

Multimodal Interaction in A Ubiquitous Environment

Mayuree Srikulwong¹

¹ Department of Computer Science, University of Bath, Claverton Down,
Bath, BA2 7AY, United Kingdom
{M.Srikulwong}@bath.ac.uk

Abstract. Result of my literature review shows the important characteristics of three distinct modalities and the factors influencing interaction in a ubiquitous computing context. Research aim is to develop the principles for the design of multimodal interactive systems in a ubiquitous environment.

Keywords: Multimodality, Multimodal Interaction, Ubiquitous Computing

1 Introduction

A ubiquitous environment has different characteristics from traditional networked computing. First, it has a high level of mobility and users are dynamic. Second, the architectural space may contain both fixed and mobile devices. Third, interaction can be ad-hoc among peer nodes. These characteristics cause new problems and challenge us to find the most appropriate forms of interaction. Such new structures of interaction should be natural and its interface should be seamless. Multimodality can be used as a bridging concept between naturalistic behavior and engagement with the system that is designed to mediate tasks.

Multimodality refers to multiple paths of communication employed by users to carry input and output. Its variations offer choices for interaction which should appropriately help, support and extend the way users communicate and perform activities in different situations. Much of past multimodality research has aimed to solve information representation focusing on the usage of visual and audio sensory channels. Recent study [1] shows that merely 1% of multi-sensory research is on haptic sense. Unlike visual channel, haptic is bidirectional, i.e. can be used as both input and output to the systems. If we use human body as a whole unit for interaction design, we will benefit from its main characteristics which are depth, height, breadth, and body posture.

2 Related Research

Brewster et al. [2] added sound into interaction action to overcome the limitation of mobile phones' screen size. Cao et al. [3] used time multiplexing crossmodal cues, i.e. vibration and sound, to enhance privacy in public spaces. O'Neill et al. [4]

proposed a novel interaction technique, called Directional Stroke Recognition (DSR), which is solely based on the direction of gesture. Nesbitt [5] has given guidelines for multimodal and tactile displays.

Although these works and guidelines are ones of the significant pieces of work in multimodality and ubiquitous computing areas, they are not unified. We still have little understanding of the relationship between these two research areas. Much of the previous multimodality research has focused on the use of visual and audio channels and the use of alternative sensory channels to enhance the usage of the visual channel. The fact that much of past research has omitted the use of haptic interaction means that opportunities to improve the level of system usability were lost. Hence, these guidelines are inadequate to be used as principles for the design of multimodal interactive systems in a ubiquitous environment.

3 Problem Statements and Research Question

3.1 Problem Statements

First, there is a lack of understanding of the relationship between the concept of multimodality and ubiquitous computing. Second, there is no written principle regarding multimodal interaction in a ubiquitous environment. Third, only visual and audio channels of human senses have been comprehensively studied.

3.2 Research Question

Consequently, my aim is to extensively study on the addition of haptic communicative channel into the multimodal interaction model. My research attempts to develop principles for the design of multimodal interactive systems for a ubiquitous environment. My contribution to the ubiquitous computing field of study is to answer the following research question:

How haptic interaction can be appropriately added into a multimodal interaction model for a ubiquitous environment in order to make the interaction transparent with a better level of efficiency and effectiveness?

In term of efficiency, I will focus on usability, more specifically, on learnability. Once the users learn how to use the system and obtain skill, the interaction between such users and the systems become transparent. This transparency, then, is achieved when the users have shifted their focus of interaction from the technologies to the tasks.

In term of effectiveness, my focal point is on cognitive awareness. The users have obtained such state of consciousness via the available feedback data, perceived it and subject to further cognitive processing. In some safety-critical situations, e.g. aviation systems, awareness breakdown can lead to devastation. As a result, an interaction design for such systems should include the notification of the undesirable state changes. This information should be *explicit* in such visually cluttered environment

to catch the users' attention, leading to higher level of cognitive processing, i.e. perception and awareness.

The main point here regards clear, explicit and informative system feedback. Past systems which omit the use of haptic channel reduce an opportunity to improve the level of learnability and cognitive awareness for particular systems. Haptic input and output could be used as an alternative as well as in combination with the other busy perceptive channels to make the interaction become transparent more rapidly and to help increasing the users' cognitive awareness of current activity.

Subsequently, my investigation will base on the following hypotheses:

Hypothesis 1: The inclusion of semantically rich haptic communication in visual/audio interaction will significantly decrease the number of breakdowns, as users will shift their focus from the technology to the task at hand.

Hypothesis 2: The inclusion of semantically rich haptic notifications about a system's change will significantly reduce the number of breakdowns, as users will always be aware of the current state of the system.

3.3 Approach and Methodology

The dissertation consists of two major parts: literature review and experiments. The result of preliminary study in the following section is based on the former. The outcome of my research, the principles for multimodal interactive system design, will base on both. Then, such principles will be used as a tool to design an instance of multimodal interactive system and be evaluated through data collected from system testing.

In order to develop the abovementioned principles, further empirical studies and a number of experiments will be carried out. Some projects would be initially run in a controlled environment to obtain understanding of characteristics, impacts and limitations of different modalities. Consecutively, some projects would be selected and evaluated in the real world in order to draw relationships of the already-understood modalities' characteristics¹ with those¹ of ubiquitous computing. Results from field tests are expected to provide rich and contextual information to complement those from the controlled experiments.

