A Comparison of Exergaming Interfaces for Use in Rehabilitation Programs and Research

Kazumoto Tanaka  
Kinki University  
kazumoto@hiro.kindai.ac.jp

Jim Parker  
University of Calgary  
jparker@ucalgary.ca

Graham Baradoy  
University of Calgary  
gbaradoy@ucalgary.ca

Dwayne Sheehan  
Mount Royal University  
dpsheehan@mtroyal.ca

John R. Holash  
University of Calgary  
rjholash@ucalgary.ca

Larry Katz  
University of Calgary  
katz@ucalgary.ca

Abstract
Exergames or active video games are video games with interfaces that require active involvement and the exertion of physical force by participants. These exergames are designed to track body motion and provide both fun and exercise for game players. Numerous video game console companies have designed exergaming interfaces that are becoming very popular. This paper examines the nature of the interfaces and explores the possibility of using these interfaces for rehabilitation programs and research. While many systems exist, this paper will focus on three major players: Sony PlayStation Move, Nintendo Wii, and Microsoft Xbox 360 Kinect. Comparisons include the technical specifications, the motion sensed by each interface, and the motion required in each therapeutic activity type. Discussion addresses the research implications of using these tools.

Author Keywords
Exergaming; game interface; rehabilitation; research

Introduction
Exergaming, which is a portmanteau word of "exercise" and "gaming", is the use of consumer video games (low cost video games) in an exercise activity (Sinclair, Hingston, & Masek, 2007). Exergaming platforms are designed to track body motion or body reactions and provide both fun and exercise for game players. The idea of exergaming is relatively old and can be found in late 1980s when Nintendo PowerPad was released. Konami's Dance Dance Revolution, released in 1998, can be said to be the first major successful application of the exergaming genre. However, the first exergaming resource to be successfully introduced into the home market was the Sony EyeToy (attachment to the playstation) released in 2004 (Larssen, Loke, Robertson, & Edwards, 2004; Wikipedia Contributors to Exergaming, 2011). The EyeToy detected a player’s kinematic information using the EyeToy camera to control the video game. Due to its two-dimensional image processing, the information that the game engine could get from the camera images was limited. In contrast, the Nintendo Wii released in 2006 utilized a motion sensitive controller that detected three-dimensional (3D) accelerations, and the 2010 version of the Wii could also detect
a player’s 3D hand posture using a three-axis gyro sensor. The robust detection of a player’s dynamic motion made the video game more exciting, and the Wii has enjoyed exceptional market success. Wii is currently leading other manufactures in sales (e.g., Sony PlayStation Move and the Microsoft Xbox 360 Kinect).

**Research in Rehabilitation with Consumer Exergaming Systems**

As consumer exergaming programs have evolved, numerous studies have researched the effectiveness of these types of games on rehabilitation or exercise. Most of the studies were conducted using the Wii or the Wii Fit (released in 2007), and utilized a new peripheral, the Wii Balance Board, for studies including: Brain function rehabilitation (Deutsch, Borbely, Filler, Huhn, & Guarrera-Bowlby, 2008; Hsu et al., 2010; Joo et al., 2010; Loureiro, Valentine, Lamperd, Collin, & Harwin, 2010; Saposnik et al., 2010); Isometric muscle strengthening (Sohnsmeyer, Gilbrich, & Weisser, 2010); Energy Expenditure (Hurkmans, Berg-Emons, & Stam, 2010); Exercising for elderly (Wollersheim et al., 2010); and Balance training (Brown, Sugarman & Burstin, 2009; Deutsch, Robbins, Morrison, & Bowlby, 2009; Kliem & Wiemeyer, 2010).

In contrast, there have been a number of studies identifying limited effects of exergaming. For example Hsu et al. (2010) looked at the potential of Wii bowling for rehabilitation in patients with upper extremity dysfunction but found that the only significant finding was a measure of enjoyment of activity when compared to a standard exercise group. Another study that applied the Wii games to exercise and the elderly showed no significant increase in physical activity due to exergaming (Wollersheim et al., 2010). One possible explanation for this outcome was that participants played the video game sitting down. Since the Wii system only detects movement through the controller, players might have adopted a motion strategy using the hand and not the body. Some of these strategies are not necessarily ideal adaptive approaches to rehabilitation (Saposnik et al., 2010). In other words, one can be successful in active game play with the Wii utilizing movement such as a wrist snap that does not engage the body. Such actions would not be successful in exergames for the Sony Move or the Microsoft Kinect, since they include cameras to evaluate movement. That notwithstanding, the motions and interactions required in commercially available exergames were not designed for rehabilitation outcomes (Anderson et al., 2010).

