Using Multimodal Interactions for 3D Television and Multimedia Browsing

David Hannah, Martin Halvey, Graham Wilson, Stephen A. Brewster
School of Computing Science, University of Glasgow,
Glasgow, United Kingdom G12 8QQ
davidh@dcs.gla.ac.uk, gawilson@dcs.gla.ac.uk,{first.last}@glasgow.ac.uk

ABSTRACT
We investigate the use of a mobile device to provide multifunctional input and output for a stereoscopic 3D television (TV) display. Through a number of example applications, we demonstrate how a combination of gestural and haptic input (touch and pressure) can be successfully deployed to allow the user to navigate a complex information space (multimedia and TV content), while at the same time visual and haptic (thermal and vibrotactile) feedback can be used to provide additional information to the user enriching the experience. Finally, we discuss our future work exploring the potential of this idea to allow multi-device and multimodal browsing of 3D TV and multimedia.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces: *Haptic I/O*

General Terms
Human Factors

Keywords
Mobile devices, 3D, TV, thermal, pressure, gesture

1. INTRODUCTION
Multimedia content is fast becoming a ubiquitous part of everyday life. There are ever growing possibilities to store, manipulate and consume media. Coupled with this growth in volume there are increasing levels of complexity associated with digital multimedia objects, this can be seen, for example, in the move from standard to high definition and now to 3D video, and this growth in complexity is likely to continue into the future. In addition, the TV is becoming a hub for interacting with much of this content. These large, high-resolution displays can be used for browsing digital photos, choosing music, playing games as well as watching movies and TV programmes. Many TV's are now connected to the Internet allowing access to online content and social media. This progression has led to the development of new products like Apple TV and Google TV which increases the volume and complexity of information accessed from the TV screen. It is important that we provide users with the best options for browsing, searching and consuming this growing and complex content. In many cases, the TV remote control is itself a limiting factor, as normally it only provides simple fixed buttons to interact with the TV, and lacks the flexibility of a mouse or gesture input common on desktop PCs and mobile devices. The aim of this work is to use other available devices, particularly mobile phones, to provide richer interactions. Initial research into using a mobile phone as a universal interaction device [1] highlighted that perhaps universal control over all appliances might not be ideal but that control over particular appliances might be beneficial. We believe that television is the ideal appliance to be controlled via one or indeed multiple mobile devices.

Currently, there are applications that run on mobile phones which provide access to content, but many of these applications simply replicate the 'look and feel' of a regular TV remote. As such, these applications do not take full advantage of all of the features that a mobile phone can offer, e.g. a local display of additional information, gestural input, etc. In this paper we present a set of prototype applications which present potential and practical uses for a mobile phone used to interact with novel visual displays of multimedia content on a 3D TV in three different scenarios: image browsing, electronic programme guide interaction and multimedia consumption. These example scenarios demonstrate how additional inputs and outputs from the phone can enrich the experience of browsing and consuming media on the TV.

2. HARDWARE
A number of different hardware devices were employed for our prototypes. We simulated a 3D TV using a PC with an Nvidia 3D Vision graphics card and active shutter glasses. Users interact with the 3D display using a Nexus One mobile phone connected via Bluetooth. The Nexus One provides a tangible controller for the 3D display. Users can use its touch screen as proxy for interacting with the data on the TV. The touch screen can also act as an additional display. Previous research [2] has identified a number of potential benefits to having secondary screen available via a phone interface for multimedia interaction, namely additional control, methods to enrich content, additional ways to share content, and finally the ability to transfer television content. Users can also use the Nexus device to perform device movement gestures, as it includes an accelerometer. Dachselt and Buchholz have previously investigated throw and tilt interactions with remote displays [3], they investigated continuous and discrete tilt gestures in a couple of media interaction environments including with a 3-D map for Google earth, however, these interactions did not include the TV or multiple types of multimedia.

In addition to the standard input and output modalities currently available on a mobile device, we added two novel modalities:

EuroITV'11, June 29–July 1, 2011, Lisbon, Portugal.

pressure input and thermal output. These have both been shown to be effective at improving interaction [4-5] and could be included in future mobile devices. Pressure input was provided by using two standard force sensing resistors and a linearising amplifier. This allowed users to push at two different points on the back of the device. In this way we have taken advantage of the additional space on the back of the devices which allows for more advanced and complex interactions with the mobile device, as outlined by Baudisch and Chu [6], which in turn provides an extra method for controlling the media on the TV screen. A custom microcontroller board that could drive four Peltier heat pumps was developed [5] (see Figure 1) to generate the thermal output. Peltiers allow for a high level of control over temperature and also allow for both heating and cooling from the same pump. Each Peltier device could be independently controlled with the temperature set anywhere within a range of -20°C to +45°C, accurate to 0.1°C. In the following sections we describe a set of prototypes which demonstrate how mobile devices could be used in conjunction with 3D TV's to create richer interactions with the television and multimedia content.

