Mobile Devices in Crossmodal Interfaces

Han Cao, Patrick Olivier, Daniel Jackson, Andrew Armstrong

Informatics Research Institute, University of Newcastle Upon Tyne, NE1 7RU, UK
Fax: +44 (0)191 2464905
`p.l.olivier@ncl.ac.uk`

Abstract. We describe a crossmodal ambient display framework with which we aim to bridge the gap between ambient display technology and personal mobile human-computer interaction through the exploitation of aspects of crossmodal cognition. We utilize this framework in the construction of CROSSFLOW, a crossmodal ambient display prototype for indoor navigation, and demonstrate a significant increase in the performance of users navigating with CROSSFLOW in comparison to their performance with a standard map. Based on our empirical studies there is support for both the utility and desirability of crossmodal ambient displays. Evaluation of the prototype has shown that crossmodal ambient displays can support faster, more accurate and less cognitively demanding navigation than can a traditional map.

Introduction

One motivation for research into ambient displays is the recognition that humans can perform multiple tasks simultaneously, and contrary to explicit interface designs which seek to support a single task, ambient and peripheral displays aim to support engagement in multiple tasks with minimal performance impairment. This has particular relevance when considered within the context of the ambient intelligence and pervasive computing research enterprises whereby ubiquitous information displays continuously display information of potential value to occupants of a space [1]. The support of users undertaking navigation in a spatial environment, through the provision of information displays that impose minimal demands on users' attention, is a canonical goal in ubiquitous computing and is our chosen application domain.

Ambient displays address a particular shortcoming of a conventional view of ubiquitous computing applications, which emphasize the use of personal computing devices that incorporate their own visual displays (e.g. mobile phones and PDAs). In such a configuration, the very act of having to visually refer to one's own display to receive information undermines the sense of immersion in a spatial environment. Indeed a characteristic of many visual display-based personal navigation solutions is that their users wander around their environment attending to their own devices and are to some degree disengaged from the very space they occupy.

Ambient Displays

An ambient display is embedded in the environment of its users and displays public information, that is, information that is not inherently tailored to any specific user. Almost by definition, most ambient displays convey general information (e.g. news, stock values, weather, traffic congestion and human activity) which can be revealed for both individual and group users through personalized mappings [6,7]. We see bridging the personal-private divide as an unresolved issue in the design of ambient displays which is partially addressed by the crossmodal modification we propose (see discussion in the sections 5 and 6).

Ambient display research has involved the development of a number of prototypes that aim to utilize highly aesthetic, and essentially peripheral, representations. For example, InfoCanvas [3], Informative Art [4] and AROMA [5], each of which incorporates abstract design elements, motivated by different styles of visual art, to represent information. With a greater emphasis on the aesthetics of everyday design, the Active Wallpaper, Water Lamp, and Pinwheels [15] artifacts all attempt to map information changes (e.g. weather, stock values) to system state changes in a "calm" and "subtle" manner, with a view to minimizing the attentional demands placed on a user engaged in some other task in an environment.

In terms of modalities applied, the majority of peripheral displays focus on visual displays. Researchers of visual peripheral displays understand and utilize the dual-task paradigm, and aim to present information in a timely manner which appropriately matches the time-sharing strategies that users engage in when performing two related tasks simultaneously. However, when the dual or multiple tasks become demanding, the visual channel is easy overloaded and errors increase. Audio Aura [8], ambientROOM [9], and AROMA [5] have explored auditory and/or olfactory, multimodal ambient displays to exploit more peripheral senses of users. Our design extends the modality dimension in which the coordinated modalities are exploited.

Cognition and Crossmodal Ambient Displays

Psychological research into attention, over many decades, has demonstrated that an information processing bottleneck implies one-at-a-time processing and an attendant limitation in the information processing capability of humans in multiple task conditions. However, overwhelming evidence demonstrates that some information from unattended sources ultimately reaches higher stages of processing [10], which presents the possibility for people to receive information efficiently in a manner that that does not require their full attention.

