

Multimodal Interaction in a Pervasive Environment

Mayuree Srikulwong¹

Abstract

Research in pervasive computing has not focused enough on the benefits that multimodal interaction has to offer. Result of my literature review shows the important characteristics of three distinct modalities and the factors influencing interaction in a pervasive computing context. The ultimate aim of this research is to develop principles for the design of multimodal interactive systems in a pervasive environment.

Mayuree Srikulwong

Department of Computer Science,

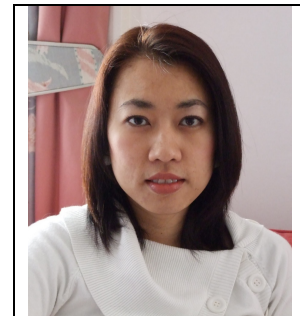
University of Bath, Bath, UK

Phone: +44 122538 6183

Email: M.Srikulwong@bath.ac.uk

Webpage: <http://www.cs.bath.ac.uk/~ms244>

PhD-Supervision: Dr. Eamonn O'Neill



Problem Statement and Research Question

A pervasive environment has different characteristics from traditional networked computing. Firstly, it has a high level of mobility and users are dynamic. Secondly, the environment may contain both fixed and mobile devices. Thirdly, 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.

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. A recent study [1] suggests that only 1% of multi-sensory research is on the haptic sense.

¹ University of Bath, Bath, United Kingdom

Unlike visualisation perception, haptic is bidirectional, i.e. can be used as both input and output to the systems. If we use the human body as a whole for interaction design, we will benefit from its characteristics of depth, height, breadth, and body posture. Applications may be designed for free movement and interactions can be naturally controlled by haptics and gesture.

Nonetheless, to provide effective multimodality, we have to carefully design the interaction process. A study [2] on mobile computing devices suggested that enhancing the buttons on computers with sound can significantly increase their usability. Nevertheless, Verrillo [9] found that heavy use of channels often reduces the usability of a user interface, as the user must expend effort to remember current states, and switch between them as necessary. This contradiction challenges interaction designers to resolve and find a balance of an appropriate use of modalities.

In the future, individuals in pervasive space might have to carry several devices and interact with a number of fixed and pervasive devices. We need to improve our understanding of how to provide an appropriate means of interaction for different situations. Alternatives or supplements for the cluttered visual and audio channels should be developed. I propose that it is high time for us to extend the boundary of multimodality to fully cover haptics and apply it to pervasive computing.

Problem statements

Firstly, there is a lack of understanding of the relationship between the concept of multimodality and pervasive computing. Secondly, there are no established principles regarding multimodal interaction in a pervasive environment. Thirdly, only the visual and audio channels of human senses have been comprehensively studied.

Research question

Consequently, my aim is to study the addition of haptic communicative channels into the multimodal interaction model. My research attempts to develop principles for human-computer multimodal interaction. This research will: (a) produce a comprehensive integrated taxonomy of multimodal interaction design in a pervasive computing context; and (b) develop principles for the design of multimodal interactive systems for a pervasive environment. My contribution to the pervasive computing field of study is to answer the following research question:

How can haptic interaction be appropriately added into a multimodal interaction model for a pervasive 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, specifically on learnability. Learnability directly affects usability. Interaction with different types of technologies requires different periods of time for learning. Less complex technologies such as the Automatic Teller Machine (ATM) requires much less time to learn than a pilot's aircraft control system. Nevertheless, this does not mean more complex systems are not learnable. Learnability is also related to task types and the level of task complexity. Once a user learns how to use the system and obtains skills to become an expert with the system, then the interaction between such user and the system becomes transparent. This transparency, then, is

achieved when the users have shifted their focus of interaction from the technologies to the tasks; that is, when the users realize “I can use this system do what I want to do.”

In term of effectiveness, my focal point is on cognitive awareness. Situation awareness normally is offered to users through system feedback. The users achieve a state of consciousness via the available data, perceiving it and subjecting it to further cognitive processing. Interaction in different environments requires different level of awareness. In some safety-critical situations, e.g. navigation and aviation systems, awareness breakdown can lead to disaster. As a result, interaction design for such systems should include the notification of undesirable state changes. This information should be *effective* in such visually cluttered environment to catch the users’ attention, leading to higher levels of perception and awareness.

