Author’s accepted manuscript

Motion Controllers, Sound, and Music in Video Games: State of the Art and Research Perspectives

Federico Visi and Frithjof Faasch

Abstract This chapter focuses on the use of motion-sensing controllers in video games, with particular focus on gameplay mechanics that link body movement, sound, an music. First, we present a review of the state of the art. This includes an overview of recent motion controllers describing the different motion-sensing technologies employed and comments on the implications such technologies have on game design and gameplay. This is followed by a set of examples of relationships between motion control, sound, music, and gameplay mechanics in recent video games. Secondly, we present a survey on recent research in expressive movement analysis and motion-based interaction, introducing the concepts of motion descriptor and parameter mapping. We then report on two serious games designed for children affected by autism using full-body motion capture and emotion recognition, and studies of player engagement in motion-controlled video games. In light of the interdisciplinary research presented, we finally discuss perspectives for motion-based game design.

1 Current state of the art

1.1 Motion-based video-gaming controllers

This section describes recent (currently available or in advanced development stage) motion-based controllers and their use in gaming. The devices are subdivided in two main categories: handheld controllers and camera-like controllers. The former includes devices that are held in the player’s hands, while the latter consists in devices

Federico Visi
Universität Hamburg e-mail: mail@federicovisi.com

Frithjof Faasch
Universität Hamburg e-mail: frithjof.faasch@web.de

that capture the player's movement without requiring direct manipulation. These two typologies differ in terms of motion-sensing technologies employed as well as in terms of interactions they can afford, which has implications for the design and implementations of motion-based gameplay mechanics. The section does not aim at providing an exhaustive list of all the available controllers. Rather, its purpose is to look at the motion-sensing technologies that are more frequently employed in popular gaming platforms and therefore characterise the design of the motion-based interactions.

1.1.1 Handheld motion controllers

Handheld motion-sensing game controllers have become more popular since Nintendo introduced the Wii Remote (Fig. 1a) in 2006. This kind of controllers often have built-in inertial sensors, such as accelerometers and gyroscopes. Accelerometers allow the tracking of translational acceleration of whatever they are attached to, and are often found in 3D configurations (i.e. able to sense acceleration along three orthogonal axes). Gyroscopes, on the other hand, allow to measure rotational velocity and are also frequently found in 3D configurations. These two sensors are used in tandem in inertial measurement units (IMUs), which are used extensively in aviation, robotics, Human-Computer Interaction (HCI), and for health applications such as gait analysis (Trojaniello et al, 2014; Visi et al, 2017c). In recent years, their increasing affordability and small size have made them a common feature of mobile and wearable devices and other consumer electronics. Due to drift error, inertial sensors do not currently allow accurate position tracking in real time. To overcome this limitation, some handheld motion controllers also include other motion-sensing technologies, such as optical or magnetic sensors. Principles for using accelerometers in game controllers, including operating mechanisms and specific parameters that directly affect gaming performance, are discussed by Champy (Champy, 2007).

When it was introduced in 2006, the only inertial sensor in the Wii Remote was a 3D accelerometer. In 2009, Nintendo released an expansion device for the controller named Wii MotionPlus (Fig. 1a), which included gyroscopes and could be attached to the original controller to allow more accurate motion tracking. The following year, Nintendo started producing a new version of the Wii Remote with built-in gyroscopes that provided the same motion-sensing capabilities of the Wii MotionPlus without the need of an expansion module. In addition to the inertial sensors, the Wii Remote has an optical sensor on the top end. This is used together with the 'Sensor Bar', a device with two groups of infrared LEDs to be placed above or below the screen. Since the distance between the two group of LEDs is known, it is possible to estimate the direction the Wii Remote is pointing to through triangulation when the Sensor Bar LED's are captured by the Wii Remote optical sensor. In addition to the Wii MotionPlus, Nintendo has introduced various other accessories and expansions for the Wii Remote. The Nunchuk connects to the Wii Remote through a cable and it is supposed to be held with the other hand by the player. It contains a 3D accelerometer similar to the one found in the Wii Remote, but no gyroscopes.

