Reality-Based Interaction
and Next Generation User Interfaces

Robert J.K. Jacob

Department of Computer Science
Tufts University
Medford, Mass. USA

CHI 2013 Workshop on Blended Interaction:
Envisioning Future Collaborative Interactive Spaces
**Background: User Interface Software Research**

**Formal specifications for user interfaces**
- Specification (UIDL) → Prototype UIMS
- State transition diagram UIMS
- Coroutine UIDL for direct manipulation

**Specifications, UIMS for next generation**
- Continuous, parallel (non-WIMP)
- Eye movements, Gesture
- Virtual reality
- Retargetable, Handheld
- Framework for next-generation...

**New interaction techniques**
- Eye movements
- 3D gesture
- Virtual reality
- Lightweight/digital library

**New interaction techniques and media**
- Tangible interaction
- Eye movements in VR
- Animatronics
- **Current: Brain-computer interface**...

Third Generation of User Interfaces?

Command Line

GUI, Direct Manipulation
- Shneiderman 1983
- Hutchins, Hollan & Norman 1986
Emerging New Interaction Styles

- Virtual, mixed, and augmented reality
- Tangible user interfaces
- Ubiquitous, pervasive, handheld, mobile interaction
- Lightweight, tacit, passive, or non-command
- Perceptual interfaces
- Affective computing
- Context-aware computing
- Ambient interfaces
- Embodied interfaces
- Sensing interfaces
- Eye-movement based interaction
Goals

• Third generation of user interfaces?
  – Or disparate developments, spreading out in many directions

• What ties them together?
  – Find common elements for understanding, discussing, identifying a 3rd generation

• A lens for analyzing, generating designs
  – Analyze/discuss designs (more later)
  – Provide insights for designers
  – Uncover gaps, opportunities for future research
  – Encourage researchers to consider explicitly
Reality-Based Interaction

• Connect many emerging interaction styles
  – Can be understood together as a new generation of HCI through RBI

• Exploit skills and expectations user already has
  – Make computer interaction more like interacting with rest of world
  – Reduce cognitive burden, training

• Leverage what users know about simple, everyday, non-digital world
  – More so than first, second generations
  – Evolution of GUI
  – “More-direct” manipulation
  – Reduce Gulf of Execution
Some Familiar Examples

• **Navigation in virtual reality**
  – Before: Learned, unnatural commands (keywords, function keys)
  – After: User's native “navigational commands” (position head and eyes, turn body, walk toward target)

• **Augmented/mixed reality, tangible interaction with objects**
  – Simple, transparent mechanical structures
  – Use knowledge of physical world

• **Cell phone, ubicomp, context-aware**
  – Do real world action, computer exploits
  – Operate system in world, UI actions + “real” actions
What is Reality?

• Obviously problematic, broad term

• Specific, narrow definition here
  – Basic aspects of simple, everyday, non-digital world
  – Not keyboards, mice
  – Not cultural, societal, political

• Union of 4 themes:
  – Naïve Physics
  – Body Awareness & Skills
  – Environment Awareness & Skills
  – Social Awareness & Skills

• Fairly basic, though may not be universal across cultures
• People have common sense knowledge about the physical world

• Example: TUI rack or slot as physical constraint on token
Body Awareness & Skills (BAS)

- People have an awareness of their own physical bodies and possess skills for controlling and coordinating their bodies
- Example: Navigate in VR by walking on treadmill
Environment Awareness and Skills (EAS)

- People have a sense of their surroundings and possess skills for negotiating, manipulating, and navigating within their environment.
- Example: Context aware/sensing system respond to user location.
Social Awareness & Skills (SAS)

- People are generally aware of others in their environment and have skills for interacting with them
- Example: VE represent user with avatar, others can respond
Some Supporting Evidence