However, the potential to change repetitive training activities which are required in rehabilitation into more “entertaining” actions using exergaming is an attractive idea. Sweetser and Wyeth (2005) have adopted and extended Csikszentmihalyi’s (1990) conceptualization of the Flow model. In their Game Flow model of player enjoyment, they formulate a set of useful design criteria for achieving enjoyment in computer games. According to the model, players should feel a sense of control over their actions in the game in order to achieve immersion. In order to use exergames in a therapeutic environment, the game system should measure the conditions of the required exercise movement and utilize the movement information to directly control the game play.
This paper describes the sensing methods and the characteristics on leading consumer exergaming systems (Wii, Wii Fit, PlayStation Move and Xbox 360 Kinect) and considers the viability of applying the various interfaces as research and training tools to measure movements in therapeutic and rehabilitative situations.

**Exergame Interfaces**

Interfaces on most video game consoles allow games to be played from almost any position: sitting, standing, or even prone. That’s because they are hand operated devices based on buttons and thumb operated levers. A game designed to promote physical activity must possess technology that detects the general motion of a human body. In the past this has meant expensive motion sensors placed at multiple locations on a player (e.g., knees, hands, head, etc). Such sensors can be electrical, in which case they are fastened to the joints and can detect their rotation, or they can be optical, as when a recognizable marker is placed on the joint and a computer/camera combination is used to view and track the markers.

The development of good exergaming programs was dependent upon the invention of new, more inexpensive ways to track body pose and position. The PowerPad used a primitive tracking scheme, a mat that was placed on the floor and recorded footsteps. The Dance Dance Revolution pad, designed for entertainment, has a similar interface to the Powerpad but became very popular. It has the side benefit of providing the user with low to medium levels of aerobic activity. A more interactive exercise experience had to await the release of the Nintendo Wii. Nintendos’ success with the Wii has led other video game console developers to join the active gaming market.

In this section, sensing devices, sensing method, and the advantages and disadvantages for each consumer game interface are described and are also summarized in Table 1.

<table>
<thead>
<tr>
<th>Sensor devices</th>
<th>Wii Remote Plus &amp; Sensor Bar</th>
<th>Wii Balance Board</th>
<th>PlayStation Move Eye &amp; Motion Controller</th>
<th>Xbox 360 Kinect sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Remote Plus: IR camera (128*96 pixels), 3D gyro sensor, 3D acceleration sensor @ 100Hz. Sensor Bar: highlighting IR LEDs.</td>
<td>Multi pressure sensors @ 100Hz</td>
<td>PS Move Eye: 640<em>480 pixels @ 60fps, 320</em>240 pixels @ 120fps with directive microphone. Motion Controller: 3D gyro sensor, 3D acceleration sensor, geomagnetic sensor.</td>
<td>Depth camera: 640<em>480 pixels @ 30 Hz, distance range 0.7m - 6.0m. RGB camera: 640</em>480 pixels @ 30 Hz. Multi-array microphone.</td>
</tr>
<tr>
<td>Sensing capabilities</td>
<td>Advantages</td>
<td>Disadvantages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D information of acceleration and rotational angular of the controller. 3D position information of the controller (limited).</td>
<td>Hand motion detection with relatively high temporal resolution. *SDKs are opened to the public (e.g., Wiimote Lib).</td>
<td>Limited detection to hand motion. Difficulty in detection of 3D hand position</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D displacement information of gravity center. Load transition.</td>
<td>Relatively high temporal resolution. SDK is opened to the public (Wiimote Lib).</td>
<td>Exercise on limited area (the board area of 511 mm*316 mm).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D information of acceleration, position and rotational angular of the controller.</td>
<td>3D hand motion recognition with high temporal, high spatial resolution and robustness</td>
<td>Hand tracking only (upper body motion can be estimated by using Inverse Kinematics, but the estimation accuracy would be worse).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D information of position and orientation of objects in the camera view field.</td>
<td>No need for holding controllers. 3D gesture recognition. 3D scene recognition. SDKs are opened to the public (e.g., OpenNI).</td>
<td>Low temporal resolution. Difficulty in occluded motion recognition. Difficulty in recognition of motion that does not change depth information (e.g., arm axial rotation).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* *Software Development Kit*

**Table 1: Comparison of Game Interfaces**

*Wii*

The Wii is a small computer based on a 700 MHz Power PC processor with 64 megabytes of external main memory, 24 Mb of graphics memory, and non-removable 512MB flash memory for game data. However, the enabling technology resided in its novel controller (see Figure 1).
Figure 1. The Wii controller, inside-out. (Top) Broken open we see a (4-layer) circuit board and some standard switches. (Below) The famous accelerometer does not occupy a central location in the controller.