Fig. 1 Handheld motion controllers: (a) Wii Remote with Wii MotionPlus expansion; (b) PlayStation Move; (c) Joy-Con; (d) Razer Hydra; (e) Vive Controller; (f) Oculus Touch.

There are other accessories for the Wii Remote such as the Wii Wheel (a wheel with a slot for a Wii Remote), the Wii Zapper (a gun-shaped shell that can house a Wii Remote and a Nunchuk), and various others including a guitar-shaped device with a slot for a Wii Remote. These do not provide additional motion sensing capabilities, but are still of interest since they are designed to provide different affordances and ergonomics, thus suggesting other uses and movements to the player.

The PlayStation Move (Fig. 1b) released by Sony in 2010 is similar to the Wii Remote as it also includes accelerometers and gyroscopes for inertial motion sensing. However, optical sensing in the PlayStation Move works differently since there is no optical sensor on the controller itself. At the top of the controller there is instead an orb that emits light of different colours using RGB LEDs. A camera placed near the screen tracks the position of the orb, and (since the size of the orb is known) the system is capable of estimating the distance between the controller and the camera by tracking the size of the orb in the camera's image plane. Before releasing the PlayStation Move, Sony updated its DualShock gamepad by adding inertial sensors to it. This is a feature found also in other more recent gamepads such, as the Steam Controller developed by Valve, although these devices will not be covered in this section since they are primarily designed as gamepads and not as motion controllers.

After the commercial success of the Wii Remote, Nintendo included motion sensors in their following two consoles. Both the Nintendo 3DS handheld console and the gamepad of the WiiU home console include a 3-axis accelerometer and a 3-axis gyroscope. However, compared to their predecessor, the games produced for these consoles employed motion control only to a limited extent.

More recently, Nintendo has introduced the Joy-Con (Fig.1c), the main controller of the Nintendo Switch game console. Released in 2017, the Joy-Con improves and extends many of the concepts and technologies that characterised the Wii Remote and also allows for a more traditional gamepad-like configuration, achieved by attaching the controllers to the game console itself or to an accessory grip. Joy-Con come in pairs, with a controller designed to be used mainly with the right hand and a second one to for the left hand ('Joy-Con R' and 'Joy-Con L' respectively). This allows the Joy-Con to be used as a two-handed motion controller, similarly to a Wii Remote paired to a Nunchuk. More compact and light than a Wii Remote, each Joy-Con contains a 6-axis inertial measurement unit (IMU) comprising both accelerometers and gyroscopes (STMicroelectronics, 2017). The direction the Joy-Con is pointing to can be tracked without the need of an additional external device (as it was the case with the Wii Remote and the Sensor Bar). This is possible after a quick calibration procedure[1] requiring the Joy-Con to rest on a flat surface (Nintendo, 2017; Kuchera, 2017). The pointer can be re-centred at any time using a specific button, suggesting that it is not necessary to point the controller at the screen to make direction tracking work (Kuchera, 2017). Additionally and unique to the Joy-Con R, an infrared depth sensor placed on one end of the controller allows to detect the distance and shapes of nearby objects, and can also be used for hand gesture recognition purposes (Takahashi, 2017).

The Razer Hydra (Fig.1d) controller developed by Sixense Entertainment in partnership with Razer and released in 2011 presents some unique features in terms of motion tracking technology. With a motion sensitive unit for each hand equipped with buttons and direction sticks, the Razer Hydra is functionally similar to other handheld motion controllers. However, it differers from the Wii Remote and the PlayStation Move as it does not use optical means for position tracking. Instead, it employs low-power magnetic fields, allowing 6 degrees of freedom motion tracking (position and orientation) of both controllers without the need of calibration procedures. The base station that senses the movements of the controllers does not require unobstructed line-of-sight, as it would be the case with optical sensors. According to its developers (Sixense Entertainment, 2011), the technology allows low-latency tracking with precision of 1 mm and 1°. However, magnetic field tracking might suffer from interference from magnetic fields emitted by nearby devices. A downside of the Razer Hydra in comparison to other handheld motion controllers is the presence of wires. At the time of writing though, Sixense has made available to developers a wireless system named STEM. It employs magnetic field motion tracking similarly to the Razer Hydra and it is mainly targeted at virtual reality (VR) applications.

