An Evaluation of Reality Tradeoffs in Tangible and Semaphoric Interaction Modalities

KATHERINE RIX, UNIVERSITY OF CAPE TOWN

As more processing power becomes available for interaction-oriented computation, new interaction modalities are introduced and old alternatives revisited. We explore the benefits of using complementary modalities in a multimodal user interface (MMUI). When considering a multimodal system, it is important to consider the appropriateness of each modality, and their respective strengths and weaknesses. In this paper we review the reality-based interaction framework (RBI) as an evaluation tool for modalities that are based on everyday interactions. We then apply these criteria to tangible and semaphoric modalities, and consider the specific characteristics of each one. Finally, we argue that there is enough potential synergy between these modalities to stimulate investigation into a semaphoric-tangible MMUI.

• Human-computer interaction ~ Multimodal systems • Human-computer interaction ~ Reality-based interations

Additional Key Words and Phrases: tangible, gestural, touchless, semaphoric, multimodal, human-computer interface

ACM Reference Format:

Katherine Rix, 2015. An evaluation of reality tradeoffs in tangible and semaphoric modalities. (April 2015, unpublished), 7 pages.

1. INTRODUCTION

Machine interfaces can be characterised by their medium of input/output, known as the *interaction modality*. Interface devices were originally mechanical and had limited modality options. Processing power was also restricted and devoted primarily to the operational task, with very little computation devoted to processing input and output. As the availability of processing power has increased, it has been possible to devote increasing amounts of it to interface-oriented computation [Myers 1998]. This has enabled the development of increasingly complex and computationally intensive interfaces that can engage new modalities, such as voice recognition [Myers 1998].

DOI:http://dx.doi.org/10.1145/0000000.0000000

New and alternative modalities are sometimes presented as wholesale replacements for traditional interfaces [Norman 2010], but they can instead be complementary. A design that combines complementary modalities is known as a *multimodal user interface* (MMUI). In general, MMUIs use computation to reduce the mental effort required to learn and use some parts the interface. They do this by using their computational resources to process everyday interactions, as with voice recognition, or less visible interactions, like changes in heartbeat. These interactions are more difficult for the machine to process, but simpler for the human user. Thus, the user is free to concentrate more on the operational task than the interface.[1] The user's ability to do this is supported by evidence that as an operational task increases in complexity, users spontaneously switch to multimodal interaction [Oviatt et al. 2004]. Despite the implied freedom for the user, MMUIs still have some cognitive overhead, and still require that the designer employs trustworthy general principles in order to achieve an effective interface [Norman 2010].

2. CONSIDERATIONS IN MULTIMODEL DESIGN

Apart from general principles, multimodal interfaces also introduce their own, specific considerations. Traditional modalities may in fact outperform “natural” modalities [Oviatt 1999; McMahan et al. 2010] on various metrics, including efficiency and expressive power [Jacob et al. 2008]. (For example, deitic (pointing) activities may suffer from user fatigue [Hincapié-Ramos et al. 2014] and limited accuracy [Norman 2010] compared to the low effort, and fine control of a keyboard and mouse [Norman 2010].) Furthermore, newer and less widespread modalities will lack a well-defined set of user norms and design standards [Norman 2010]. Interaction between modalities must also be considered, as an output modality may inform the user's choice of input modality [Bellik et al. 2009]. Finally, in cases where the system still includes a graphical modality, the GUI must remain an intuitive and integrated part of the interface [Profanter 2014].

When appropriately selected, multimodal interfaces offer significant benefits. The system requires less interface knowledge, making it more accessible to domain-knowledge experts without interface-specific skills. (For example, to program industrial robots, the user must be trained in the use of the programming device. Profanter [2014] designed an interface that allowed users without training to engage in the programming task.) When processing multiple input signals, signal fusion can

---

1 When the MMUI reduces or removes the need for handheld devices or deskbound use, especially by offering voice and gesture modalities, it may be referred to as a *natural user interface* (NUI).

simplify interaction for the user interaction and help the system resolve ambiguous input cases [Bolt 1980; Oviatt 1999]. Finally, the user can have a more engaging experience [Underkoffler and Ishii 1999].

