, 1975; Goldin-Meadow, 1979) a high level of similarity in the languages they develop.

Plains Sign Language. The most sophisticated and successful gestural language is the American Indian Sign Language, or Plains Sign Language (PSL), which has received significant study (Clark, 1959; Mallery, 1972; Skelly, 1979; Taylor, 1975; Tomkins, 1969; Umiker-Sebeok and Sebeok, 1978). PSL is at least 500 years old, and at its peak was used by hundreds of thousands of users, who spoke over 500 spoken languages from over 70 linguistic families. The language does not owe its grammar to any particular spoken language. Indeed, it evolved in order to facilitate communication amongst tribes with no common language.

Start Signal. In all of these systems, people send signals that they are about to communicate. However, these are *extra*-linguistic cues such as body stance, eye gaze, and social context. In the case of an Embodied interface, the need to detect a start symbol is particularly acute, just as it is in computerised systems with audio input, as there is no restricted location that must be used for the communication. Computers aren't as good at detecting extra-linguistic cues as humans. Therefore, if an explicit linguistic start symbol is employed, it should try to inherit the desirable properties of the extra-linguistic cues. Namely, the gesture for the start symbol should flow easily into, yet be distinct from, the gesture space used for other symbols, something like the wrist-flick suggested by Sturman et al. (1989), or the squeeze used by Harrison et al. (1998).

Stop Signal. In all of these systems, an explicit stop signal is rarely employed: there is no analogue to the "over" used by simplex communication protocols. Instead, a temporal pause indicates that a sentence is complete. In the rare cases when an explicit

stop symbol is used, it's simply to provide emphasis (Mallery, 1972). We propose, therefore, that interfaces that support multi-gesture sequences should also use this protocol: a sentence should be considered complete after a temporal pause. This implies that an interface should not act immediately upon receipt of a gesture, but rather delay to make sure the sequence is complete. In this case the device must give continual feedback so the user knows which parameters have been received. This is particularly significant in light of the preferred gestural sequencing (see below), in which the user may optionally transmit modifiers by successive gestures.

Appropriate Linguistic Units. Earlier, we proposed that each gesture in an interface should represent known linguistic units such as nouns, verbs, adverbs, and adjectives. One can imagine systems in which this is not the case, in which the gestural units support some other structuring. However, even in home systems, where the children don't know any spoken language, individual gestures still represent linguistic units of these types. It appears that there is something deeply natural about conceptualisation at this level. Therefore, we propose that an interface also be structured to support this level of conceptualisation.

Gestural Sequencing. A complicated command can be given by arranging its component gestures in a temporal sequence. We now discuss design principles for the ordering of this sequence. Consider the home systems and PSL, in which grammars have evolved from scratch. Perhaps the dominant “meta-rule” of these grammars is to transmit a sentence in the fewest gestures possible. Specifically, they are structured such that the most significant information is transmitted first, with successive gestures progressively refining the sentence into finer and finer detail. The sentence is ended whenever a sufficient amount of detail has been transmitted. The literature refers to this as a “taxonomic” (Taylor, 1975) or “telegraphic” (Skelly, 1979) protocol. To computer scientists, this is recognisable as a “big-endian” communications protocol, in which the most significant bits are transmitted first.

If we consider a gestural sentence as indicating that an operation *op*, plus its parameters, affect an embodied virtual object *affectee*, then home systems have the basic grammar (paraphrasing Goldin-Meadow, 1975):

[*affectee*] [*op*] [*param*₁] [*param*₂] ... [*param*_{*n*}] STOP

Every element in this grammar is optional. Gestural systems make heavy use of defaults, only expressing the information that is necessary. This suggests that an embodied interface should be similarly structured.

Each parameter consists of one noun phrase. In both home systems and PSL, the modifiers to a noun phrase occur after the noun, in accordance with the “big-endian” principle: “big house” is transmitted as “house big”.

Putting these implications together, consider an interface that is used to manipulate graphical images. The command “using a Gaussian filter of width 3 (the default), scale image A about the x-axis by 120%” should be transmitted as: image-A scale 120 X-axis gaussian. The affectee “A” comes first, then the operation “scale,” then the parameters.