Figure 1: Nexus One phone, FSR's and microcontroller are attached to the back of the device. On the left is the Peltier Microcontroller with two Peltier heat pumps detached from the back of the phone.

3. INTERACTION WITH MEDIA

In this section we describe three example application scenarios and demonstrate how the gestural and pressure input as well as thermal output could be used to enhance user interaction in these media interaction scenarios.

3.1 Browsing Image Collections

Many displays offer functionality to browse image collections e.g. table top devices, laptops, mobile phones etc. TVs are also commonly used to display digital photo collections to friends and family. The use of a mobile device to interact with these types of displays has previously been investigated, previous research has looked at the use of a mobile device to upload and interact with images on a shared display [7] or applications for Web browsing, voting or interaction [8].

3D displays allow for some alternative views of multimedia content. Figure 2 shows an example of a visualisation which allows the user to view a collection of images in a similar way to the iTunes Cover Flow, but with the content stretching back into the screen in 3D. This system can be used in two ways. First, as a slideshow where the images automatically scroll or second using user-controlled free scrolling. In slideshow mode, information about the current image is displayed on the screen of connected phones, along with a local version of the image for closer inspection. In the free scrolling mode swipe gestures on the touch screen of the phone are used to navigate through the collection on the TV. This could be annoying for others watching the TV as it disrupts their viewing of the slideshow. However, one other benefit of using a mobile phone as a remote is that there are likely to be many remotes rather than just one as most viewers are likely to have a phone, potentially allowing multiple viewers to have input into controlling the slide show. In part to address the problem of annoying other viewers, in slideshow mode it is also possible to 'disconnect' from the show and browse the photos on the phone as in free scrolling mode; in this case the TV will still show images every few seconds but they are not pushed to the phone. This allows individual users to explore the collection in different ways, or focus on a particular image for longer without annoying other viewers. Users can then re-sync to the slideshow if they wish.

Figure 2: A 3D visualization of a collection of images.

This interaction example demonstrated some of the flexibility of using a phone as a remote controller; it can duplicate the TV screen, or individual users can have individual displays of information, allowing multiple views of the same data through multiple devices, realising some of the possibilities outlined by Cesar et al. [2]. The technique proposed could also be effective for public displays where slide shows or a rotating set of adverts are shown, for example. Silfverberg *at al.* [9] have performed a study of controlling the cursor on a public display using the joystick on a handheld device. Typically the phone is only used as an input to the wall display but in our case could also be used as a local display.

Figure 3: Pressure input on the mobile device causes images on the 3D display to kink out to allow images at the back to be viewed more clearly.

One drawback of the visualisation outlined in Figure 2 is that while it takes advantage of the 3D space it can be difficult to view images other than the image currently at the top of the "queue". The use of pressure could be used to overcome this problem and improve the user experience while using the TV display to view images. Figure 3 shows a version of one of Ramos *et al.*'s [10] pressure widgets in use. Here, the stream of photos on the TV kinks in 3D when pressure is applied on the phone, with more pressure resulting in more marked kink; images further down the stack can then be more easily seen. This interaction exploits both the easy interaction with the mobile device and the additional visual space allowed by the 3D display to show more information.

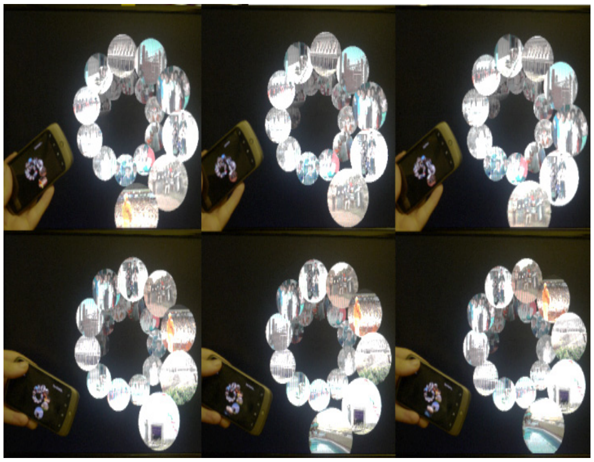

Figure 4: Gestures with a phone to browse through spiral representation of images.

Figure 4 shows an alternative solution to the "shadowing problem" outlined above. The same collection of images is arranged in spiral visualization with the current image in front and the tail of the spiral going into the screen in 3D. As the user rotates the phone accelerometer data is filtered and commands sent to the TV to rotate the spiral forwards or backwards, with different photos being brought into focus at the top of the spiral. Gestures performed by moving the phone as opposed to gestures performed on the phone touch screen can create some new issues. When moving the phone to generate input, the phone screen can no longer be seen easily and any extra information presented on it cannot be used. Care must be taken to use the appropriate type of gesture and local display so that both are used effectively.