Tripathi [2008] suggests that multimodal interfaces should complement user activity, and that in effective systems “the interface and human must share the same real-world model for effective reference” [Tripathi 2008, page 10]. Similarly, Jacob et al. [2008] note a trend to include commonplace, non-digital interactions in machine interfaces. They introduce “reality-based interaction” (RBI) as descriptive framework that captures modalities that fall into this trend, including direct manipulation, location awareness, tangible interfaces and others. They also present a tradeoff decision model, where the designer might choose to sacrifice the some of the benefits of a reality-based interaction for other metrics, such as expressive power or efficiency. These tradeoffs are captured in the following diagram:

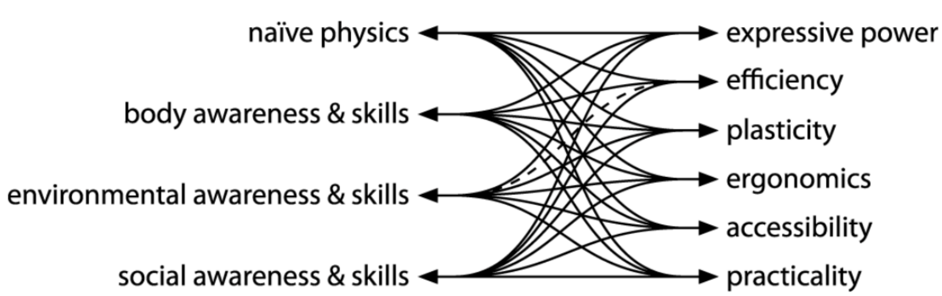

Fig. 1. Tradeoffs between the user's reality-based awareness and skills (on the left) and other interface metrics (on the right). (From Jacob et al. [2008], where the authors were highlighting the tradeoff between environmental awareness and efficiency.)

Such tradeoffs would certainly be present in multimodal systems where some modalities are not reality-based. The suggestion is to maximise RBI themes, and employ non-reality interactions only when there is a strong reason to do so [Jacob et al. 2008]. Common user tasks should use everyday actions, and tasks with special requirements may employ more abstract actions. Tapping a button widget on a touchscreen would be a common task, employing an everyday interaction. Using a soft keyboard to enter a wildcard-based search string to match a subset of files would be a special task, consciously trading everyday interactions for the expressive power and efficiency of text-based search [Jacob et al. 2008].

Having reviewed the descriptive power of the RBI framework, we can now apply these analyses to two potentially complementary modalities: tangible and gestural interactions.

3. THE TANGIBLE MODALITY

A tangible interface presents users with physical objects that they can manipulate. As they manipulate the objects, they are simultaneously modifying the system's internal model [Ishii 2008]. Thus the tangible objects are a combination of input, output and data. The overarching goal of tangible interfaces (TUIs) is to make the computing system both invisible and ubiquitous [Ishii 2008].

This complete embodiment of input/output/data in individual objects is difficult to achieve in practice, and many TUIs approximate this ideal through additional modalities. For example, the primary interface for Urp, a city-planning application, was miniature buildings placed on a physical tabletop. Tangible manipulation of these objects was an appropriate fusion of input, output and data for positioning buildings on the map, but light projection was used as a supporting modality to introduce additional data into the scene [Underkoffler and Ishii 1999].

The core interactions in the Urp system can be classified as reality-based. Jacob et al. [2008] offer the following analysis: Placing a building relies on the user's knowledge of everyday physics and their physical space: they pick up the model and place it where they want on the table. Viewing the plan from different angles relies on body awareness and physical space control: the user moves their body in the space around the table. The system makes some reality tradeoffs for more extraordinary actions. To adjust a building's material, the user touches it with a material wand. Here reality is traded for expressive power. There is no possible action that changes the shape of a building. Here expressive power is traded for reality.

More recently, the increased accessibility of 3D printing has made tangible interfaces a practical option in environments with more limited equipment options, for example augmented reality [Gillet et al. 2004] and digital tabletop [Shaer et al. 2014] systems. Here the limited expressive power of the tangible modality was compensated for by digital GUI modality. The 3D printed models supported user exploration of complex visualizations in molecular biology [Gillet et al. 2004] and synthetic biology [Shaer et al. 2014]. Both studies identify the potential applications in both education and research, and both note that the physical object improves the connection between user intent and the computational task.

