

# **Towards a conceptual framework for touchscreen 360° video on small mobile screens: praxis in action**

Joe Citizen

## **Abstract**

This paper describes a praxis approach to the conceptual framework required for producing touchscreen 360° video on small mobile screens, using a Point Grey Ladybug 3 camera. It is specifically concerned with how theory can help inform practice within a tertiary education environment, through a survey of the available literature of this emerging field. It concludes with a brief summary of a 360° video project shot as a result of these preliminary findings, and offers some possibilities for further development.

## Contents

As a tertiary educator teaching in New Zealand, is it possible to remain current during an era of technological convergence, and to do so in a cost-effective manner? Can theory be used to inform practice so that production costs are kept to a minimum? Is it possible as Laudrillard (2002) puts it, to have “innovation through technology” by using the available resources of a tertiary degree programme? Using an action research model, this paper presents some preliminary findings of a theoretical framework that is being used to inform the production methodology of a working prototype – an interactive drama that uses 360° video played through small touchscreen mobile devices such as an Apple iPad or similar. It is an experiment located within a larger multidisciplinary collaborative project concerned with exploring some possible convergences between the codes and conventions of moving image, interactivity and immersion.

The desire for immersion has a long history perhaps best described by Grau (2003) in his seminal work *Virtual art: From illusion to immersion*. For him, immersion must allow the user to be fixed within a “hermetically closed-off image space of illusion” (p.5). By using a Point Grey Ladybug 3 video camera that uses six cameras to “enable the system to collect video from more than 80% of the full 360° sphere” (2012) it appears to be possible to record such a hermetically closed off space. Playback of 360° video to an immersive installation however appears both costly and time consuming, as current resources do not currently provide for fully immersive projection. Instead, a Flash based player will be used to stream together still panoramic images (See Figure 1).



Figure 1. Flash based player used to stream together still panoramic images

Since a user’s view looks similar to many first person videogames, an exploration into the codes and conventions of videogame interactivity is useful. Whilst it is beyond the scope of this paper to describe the historical debate between ludologists and narratologists in the production of meaning for videogames, it is still worthwhile to consider how ‘middle ground’ responses may inform the production of meaning for interactive 360° panoramic video. Clarke and Mitchell identify how game designers

unlike filmmakers cannot use a close up to reveal detail, but must rely instead on “one narrative technique: that which we refer to in film as *mise en scene*” (2001, p. 85).

Henry Jenkins asks us to consider how meaning can be understood through “game design as narrative architecture” and the possibilities of “embedded narratives” (2002) as ways in which users can refer to their own inter-textual knowledge. Some videogames attempt to use an interactive narrative structure, which attempt to provide users with a sense of agency in order to provide emotional responses in reaction to choices they have to make in the course of playing the game. As Kennedy (2011) points out, many of these narratives either use a decision tree model or a complex artificial intelligence engine, usually based around character interactions with reasonably limited and linear outcomes. As he puts it, this “creates the verisimilitude of choice, but in reality it is still a linear experience, and the viewer-user is aware of this”. But do all interactive narratives have to be based around these models? If, as Jenkins suggests, inter-textuality can reference layers of meaning according to the knowledge of the user, then are there specific genres or hybrid genres that evoke particularly rich layers of user’s knowledge? Can users then map narratives with some level of agency?

Jenkins himself explicitly refers to the example of the detective story, which as a genre refers both to jumps in time and to uncovering clues in the pursuit of solving a mystery. Using this as a model, the conceptual framework for an interactive narrative script can begin to be conceived. Within a 360° sphere the possibilities for hiding clues are multiplicitous, but unlike a traditional frame the user is able to scan the environment at will. This ability to move within the sphere exerts an interactive agency not available from cinematic drama, so genre and *mise en scene* alone are not enough to maintain the interest of what Darley terms a “fascinated spectator” (2000), which refers to the need for contemporary audiences to be constantly dazzled by spectacle in cinema. This bedazzlement cannot occur by players simply doing as they please in games however – their agency is limited by the rules. As Bunting, Hughes and Hetland (2012) put it, it is “the gamemaker’s role is to create a canvas upon which all players’ various game-stories can be told”. It will be necessary for our prototype users to have a sense of creating their own stories, not simply through referencing their inter-textual knowledge, but by their different experiences of navigating the game space.

