PLAYGROUND GAMING WITH WII, KINECT AND PROCESSING

Grethe Mitchell & Andrew Clarke

This paper is a case study in the design and development of the “Game Catcher” - an interactive application produced as part of an AHRC “Beyond Text” Major Grant project that uses open source software (Processing, libfreenect) and hacked games hardware (Kinect and Wiimote) to create a low-cost motion tracking system that allows the recording, playback and analysis of children’s playground games in 3D.

Fig 1. Hand positions in a clapping game (© Grethe Mitchell & Andy Clarke)

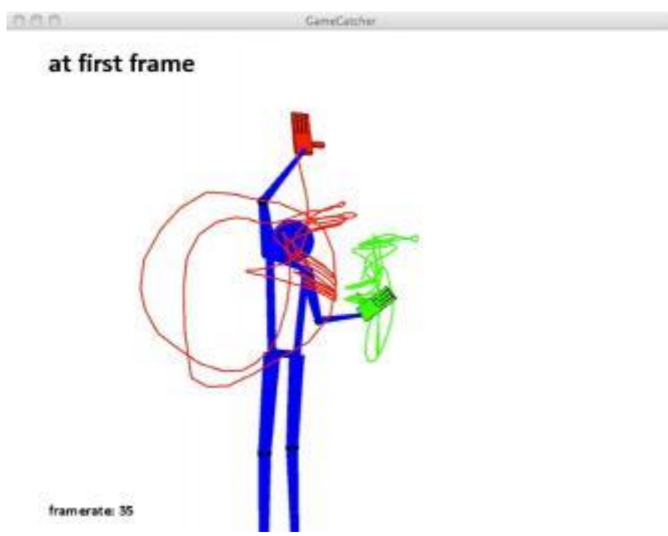

Fig 2. Visualisation of the path of the hands over time (© Grethe Mitchell & Andy Clarke)

INTRODUCTION

This paper is a case study in the development of the Game Catcher, a low-cost markerless motion tracking system produced as part of “Playground Games and Songs in the Age of New Media”, a two-year project funded by a large grant from the AHRC as part of the Beyond Text Programme.
The Game Catcher uses open source software (Processing, libfreenect, DarwiinRemote, OSCeleton, Simple-OpenNI) and hacked games hardware (both the Microsoft Kinect sensor and Nintendo Wiimote) to

create a low-cost markerless motion tracking system that allows the recording, playback and analysis of children's playground games and movement in 3D.

The following paper focuses on some of the technical aspects of the project, particularly the advantages of using videogame hardware and open source software to produce this application. Readers requiring more information on the archiving and visualisation aspects of the application are directed to our other papers [1] [2] and to the project report. [3]

The development of the Game Catcher was supervised by Grethe Mitchell of the University of East London (now at the University of Lincoln), and the final version of the application was developed by Andy Clarke.

PROJECT AIMS/CONTEXT

Recording movement for archiving and analysis is a common task in the Arts and Humanities. Approaches include written descriptions and drawing. Video recording technology is also commonly used, but although it is cheap and easy to use and therefore has advantages at the recording stage, it is more cumbersome when it comes to analysis. For instance, it may leave some details lost (though being out of shot, filmed from too far away, or obscured (either by the subject's other limbs or by other participants). Motion capture has many advantages over video, in that it allows the movement to be viewed from any angle or distance, not just the one that it was recorded from. There are, however, also negative aspects – such as price, availability, and the complexity of configuration and operation – which make them generally unsuitable for use in Humanities research projects.
We felt that a low-cost markerless motion tracking system would be able to combine the price/ease of use of video with the benefits of motion capture. Also, having researched and written about the appropriation of videogame hardware and software, [4] we felt that the new forms of motion sensitive videogame hardware (such as the Wiimote) could be adapted to form the basis for such a system, particularly as this approach would build upon the other advantages of videogame controllers such as an extremely favourable price to performance ratio, easy availability, robustness, familiarity with users, and so on.

The Game Catcher was developed as a fully functional proof of concept application, designed to address the preservation, visualisation and archiving of movement and to fill this price/performance gap between video and motion capture. In parallel, we also wanted to create a playable game which would allow a child to record their (movement-based) game and then play "against" this recording. Our aim here was that the computer game and the research tool could form a virtuous circle, with the playability acting as a reason for children to keep playing with the Game Catcher, thereby also providing more data for the researchers. The process of making the Game Catcher playable as a game also made us look at the issues involved in bringing "real world" playground activities into a computer environment, and made us think more consciously about what makes real life playground activities – and computer games – enjoyable.

TECHNOLOGY

To fulfil its stated aims of allowing the recording and playback of clapping games, the Game Catcher had to perform two different tasks: (a) track the position of the hands in 3D space; and (b) detect the orientation of the hands. From this information, the path of the hands could be determined (and reconstructed in "playback" mode) and claps could be identified – whether they were player to player, player to computer, or computer to computer claps. This all had to done with a sufficiently high frame rate and accuracy, and sufficiently low latency, for the data to be useful as research material and enjoyable during play. In addition, the Game Catcher had to be usable by researchers "in the field" – meaning that it had to be simple, robust, not require an excessive amount of technical knowledge to set up and use, and provide resilient tracking under a variety of lighting conditions and environments.