Looking like a television remote control, the Wii controller, Wii Remote or Wiimote, uses a wireless Bluetooth connection, contains a small speaker that can reproduce sounds from the main sound system, and has a small motor that is used to convey rumble impacts (Lee, 2008).

The most important new aspect to the controller is that of motion sensing. The motion of the controller is detected by an accelerometer; acceleration of the controller is sent to the Wii console and can be used in a dead-reckoning system to infer velocity and position. The controller uses an ADXL330 chip, which is really three accelerometers placed at right angles to each other so as to measure accelerations on 3D space, but the dead reckoning may less accurate because the accelerometers need to be calibrated frequently.

The Wii MotionPlus released in 2009 is an expansion device for the Wii Remote that allows it to more accurately capture complex motion by using gyro sensors (Vaughan-Nichols, 2009). The data from these sensors are used to recognize 3D rotational motion of the controller (i.e., player’s hand).

The Wii also utilizes an IR LED unit called Sensor Bar to estimate controller’s position. The unit has IR LED clusters in both sides respectively. The controller also has an IR camera that can capture the 2D position information of the two LED clusters. The distance between the Sensor Bar and the controller can be estimated using the distance between the LED clusters and LEDs image (see Figure 2), and thus 3D position information of the controller can be estimated. However, it is difficult to apply the 3D estimation method for the game control due to the limitation to the controller position and posture.
Figure 2: The Wii controller and the Sensor Bar. A depth estimation by comparing the distance between the two LED clusters with the distance between the cluster’s images is available when the IR camera faces the LEDs.

The most important capability of a game platform in support of exergames is the speed with which it can respond to commands and position changes. The term latency refers to the delay between when a change occurs and when the computer recognizes that change. Although Nintendo does not open the latency of the Wii system, several researchers have investigated the Wii performance. Regarding report frequency, the computer which the Wii controller is connected to via Bluetooth receives input reports 100 times per second (Lee, 2008). As to the latency on the Wii game system, the average of measured latencies is 143 milliseconds (Forums - Gaming Discussion - Motion Controller Lag Comparison).

Finally, several SDKs which provide an application programming interface (API) (e.g., WiimoteLib at CodePlex) are opened to the public for developing applications which utilizes the controller. In addition, the Wii can be programmed in Flash and the resulting game can be downloaded using the wireless network connection, meaning that it is relatively easy to make game for the Wii.

**Wii Fit**

The Balance Board is the interface of the Wii Fit and contains multiple pressure sensors in it (Clark et al., 2010). 2D information of player’s gravity center and load transition are calculated by using the sensor data. The data are transmitted at about 100 Hz (Tietäväinen, 2009), and it is relatively high temporal resolution. The SDK WiimoteLib can be used for building applications. The disadvantage is that the exercise is limited to the board area of 511 mm*316 mm.
PlayStation Move
The PlayStation Move has the Move Eye and the Motion Controller (or wand, see Figure 3) as its interface. The developer Sony Computer Entertainment announced that the Move Eye is an RGB camera (640*480 pixels @ 60fps / 320*240 pixels @ 120fps) with directive microphones.

Figure 3 : The Move wand, showing the illuminating sphere used to track it. These vary in color, and allow the wand to be used as a pointer.

The camera is utilized to detect an illuminating sphere attached to the wand for tracking the wand (Wikipedia Contributors to PlayStation Move, 2011). The game system knows the size of the sphere and thus can determine the wand's distance from the camera by comparing the size with the sphere's image size (see Figure 4), thus enabling the wand's position tracking in three dimensions. The 3D coordinates of the position can be calculated with special, high resolution based on sub-pixel image processing (Wikipedia Contributors to PlayStation Move, 2011). Furthermore, according to the report of the latency on the PlayStation Move system, the average of measured latencies is 115 milliseconds (Forums - Gaming Discussion - Motion Controller Lag Comparison). The Move clearly has the fastest response.