Individual acts of navigation therefore become critical factors in informing these experiences. The fingertip control available through small mobile touchscreen devices allows users to quickly and continuously scan the available 360° sphere, and enable zoomed in inspection. The placement of clues within the mystery story world is not necessarily limited to the foreground conventions of dramatic film, for although 360° video does not enable close up shots without camera movement, the user can reveal deeper layers within the image through pausing and zooming functions commonly used by iPads and other tablets. This experience of touching in order to see new information, challenges the primacy of vision as the way by which media is known and understood, or as Verhoeff says “it transforms the practice of screening as tactile activity into a haptic experience of this practice” (2012). To some extent then, the ability to manipulate imagery through touch is to allow a subjectively

embodied experience to occur in the process. This claim however is subject to some debate – can embodiment be achieved by fingertip control alone? Farrow and Lacovides (2012) maintain that for “a convincing and immersive experience, one should be more or less unaware of the way in which it is being mediated”. Can small mobile touchscreens ever truly be immersive, if they are always constantly being engaged with in the context of our wider environments? Perhaps some sense of embodiment within the 360° virtual space of a small mobile touchscreen could be mediated through sound however, for as Dyson argues, a shift from the visual to the sonic may be partially engendered by virtual audio technology:

*...[It] acknowledges and territorializes the presence of the body in the environment in a manner that the visual aspect of VR is only just beginning to encompass. Like the visual component of VR, virtual audio technology creates an interactive aural space or “sound field,” in which sounds are heard as if they occurred in the lived environment. If digital audio transforms the ear of the listener into a hearing eye [emphasis in original], virtual audio threatens to shift this visual orientation. (2009, p.138)*

If virtual audio technology can help to produce a sense of embodiment, and therefore help with orientation and navigation, the relationship between recording and playback will need further exploration. This perhaps highlights the software – hardware interdependence of this type of media at this time. The Ladybug 3 is not a typical ‘run and gun’ camera, it is entirely software dependent with regards to both its recording and playback. Small mobile touchscreens, however, do allow 360° video to be played using a Flash based player. Within the inter-relational space afforded by spherical representation the user is located at being at its centre, but is simultaneously denied the full panoramic gaze. Instead ordinary framic space is seen, with which the user then physically interacts with to see the rest of the sphere. Consequently, although 360° video promises the collapse of the subject-object space back into a field of relations, the physical navigation of the user reanimates it through temporal dislocation. This allows the continuity of space to be interrupted by the continuity of time, as each new view is seen through a frame navigated sequentially through the user’s actions.

Additionally, the use of the murder mystery genre as a narrative convention also creates linear time limits to this negotiation of space. Users must mentally reconstruct their stories within the available time or play again. Knowing this, how can users be directed or misdirected within the timeframe of the story space, so that both dramatic tension and journey agency are maintained? Can navigating the touchscreen interface also be used to focus the user towards preferred meanings? It is Bruno (2010) that reminds us that the panoramic view is akin to the experience of walking through a city and turning one’s head to view different perspectives. The navigation of the 360° sphere could therefore privilege certain perspectives over others, through the object ‘centring’ technology that many touchscreen photographic viewers already use. Because the Ladybug 3 camera uses independent cameras to capture video and can export these views separately, then re-stitching and centring could occur in the player. By considering each of these views as a sector, object-orientated events can be triggered, and by moving the camera new panoramic views will be revealed or hidden.