More recently, empirical research in cognitive neuroscience has given rise to the notion of crossmodal attention, a term used to refer to capacities and effects involved in the process of coordinating (or 'matching') the information received through multiple perceptual modalities [11]. Recent studies reveal extensive crossmodal links in attention across the various modalities (i.e. audition, vision, touch and proprioception). More significantly for ambient display design it has been demonstrated that some crossmodal integration can arise preattentively [12].

In a crossmodal ambient display, information (which may contain both public and private components) is displayed throughout the whole physical extent of a shared information space. In contrast to a conventional ambient display, the crossmodal ambient display temporally cycles through information of potential interest to users in a space. In the case of our indoor navigation application, directions to different locations in the space (including exits) are projected on the floor of an environment one-at-a-time on a fixed time cycle. For example, in time slot 1, directions to destination A are displayed at all locations in the physical space, in time slot 2, directions to destination B are displayed, and so on until the directions to destination A are repeated. The user identifies (or decodes) which time slot in the cycle is relevant to their own request for directions through the utilization of a crossmodal cue (e.g. a sound or vibration) issued by his/her personal mobile computing device. That is, either in response the user's request for directions or on entry to the physical space, the user's device communicates with the ambient display infrastructure to establish the schedule of time slots when the different directions will be displayed. In other words, the personal mobile device displays private information cues, that only individual users can perceive, that allow users to decode the ambiently displayed public and/or personal information.

Note that the directions displayed at a location depend on the direction of the destination from that location. Again we can contrast this with traditional hand-held notions of navigation, whereby there is a requirement to track the position of a user and present directions salient to the specific location of the user. We can contrast these two configurations in terms of the multiplexing of information displayed. In traditional mobile device applications, incorporating tracking, information is spatially multiplexed. That is, the position of a user is known and information specialized to the location of the user is displayed on the user's device. In the crossmodal scenario information is temporally multiplexed and information relevant to a location is displayed at all locations (in our case through projection on the floor of the environment) at a specific time.

We explored our design idea through a crossmodal ambient display prototype named CROSSFLOW, for indoor navigation, and evaluated it in a dual task experiment with nine participants. In accordance with our design hypothesis as to the use of a cognitively well founded coordination of modalities, we found that the participants had significantly higher performance when using CROSSFLOW for indoor navigation as compared to a traditional map. Furthermore users of CROSSFLOW also performed better in terms of performance of the primary task of the dual task paradigm, implying that it had significantly lower attentional requirements. These results have implications for the design and evaluation of novel navigation tools, information displays and multimodal user interfaces.

Indoor Navigation with CrossFlow

CrossFlow is a prototype indoor navigation system based on our notion of a cross-modal ambient display design and embodies our framework for integrating public ambient displays and personal cues across modalities. The prototype is designed for use by a user with a mobile phone inside large unfamiliar buildings. The time-multiplexing technique described in the previous section prompts users as to which directions correspond to their destination of interest. Advantages of such an indoor navigation system include low cost, no requirement for sensing or tracking of the users, and the maintenance of user privacy.

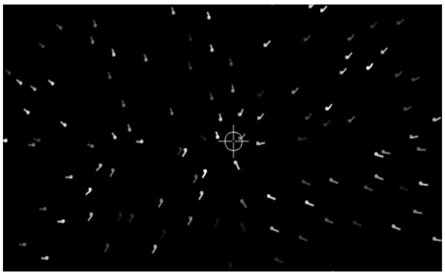

Fig. 1. “Fish-like” pattern in CrossFlow

Fig. 2. Arrow pattern

Ambient display design

CrossFlow uses aesthetically pleasing ambient displays combined with a crossmodal cue on a user’s personal mobile device to provide direction information. The ambient display for directions was designed to be as peripheral and calm as possible. Figures 1 and 2 show two examples of the raw display. Figure 1 shows an animated “fish-like” pattern, the elements of which orientate themselves and flow in the direction of the destination, and figure 2 shows an arrow-like pattern that does not translate but animates between different directions (see accompanying video). Once projected the elements of the design (individual fish or arrows) are approximately the size of a hand, and have a visual intensity that integrates with the floor giving the appearance of a sparkling carpet. Figure 3 shows a close-up of the projection on the floor of the experimental region.

