Audioclouds: Three-Dimensional Auditory and Gestural Interfaces for Mobile and Wearable Computers

Stephen Brewster and Roderick Murray-Smith
Department of Computing Science, University of Glasgow

1. DESCRIPTION OF PROPOSED RESEARCH

Mobile telephones, Personal Digital Assistants (PDAs) and handheld computers are currently one of the fastest growth areas of computing. As the Government's Foresight Committee reports, this growth will extend into more sophisticated, fully wearable computers in the near future [14]. One problem with these devices is their limited input and output capabilities. There is a very limited amount of screen space on which to display information. Small displays easily become cluttered with information and widgets. Input is limited, with small keyboards or simple handwriting recognition the norm. Speech-recognition is also not always an ideal option, even if recognition rates can be improved further.

With the imminent dramatic increase in bandwidth available to mobile devices, one question must be answered: how can this be used by people on the move, given the limited interaction they have with their devices? Most input methods are slow and hard to use when mobile. Current interfaces limit use because walking or running, driving or navigating all require a large amount of visual attention and adding to this with a complex graphical display can cause problems [8]. The research proposed here is therefore timely – it will investigate *3D auditory* and *gestural* interfaces as new presentation methods and interaction techniques that will allow richer and more complex interactions with mobile devices and mobile services when on the move, opening up the possibilities for using mobiles in new and more powerful ways.

In future interactive systems the user will be in constant, tightly coupled interaction with the computing system, and this interaction will not be only an exchange of discrete messages, but a continuous interaction [11]. For example, users can use simple head gestures (nodding at an item in the 3D audio space in which they are interested) to select things, or hand gestures to move audio objects around. Such tight coupling, in continuous-time, and real space requires insight from control theory, and pattern recognition/statistics.

Our proposal is innovative in two ways: *1*. We bring the areas of statistics and control theory together with multimodal HCI. *2*. We will explore a powerful new paradigm for interacting with mobile computers, based on novel techniques using sound and gestures. Virtual three-dimensional (3D) sound through headphones will be used in a novel way to create a 'personal audio space' around a user to give a larger area in which to display auditory information than standard mobile devices. Head/hand trackers (developed for virtual environments) will be used to allow gestures as input to mobile devices. Modern non-parametric statistical dynamic models will be used for high-performance gesture recognition. The main aims of this project are to:

- Investigate the possibilities for using 3D sound to create a larger space in which to display information on mobile computers;
- Investigate gestural interaction – based on head, hand and device gestures – to improve input and to control of the large audio space;
- Develop new interaction techniques combining sound and gestures that are usable on the move;
- Conduct rigorous evaluations of these new techniques to ensure they are usable in mobile contexts.

This research is interdisciplinary, bringing together specialists in multimodal HCI and mathematical modelling. Our own work provides solid foundations for this project. Brewster is an active researcher in multimodal interaction and has shown that non-speech audio and basic gestures on mobile devices can improve usability [28] and that 3D sound can benefit desktop interactions [34, 35]. Murray-Smith has a background in modelling dynamic systems [25], human control behaviour [23, 24, 26] and linking these to modern statistics [32].

This proposal is of key importance to the UK for it to maintain its strong position in the area of HCI and design for mobile computers. The Government's Foresight Committee [14] suggests *"Ease of interaction i.e. the use of intuitively obvious methods for controlling and interacting with systems, such as voice, touch, gesture, writing, image and so on, in addition to today's keyboard and mouse"* are specific technology opportunities available to the UK. The report is also clear that user needs and ease of interaction are priorities. The project will therefore be addressing problems that are directly relevant to the future of UK science and industry.

1.1 Programme Summary

We are requesting 1 RA, 1 RS, equipment and travel for a period of 3 years. Our progress will be measured against the following milestones:

- *Year 1:* Produce prototype of 3D audio system on a wearable computer. Demonstrate basic head gesturing and produce paper describing initial evaluation.
- *Year 2:* Produce paper describing theory and details of the gesture recognition system, plus its detailed testing with head and hand gestures. Publish benchmark data sets. Produce paper on interaction techniques combining audio and visual presentation. Demonstrate prototype systems for generating sound from gestures.
- *Year 3:* Papers on the results of detailed evaluations of the complete 3D audio and gestural system. Produce final demonstrator applications and design guidelines for using new interaction techniques. Produce final report and host workshop to disseminate results and knowledge gained.