Since the touchscreen interface allows for sound and vibration to be used as interactive feedback, then any navigational interface needs to consider how they can be used to inform spatial distance within the nominal 360°. This has particular ramifications for dialogue – in a city we turn our heads towards louder or quieter noise to determine our proximal relationship. Due to the simultaneity of events occurring within a 360° sphere, image centring technology could trigger events to enable both sequential and simultaneous dialogue, depending on which camera view a user was currently looking at. Similarly, the use of the presence or absence of light and/or moving bodies as a time-based environmental factors, can direct attention within image-centred compositions. The choreography of performance in relation to the camera therefore becomes critical. Bruno reminds us too that filmic acts of movement through a city's architecture is also a psychological journey: "Like the city, motion pictures move, both outward and inward: they journey, that is, through the space of the imagination, the site of memory, and the topography of affects". Similarly, traversing the architecture of the story world doesn't just reveal new vistas, nor simply reference the mise en scene conventions of genre, but is capable of accessing different conceptions of time, memory and affect.

Film tends to link spatial movement and conceptual space through the use of editing, but navigation of first person videogames conventionally avoid cuts without the use of rhizomatic hyperlinks or portals. Clarke and Mitchell (2001) note, however, that this is a convention, and one that is "particularly strange given the discontinuity of the player's experience" (p. 86). Bruno argues that Eisenstein's concept of montage came from architecture:

*"[A]n architectural ensemble... is a montage from the point of view of a moving spectator. ... Cinematographic montage is, too, a means to 'link' in one point – the screen – various elements (fragments) of a phenomenon filmed in diverse dimensions, from diverse points of view and sides." (1980, pp. 16-17)*

Since each camera's perspective within the 360° sphere can be mapped and centred, can continuity and discontinuity editing still occur? Is it therefore possible to either construct sets that graphically match one another for each centred frame, or cut into new action regardless in order to create symbolic meaning? The centring capacity that many small mobile touchscreens can be programmed to use, may allow 360° video to be navigated as a series of views, whilst retaining its videogame-like appearance. By choosing to reference a genre that allows jumps in time and the uncovering of a mystery, users can utilise the zoom and panning functions of touchscreens to seek out clues. These clues are able to allow users to reference their own intertextual knowledge and thus impart deeper layers of meaning. The action of navigation will additionally limit the number of sequential views that can be experienced within a given timeframe. Narrative closure will only occur however when the user is able to link the available clues presented out of chronological time, through their own imaginative reconstruction of events.

These theoretical underpinnings of a production methodology are in the process of being tested. The prototypical model used for the experimental drama with the working title *Revolutionary Acts* was shot on the grounds of Wintec on the 4th July

2012 (See figure 2). The working methodology considered each camera view as a potential image-centred sector through which preferred meanings could be produced for the direction or misdirection of potential users, in relation to an individual's touchscreen navigation.



*Figure 2. Revolutionary Actswas, July 2012*

Consideration of the virtual audio/player relationship led to thinking about a script for a navigable environment rather than for a sequence of events. This in turn led to the need to conceptualise as to how dialogue and sound design might work together, as sound now becomes a critical determinant for both orientation and the production of meaning for individual users. Dialogue was recorded using radio lapel microphones onto eight separate tracks to enable either sequential or simultaneous playback. Each individual sound level can now be independently adjusted depending on the 'front' view of a user, in relation to sector based events.

Since 360° video records everything except that which is immediately below it, the lighting design also needed to be reconsidered. The stitching between each camera on the Ladybug 3 means that the contrast ratios need more attention than a single lens. One option was for the illuminants to be built into a set, but due to the zero budget available for this research another solution was required. Consequently, the murder mystery was located on a film studio, a rationale that enabled the use of practical lighting and allowed for consistent contrast ratios through the use of illuminants being set to the same intensity against black background. This conceit also allowed for a wide range of lighting to be tested against our operational concerns of resolution, frame rate, flicker and flare.