Each parameter transmits itself in increasing level of detail. The gestural sequence could end at any point.

5. CONCLUSIONS

This paper has proposed that recent user interface designs have represented isolated probes into a rich, largely unexplored paradigm for user interface design: that of Embodied Interfaces. Furthermore, we feel that this paradigm can be placed on a continuum with other paradigms, and that this points the way towards future paradigms whose goal will be making the user interface more and more invisible. We believe that it is important to begin to try to understand the paradigm to lay the groundwork for more disciplined design. We have presented a conceptual framework in which to place the various aspects of design in this direction, and then attempted to enumerate a set of design principles within that framework. From this investigation, we have come to appreciate that different theoretical viewpoints are needed to encompass the range of issues confronting designers.

This is only the first step in understanding the design of directly manipulable computational artefacts. We will continue to explore the design space by implementing and testing more devices to gain experience in designing, using, and assessing new user designs within this space.

At the same time, we want to broaden and deepen our analysis and understanding of the nature of the design paradigm. For example: just how important is specificity of technique, as opposed to generality? How are multimodal interaction techniques (Oviatt and Wahlster, 1997) best integrated? The answers to these can only be found by proposing different theoretical formulations.

This paper attempts to identify and understand this new user interface design paradigm. We hope that it provokes others to think systematically about where current research in user interface design is taking us, and what the design principles might be.

ACKNOWLEDGMENTS

The breadth of the Bibliography reflects the debt the authors owe to Lisa Alfke and Maia Pindar of the PARC information center, whose tireless scouring of the nation's libraries is appreciated greatly. We thank the members of the PARC Personal Document Reader effort, led by Polle Zellweger, and the "eXtreme team" (Roy Want, Anuj Gujar, and Carlos Mochon) for their support. Thanks also to Rob Burtzlaff, PARC patent attorney, for his help in clearing this paper for release.

REFERENCES

- Barakat, R.A. "The Cistercian Language: A Study in Non-verbal communication". Cistercian Publications. Kalamazoo MI, 1975.
- Butterworth, R. and Flodin, M. "The Perigee Visual Dictionary of Signing", Perigee, New York, 1991.
- Buxton, B. and Myers, B. "A Study in two-handed input", Proceedings of CHI '86, pp. 321-326.
- Card, S.K., Mackinlay, J.D., and Robertson, G.G. "A Morphological Analysis of the Design Space of Input Devices". ACM Transactions on Information Systems, (2), April 1991, pp. 99-122.
- Card, S.K., Moran, T.P., and Newell, A. "The Psychology of Human-Computer Interaction". Erlbaum Associates, 1983.
- Clark, W.P., "The Indian Sign Language". Hammersly, Philadelphia, 1885. Reprinted 1959 by Rosicrucian Press, San Jose.
- Feldman, H., "The Development of a Lexicon by deaf children of hearing parents, or, there's more to language than meets the ear". PhD. thesis, University of Pennsylvania, 1975.
- Fitzmaurice, G. "Situated Information Spaces and Spatially Aware Palmtop Computers", CACM, Vol. 36, 7, July 1993, pp.38-49.
- Fitzmaurice, G., Ishii, H., and Buxton, W. "Bricks: Laying the Foundations for Graspable User Interfaces", Proceedings of CHI'95, pp. 442-449.
- Goldin-Meadow, S., "Structure in a manual communication system developed without a conventional language model: Language without a helping hand". In H. Whitaker & H. A. Whitaker (Eds), Studies in Neurolinguistics, Vol. 4, New York, Academic Press, 1979, pp. 125-209.
- Goldin-Meadow, S., "The Representation of Semantic Relations in a Manual Language Created by Deaf Children of Hearing Parents: A Language you can't dismiss out of hand". PhD. Thesis, University of Pennsylvania, 1975.
- Guiard, Y. and Ferrand, T. "Asymmetry in Bimanual Skills", in "Manual Asymmetries in Motor Performance", Elliot & Roy (eds), CRC Press, Boca Raton FL, 1995, pp. 176-195.
- Hall, E.T. "The Hidden Dimension". Doubleday, 1966.
- Harrison, B.L., Fishkin, K.P., Want, R. Gujar, A., and Mochon, C.. "Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces". Proceedings of SIGCHI '98 (to appear).
- Hinckley, K., Pausch, R., Goble, J.C., and Kassell, N.F. "Passive Real-World Interface Props for Neurosurgical Visualization", Proceedings of CHI'94, pp. 452-458.
- Ishii, H. and Ulmer, B. "Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms". Proceedings of CHI'97, pp. 234-241.
- Mallery, G. "Sign Language Among North American Indians, compared with that among other peoples and deaf-mutes". Mouton, 1972 (reprint).
- McNeill, D. "Hand and Mind: What Gestures Reveal about Thought". University of Chicago Press. 1992.
- Moran, T.P. "The Command Language Grammar: a representation for the user interface of interactive computer systems". International Journal of Man-Machine Studies, 1981, 15, 3-50.
- Moran, T.P. "Getting into a system: external-internal task mapping analysis". Proceedings of SIGCHI '83.
- Morris, D., "Manwatching: A field guide to human behavior". Harry N. Abrams, New York, 1977.

