Gesturing at Architecture: Experiences & Issues with New Forms of Interaction

Abstract
When developing public installations, interaction designers are able to utilise increasingly natural modes of expression such as speech, gesture and touch. Conversely the resulting installations often place users in situations where they are confronted with entirely unnatural forms of interaction. How do we establish an understanding of peoples’ behaviour in such situations, and what bearing could this have on the design of better interactive experiences? This paper addresses these questions, drawing upon a study of a high profile installation that invited members of the public to control the lights on the London Eye using hand movements and heart rate measurements.

Author Keywords
Media & art; gestural interfaces; architecture; London Eye; Microsoft Kinect; mood; emotion

ACM Classification Keywords
H.1.2 [User / Machine systems]: Human information processing; H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; I.2.10 [Vision and scene understanding]: Video analysis.

General Terms
Design; Experimentation; Human Factors
Introduction
The last decade has heralded a new era for human-computer interfaces, characterised by technologies that seek to reflect and imitate our evolved interactions with other humans and our physical environment. Where we once used clunky joysticks and keyboards, computers are now able to recognise more natural forms of input such as touch, gesture and speech. What’s more, the hardware associated with these interfaces is becoming increasingly miniaturised and affordable. Consequently such technologies are particularly amenable for applications in public spaces, where they can be deployed in short time frames, and with little impact on existing infrastructures. The versatility afforded by this convenience has led interaction designers to come up with increasingly innovative interactive experiences (for an example see 'Perspective Lyrique' by creative collective 1024 Architecture). For those partaking in such experiences the novelty of the interaction is what provides much of the excitement and intrigue. However, for the psychologist, social scientist and interaction designer the obscure nature of these new forms of interaction poses many interesting questions.

We believe that one of the best ways to address these questions is to encourage interaction researchers and designers to undertake collaborative studies involving the collection and analysis of real-world data from public installations. During the summer of 2012 we had the opportunity to be involved in such a study when we worked with Cinimod Studio in London to design and analyse a high profile interactive installation called the Mood Conductor.

---

Background
The Mood Conductor combined both gestural and physiological interfaces in order to allow a person to interact with an architectural landmark. We briefly discuss these modes of interaction below.

Gestural Interfaces
When we talk of gestures in human-computer interaction we are loosely referring to the use of finger, hand or bodily movement as a means of input. Touch screens act as two-dimensional gestural interfaces allowing simple gestures to be performed, such as pinching two fingers together to zoom in on an image. By sensing motion in three-dimensional space more complex and expressive gestures can be performed. In 2010 the release of the Kinect brought motion sensor based interaction into the homes of millions of people around the world, allowing users to interact with computer games using free bodily motion. Since then the development and release of open source drivers has led to it being widely adopted by third party developers in a range of applications including art, advertising, entertainment and healthcare (see the Kinect Hacks website for examples). A problem with the Kinect as a gestural interface is that it lacks the resolution to accurately detect small movements. This issue could be addressed by the LEAP motion sensor, purportedly capable of tracking finger movements to a precision of 0.01 millimetres. Due for release early in 2013, the LEAP sensor is evidence that gestural interfaces are here to stay and will have an important role in the human-computer interfaces of the future.

---

1 http://www.1024architecture.net/en/projects/
2 http://cinimodstudio.com/
4 http://www.kinechacks.com/
5 https://leapmotion.com/
Physiological Interfaces
The electrical activity of the brain, heart and muscles, as well as biochemical activity such as perspiration, can all be measured using biosensors, which have become increasingly low-cost and unobtrusive in recent years. Since many aspects of physiology are closely linked to cognitive processes it is possible to relate physiological measurements to felt experiences such as emotion and mood. This has led to the use of physiological signals in domains of interactive art and media such as computer gaming [1], musical performance [3], and film [2].

Architecture as a Medium for Interaction
The use of architecture\textsuperscript{6} as a component in digital art is a concept that has been popularised in recent years by the practice of projection mapping. Projection mapping involves manipulating 2D projections such that they appear to be superimposed on the surfaces of 3D shapes. When coupled with well-designed content the effect can be incredible, transforming static buildings and objects into surreal, colourful and animated forms. There are, however, only a few examples of projection mapping projects that feature an element of live human control or interaction (see works by the companies Seeper and YesYesNo)\textsuperscript{7}.

Fixed lights, soundscapes and movement also present opportunities for people to control or interact with physical structures. As concrete and glass buildings strive for city skyline dominance, new technologies hold fantastic potential in enabling people to use these structures as instruments for human and artistic expression.

The Mood Conductor
Overview
To coincide with the London 2012 Olympics, the commercial sponsors of the London Eye, ran a public relations campaign titled 'Energy of the Nation'\textsuperscript{8}. As part of the campaign Cinimod Studio were commissioned to develop an installation which would allow individuals to represent their mood by taking control of the lights on the London Eye. The installation was named The Mood Conductor and over the course of three weeks it invited members of the public to use the motion of their hands to control the 320 lights that line the rim of the Eye (Figure 1). Participants were also given the option of wearing a pulse monitor that enabled their heartbeats to be represented as pulsing sections of light at the top of the Eye.