4 Proposed Work

4.1 Preliminary Results

The initial study has focused on the finding of factors and characteristics of components existing in a ubiquitous interaction space. The results of my literature review are presented as factors of the taxonomy of multimodal interaction design for a ubiquitous environment in table 1 – 4 and figure 1.

¹ See 4.1 Preliminary Results

The taxonomy consists of eight factors which are information, usage modes, modality, user contexts, user type, degree of publicness, level of mobility and level of user control factors. This preliminary review provides an overview picture of what factors influencing the design of multimodal interactive systems in a ubiquitous setting.

Table 1. Information factor.

Information Type	Data Structure
Meta Data	Linear
Objects	Circular
Attributes	Tree
Notification	Graph
	Object-oriented
	Relational

Table 2. Usage mode factor

Usage Mode
Alternate
Synergy
Exclusive
Concurrent

Table 3. Modality factor

Modality	Interactivity type	Interaction activity	Differential property
Visual	Abstract Signal, Icon, Textual, Graphic, 3D, Animation	See, Look, Explore, Navigate	Size, Color, Texture, Orientation, Shape
Audio	Abstract Sound, Natural Sounds, Speech	Listen, Navigate	Frequency, Wave Length, Period, Amplitude, Velocity, Speed
Haptic	Abstract mechanical pattern, Natural pattern	Touch, Feel	Resistance, Capacitance, Position, Velocity, Acceleration, Force, Torque, Pressure, Types of grasp, Temperature of contacting body part, roughness, vibration pattern, hardness, stiffness, weight

Table 4. User context factor

User Context
User preference
User intent
User Previous experience
User needs and desire
User physical status

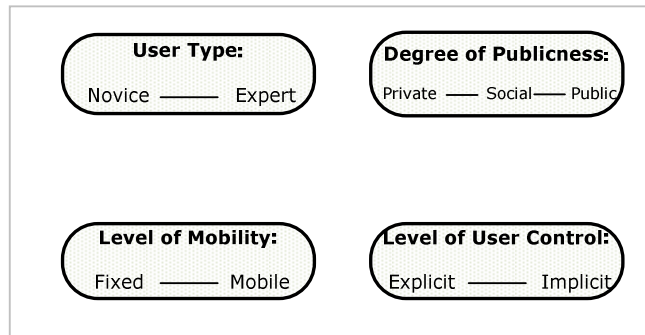


Fig. 1. User type, Degree of publicness, Level of device mobility and Level of user control factors

4.2 Future steps

The following steps are to perform a series of empirical studies and experiments in order to find relationships among all factors in the taxonomy. In other words, how to combine multiple sensory usages and find a balance between them? In which situation could we use different channels? What are the users' available communicative channels in different situations? How would the multimodal interactive systems *look*, *sound* and *feel* like?

A primary set of experiments will be to test different combinations of modalities for two types of applications, an interpersonal communication application and an aviation control system.

The first application is to add alternative channel, i.e. haptic, to convey communicative meaning for interpersonal communication in a ubiquitous context. The chosen scenario is the communication between people at home with a fixed device, i.e. picture frame equipped with a set of pre-stored images with related tag information, and family members on the move with their GPS enabled mobile devices. The main objective of this experiment is to investigate the users' understanding of haptic meaning.

The second application is to manipulate modalities in a visually-cluttered environment, i.e. an aircraft cockpit wherein interaction is safety-critical [6]. The chosen scenario is an aircraft descent scenario with interventions from the autopilot. The main objectives of this experiment are: to investigate the pilot's attention drawn from rapidly changing cockpit (multimodal) interface by the autopilot and to discover the best perceptible modality for improving the overall safety for this environment.

For each testing, both quantitative and qualitative data will be collected. Quantitative data (e.g. time taken, number of errors and number of reported observations) will be gathered automatically by the application and analysed relating to learnability and awareness. Qualitative data will be collected through a video recording, post-session interview and questionnaires to ensure the consistency of quantitative results and to capture user preference and acceptance.

Further sets of experimental projects include the applications which have different degrees of publicness and levels of user control² in order to examine the attributes of each factor in the taxonomy and relationships among them.

5 Conclusion

Preliminary results of my study are based on literature review. Future works include the development of experiments and the validation of the hypotheses. The principles produced will be used to support system designers in making decisions about the usage of three modalities (i.e. visual, audio, and haptic) in a ubiquitous computing context.

Acknowledgments. I would like to thank my supervisor, Dr.Eamonn O'Neill, for guidance and support and to HCI@Bath for comments.

References

1. Aziz, F.A., Nicholas, L.: Simulated Lifting with Visual Feedback BCS-HCI2006, Vol. 2, London (2006)
2. Brewster, S.: Overcoming the Lack of Screen Space on Mobile Computers. *Personal Ubiquitous Computing* 6 (2002) 188-205
3. Cao, H., Olivier, P., Jackson, D., Armstrong, A., Huang, L.: Enhancing privacy in public spaces through crossmodal displays. (2005)
4. O'Neill, E., Kaenampornpan, M., Kostakos, V., Warr, A., Woodgate, D.: Can we do without GUIs? *Gesture and speech interaction with a patient information system. Personal and Ubiquitous Computing*, Springer-Verlag (2005)
5. Nesbitt, K.V.: A Framework to Support the Designers of Haptic, Visual and Auditory Displays.: *GOTHI 2005, Canada* (2005)
6. Hourizi, R., Johnson, P.: Towards an explanatory, predictive account of awareness. *Computers & Graphics* 27 (2003) 859-872

² See Figure 1