Figure 4: Depth estimation based on the sphere size. The sphere’s image size is calculated by using sub-pixel image processing so that the size can be estimated with high accuracy.
The 3D coordinates of the position can be calculated with high special resolution based on sub-pixel image processing (Wikipedia Contributors to PlayStation Move, 2011). Furthermore, according to the report of the latency on the PlayStation Move system, the average of measured latencies is 115 milliseconds (Forums - Gaming Discussion - Motion Controller Lag Comparison). The Move clearly has the fastest response.

The wand contains a three-axis accelerometer, a three-axis gyro sensor and a geomagnetic sensor. The accelerometer and the gyro sensor are used to track rotation in overall motion and can be used for dead reckoning in cases when the camera tracking is insufficient, such as when the wand is obscured behind the player's back. This deals effectively with what has been a typical problem of visual tracking, that of the occlusion of the target. To correct cumulative errors on these sensors, the geomagnetic sensor is used for calibrating the wand's orientation against the Earth's magnetic field (Wikipedia Contributors to PlayStation Move, 2011). Consequently, the sensor fusion method makes it possible to recognize the wand's position and orientation robustly and accurately.

The biggest difference between the PlayStation interface and the Wii interface is the PlayStation’s ability to detect 3D positioning. However, the interface extracts hand motion information similar to the Wii. Upper body motion can be estimated by using Inverse Kinematics, but the estimation accuracy is worse. Finally, the SDK is not opened to the public thus it is difficult to make original applications for rehabilitation or exercising.

**Microsoft Xbox 360 Kinect**

The Xbox 360 Kinect is the best example of a combination of modern hardware and software technologies. The Kinect utilizes the Kinect sensor device (see Figure 5) that comprises a depth camera (monochrome camera), an RGB camera, an infrared laser speckle pattern projector and multi-array microphone (Wikipedia Contributors to Kinect, 2011). The novel aspect of the Kinect is the depth detection. The device enables players to control the Xbox 360 games without the need to hold any controllers due to whole body motion recognition in three dimensions.

![Figure 5: The Kinect sensor and a projected speckle pattern. The speckle pattern is invisible due to a projection of infrared rays.](image-url)
The depth camera that has been developed by Israeli company PrimeSense, captures a depth map of the camera view field based on a speckle-based depth mapping method. The depth mapping method utilizes reference images of the speckle pattern that have been captured in advance on reference surfaces at a range of different distances from the projector. The image of the speckle pattern that is projected onto an object is compared with the different reference images in order to identify the reference pattern that correlates most strongly with the speckle pattern on the object, and thus to estimate the distance of the object from the projector (García, et al., 2008; United States Patent US7433024B2). The image processing is executed by a built-in processor, and then depth images come out. By using both the depth image and the RGB image, another image processing system in the game console can extract 3D information of position, displacement, velocity and orientation of objects.

The Kinect interface has large advantages over the other interfaces that utilize motion controllers. First, the interface enables the game system to recognize a player’s various motions. Furthermore, the interface senses not only players but also other objects in the field. That means that it provides enough data for making an Augmented Reality (AR) environment that allows players to interact with virtual objects in the real world. For example, when the coordinates of a floor in a field are recognized by a game system, a virtual soccer ball can be put on the floor in front of a player in the field. For building such application, there is an open source framework - the OpenNI (Open Natural Interaction) – which allows for communication with devices (e.g., capturing the depth data) and for realizing scene recognition (e.g., gesture recognition). The Kinect has a number of other interesting features such as audio input for speech recognition, but these do not bear on its ability to support exergames specifically.

A problem with the Kinect is the inherent limitation on image processing. That is the difficulty in occluded motion recognition. In addition, the camera’s temporal resolution of 30 Hz (Wikipedia Contributors to Kinect, 2011) is relatively low compared with the Wii and the PlayStation Move. According to the report of the latency on the Kinect game system, the average of measured latency is 218 milliseconds (Forums - Gaming Discussion - Motion Controller Lag Comparison). The average latency is nearly a quarter of a second, much too slow for games like hockey. Another disadvantage is the difficulty in the recognition of motion that does not change depth information (e.g., arm axial rotation).