Design & Implementation
A single computer was used as the system controller, taking the heart monitor and motion sensor inputs and generating the lighting control signals for the Eye. We wanted the interaction to be simple, given the short amount of time which each participant would have to familiarise themselves with the installation. Hand tracking was implemented using a Kinect motion sensor positioned in front of the participant. The angles between each hand and the participant’s torso were then used to control the positions of two segments of lighting content on the Eye (Figure 2). Three separate styles of content were developed to reflect different mood states (Figure 3). A particular content style was chosen for the duration of each participant’s turn based upon his or her initial (first 10 seconds) heart rate and hand movements (Table 1).

\textsuperscript{6} Term used loosely to include physical structures and landmarks
\textsuperscript{7} \url{http://seeper.com/} & \url{http://yesyesno.com/}
\textsuperscript{8} \url{http://www.edfenergy.com/energy-of-the-nation/}
Table 1. Content styles, their intended mood and the selection criteria based upon initial pulse rate and hand motion

<table>
<thead>
<tr>
<th>Content Style</th>
<th>Intended Mood</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave</td>
<td>Calm &amp; relaxed</td>
<td>Low</td>
</tr>
<tr>
<td>Spectrum</td>
<td>Flamboyant &amp; creative</td>
<td>Med</td>
</tr>
<tr>
<td>Fire</td>
<td>Energetic &amp; aggressive</td>
<td>High</td>
</tr>
</tbody>
</table>

The heart monitor was custom built and consisted of an ear clip (placed on the ear lobe) and a matchbox-sized transmitter that was capable of sending data wirelessly to our system controller.

Analysis & Results

Over 1 million instances of data (captured 25 times per second) were collected from roughly 800 participants. The data comprised joint coordinates, pulse rate and content style. We carried out a broad analysis of the data that included spatial and kinematic characteristics of hand movements as well as behavioural observations relating to the types of gestures used and the influence of the lighting content.

In order to visualise the spatial characteristics of hand movements we generated histogram images. In Figure 4 the brightness of each pixel of the image is positively related to the amount of hand movement at that particular X-Y coordinate. Using a Mixture of Gaussian clustering algorithm we were then able to outline distinct areas where hand movement was most prominent (Figure 5).

We then compared the average hand velocity and heart rate for participants, separated by the content style they were interacting with (Table 2). The aim of this analysis was to explore whether the content style had an effect on behaviour. In particular we see that people interacting with the ‘Spectrum’ content style exhibit faster average hand movements and higher average heart rates.

<table>
<thead>
<tr>
<th>Lighting Content</th>
<th>Mean Hand Velocity (m/s)</th>
<th>Mean Heart Rate (bpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave</td>
<td>1.49</td>
<td>94.4</td>
</tr>
<tr>
<td>Spectrum</td>
<td>1.85</td>
<td>102.1</td>
</tr>
<tr>
<td>Flame</td>
<td>1.52</td>
<td>96.7</td>
</tr>
</tbody>
</table>

Table 2. For each content style the average hand velocity and heart rate is shown for all participants interacting with that content.

By using the joint coordinates to reconstruct and visualise hand movements we manually analysed the gestures which participants on the Mood Conductor performed. The result was a catalogue of 10 of the most commonly observed gestures, each with a short description and illustration (see Figure 6 for three examples).

In addition to data-based analysis we also used qualitative techniques, such as informal interviews and video recordings, to analyse peoples’ experiences with the installation. One of the most interesting observations was how, in the absence of any interaction instructions, people formulated their own rules and ideas about how the installation worked. When asked how it felt, participants reported a great sense of power when controlling the lights on the Eye.
Discussion

The use of histogram images (Figure 4) showed that the majority of hand movements occurred along a circular pathway centred on the participant’s torso. It may seem trivial to conclude that this was related to the circular shape of the London Eye. However, it leads us to question the extent to which more complex shapes might influence the perceived interaction space in situations where the gestural interface allows free movement.

Using quantitative measures (Table 2) we found that participants interacting with the ‘Spectrum’ content behaved differently to those interacting with the other two content styles. It is not possible to draw any definite conclusions as to why this was, however the ‘Spectrum’ content did exhibit more colour variation than the other two content styles. Despite lacking specificity, such quantitative measures could be adopted as tools for evaluating links between content and behaviour in future installations.

By cataloguing gestures (Figure 6) we were able to summarise how people interacted with the Mood Conductor. We were surprised by the frequency at which gestures re-occurred between nights and participants, especially in the absence of any instructions on how to interact with the installation. There was a tendency for gestures to be performed in the coronal plane, with relatively short and repeatable movements, which exhibited high degrees of rhythm, synchrony and symmetry. Qualitative measures such as cataloguing could enable interaction designers to develop their designs with knowledge of the types of gestures that may be used in a particular interaction situation.

Through our involvement with a public installation it was possible to obtain a large amount of data collected from a non-laboratory environment. The downside is that it is difficult to obtain any control data in order to carry out rigorous statistical analysis. This issue limited the extent to which we could investigate heart rate related hypothesis.

In conclusion, our study demonstrated the benefits of collecting and analysing data from interactive public installations. It is likely that such installations will become increasingly commonplace over the coming years. In response, interaction researchers and designers should engage in collaborative work more frequently, progressing our understanding and improving the quality of interactive experiences.

Acknowledgements

This work is supported by the Media and Arts Technology Programme, an RCUK Doctoral Training Centre in the Digital Economy.

References