By considering how an individual's navigation of the world of the story intersects with their own production of meaning, rethinking the choreography of the camera with performers and architecture was required. *Revolutionary Acts* solved this problem by placing the Ladybug 3 camera on a dolly and pulling it through the environment. Whilst this is comparable to the convention of seeing the hand that holds the weapon found in some first person videogames, it does not really allow for the possibilities of exploring montage. Other approaches could include the use of steadicam rigs, wheeled tripod dollies, or similar. Actors were asked to ignore the presence of the

Ladybug 3 in the same way as the camera is usually treated in conventional drama. This however emphasised the return to a field of relations within the context of the shoot – instead of being pointed at them, the camera was now very much part of the environment. The on-set boundaries for this production between cast and crew consequently became considerably less distinct.

As the project is now in the software testing and building phase it is anticipated that further research will be required. Nevertheless, this praxis approach has yielded a prototypical methodology – one that seeks to reconsider some potential shifts from a supposed visual dominance in immersive environments to sonic, haptic and navigational concerns. These potential shifts however also need to be considered in relation to the size of small mobile touchscreens, and their ability to provide an interface for 360° video and locational sound. The ramifications for production are far reaching too, as the relationships between sound, lighting, choreography, performance and environmental design are re-examined.

## References

- Bruno, G. (2010). Motion and emotion: Film and haptic space. In *Revista Eco-Pos*, 13(2). Retrieved June 14, 2011, from Da Universidade Federal Do Rio De Janeiro: [http://www.pos.eco.ufrj.br/ojs-2.2.2/index.php?journal=revista&page=article&op=view&path\[\]=373](http://www.pos.eco.ufrj.br/ojs-2.2.2/index.php?journal=revista&page=article&op=view&path[]=373)
- Bunting, B. S., Hughes, J., and Hetland, T. (2012). The player as author: Exploring the effects of mobile gaming and the location-aware interface on storytelling. In *Future Internet* (4), 142-160. doi: 0.3390/fi4010142
- Clarke, A. & Mitchell, G. (2001). Film and the development of interactive narrative. Proceedings in the international conference of virtual storytelling: Using virtual reality technologies for storytelling. *Lecture Notes in Computer Science*, 2197, 81-89.
- Darley, A. (2000). *Visual digital culture: surface play and spectacle in new media genres*. London: Sage.
- Dyson, F. (2009) *Sounding new media: Immersion and embodiment in the arts and culture*. Berkley: University of California Press.
- Farrow, R. and Iacovides, I. (2012). 'In the game'? Embodied subjectivity in gaming environments. In *Proceedings in the 6th international conference on the philosophy of computer games: The nature of player experience* (pp. 29-31). Madrid, Spain.
- Grau, O. (2003). *Virtual Art: From illusion to immersion*. Cambridge, MA: MIT Press.
- Jenkins, H. (2002). Game design as narrative architecture. In Harrington, P. and Frup-Waldrop, N. (Eds.). *First person*. Cambridge, MA: MIT Press.
- Kennedy, J. (2011). Triggering core emotional experiences from interactive narratives. *Journal: Creative Technologies* (2). Retrieved November 9, 2011, from <http://journal.colab.org.nz/article/13>
- Laurillard, D. (2002). *Rethinking university teaching: A conversational framework for the effective use of learning technologies*. (2nd ed.). Milton Park, UK: RoutledgeFalmer.
- Point Grey Research, Inc.* (2012). Retrieved June 18, 2012, from [http://www.ptgrey.com/products/ladybug3/ladybug3\\_360\\_video\\_camera.asp](http://www.ptgrey.com/products/ladybug3/ladybug3_360_video_camera.asp)
- Verhoeff, N. (2012). *Mobile screens: The visual regime of navigation*. Amsterdam: Amsterdam University Press.



## **List of Figures**

<i>Figure 1.</i> Flash based player used to stream together still panoramic images .....	2
<i>Figure 2.</i> Revolutionary Actswas, July 2012 .....	6

**The Journal of Creative Technologies (JCT)**

**Issue 3: Interactivity, 2013**

ISSN: 2230-2115

Colab, Auckland University of Technology, New Zealand

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

*JCT* is a research communication platform published by Colab at the Faculty of Design and Creative Technologies, Auckland University of Technology.

<https://ctechjournal.aut.ac.nz>