[1] This calibration procedure suggests that direction tracking is achieved by using drift-corrected orientation data obtained from the built-in IMU. A similar calibration procedure was previously used also in games employing the Wii MotionPlus expansion for the Wii Remote.

Fig. 2 External Camera-like Motion Controllers: (a) Kinect (version 1); (b) Kinect (version 2); (c) PlayStation Camera (2016 version); (d) A Leap Motion Controller used as a tabletop device with a laptop; (e) A Leap Motion Controller mounted on a head-mounted display; (f) Soli Development Kit.

Recent handheld controllers dedicated to virtual reality such as the Vive Controllers (Fig. 1e) and the Oculus Touch (Fig. 1f) are functionally similar to other motion controllers described here as they are also equipped with buttons and sensors. However, since virtual reality applications often require precise and consistent positional tracking, these controllers rely on multiple external infrared optical sensors, which are used also to track the position of VR head-mounted displays (HMDs).

A new device currently being developed by Valve might introduce some novelty in the panorama of handheld motion controllers. Known as the Knuckles, these devices add finger tracking capabilities to inertial and positional tracking common to other controllers. The Knuckles are strapped on the hands, allowing the user to move all the fingers without risking of dropping the devices (Yang, 2017b). Several capacitive sensor placed on the handgrip track the movement of individual fingers (Yang, 2017a), potentially allowing detailed smaller-scale interactions. At the time of writing, Valve is distributing the controllers to a limited number of selected developers.

1.1.2 External camera-like motion controllers

This section presents an overview of motion controllers that do not require the user to directly hold the device in order to work. These usually consist in devices that work similarly to a camera: they are placed nearby and pointed to the players in order to track their movements. However, these devices often employ other technologies in addition to optical sensing. The Kinect developed by Microsoft is probably one of the most popular camera-like motion controller for gaming applications. The first version introduced in 2010 (Fig. 2a) includes an RGB camera and an active depth sensor (thus resulting in what is also known as an RGB-D sensor). The latter works by projecting an invisible pattern through an infrared projector, which is then captured by a monochrome infrared sensor. Depth data is reconstructed by analysing how the projected pattern is reflected by surrounding objects and bodies. By combining the information obtained from its sensors, the Kinect is able to to track the position of specific points on the player's full body, called joints. The second iteration of the controller, Kinect 2.0 (Fig. 2b), was released in 2013 and brought many improvements over the first Kinect, particularly in terms of tracking accuracy, resolution, and field of view (Wasenmüller and Stricker, 2017).

After releasing the simpler PlayStation Eye (a device similar to a webcam with limited motion tracking capabilities based on computer vision and gesture recognition algorithms) in 2007, Sony released the PlayStation Camera, which – differently from the Kinect – uses stereoscopic vision to track depth. The first version released in 2013 was followed by an update in 2016 (Fig.2c), which included slight improvements (Halston, 2016).

Released to the public in 2013, the Leap Motion Controller (Fig.2d) employs technologies similar to the Kinect, this time for tracking hand movements instead of full bodies. Despite a smaller field of view and capture volume compared to the Kinect, the Leap Motion Controller's higher accuracy makes it a more suitable device for precise finger tracking (Weichert et al, 2013). Originally, the controller was designed to be placed horizontally in front of a screen, however other ways of using the device have been proposed by the research community (Brown et al, 2016) and the developers of the controller are offering development kits for virtual reality applications that use the Leap Motion Controller as a hand tracking device placed on the back of head-mounted displays (Fig.2e) (Holz, 2014).

Still under development at the time of writing, Google's SOLI (Fig.2f) (Google ATAP, 2017) employs an electromagnetic wave radar to capture very fine movements. The small form factor and high temporal resolution may lead to gaming applications, and researchers are currently exploring its potential for musical interactions (Bernardo et al, 2017).