Thus far we have described the implementation of a number of demonstration applications that take advantage of many features of a mobile phone to allow browsing of relatively simple (in terms of what is presented on the screen) image layouts. The input methods described above can also be used to navigate through more complex structures than the linear ones presented above. Figure 5 shows a hierarchy of images represented as a cone tree in 3D [11]. This representation allows the viewer to see more images in their relative positions within a file structure and also displays the relationships between images in the hierarchy. The user uses pressure input to traverse up and down the hierarchy, by squeezing on one side of the device to zoom in and the other to zoom out. When zooming in, the lower level elements becoming gradually more visible as they expand to populate the top layer shown in Figure 5. By rotating the phone itself the display is rotated to bring alternative sub-trees to the front of the visualization. Zooming and rotating can easily be carried out in parallel for a rapid and smooth interaction. This same kind of interface could potentially also be used for browsing other hierarchical content such as music or files.

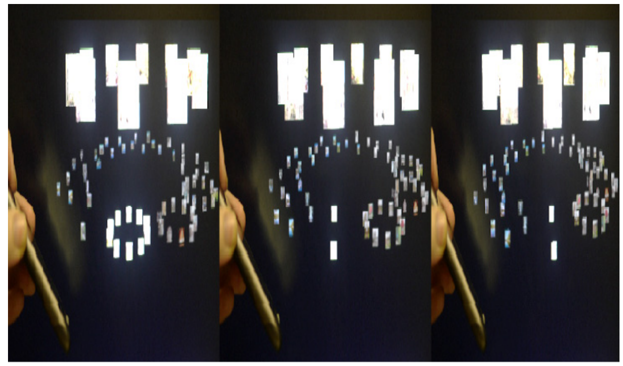

Figure 5: Gestural input with the phone to browse through a cone tree of images in 3D on the TV screen.

3.2 Electronic Programme Guides

Richer interactions can also be used to control an electronic programme guide (EPG). EPG's contain a lot of dense information and, when many TV channels are available, can often be difficult and slow to navigate using a regular TV remote. There are already applications to allow users to browse EPG's on additional devices e.g. tablets or phone; however these are not always connected to the TV. This is another domain in which alternative methods of input e.g. gesture and pressure could be utilised to provide a better user interaction.

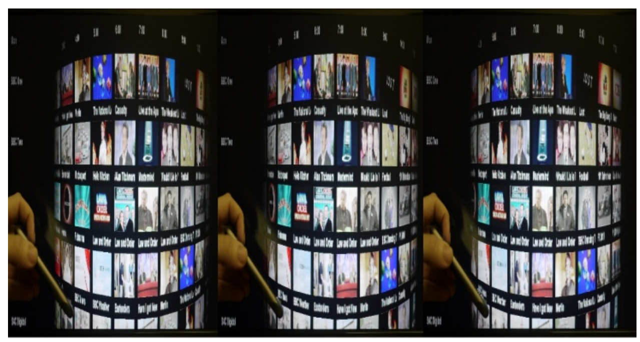

Figure 6: Three views of an EPG displayed as a fisheye on a cylinder in 3D.

Figure 6 shows a possible 3D representation of an EPG. The fisheye view in 3D gives the user an idea of programmes close to the time being viewed; users can see and compare content easily and efficiently as the visualization makes full use of the TV screen. The program information is modelled as a spiral viewed from the side, where a 24 hour period is displayed on a full 360° rotation of the spiral, with channels on the y axis. The spiral view means that there are no discontinuities between the days; the next day joins smoothly to the previous one. In many EPGs different days are presented on different screens, making the transition from one day to the next more awkward. The viewer can preview what is on earlier or later by rotating the phone to rotate the spiral. Tilting the phone moves up and down through the channels. By squeezing the phone the user can control the scaling of the view, for example zooming in to a particular time/channel area. Pressing harder causes the view to "pop" through to the following day at the same time, allowing rapid skipping through different days. A particular programme can be selected by tapping on the phone screen. This brings up a page about the show and allows the it to be played on the 3D TV. In addition to this view on the 3D TV display, the user can also use the phone to provide a source of additional information or indeed alternative views of the same data. The intention is to allow a person to get more information about a particular show while still looking at the overall view on the 3D display or indeed to allow a user to view information for different days, times etc. without disturbing other people looking at the 3D EPG.