- Oviatt, S. and Wahlster, W., eds. Special Issue on Multimodal Interfaces, Human-Computer Interaction, 1997, vol. 12, nos. 1-2, pp. 1-226.
- Payne, S. and Green, T. "Task-Action Grammars: A Model of the Mental Representation of Task Languages." Human-Computer Interaction, 1986, vol. 2, no. 2, pp. 93-133.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., and Carey, T. (1994). Human-computer interaction. New York: Addison-Wesley.
- Rekimoto, J. "Tilting Operations for Small Screen Interfaces". Proceedings of UIST '96, pp.167-168.
- Skelly, M. "Amer-Ind Gestural Code Based on Universal American Indian Hand Talk". Elsevier, 1979.
- Small, D. and Ishii, H. "Design of Spatially Aware Graspable Displays". Extended Abstracts of CHI'97, pp. 367-368.
- Smith, G.C. "The Hand that Rocks the Cradle". *I.D.*, May/June 1995, pp. 60-65.
- Sturman, D., Zeltzer, D., and Pieper, S. "Hands on Interaction with Virtual Environments". Proceedings of UIST'89, pp. 19-24.
- Sturman, D. "Whole-hand input". PhD Thesis, MIT, 1992.
- Taylor, A.R. "Nonverbal Communications Systems in Native North America", *Semiotica*, 1975, 13, 4, pp. 329-374.
- Tomkins, W. "Indian Sign Language", Dover, New York, 1969. Reprint of 1931 edition.
- Umiker-Sebeok, D.J. and Sebeok, T.A. introduction to "Aboriginal Sign Languages of the Americas and Australia", Plenum Press, New York, 1978. 2 volumes.
- 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

BIOGRAPHY

Kenneth P. Fishkin is a research scientist at Xerox PARC. Since joining PARC in 1991, his research interests have included colour, 2D computer graphics, the Magic Lens user interface, and augmented reality. He holds BS degrees in Mathematics and Computer Science from the University of Wisconsin-Madison, and an MS degree in Computer Science from the University of California-Berkeley.

Thomas P. Moran is Principal Scientist and Manager of the Collaborative Systems Area at Xerox PARC. He was the first Director of the Rank Xerox EuroPARC laboratory in Cambridge. He is the founding and current Editor of the journal *Human-Computer Interaction*. His research with Card and Newell on the theoretical foundations of HCI resulted in the seminal book, *The Psychology of Human-Computer Interaction* in 1983. He has developed analysis tools and theoretical frameworks for HCI (e.g. CLG and QOC), innovative interactive systems (e.g. NoteCards and Tivoli), and collaborative multimedia systems (e.g. RAVE and meeting capture tools).

Beverly L. Harrison is a research scientist at Xerox PARC. Her research interests include augmented reality, the design/evaluation of physical-virtual interfaces, and the

design of media technologies. Prior to joining PARC she worked in product and UI design for Bell-Northern Research and has done extensive consulting in UI design and usability. She holds a Ph.D. and M.A.Sc. in Industrial Engineering (Human Factors Engineering, U. Toronto) and a B.Mathematics (Computer Science, U. Waterloo).