There are some evident functional differences between camera-like and handheld motion controllers, the most evident of which is that the former does not require the player to manipulate a device in order work. However, differences go beyond that, as different systems represent movement through motion data in different ways. This has some crucial implications on how meaningful movement interactions are designed, as it will be described in section 2.1.

1.2 *Examples of relationships between motion control, sound, music, and gameplay mechanics*

Following the review of the hardware devices available, we present and discuss some gameplay mechanics involving motion, sound, and music, with particular attention to the representation of musical instruments in video games and the design strategies adopted for implementing musical interactions.

Rhythm games or rhythm-action games are 'video games in which the player must respond in some way to the rhythm or melody being presented, either through repeating the same melody or rhythm by pressing buttons (with hands or feet), or kinetically responding in some other way to the rhythm, often using specially designed controllers' (Collins, 2008, p. 74). The first examples of this game genre date back to the late 1970s (Collins, 2008). The category gained renewed popularity between the late 1990s and the 2000s, with hits such as Dance Dance Revolution and Guitar Hero. Both games involved the use of specially designed controllers: the former a dance pad with panels the player is supposed to step on following the dance patterns presented on the screen, and the latter a guitar-shaped controller (more about this type of controllers in section 1.2.1). More recently, music-themed rhythm games started taking advantage of motion controllers. The dance game Just Dance introduced by Ubisoft in 2009 eschewed the use of dedicated devices in favour of the Wii Remote. In the game, players follow the dance sequences performed on screen by an animated dancer while holding a Wii Remote. They receive points for how accurately they mimic the dance moves of the on-screen characters. The first episode received generally unfavourable reviews from the critics (CBS Interactive, 2017b), as its gameplay was considered excessively basic and motion detection was deemed imprecise. However, the game was a commercial success and users praised its simplicity and multiplayer mode that allowed four people to play at the same time (CBS Interactive, 2017b). The third instalment of the series released in 2011 introduced the support for Kinect and PlayStation Move, while later episodes gradually improved motion tracking and also included support for Nintendo Switch and smartphones equipped with motion sensors. Following the success of Just Dance, in 2010 Harmonix (the development company behind other well-known music-themed rhythm games such as Guitar Hero and Rock Band) released Dance Central. Available exclusively for Microsoft's Xbox 360, Just Dance takes advantage of the full-body motion capabilities of the Kinect. According to critics (CBS Interactive, 2017a), this resulted in a rewarding gaming experience, with more accurate motion tracking and variable difficulty, allowing to go beyond casual playing and increasing the longevity of the game. Compared to Just Dance, Dance Central offers the player more possibilities to improve their skills. Full-body motion tracking allows the game to show which dance moves can be improved in detail and a training mode allows to break dance sequences down to individual moves and practice at regular speed or in slow motion.

Close relationships between music, gameplay mechanics, and motion control are not found exclusively in rhythm games. For example, in some stages of Nintendo's

Super Mario Galaxy, the player controls Mario while the character is balancing on a sphere. By tilting the Wii Remote, the player makes Mario walk on the sphere. The direction and angle of the tilt correspond to the direction and speed of Mario's walk. The instability of this balancing act is underlined by the music, which changes dynamically according to the sphere movements. The speed of the melody follows the velocity of the sphere, and so does the timbre of the whole tune. This results in the music sounding as if it was being played back by a tape recorder at a speed that constantly changes with the dynamics of the movements of the character on screen. This contributes, albeit subtly, to the sense of unsteadiness stemming from Mario's movements, and helps the player to anticipate and follow the dynamics of the interaction.

Multimodal relationships and links between music and gameplay are more pronounced in Child of Eden, developed by Q Entertainment and published by Ubisoft in 2011. It is compatible with Kinect and PlayStation Move, but can also be played using traditional controllers. In the game, the player aims at various targets using one or more on-screen pointers (the 'Reticles'). When using motion controllers, the player controls the Reticles with the movement of their hands. Different reticles have different gameplay mechanics: the blue one has a lock-on function and the player can release a laser shot by quickly flicking their hands when locked on a target, while the other reticle is characterised by a rapid fire function that is more effective with certain targets. On-screen visuals are synchronised to the music and hitting a target often results in sounds and musical effects that match the key or tempo of the backing soundtrack. If using a Kinect, the player can swap the control of the two Reticles between left and right by clapping their hands. Since all action on screen is synchronised to a beat-heavy music, this and other gestures involved in controlling the Reticles to hit targets contribute to build up a very rhythmical, fast-paced gameplay tightly coupled with sounds, music, and visual effects.