Processing was chosen as the programming language for the Game Catcher because of its ease of development, its focus on creative computing, and its extensive collection of pre-existing contributed libraries (and active development of new ones). If we had needed a C-based alternative, we would have probably used OpenFrameworks, for the same reasons.

During the course of developing the Game Catcher, we used a number of different solutions before adopting a "best of breeds" approach which used the Kinect sensor to track hand position and Wiimote controllers to track hand orientation. This proved ideal, exploiting the strengths of each system.
One of the issues that we came across (and perhaps failed to fully appreciate in drafting the proposal) was the difference between a controller like the Kinect that does motion tracking and one like the Wiimote which is merely motion sensitive. The Kinect detects hand movements by tracking changes in their position whereas the Wiimote only sensing changes in acceleration and therefore has no idea of where the hand actually is.
The Nintendo games are designed to minimise or conceal these limitations, but no such workaround was possible with the Game Catcher. As a result, it was necessary to adopt a hybrid approach which used the Wiimotes for the task that they were most suitable for (determining the hand orientation) and explored alternative solutions to the issue of tracking where the hands were in 3D space. This gave the Game Catcher a theoretical performance of around 3mm in the XY axis, 1cm in the Z axis and 1° in orientation. In practice, unavoidable system noise in the Kinect depth map reduces the XYZ accuracy slightly from these ideal figures, though it still remains well within acceptable levels. The accuracy with which depth is measured varies, and when an object is far from the Kinect, its movement is measured in larger steps than when it is close. The figure for orientation is likewise the raw measurement and in some hand orientations the Game Catcher makes the hand snap to 90° to avoid gimbal lock issues.

The authors have previous experience in video tracking, having used applications such as STEIM BigEye and Danny Rozin's TrackThemColors Director xtra within the context of dance and technology. As a result, they were aware of the shortcomings of video tracking – being too easily affected by outside conditions such as the brightness and colour temperature of ambient lighting or the clothing of participants.

As a result, we rapidly switched to an innovative approach which used Wiimotes not only to sense the orientation of the hands, but also to track their position. We attached an infra red LED to the Wiimotes in the player's hands and used a third Wiimote as a camera pointing at the player to track the position of these LEDs. The advantage of this approach is that it is very fast as the Wiimote has a dedicated chip optimised to do this image analysis in hardware. The tracking is also robust and resilient to lighting conditions, etc. as it is tracking infra red light, rather than visible spectrum. The LED point source does not allow tracking depth and we therefore planned to use triangulation to determine distance, an approach demonstrated by researchers at the University of Cambridge. [5]

Although the Kinect was hacked on its first day, it took a while before Processing libraries by Daniel Shiffman and Paul King became available. These only provided access to the raw depth map, not to the player skeleton, so bespoke hand tracking code was necessary. This approach provided an extremely high framerate – up to 100fps, even when tracking two hands – though self-occlusion was an issue, with the tracking becoming confused and tracking the wrong body part when, for example, a hand was temporarily hidden behind the other arm.

Because of this, we switched to OSCeleton as soon as it was available as this provided access to the player skeleton. OSCeleton ran as a separate command line program, communicating with Processing using the OSC protocol. This added additional complexity in setup/use and was felt, in the end, to be inappropriate for the Game Catcher if it was, as intended, to be useable by researchers in the field. For instance, when OSCeleton experienced a segmentation fault, this could leave both it and the Game Catcher frozen. We therefore switched to the Simple-OpenNI library.
As both OSCeleton and Simple-OpenNI can track multiple users, we developed and tested a multiplayer version of the Game Catcher, focused on other movement-based playground games such as skipping, hopscotch, dance routines etc.

ISSUES

There were a few issues with the Wiimote from a technical point of view. It does not, on its own, track rotation about the Y axis and suffers from a gimbal lock problem when pointed vertically upwards. This meant that we could not distinguish between two key hand positions: the palm out, fingers up (position 3 in large image) and the similar position with the palm facing sideways (position 5). In addition, the rotation could flip uncontrollably by 180° when the hand was in these positions.

Normally, the Wiimote relies on the Sensor Bar and/or the Wii Motion Plus to solve these problems, but neither was suitable for the Game Catcher. The first would require the user to keep their hands pointing at the bar, which clearly unsuitable for a clapping game, while the latter added too much additional bulk and weight.

An alternative approach was used. By paying attention to the previous position/orientations of the hands and the limits of human anatomy/physiology, we could add an additional level of interpretation to the raw data received from the Wiimotes and eliminate hand orientations which were either physically impossible or made no sense given the context of a clapping game. For instance, when the hands are vertical (fingers pointing up) in a clapping game, it is unlikely that the player's palms are facing their body as this is not a common move and even if they were in this "palms facing body" pose, they would have gone through some clearly identifiable intermediate positions to have arrive here.