1. Survey of Published Literature

2. CHI Workshop

3. Informal Field Study
1. Survey of Published Literature

- Retroactively observe designers doing this implicitly
  - When a display surface can sense touch, selecting items by tapping with your finger or a pen is immediately appealing, as it mimics real world interaction [48].
  - For example, in a photo browsing and sorting application, it is natural and convenient to move and rotate virtual photos as if they were real photos lying on the surface, and to support other operations that may be physically impossible but desirable and plausible anyway, such as resizing [32].
  - By moving their two fingers apart diagonally, the user controls the zoom level of the lens visualization ... The amount of zoom is calculated to give the appearance that the tabletop is stretching under the user's fingers. There is an illusion of a pliable rubber surface [17].
  - In this paper, we explore the use of a novel wearable eye pointing device ... Users are also very familiar with the use of their eyes as a means for selecting the target of their commands, as they use eye contact to regulate their communications with others [38].
  - We introduce ViewPointer, a wearable eye contact sensor that detects deixis towards ubiquitous computers embedded in real world objects [38].
  - Systems such as PlayAnywhere are natural platforms for exploring ways to blur the boundary between the virtual, electronic office document and the real thing, as well as scenarios that exploit the natural and familiar feel of manipulating and drawing on paper [32].
  - In eyeLook we modeled our design strategy on the most striking metaphor available: that of human group communication [12].
  - By incorporating eye contact sensing into mobile devices, we give them the ability to recognize and act upon innate human nonverbal turn taking cues [12].
  - When the user is finished examining the details of the underlying dataset, he simply lifts his fingers off the table. At this point, DTLens responds by resetting the local zoom level to its original level. This transition is animated over a period of one second to preserve the illusion of a pliable surface returning to its original state [17].
  - Embodying an agent grounds it in our own reality [29].
  - We developed a new graspable handle with a transparent groove. Our graspable handle enables the user to perform a holding action naturally-the most basic action when physically handling a curved shape in the real world [4].
  - Real appliance's controls are used normally but the user's actions involving these components (looking at a part of the interface, touching a button, etc.) are taken as inputs to the wearable computer which in turn modifies the user's view of the real-world [33].
  - User interface actions are intended to be as natural as possible through the use of a variety of visual affordances. Some of these affordances are derived from equivalent, purely physical interactions that occur with printed photographs. ... To maintain the link with the physical world, users interact only with photographs - there are no buttons, menus or toolbars to be navigated [3].
  - The nature of a tabletop interface makes it very natural to use in a social setting with two or more people [3].
  - By keeping the gesturing behavior more naturalistic we are designing from a more 'mixed ecology' perspective - designing the gesture system such that it approximates natural interactional behaviors as closely as possible [26].
  - ...
2. CHI Workshop

- "What is the Next Generation of Human-Computer Interaction?"
  - Look for common ground
  - Begin with same questions, look for answers

- Review discussions, breakout groups for support or contradiction
  - Most themes identified were closely connected to RBI
  - Expressed in variety of different terminologies
3. Informal Field Study

- Interviewed researchers at MIT Media Lab
  - Had not introduced RBI to them
  - 2 examples...

- Engine-Info: James Teng, Ambient Intelligence Group
  - BAS
  - EAS

- Connectibles: Jeevan Kalanithi, Object-Based Media Group
  - SAS
Implications for Design

• Distinguish 2 claims
  – RBI = Good characterization of next generation
  – RBI = Good UI

• Base on pre-existing real world knowledge and skills
  – Reduce mental effort (already possess some skills)
  – Casual use: speed learning
  – Info overload, time pressure: improve performance
  – NP may also encourage improvisation, need not learn UI-specific skills

• But copy of reality is not enough
  – Make the tradeoff explictly
Reality...Plus Artificial Extensions

• Exact duplicate of real world?

• Real plus extensions
  – Desktop GUI plus "find" command
  – Interact normally plus can turn on X-ray vision
  – Walk and move normally in VR plus can fly by leaning
  – Grasp and move tangible architectural model plus see effect on wind
• Claim: Give up reality only explicitly, only in return for desired qualities
  – *Expressive Power*: Users can perform variety of tasks within application domain
  – *Efficiency*: Users can perform task rapidly
  – *Versatility*: Users can perform many tasks from different application domains
  – *Ergonomics*: Users can perform task without physical injury, fatigue
  – *Accessibility*: Users with varying abilities can perform task
  – *Practicality*: System is practical to develop and produce
Example

• Use conventional walking gesture for walking
  – Give up the reality of the walking command carefully
  – Only if gain added efficiency, power, etc (speed, automatic route finding)

• No conventional gesture for flying, x-ray vision
  – Degrees of realism (x-ray by focus vs. by menu pick)
  – Prefer analogies of realistic for the additional functionality
Case Study 1: URP

- Classic TUI
- Underkoffler & Ishii CHI 99
- NP
- EAS
- BAS
- (SAS)
Case Study 2: Apple iPhone

- Commercial product

- NP
- EAS
- BAS
Case Study 3: Electronic Tourist Guide

- Mobile, context-aware
- Beeharee & Steed, PUC 2007

- EAS
- BAS
Case Study 4: Visual-Cliff Virtual Environment

- Virtual reality
- Slater, Usoh, & Steed, TOCHI 1995; Usoh et al, SIGGRAPH 99
- NP
- EAS
- BAS
More Characteristics of Next Generation