At any point in time, and at every location in the space, directions are displayed for a user to follow in order to reach a particular destination. Thus in figure 1 the flowlines configure themselves to form paths (around any attendant obstacles) towards the crosshair indicating the location of the destination. The pattern changes every 800ms; during each time slot all elements in the projected display have the appearance of swimming towards the same destination along the designed paths. Similarly, for the design in figure 2 the entire arrow set points along the designed paths to the same destination. During the next time slot, the elements undergo an animated transition in configuring themselves for the next destination.

Fig. 3. Projected “fish-like” design (on floor)

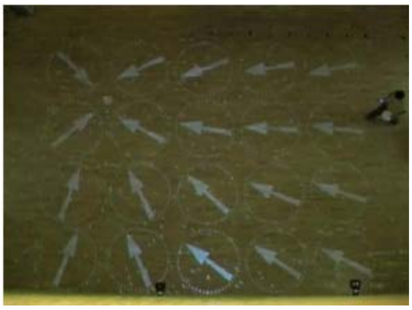

Fig. 4. Influence arrows for configuration

The configuration of the pattern is achieved using configuration files to set parameters that control the design, size, density, and dynamic properties of the individual elements (rate of movement and visual persistence), and the duration and number of time slots of the display. Influence arrows are used to interactively configure the direction of motion during each time-slot. Figure 4 shows a set of influence arrows which a designer interactively manipulates to control the direction of flow at locations in the environment. Influence arrows give the designer the flexibility to specify local flow tendencies for the pattern, which are aggregated for the final pattern. Thus a designer will configure the influence arrows to steer the flow around obstacles and away from sites that are not intended to lie on the path to a destination. Influence arrows may be interactively added, scaled, and rotated to attain the desired pattern of flow, and a key press binds the configuration to a time slot. In figure 4 we can see that a convergent pattern has been specified where all flow is directed towards the location of a white disc placed on the floor.

Crossmodal cues

The second element of a crossmodal ambient display is the design of the crossmodal cue on the personal device. A personal mobile device, in this instance a Microsoft smart phone, issues a crossmodal cue in the form of one or both of: (a) a vibration for the duration of the corresponding time-slot; (b) an audible high pitch sound coordinated with the onset of the timeslot. We have yet to study empirically the impact of the different cues on the effectiveness of the display and in our initial evaluation we use both modalities simultaneously. The crossmodal cue causes the user to pay attention to the directions shown by the public ambient display in the corresponding time-slot and induces a subtle switch of the user’s attention. The personal mobile device connects to a central server to synchronize the time-slots at the beginning of navigation and to receive the schedule of time-slots and their mappings to the numerical keys of the navigation interface on the mobile device (which simply asks the user to press the number corresponding to the destination required). When the user selects a new destination the personal mobile device presents a cue in a different time-slot corresponding to the display of directions to the new destination.

Preliminary Evaluation

We ran a preliminary study to explore the effectiveness and efficiency of the crossmodal ambient displays system as a personal navigation tool. The goal of the study was to inform our understanding of crossmodal ambient displays, in particular: (1) to determine if the use of a crossmodal ambient display system can improve human performance for a navigation task in comparison to the use of a standard map; (2) to provide an understanding of the impact of the use of a crossmodal ambient display on the performance of a primary task; and (3) to explore the notion of ambience through the measurement of subjective reports of mental workload.