3.3 Consuming Media

Finally the use of multimodal interactions through a mobile device in this context is not limited simply to being that of a controller. The mobile device can also be used to provide output to enhance the user experience and also provide additional information for the user. For example, in our prototype demonstration it is possible for the vibrotactile motor in the phone to vibrates in relation to TV predefined content, e.g. when a train goes by on screen, to enhance the experience of the viewer [12]. Thermal output is also quite suited to providing additional information for a media presentation or video. Thermal output has particular hedonic features that may not be available in other modalities, as has already been highlighted by Nakashige *et al.* [13] amongst others. In our implementation heating and cooling effects have been added to a number of videos, e.g. the device warms up for desert scenes or for happy moments and cools for arctic or sad scenes. To work in real situations some metadata would be need to be sent to the phone to trigger these effects, but our hardware shows that this is possible.

4. CONCLUSIONS AND FUTURE WORK

In this paper we have described an implementation and potential uses for using a mobile phone as a remote control which has richer input and output for a 3D TV that can present a range of different media types. Many of the viewers of the TV will have these devices with them and they can be used to improve the interaction with the rich content on the TV, which is often limited by the remote control. The use of the additional device allows multiple visual representations of the same data, allows for using more interactions through a variety of modalities for interacting with complex data and also allows for augmentation of visual data with additional multimodal information to enhance user experience and understanding. The next major stage of this work will be to demonstrate how these types of interactions and information display would operate in real life. This will involve evaluating the performance of users in using the different interface configurations to browse and interact with multimedia. Given the success of current gestural gaming platforms (e.g. Nintendo Wii and Microsoft Kinect) and the increase in the use of rich touch interactions on mobile devices, the integration of these techniques into a mobile phone and the use of the phone as a controller are the next steps forward. We believe that these techniques provide a mechanism for increasing the accessibility of 3D and other visual devices, as well as providing one simple and easy to use device which could replace the myriad of different controls that are currently used.

5. Acknowledgements

This research has been funded by the Industrial Members of MobileVCE (www.mobilevce.com), with additional financial support from EPSRC grant EP/G063427/1.

6. REFERENCES

[1] Roduner, C., Langheinrich, M., Floerkemeier, C. and Schwarzentrub, B. Operating Appliances with Mobile Phones - Strengths and Limits of a Universal Interaction Device. In Proceedings of the International Conference on Pervasive Computing (2007)

[2] Cesar, P., Bulterman, D. C. A. and Jansen, A. J. Usages of the Secondary Screen in an Interactive Television Environment: Control, Enrich, Share and Transfer Television Content. In Proceedings of the 6th European Conference EuroITV (2008)

[3] Dachselt, R. and Buchholz, R. Natural Throw and Tilt Interaction between Mobile Phones and Distant Displays. In Proceedings of the Conference Extended Abstracts on Human Factors in Computing Systems (2009)

[4] Wilson, G., Stewart, C. and Brewster, S. Pressure-Based Menu Selection for Mobile Devices. In Proceedings MobileHCI International Conference on Human computer interaction with mobile devices & services (2010).

[5] Wilson, G., Halvey, M., Brewster, S. A. and Hughes, S., A. Some Like it Hot? Thermal Feedback for Mobile Devices. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (2011)

[6] Baudisch, P. and Chu, G. Back-of-Device Interaction Allows Creating Very Small Touch Devices. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (2009).

[7] Cheverst, K., Dix, A., Fitton, D., Kray, C., Rouncefield, M., Sas, C., Saslis-Lagoudakis, G. and Sheridan, J. G. Exploring bluetooth based mobile phone interaction with the hermes photo display. In Proceedings MobileHCI International Conference on Human computer interaction with mobile devices & services (2005).

[8] Paek, T., Agrawala, M., Basu, S., Drucker, S., Kristjansson, T., Logan, R., Toyama, K. and Wilson, A. Toward Universal Mobile Interaction for Shared Displays. In Proceedings of the ACM Conference on Computer Supported Cooperative Work (2004).

[9] Silfverberg, M., MacKenzie, I. S. and Kauppinen, T. An isometric joystick as a pointing device for handheld information terminals. In Proceedings of the Graphics interface 2001 (2001).

[10] Ramos, G., Boulos, M. and Balakrishnan, R. Pressure Widgets. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (2004).

[11] Robertson, G. G., Mackinlay, J. D. and Card, S. K. Cone Trees: Animated 3D Visualizations of Hierarchical Information. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (1991).

[12] Krol, L. R., Aliakseyeu, D. and Subramanian, S. Haptic feedback in remote pointing. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (2009).

[13] Nakashige, M., Kobayashi, M., Suzuki, Y., Tamaki, H. and Higashino, S. "Hiya-Atsu" media: augmenting digital media with temperature. In Proceedings of the Conference Extended Abstracts on Human Factors in Computing Systems (2009).