• Higher-level
  – *Reality-Based Interaction*
  – Lightweight, Non-command

• Lower-level
  – Continuous + Discrete
  – Parallel, Highly-Interactive

• Plus maybe
  – Smaller, Bigger, Retargetable
  – Discourse Properties
Lightweight, Non-command

- Emerging common thread, variously called:
  - Passive
  - Context
  - PUI
  - Tacit (Nelson)
  - Noncommand (Nielsen)
  - Affective computing (Picard)
  - Ambient media (Ishii)

- Get more information from user, without much effort from user
  - User not really give explicit commands
  - System observes, guesses, infers, takes hints
Lightweight Inputs

• Inputs:
  – Physiological sensors, affective
  – User behavior
  – Context information, e.g. GPS
  – User commands
  – Real-world actions

• But:
  – All are noncommittal, weak inputs, must use judiciously
  – Midas touch
Related Taxonomies and Frameworks

- **Individual classes of new interfaces**
  - Dourish 2001
  - Fishkin 2004
  - Fishkin, Moran, Harrison 1998
  - Hornecker & Buur 2006
  - Nielsen 1993
  - Ullmer & Ishii 2001

- **New issues for non-WIMP, considered more generally**
  - Belloti et al. 2002
  - Benford et al. 2005
  - Coutrix & Nigay 2006
  - Dubois & Gray 2007
  - Klemmer, Hartmann, Takayama 2006

- **Specific new interaction styles**
  - Beaudouin-Lafon 2000
  - Hurtienne & Israel 2007
  - Rorher 1995
  - Weiser 1991

- **Methodology for discussing tradeoffs**
  - QOC, MacLean et al. 1991

- **Direct Manipulation/GUI generation**
  - Shneiderman 1983: Identify
  - Hutchins, Hollan, and Norman 1986: Explain
Project: Senseboard

• **TUI**
  – Augment physical objects with digital meaning
  – Combine physical and digital representations to exploit advantages of each

• **Evolution of GUI**
  – Increase realism
  – “More-direct” manipulation
Blend Benefits of Physical and Digital

- **Physical**
  - Natural, free-form way to organize, group
  - Rapid, fluid, 2-handed manipulation, handfuls
  - Collaboration

- **Digital**
  - Selectively reveal details
  - Display alternate views
  - Sort, Search
  - Save, Restore alternate arrangements
  - Export, Backup

- **Current:** one or other set of advantages exclusively
Interface Design

• Data Objects
  – Starting point: Just grab and move
  – Pucks represent operands straightforwardly

• Command Objects
  – Operators = special puck, unique shapes
  – Tool, operator, “stamper”

• Syntax
  – Flat pucks represent data
  – Tall, special shape pucks represent commands
  – Place command over data puck
View Details Command

- Temporarily overcome size limit of data pucks
- Temporarily obscure adjacent pucks
- Command puck physically obscures cells below
  - Physical way to tell user temporary information is placed over those cells
  - Cells below still present, temporarily obscured
• Paper conflicts shown graphically

• Benefit of computer augmentation

Red line = author conflict
Yellow = topic conflict
Experiment

- Quantify costs, benefits possible from TUI
- Compared to GUI:
  - Benefits of natural interaction with real physical objects, their affordances and constraints
  - "Tangible thinking"
- Compared to pure physical:
  - Computer augmentation, display conflicts as user interacts: Expect performance benefit
  - But: imperfections in how TUI simulates physical: Expect performance penalty
Paper Condition

- Conventional paper sticky notes
- Use same vertical board
- Task designed not to require pressing or stamping
Reduced-Senseboard Condition

- **TUI simulation of world imperfect**
  - Latency
  - Projector misregistration
  - Lower resolution
  - Puck loses its display when off the board

- **Measure its performance cost**
  - No constraint checking = “reduced”
  - Expect: worse than paper, worse than regular Senseboard
  - Use to tease apart components of performance
Results

• Completed nearly all tasks correctly (99%)
  – So use time as single performance measure

• Data suggest expected trends
  – Weak significance, ANOVA condition $F(3,36) = 2.147$, $p=0.11$
Experiment Discussion

• Use Reduced-Senseboard condition to decompose small TUI-vs.-paper into 2 larger components:
  – Cost of imperfect TUI simulation of paper
  – Benefit of augmentation