4. THE SEMAPHORIC MODALITY

A semaphoric interface allows users to interact with the system by making particular gestures with their hands or bodies, and without directly manipulating a device [O'Hara et al. 2013]. These systems typically perceive these semaphores via a camera or other vision sensor. (Some systems classify finger gestures based on forearm electromyography [Kerber et al. 2015; Lu et al. 2014].) When applicable to the task at hand, gestural[2] interfaces do not only lower the barriers to using a system, they also inform the user's relationship with the content [O'Hara et al. 2013]. In addition to user neutrality, which most interfaces take for granted, gestural interactions must also account for position and hand neutrality [Norman 2010].

The core interactions of semaphores can be classified as reality-based. In the context of the RBI framework, they clearly build on the user's body awareness and control. However, current implementations [Kerber et al. 2015; Spano et al. 2012] offer a limited set of recognised gestures (Lu et al. [2014] are an exception), and frequently require that the user hold a pose [Spano et al. 2012] or otherwise wait [Kerber et al. 2015] for feedback. Semaphoric interactions may be able to exploit speed of movement in their feedback, which can somewhat alleviate user perception of this weaknesses [Spano et al. 2012]. Overall, though, these interactions appear to sacrifice efficiency for reality [Kerber et al. 2015; Lu et al. 2014].

Extended or continuous use of gestural interfaces, including semaphoric ones, may also lead to fatigue [Hincapié-Ramos et al. 2014], a factor that users anticipate and expect the system to account for [Kim et al. 2013]. Hincapié-Ramos et al. [2014] introduced *consumed endurance* as a metric to measure fatigue and evaluate mid-air (touchless gesture) interfaces. If fatigue is successfully avoided, users anticipate strong engagement with semaphoric interfaces [Kim et al. 2013], and even a limited set of gestures can provide a command set suitable for simple tasks [Kerber et al. 2015].

5. DISCUSSION

Since tangible interfaces already involve hand movement, semaphores that are hand-based present themselves as a potential complementary modality. As with other reality-based interactions, tangible objects and semaphores offer both RBI themes and tradeoffs.

2 “Gestural” is a catchall term that includes interfaces that semaphoric, deitic (pointing) and direct touch modalities [O'Hara 2013].

These are summarised in the tables below:

Table I. Presence of RBI themes in tangible and semaphoric modalities

Modality	Tangible [a]	Semaphoric
Naive physics	✓	-
Body awareness & skill	✓	✓
Environmental awareness & skill	✓	-
Social awareness & skill	✓ [b]	-

[a] As discussed in Jacob et al. [2008]
[b] In multiuser contexts.

Table II. RBI tradeoffs made by tangible and semantic modalities

Modality	Tangible [a]	Semaphoric
Expressive power	↓ [b]	- [c]
Efficiency	-	↓
Versatility	↓	-
Ergonomics	-	- [d]
Accessibility	-	-
Practicality	-	-

[a] As discussed in Jacob et al. [2008]
[b] May be somewhat accounted for by reality tradeoffs.
[c] Sufficient for simple tasks.
[d] Fatigue *must* be accounted for, otherwise ergonomics will be traded off to the point of impracticality.

It will be noted that counter-measures are available for at least some of the tradeoffs in each modality. We can also see that their reality-based tradeoffs do not coincide, which suggests the potential for each modality to complement the other on these axes.

We argue that since tangible and semaphoric modalities are somewhat complementary in nature, and do not have coinciding tradeoffs, it is worth investigating the implementation of a semaphoric-tangible interface. To maximise the efficacy of this MMUI, this interface should be founded on a shared real-world model.

6. CONCLUSIONS

We have seen that tangible interfaces often require at least one additional modality to account for their inherent limitations in expressive power and versatility. Multimodal and reality-based design guidelines suggest that a companion interface should: be based in the same real-world model, be somewhat aligned with existing user activity and complement these tradeoffs. We have demonstrated that this is sufficiently true for tangible and semaphoric interactions to justify further investigation into a semaphoric-tangible MMUI.

REFERENCES

Yacine Bellik, Issam Rebaï, Edyta Machrouch, Yasmin Barzaj, Christophe Jacquet, Gaëtan Pruvost, and Jean-Paul Sansonnet. 2009. Multimodal interaction within ambient environments: An exploratory study. In *Human-Computer Interaction* (INTERACT 2009), 89-92.