**Discussion**

As can be seen, each of the exergaming systems described above has advantages and disadvantages for researchers interested in using exergaming technology for rehabilitation and for research. Existing games do not allow for easy access to user data for research or feedback for rehabilitation purposes. Therefore, it is necessary to customize applications. The need for controllers and interfaces that report position and motion should be clear. The nature of the sensors and software involved dictate how accurate the results are and how quickly they can be acquired, which would seem to be the key aspects of controllers for exergames.

All three consoles can sense position and motion of the controller or player. By using the controller as a marker the Playstation and the Wii simplify the software needed to collect the data
and can therefore, achieve higher speeds. The use of more complex vision-based software by the Kinect means that more of the acquisition is accomplished by the software, while relieving the players of the need to hold a specific device or mark themselves in any way. This is ultimately freeing, and is a natural interface for exercise play. Being connected to a machine limits motion. The software technology used in the Kinect has built-in limitations, but does have the potential to handle more people at the same time, and to increase the speed at which it operates. Technology of this type will define the preferred interfaces for future activity-based games and performance capture.

Speed and accuracy of these devices are sufficient in each case for creating games with exercise potential. Some sport-based games require speeds that are greater than what can be achieved by some of the platforms, but designing activities that provide an aerobic experience for a player is not really restricted by the devices themselves. The speed of response and the degree of positional accuracy of the controllers and corresponding consoles equals or better that of a human being in nearly every aspect. A game requires quick and simple access to the information being sent to it by the user interface. In a game that requires real-time response, this means that the game needs to respond as fast as a person, or within about a quarter of a second. All of the devices presented here can accomplish this in most cases. Software-based methods, such as the Kinect, can be more variable in this because the software can take longer in some specific complex circumstances. Nonetheless, the majority of activities do not push the limit of the device’s speed.

The issues surrounding the design of exercise activities and games involve the developer’s access to the devices. A significant problem for casual developers is access to the interfaces and the data they transmit. Game developers are licensed by the console makers to create games for their device, and the license is expensive. There is also the need to purchase a development system for the console, and these are not only expensive but require specialized expertise and programming ability. Sony, Nintendo, and Microsoft look carefully at the companies who apply to be developers to ensure that they meet a set of basic criteria, and not everyone who applies will be accepted. Some academic labs and small-scale developers cannot afford the license and development system, even if the console makers approved them. However, prices continue to decrease with the passage of time, and other options are appearing. A version of the Kinect developer’s kit is available for download by anyone who wants it (http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/download.aspx), and Microsoft has had a version of their Xbox development system publically available for some time; called XNA, it allows independent developers to create games, and even offers a web site and download facility where the games can be accessed by the public.

Sony has announced the Move.Me project (http://code.google.com/p/moveme), a public release of a software development kit (SDK) for the PS Move controllers. It requires programming skill in C or C#, and requires a PS3 and a PC computer. The goal is to allow academics and the public to experiment with the controller, but no facility is available to distribute games. It should be possible to build small numbers of installations using Move as the basis using this facility.

The price of a Nintendo Wii developer’s kit has dropped to $1700 for qualified groups. The Wii has been open to Flash development for some time, and that includes access to the accelerometer
and position data from the controllers of course. Since the Wii has a browser and a Flash implementation, such games can be downloaded and played with relatively little fanfare. Still, Flash is a very high-level tool and would be suitable for very high-speed games. The Wii controller is a Bluetooth device, and as such, it can be seen by any PC having Bluetooth capability (e.g., http://www.wiiprojects.org/bluetooth.html). This means that PC games can be created that make use of the Wiimote in the absence of a Wii.

All things considered, the Wiimote would appear to be the more flexible and useful device as a controller. Being accessible using Bluetooth means that raw accelerometer data can be accessed in real time by any computing device, therefore, the Wiimote could be used as the basis for novel exercise applications using PCs. The Kinect has advantages as a game device; for example, new games can be developed for the Xbox using an official Kinect SDK and distributed to Xbox users on the Internet. The Kinect also represents the future of game interfaces in that it uses vision rather than “instrumenting” the player. The Sony Move has some technical advantages as a controller but access to it is restricted by a lack of tools and distribution to the point where it is the least suitable of the three for exergame development work, except by licensed developers.

Other interfaces that could be used in exercise-based games have yet to be fully explored. The use of heart rate is being examined (Parker, Baradoy, & Katz, 2011) and galvanic skin response has been studied, as means to communicate emotional responses (biofeedback) to a computer. More complex biological measures, such as electrocardiogram signals or oxygen consumption devices, could certainly have applications here, but reliable devices are too expensive for home use.

References