Experimental design

The study used a within subject design. We utilized one independent variable, type of task and navigation tool used, with two levels of treatment: (1) navigation using a map in a dual-task condition (answering arithmetic questions and navigating), (2) navigation using CROSSFLOW in a dual-task condition (answering arithmetic questions and navigating). We also measured user performance on the primary task (answering arithmetic questions) in the absence of navigation. The dependent variables included time of completion, time taken per arithmetic question, accuracy of completion of the primary task, navigation errors, and perceived mental workload.

The 9 subjects, both male and female, were aged between 20 and 30 years old and had no discernable visual or physical impairment. The mathematical ability of each participant was elicited through a screening questionnaire prior to the study. An initial evaluation of performance on arithmetic questions yielded a mean time per question of 4.0 seconds and a mean accuracy of 97%.

The primary task involved answering a set of arithmetic questions posed one-by-one by the experimenter whilst the subject undertook the navigation task. Each subject was videoed and their responses for the questions were recorded during the experiment by the experimenter. The secondary navigation task was for the subject to find 5 targets (out of 15 targets displayed) with the aid of either a map or CROSSFLOW. Around the experimental area there were 15 containers positioned at different locations within the projected image. 5 out of the 15 containers held navigational information for the user.

In the case of the map, the subject was given the name of the first target container, and subsequent target containers contained the name of the next destination. In the CROSSFLOW condition, the containers provided the number to be entered into the phone for the next target. The 10 containers that were not valid destinations contained the statement “this is an incorrect location”. The use of these “distracter” targets was with a view to adding a degree of real world complexity and ambiguity for both the map and CROSSFLOW conditions. As the experiment was conducted in a relatively small space, as compared to an airport, it was necessary to have a significantly denser array of locations for the user to navigate between (see later for a discussion of this aspect of the experimental design).

Fig. 5. Map condition

Fig. 6. CROSSFLOW condition

Procedure

In an initial phase, each subject answered 18 arithmetic questions as a briefing task. The briefing task was intended to evaluate the baseline performance of each user with regards to the primary task. Subjects were asked to try to answer all the arithmetic questions correctly, and no time limit was enforced. On completion of the briefing task each subject navigated in the experimental area (10.0 x 6.5 meters) to find 5 targets out of the 15 targets under both conditions. Figure 5 and 6 show the experimental area with a sample distribution of the targets used in the two conditions. The order of presentation of the conditions for each subject was randomized, as was the set of destinations used for a subject.

Condition 1 was undertaken with the aid of a simple map and subjects gave spoken answers to the arithmetic questions posed by the experimenter. On finding a target, the subject read the name of the next target from the container and continued. The subjects were told that answering the arithmetic questions was the primary task which would not stop until all designated 5 targets were found. No time requirement was placed on the subjects. In the second condition, the primary task was the same and navigation task was completed with the CROSSFLOW system. Each subject used a SPV E200 smart phone and the five 800ms time-slots, corresponding to the five destinations, cycled every 4 seconds and utilised both an auditory and vibration cue as described in the previous section. On discovery of each target, the subject selected the number of the next target on the keypad of phone.

Five measures were collected: (1) the completion time for the navigation task in each condition; (2) the time spent answering each of the arithmetic questions in the briefing phase and in the two conditions; (3) the accuracy of arithmetic question answers; (4) the number of navigation errors; and (5) a subjective measurement of mental workload using the NASA Task Load Index (NASA TLX) [14].

Table 1. Mean performance measures for the map and CrossFlow conditions (9 subjects)

Condition	**Map**	**CrossFlow**
Total time (secs)	133	80
Time per question (secs)	8.5	6.1
Questions correct (%)	84	98
Navigation errors	1.2	0.4
NASA TLX score	79	60

Results

Our hypotheses were supported in terms of primary and secondary task performance, total completion time and effect on mental workload. The descriptive statistics reveal that in contrast with a map the use of CrossFlow resulted in better performance in all aspects. Table 1 presents the mean measures for the two conditions.