Richard A. Bolt. Richard A. Bolt. 1980. "Put-that-there": Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques* (SIGGRAPH '80). ACM, New York, NY, USA, 262-270.

Alexandre Gillet, Michel Sanner, Daniel Stoffler, David Goodsell, and Arthur Olson. 2004. Augmented reality with tangible auto-fabricated models for molecular biology applications. In *Proceedings of the conference on Visualization '04* (October 2004), IEEE Computer Society, Washington, DC, USA, 235-242.

Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (CHI '14). ACM, New York, NY, USA, 1063-1072.

Hiroshi Ishii. 2008. The tangible user interface and its evolution. *Commun. ACM* 51, 6 (June 2008), 32-36.

Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08). ACM, New York, NY, USA, 201-210.

Frederic Kerber, Pascal Lessel, and Antonio Krüger. 2015. Same-side Hand Interactions with Arm-placed Devices Using EMG. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '15). ACM, New York, NY, USA, 1367-1372.

Ga Won Kim, Ji Jyoun Lim, and Myung Hwan Yun. 2013. Identifying users' perceived values toward reality-based interaction, using semantic network analysis. *人間工学* 49, Supplement (2013). JES, Japan, S443-S446.

Zhiyuan Lu, Xiang Chen, Qiang Li, Xu Zhang, and Ping Zhou. 2014. A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices. In *IEEE Transactions on Human-Machine Systems* 44, 2 (April 2014). IEEE, 293-299.

Ryan P. McMahan, Alexander Joel D. Alon, Shaimaa Lazem, Robert J. Beaton, David Machaj, Michael Schaefer, Mara G. Silva, Anamary Leal, Robert Hagan, and Doug A. Bowman. 2010. Evaluating natural interaction techniques in video games. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)* (March 2010). IEEE, 11-14.

Brad A. Myers. 1998. A brief history of human-computer interaction technology. *interactions* 5, 2 (March 1998), 44-54.

Donald A. Norman. 2010. Natural user interfaces are not natural. *interactions* 17, 3 (May 2010), 6-10.

Kenton O'Hara, Richard Harper, Helena Mentis, Abigail Sellen, and Alex Taylor. 2013. On the naturalness of touchless: Putting the “interaction” back into NUI. *ACM Trans. Comput.-Hum. Interact.* 20, 1, Article 5 (April 2013), 25 pages.

Sharon Oviatt. 1999. Ten myths of multimodal interaction. *Commun.* ACM 42, 11 (November 1999), 74-81.

Sharon Oviatt, Rachel Coulston, and Rebecca Lunsford. 2004. When do we interact multimodally?: cognitive load and multimodal communication patterns. In *Proceedings of the 6th international conference on Multimodal interfaces* (ICMI '04). ACM, New York, NY, USA, 129-136.

Stefan Profanter. 2014. *Implementation and Evaluation of multimodal input/output channels for task-based industrial robot programming.* Master's thesis. Technische Universität München. Retrieved March 28, 2015 from https://arxiv.org/pdf/1503.04967v1.pdf

Orit Shaer, Consuelo Valdes, Sirui Liu, Kara Lu, Kimberley Chang, Wendy Xu, Traci L. Haddock, Swapnil Bhatia, Douglas Densmore, and Robert Kincaid. 2014. Designing reality-based interfaces for experiential bio-design. *Personal and Ubiquitous Computing* 18, 6 (2014), 1515-1532.

Lucio Davide Spano, Antonio Cisternino, and Fabio Paternò. A compositional model for gesture definition. In *Human-Centered Software Engineering* (2012). Springer Berlin Heidelberg, 34-52.

Priyamvada Tripathi and Sethuraman Panchanathan. 2008. Implication of multimodality in ambient interfaces. In *Proceedings of the 2008 Ambi-Sys workshop on Haptic user interfaces in ambient media systems* (HAS '08). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Article 10, 10 pages.

John Underkoffler and Hiroshi Ishii. 1999. Urp: a luminous-tangible workbench for urban planning and design. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '99). ACM, New York, NY, USA, 386-393.

Received April 2015; revised April 2015; accepted April 2015