The main point here regards clear, explicit and informative system feedback. Systems which omit the use of haptic interaction channel miss an opportunity to improve the level of learnability and cognitive awareness. Haptic input and output can be used as an alternative as well as in combination with the other busy perceptive channels to reduce learning time on task, and make the interaction between human and the system become transparent more rapidly. Furthermore, it is believed to help increasing the users’ cognitive awareness of current activity [4].

Subsequently, my investigation will be based 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 be helped to shift their focus from the technology to the tasks.

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

Approach and Methodology

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

In order to develop these principles, further empirical studies and a number of experiments have to be carried out. Some projects will be initially run in a controlled environment to obtain understanding of characteristics, impacts and limitations of different modalities. Subsequently, some projects will be selected and evaluated in the field in order to draw out the relationships between the already-understood modalities’ characteristics² and those² of pervasive computing.

² See Preliminary Results section

Alternatively, this research could take either an ethnographic perspective or theoretical approach. But it would not be suitable due to the novelty of multimodal technologies. The experimental approach will allow me to distil attributes of modalities in the empirical tests to identify the best practices and ensure the quality and usefulness of the results for the principles. Results from field tests are expected to provide rich and contextual information to complement those from the controlled experiments.

Related Work

Brewster et al. [3] added sound to interaction to overcome the limitation of mobile phones' screen size. Cao et al. [4] used time multiplexing crossmodal cues, i.e. vibration and sound, to enhance privacy in public spaces. O'Neill et al. [8] proposed a novel interaction technique, called Directional Stroke Recognition (DSR), which is solely based on the direction of gesture. Nesbitt [6, 7] et al. has given guidelines for multimodal and tactile displays of abstract data in the virtual environment.

Two observations can be drawn from the above literature. The first is that 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 second is that much of the past pervasive computing research has not yet fully focused on the benefits of multimodality. Although there are significant pieces of work in multimodality and pervasive computing, they are not unified. We still have little understanding of the relationship between these two research areas. The fact that much 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 pervasive environment.

My purpose is to relate these two disciplines, i.e. multimodality and pervasive computing, and find the relationships among them in order to develop principles that will be used to support system designers in making decisions about the use of three modalities (i.e. visual, audio, and haptic) in a pervasive computing context.

Preliminary Results

My initial study has focused on finding factors and characteristics of components in a pervasive interaction space. The results of my literature review are presented as factors of the taxonomy of multimodal interaction design for a pervasive environment in table 1 – 4 and figure 1. The taxonomy consists of eight factors which are usage modes, information, modality, user contexts, user type, degree of publicness, level of mobility and level of user control factors. These factors would influence the design of usable multimodal interactive systems.

Usage Mode
Alternate
Synergy
Exclusive
Concurrent

Table 1. *Usage mode factor*

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

Table 2. *Information factor*

Modalities	Interactivity type	Interaction activities	Differential properties
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 3. *Modality factor*

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

Table 4. *User context factor*

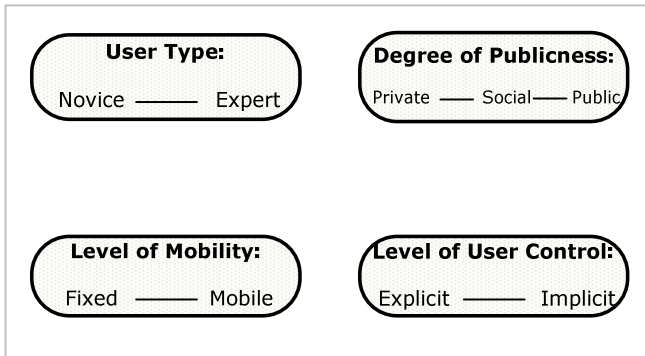


Figure 1. *User type, Degree of publicness, Level of device mobility and Level of user control factors*

Table 1 shows the information factor which consists of information types and data structures which influence interactivity types. Table 2 lists different usage modes which might be influenced by user preferences, task types or the limitation of the devices. Table 3 shows the modality factor which contains interactivity types, interaction activities and differential properties of each modality. Interactivity types refer to a variety of media types belong to each modality. Interaction activities refer to task types that the users carry out when dealing with information. Differential properties are a set of attributes which characterise each piece of information of different modalities. Table 4 shows user context factors which influence their interaction. Figure 1 presents the other four factors as spectra. These factors are user type, degree of publicness, level of device mobility and level of user control factors.