Comparison of primary task performance

The performance for the primary task was compared across the briefing task, and the two dual-task conditions: navigating using the map and navigating using the cross-modal ambient display system. Two aspects of the performance on the primary task were considered: (1) the time taken to answer arithmetic questions; and (2) the percentage of correctly answered questions. With respect the time taken, a one-way repeated measures ANOVA revealed very significant differences between performance without and with the navigation task (F(2,16)=42.28; df=2,16; p<0.001). A post-hoc paired samples t-test further show that the average time taken to answer an arithmetic question in the dual-task condition decreased very significantly from using the map to using CrossFlow, t(8)=6.60, p<0.001. The mean time using CrossFlow was 28% quicker than when using the map.

To compare accuracy, a one-way repeated measures ANOVA revealed that the differences were significant between conditions (F(2,16)=4.89; df=2,16; p=0.022). A post-hoc two-tailed paired samples t-test shows that the difference of the accuracy of processing arithmetic questions was only marginally significant between the map and CrossFlow conditions, t(8)=-2.26, p=0.054, with the mean accuracy using CrossFlow being 17% higher than for the map.

Comparison of secondary task performance

The performance on the secondary (navigation) task was compared for the map and CrossFlow condition according to two criteria: (1) total time spent finding 5 destinations (total time in the dual-task condition); and (2) number of navigation errors in discovering the 5 destinations. Navigation errors were recorded formally when subjects addressed the wrong location, i.e. subjects incorrectly identified a distracter location as the next destination and when users returned to a previous prior destination in order to ascertain the location of the next target. In the map condition subjects averaged 1.2 navigational errors and for CrossFlow the average was 0.4. In is apparent that for such a small scale experiment (both in terms of the spatial scale of the navigation problem and the number of subjects) few conclusions can be drawn from such a low error rate. As for total time taken, a paired samples t-test showed that the total time spent on the whole experiment in dual-task condition decreased significantly from using the map to using CrossFlow, t(8)=3.457, p=0.009.

Comparison of judgments of mental workload

A paired samples t-test was conducted on the subjective judgments of the subjects in each of the two conditions using the NASA TLX rating of mental workload. The results show a significant reduction in the perceived mental workload when using CrossFlow as compared to the map, t(8) = 6.24, p < 0.001.

Discussion

The experiment compared the effectiveness of CrossFlow with a traditional map for navigating an indoor environment. The results indicated that subjects using CrossFlow as a navigation tool performed better on both the primary (arithmetic question answering) and secondary (navigation) tasks. This can be explained in terms of the ambient nature of CrossFlow as supported by the NASA TLX reports, and we observe that the attention bottleneck effect is apparent for subjects in the map condition. Although the experiment supports the utility of the crossmodal ambient display system, informal observations of subject behavior should be incorporated in future design iterations of both the system design and the experimental design. Another artefact of the small experimental area is that the dense array of destinations meant that the directions indicated by the display appeared vaguer than would have been the case in a larger area with larger targets. In a number of cases a subject needed to step out of the experiment area in order to gain some perspective on the display and find the next target. Finally, the aesthetics of the experience was not addressed in the experimental set-up, though informal subject feedback received after the study revealed that subjects felt at ease with the system and most found CrossFlow fun to use and helpful.

Closing remarks

Serious notice must be taken of the physical configuration of our preliminary evaluation and its ecological validity. In the real world navigation tasks take place over significantly larger distances, and landmark identification is a significantly smaller component (in terms of time taken) of the navigation task than in our experimental set-up. Furthermore, people do not generally navigate while performing mental arithmetic and the spaces they occupy are usually populated by other people undertaking a range of activities. We intend to address these observations in a multi-user study on a larger spatial scale. The small distances involved in our preliminary study are potentially detrimental for CROSSFLOW too as the shorter the distances are between the destinations, the greater the impact of the user having to wait for the time-slot corresponding to the next destination. In the worse case users will have to wait 4 seconds for the full cycle of directions, and this time is comparable to the time required to move between destinations. However, this is clearly going to be less time that users will typically require to identify a destination on a map and to decide upon appropriate landmarks by which to navigate. The use of crossmodal perception to index temporally multiplexed information has significant potential for applications other the navigation. Although the requirement for floor projection (in this configuration) is onerous, this is outweighed by the fact that there is no need for tracking, which is particularly difficult for indoor environments. We also see potential for crossmodal displays (ambient or otherwise) to address the public-private divide through the display of public, but anonymized, information.