This "filtering out" of non-used positions in the electronic environment is one of our "design solutions" of the development, based on the "perceived affordances" [6] of the system and the properties and affordances of the clapping movement ("real" and electronic), as well as on the affordances of the "real world" environment.

The other main technical issue that we faced was the relative weakness of Processing in rendering certain types of 3D graphics. In computer games such as the Wii Sports boxing game, the player's avatar is seen from behind in a ghostly form. This is an established convention which we wanted to use in the Game Catcher when the player is clapping with their recording, but we found Processing unable to render the semi-transparent 3D graphics needed for this effect. A workaround, using a wireframe avatar was felt to be sub-optimal.

The Wiimotes are convenient for holding, but are a little large to be comfortably mounted on gloves (so as to track clapping, for instance). Having done our prototyping using the Wiimotes, we are now looking at using the Seeeduino Film to engineer a smaller alternative.

VISUALISATION AND ANALYSIS

The raw movement data is saved in a plain text format and records the X/Y/Z/roll/pitch/yaw of every body part in every frame along with a timestamp (a sound file is also recorded of the clapping rhyme/song which plays back in sync with the movement). This data file is used in replaying the game, and is also designed to be easy to analyse. The Game Catcher provides an example of this analysis, having a built-in “visualisation” mode (see small image) which shows the path of the hands throughout the entire game. This image can be rotated and viewed from any angle and played, paused and rewound at will.

A natural extension of this would be to be able to view two recordings side by side for comparison or to superimpose the paths of one recording over another. Both of these would assist greatly in identifying similar clapping games or spotting when a variant of a game has emerged. Currently these are difficult, time-consuming, tasks relying on the manual comparison of video recordings, which may have been taken at different angles or be of clapping games performed at different tempos.

Computer analysis is also possible. Identifying clapping games from the analysis of video footage is complex, [7] but doing the same from motion data (such as that collected by the Game Catcher) is feasible as claps can be identified from changes in the velocity of the hands, changes in direction, or proximity to one another. These same principles and solutions can be applied to many different movement-based activities.

CONCLUSIONS

The Game Catcher is a fully functional prototype, working both as a computer game and as a research tool. It has proven the viability of a tracking solution “in-a-suitcase”, an exciting development as it means that ephemeral movement data can be captured, analysed and conserved at low cost, without extensive technical knowledge.

This has interesting outcomes for research in the Arts and Humanities. In terms of development, the Game Catcher allows for a low-cost visualisation mechanism allowing performers of all types (dance, music, theatre, circus etc) to see the shape and position of their activity in 3D space, thus opening up new avenues for artistic development and performance improvement.
In terms of research, the Game Catcher is particularly interesting to the arts and humanities disciplines, where research often takes place in-situ, rather than in a pre-determined, specially designed space as is the case for some medical or sports movement research, or in the entertainment sector (e.g. film and videogame production) where high-end (and expensive) motion capture studios are commonly used.

The above outcomes – and overall process of developing the Game Catcher – have led us to consider more deeply how movement is studied across disciplines, and whether these techniques are shared. From our research, it would seem that the techniques for recording and documenting movement are patchy, with bodies of knowledge siloed within certain fields and little known (or little used) outside of that particular field. For instance, formal movement notation systems such as Labanotation are used within dance, but not outside of it, even though it could have application elsewhere. Commercial motion capture systems are likewise used in the entertainment industry and in high-end medical or sports research and development, but are not generally used in the arts and humanities.

A follow-on project by one of the authors (Grethe Mitchell), also funded by the AHRC, is addressing this issue. It will organise a symposium and book allowing the exchange of best practices across all disciplines with regard to the recording, analysis and archiving of movement-based activities (in all their forms).

References and Notes:

1. *G. Mitchell and A. Clarke, "Playing, Recording and Archiving Playground Games," in Museums at Play: Games, Interaction and Learning, ed. Katy Beale, (London, Museums Etc., 2011)*
2. *G. Mitchell and A. Clarke, "Capturing and Visualising Playground Games and Performance: A Wii and Kinect Based Motion Capture System," Proceedings of Electronic Visualisation and the Arts, London, (July 2011).*
3. *A. Burn, J. Marsh, G. Mitchell, J. Robinson and R. Willett, "Children's Playground Games and Songs in the New Media Age (2009-2011 Project Report)," Beyond Text's Web Site, 2011, http://projects.beyondtext.ac.uk/ (accessed June 12, 2012).*
4. *A. Clarke and G. Mitchell, Videogames and Art (Bristol, Intellect: 2007).*
5. *S. Hay, J. Newman and R. Harle, "Optical Tracking Using Commodity Hardware" (paper presented at IEEE International Symposium on Mixed and Augmented Reality 2008, Cambridge, September 15-18, 2008).*
6. *G. Mitchell, "Porting Playground Games into a Computer Game Environment: Game Catcher Concepts, Aims and Issues" (paper presented at Children's Playground Games in the Age of New Media Interim Conference, London, February 25, 2010).*
7. *P. Wang, G. D. Abowd and J. M. Rehg, "Quasi-Periodic Event Analysis for Social Game Retrieval," Proceedings of IEEE International Conference on Computer Vision, Kyoto, (2009).*