This initial review provides an overview of factors influencing the design of multimodal interactive systems in a pervasive setting.

Conclusions and Future Steps

Future Steps

The following steps are to perform a series of empirical studies and experiments in order to find relationships amongst the factors in the taxonomy. In other words, how so we combine multiple sensory channels 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?*

A first 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 purpose of these experiments is to examine the effects of modalities on learnability and awareness under two different circumstances with different degrees of factors' attributes in the taxonomy.

The first application is to add a channel, i.e. haptic, to convey communicative meaning for interpersonal communication in a pervasive context. The chosen scenario is communication between people at home with a fixed device, i.e. a 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

general location of the phone, e.g. at home, causes the pictures to change in the digital picture frame. Haptic input to the picture frame when a certain picture is shown will generate a semantic vibration pattern on the family members' phones conveying communicative meaning to them. The main objective of this experiment is to investigate the users' understanding of haptic meaning. The application is also expected to promote communication, by knowing the general location of the person, one might decide to make a phone call.

The second application is to manipulate modalities in a visually-cluttered environment, i.e. an aircraft cockpit wherein interaction is safety-critical [5]. The chosen scenario is an aircraft descent scenario with interventions from the autopilot. In the course of the experiment, the users will be asked to perform a series of instructions to affect the descent and ensure that the aircraft is traveling towards the airport. During this scenario, the autopilot will intervene and the aircraft will start to move away from the airport. To report this intervention to the users, the system will generate different display modalities. The main objectives of this experiment are: to investigate the effect of drawing pilot's attention from the rapidly changing cockpit (multimodal) interface shortly before the autopilot makes alterations to the course of the flight; and to discover the best perceptible modality for improving the overall safety for this highly situation-aware environment.

A variety of modalities and their combinations will be used in different conditions of both experiments to test the hypotheses. For each test, 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 in relation to learnability and awareness. Qualitative data will be collected through video recording, post-session interviews and questionnaires to validate the quantitative results and to capture user preference and acceptance.

Further sets of experimental projects include applications which have different degrees of publicness and levels of user control³ in order to examine the attributes of each factor in the taxonomy and relationships amongst them.

Conclusion

Preliminary results of my study are based on a literature review. Future work includes the development of experiments and the validation of the hypotheses.

Despite the fact that I have directions and experimental plans for my study, I am at the early stage of my research. The issues on which I would appreciate guidance include scenarios for the experiments, justification of experimental results, and the unification approach for linking factors influencing multimodal interaction to a pervasive environment.

Participation in Pervasive 2007's Doctorial Colloquium will: give me the opportunity to have my research critically appraised by a number of experts in the field of HCI; provide me with support and guidance from people with similar research interests; and boost my knowledge regarding recent developments and future advances in multimodality and pervasive computing.

³ See Figure 1

References

- [1] AZIZ, F.A., NICHOLAS, L.: Simulated Lifting with Visual Feedback BCS-HCI2006, Vol. 2, London (2006).
- [2] BREWSTER, S.A., CRYER, P.G.: Maximising screen-space on mobile computing devices. CHI '99 extended abstracts on Human factors in computing systems. ACM Press, Pittsburgh, Pennsylvania (1999).
- [3] BREWSTER, S., LUMSDEN, J., BELL, M., HALL, M., TASKER, S.: Multimodal 'eyes-free' interaction techniques for wearable devices. Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, Ft. Lauderdale, Florida, USA (2003).
- [4] CAO, H., OLIVIER, P., JACKSON, D., ARMSTRONG, A., HUANG, L.: Enhancing privacy in public spaces through crossmodal displays (2005).
- [5] HOURIZI, R., JOHNSON, P.: Towards an explanatory, predictive account of awareness. *Computers & Graphics* 27 (2003) 859-872.
- [6] NESBITT, K.V.: Structured Guidelines to Support the Design of Haptic Displays.: GOTH1 2005, Canada (2005).
- [7] NESBITT, K.V.: A Framework to Support the Designers of Haptic, Visual and Auditory Displays.: GOTH1 2005, Canada (2005).
- [8] O'NEILL, E., KAENAMPORN PAN, 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).
- [9] VERRILLO, R.T.: The effects of aging on the sense of touch. In: Verrillo, R.T. (ed.): *Sensory research: Multimodal perspectives*. Lawrence Erlbaum Associates, Inc., New Jersey, USA (1993) 285-298.