Acknowledgments

We would like to thank Andrew Monk (University of York) for kindly lending us the mobile phones used in the study (these were in turn donated by Microsoft Research Cambridge). We would also like to acknowledge Antonio Krüger (University of Münster) with whom we conceived the original notion of using crossmodal displays for indoor navigation.

References

1. Weiser, M. “The computer for the 21st century.” *Scientific America*n, 265(3):94-104, 1991.
2. Ishii, H. and Uller, B. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, United States, March 22-27, 1997). S. Pemberton, Ed., ACM Press, New York, NY, 1997.
3. Plaue, C., Miller, T., and Stasko, J. Is a picture worth a thousand words? An evaluation of information awareness displays. In *Proceedings of Graphics Interface* (London, Ontario, Canada, May 17-19, 2004). ACM International Conference Proceeding Series, vol. 62. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, 2004.
4. Redström, J., Skog, T., and Hallnäs, L. Informative art: using amplified artworks as information displays. In *Proceedings of DARE 2000 on Designing Augmented Reality Environments* (Elsinore, Denmark). ACM Press, New York, NY, 2000.

5. Pedersen, E. R. and Sokoler, T. AROMA: abstract representation of presence supporting mutual awareness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, United States, March 22-27, 1997). S. Pemberton, Ed., ACM Press, New York, NY, 1997.
6. Prante, T., Stenzel, R., Röcker, C., Streitz, N., and Magerkurth, C. Ambient agoras: InfoRiver, SIAM, Hello Wall. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24-29, 2004). ACM Press, New York, NY, 2004.
7. Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (Santa Fe, NM, USA, October 24-27, 2004). ACM Press, New York, NY, 2004.
8. Mynatt, E. D., Back, M., Want, R., Baer, M., and Ellis, J. B. Designing audio aura. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Los Angeles, California, United States, April 18-23, 1998). C. Karat, A. Lund, J. Coutaz, and J. Karat, Eds., ACM Press/Addison-Wesley Publishing Co., New York, NY, 1998.
9. Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., and Yarin, P. ambientROOM: integrating ambient media with architectural space. In *CHI 98 Conference Summary on Human Factors in Computing Systems* (Los Angeles, California, United States, April 18-23, 1998). ACM Press, New York, NY.
10. Luck, S. J. & Vecera, S. P. Attention. In H. Pashler (Series Ed.) & S. Yantis (Volume Ed.), *Stevens' Handbook of Experimental Psychology: Vol. 1. Sensation and Perception* (3rd edition), pp. 235-286. Wiley, New York, 2002.
11. Driver, J. & Spence, C. Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*(7), pp. 254-262, 1998.
12. Driver J. & Spence C. Crossmodal attention. *Current Opinions in Neurobiology*, Apr; 8(2):245-53, 1998
13. Kray, C., Kortuem, G., and Krüger, A. Adaptive navigation support with public displays. In *Proceedings of the 10th International Conference on Intelligent User Interfaces* (San Diego, California, USA, January 10-13, 2005). ACM Press, New York, NY, 2005.
14. Hart, S. Background Description and Application of the NASA Task Load Index (TLX). In *Department of Defense Human Engineering Technical Advisory Group Workshop on Workload*. Newport, RI, 1987.
15. Wisneski, C., Ishii, H., Dahley, A., Gorbet, M. G., Brave, S., Ullmer, B., and Yarin, P. Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information. In *Proceedings of the 1st International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture* N. A. Streitz, S. Konomi, and H. J. Burkhardt, Eds. Lecture Notes In Computer Science, vol. 1370. Springer-Verlag, London, 1998.