• Measure value of “natural” interaction [Paper]
  – Minus: cost of simulating it [Reduced]
  – Plus: benefit of “artificial” additions
Project: Tangible Video Editor

- New implementation approach for tabletop TUI
- Extends workspace into the whole room
- Uses physicality to communicate syntax (clips, transitions)
Project: TERN TUI for Children
Project: X-ray Vision

- 1. Entire virtual room
- 2. Portion of virtual room
  - No object currently selected
- 3. Stare at purple object near top
  - Internal details become visible
Project: Pre-screen Projection

- Scene displayed on physical screen
- But dynamic perspective from user's viewpoint as if in front of screen
- Move head naturally to pan and zoom
- James Templeman, NRL
Project: Experiment on RBI

• Compare interaction styles, not UI designs
  – Same task (3D assembly)

• Design 4 user interfaces for doing it
  – As similar as possible
  – Differ only in interaction style
Project: Eye Movement-Based Interaction

- Highly-interactive, Non-WIMP, Non-command, Lightweight
  - Continuous, but recognition algorithm quantizes
  - Parallel, but implemented on coroutine UIMS
  - Non-command, lightweight, not issue intentional commands

- Benefits
  - Extremely rapid
  - Natural, little conscious effort
  - Implicitly indicate focus of attention
  - “What You Look At is What You Get”
Issues

• Midas touch
  – Eyes continually dart from point to point, not like relatively slow and deliberate operation of manual input devices
  – People not accustomed to operating devices simply by moving their eyes; if poorly done, could be very annoying

• Philosophy
  – Use natural eye movements as additional user input
  – vs. trained movements as explicit commands
### Natural Eye Movements

- **A taxonomy of approaches to eye movement-based interaction**

<table>
<thead>
<tr>
<th>Unnatural (learned) eye movement</th>
<th>Unnatural response</th>
<th>Natural (real-world) response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most work, esp. disabled</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Natural eye movement</td>
<td>Jacob</td>
<td>Starker &amp; Bolt, Vertegaal</td>
</tr>
</tbody>
</table>
Project: Brain–Computer Interface

• Plans
  – Use with mouse, keyboard, not primarily for disabled users
  – Mainly use fNIRS, not EEG
  – Input to adaptive brain-computer interface

• As with eye movements
  – Natural input
  – Unnatural response
Detecting different levels of mental workload

• Conditions
  – Rest
  – Low: 2 colors
  – Medium: 3 colors
  – High: 4 colors
### Brain State Measurements with fNIRS

- **Signals we have measured to date**
- **How to use them in lightweight, adaptive user interfaces?**

<table>
<thead>
<tr>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>General short term memory workload</td>
<td>(Hirshfield et al., 2007)</td>
</tr>
<tr>
<td>Separate spatial/verbal, syntactic/semantic workload</td>
<td>(Hirshfield, Solovey, Girouard, Kebinger, Jacob, et al., 2009)</td>
</tr>
<tr>
<td>Difficulty level of game play</td>
<td>(Girouard et al., 2009)</td>
</tr>
<tr>
<td>Cognitive multitasking</td>
<td>(Solovey et al., 2011)</td>
</tr>
<tr>
<td>Preference measurement in prefrontal cortex</td>
<td>Peck et al., in progress</td>
</tr>
<tr>
<td>Technical guidelines for HCI</td>
<td>(Solovey et al., 2009)</td>
</tr>
</tbody>
</table>
Human-Robot Task

- Dual task scenario, 2 robots, objective performance measure
- Adjust 1 robot’s automation level based on user’s brain state
- Robot simulator, robot’s eye view on screen
Results

- Autonomy helpful *when matched to cognitive state*
  - ADA > NON > MAL

- Measures:
  - Quantitative performance improvements (shown)
  - Lower task load
  - Improved perceptions of robot
  - Successful classification
  - Task independence

![Bar chart showing # Targets Found for ADA, NON, and MAL.]
Another UI: Level of Detail

- Adjust visual representation of data
- GPS slowly fades noncritical information (surrounding streets) in and out depending on driver’s workload
Focus vs. Context

- Vary use of display real estate for focus or context
Project: UIMS for VR, Non-WIMP

- Handle continuous explicitly in language
- Could handle with events as usual, but wrong model for non-WIMP
- Want continuous as first-class element of language
## Implications for Software

<table>
<thead>
<tr>
<th>Current</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Single-thread I/O</td>
<td>- Parallel, asynchronous dialogues; may be interrelated</td>
</tr>
<tr>
<td>- Discrete tokens</td>
<td>- Continuous inputs and responses (plus discrete)</td>
</tr>
<tr>
<td>- Precise tokens</td>
<td>- Probabilistic input, not easily tokenized</td>
</tr>
<tr>
<td>- Sequence, not time</td>
<td>- Real-time requirements, deadline-based</td>
</tr>
<tr>
<td>- Explicit user commands</td>
<td>- Lightweight, passive monitoring</td>
</tr>
</tbody>
</table>

*Easier to use -> harder to program*
Continuous + Discrete

- Discrete: Current, GUI

- Continuous + discrete:
  - Grasp, move, release object in VR or TUI
  - Airplane (flight simulator) controls
  - View/manipulate molecular model
  - Virtual environment controls
  - Bicycle controls, feedback
  - Eye movement-based interaction
  - Conventional control panel
  - Scrollbar (conventional GUI)
Software for Emerging Interaction Style

- Language to describe and program fine-grained aspects of non-WIMP interaction
  - Basis: Essence of non-WIMP = set of continuous relationships, in parallel, most are temporary

- Combine data-flow component for continuous + event-based for discrete
  - Discrete can enable/disable the continuous links

- Separate non-WIMP interaction into 2 components
  - Each can exploit existing approaches
  - Provide framework to connect the 2

- Keep model simple enough for fast run-time
  - Support VR interfaces directly
• **Grab and drag object with hand in 3-D**
  
  – Common, simple interaction in VR
  
  – Diamond cursor permanently attached to user's hand
  
  – Ugly object can be grabbed and moved
UIDL for Grab in 3-D

- Grab object by holding Button 1
- While button held, object position = cursor position
- When button released, relationship ceases (but cursor still follows user's hand)
Hinged Arm

• User can grab arm and move in 3D
  – Left end always fixed to base column, as if hinged
  – Arm pivots to follow hand cursor
UIDL for Hinged Arm

- State change when user grabs arm (activates \textit{linkc1}), and releases arm (deactivates \textit{linkc1})
- Hand (\textit{polhemus1}) always drives cursor position
- \textit{Linkc1} connects cursor position to arm rotation continuously
  - But active only while user grasping arm
Two-Jointed Arm

- User can grab and move the first (proximal) segment of the arm as in previous example
- Second (distal) segment hinged at tip of proximal
- User can grab distal and rotate wrt tip of proximal
UIDL for Two-Jointed Arm

- *Linkc1* active when hand cursor controlling rotation of proximal segment (*GRASPED1* condition)

- *Linkc2* active when hand controlling distal (*GRASPED2*)

- Language clearly shows: Depending on state, hand position sometimes controls *rot1* and sometimes *rot2*
• Two instances of two-jointed + 24 of one-jointed

• One-jointed arms point to proximal/distal tips of two-jointed, can turn on/off in groups

• Use to demonstrate performance
Technology for TUI

• **Sensors**
  – Produce signal in response to change in surroundings

• **Actuators**
  – Produce physical change in response to signal

• **Microcontroller**
  – Communicates with devices and with main computer

• **Other technologies**
  – RFID
  – Computer vision
  – ...

• **Wider range of I/O events than GUI**

• **Development methodology, tools**
  – None!
• A Visual Language for Modeling Tangible User Interfaces [Orit Shaer]

• TAC paradigm
  – Each TUI consists of token within a constraint
  – Same object may sometimes be token, sometimes constraint

• Two tier model fits well
  – Dialogue (states, storyboard)
  – Interaction (especially continuous)
Conclusion: Reality-Based Interaction

• **Common language, unifying framework**
  - For large subset of seemingly divergent research
  - Understand, compare and relate new interaction styles
  - Bridge gaps between research areas

• **Implications for design, analysis of new interfaces**
  - Lens to analyze, compare alternative designs; Evaluate tradeoffs
  - Say something intelligent about proposed new interface, from principled framework
  - Descriptive: Leverage to move forward, predict, understand
  - Generative: Opportunities inspired by gaps, sweet spots suggested by framework

• **Community of HCI researchers**
  - Think explicitly about connecting our research to others in next generation HCI
With Thanks to

• HCI Group at Tufts

• Michel Beaudouin-Lafon, Wendy Mackay, and the HCI group at Université Paris-Sud

• Sponsors:
  – NSF
  – AFOSR
  – NSERC
  – NEH
  – ONR
  – NRL