Another example of gameplay mechanics with strong interplay between music and motion can be found in the game Johann Sebastian Joust[2], part of the game collection Sportsfriends released by the independent developer Die Gute Fabrik in 2014. Johann Sebastian Joust offers some rather distinctive gaming dynamics, which perhaps even stretch beyond the conventional definition of video game. It is a local (i.e. not networked) multiplayer game with no on-screen visuals specifically designed for handheld motion controllers. Its designers describe it as 'a duelling game, controlled by the tempo of the music'[3].

The main gameplay mechanics are effectively summarised by the developers on the game's web page:

> The goal is to be the last player remaining. When the music – J.S. Bachs's Brandenburg Concertos – plays in slow-motion, the controllers are very sensitive to movement. When the music speeds up, the threshold becomes less strict, giving the player a small window to dash at their opponents. If your controller is ever moved beyond the allowable threshold,

---

[2] `http://www.jsjoust.com`

[3] From the video trailer of the game (at 00:13) on Die Gute Fabrik Vimeo channel: `https://vimeo.com/31946199`

> you're out! Try to jostle your opponents' controllers why protecting your own. (Die Gute Fabrik, 2014)

Interestingly, the developers used, perhaps unwittingly, the term 'slow-motion' instead of slow tempo, possibly anticipating the effect that the music will have on the movements of the players. In the game, the music is front and centre, as it regulates the pace of each match by letting the players know which rules are currently in play, that is, how tolerant to sudden movements their controller is. As the tempo of the music increases, so does the tolerance of the motion controllers, which allows players to move more rapidly and possibly attempt to reach and set off other players' controllers. If the controller of one of the player is shaken beyond the acceleration threshold, he or she is eliminated from the game, living the remaining players 'jousting' for the final victory. Albeit very essential, the game provides an effective example of successful gameplay mechanics involving no screen, only music, sound, and motion controllers.

1.2.1 In-game musical instruments

An *in-game musical instrument* is a representation of a musical instrument *within* the game world that the character controlled by the player can interact with. This has conceptual and practical differences from music rhythm games involving instrument-like controllers such as Guitar Hero. Such games employ controllers shaped to resemble a musical instrument usually equipped with a set of buttons and other simple input devices placed in a way to induce some instrumental-like gestures. For example, the Guitar Hero controller is shaped like a smaller, stringless guitar with a set of buttons on the fretboard and a two-way directional pad placed in the middle of the guitar body, used to simulate strumming . The main goal of musical rhythm games is to follow visual cues presented on screen and push the corresponding button on the controller in time with the music in order score points. On the other hand, in-game musical instruments are objects represented in the game that the player can use and interact with to produce music with various degrees of freedom.

The action-adventure games of the 'Legend of Zelda' series often include some form of musical interaction. The sophistication of such interactions has varied through the years: from simply pressing a single button to play tones with a recorder in the first episode of the series released in 1986, to a set of five tones playable with the Nintendo 64 controller in 'Ocarina of Time' (1998), to a mini rhythm game to control the direction of the wind with a conductor's baton in 'The Wind Waker' (2003). The latter was further refined with motion controls in the remastered version released for the WiiU in 2013. Moreover, in the episode 'Spirit Tracks' published in 2009, the player can play a pan flute by touching the display of the Nintendo DS while simultaneously blowing in the microphone of the device. Finally, the harp in 'The Legend of Zelda: Skyward Sword' is an effective example of motion-controlled in-game musical instrument. In the game developed and published in 2011 by Nintendo for the Wii console, the player takes the role of the series protagonist, Link.

The game makes extensive use of the Wii MotionPlus expansion to the Wii Remote (see section 1.1.1) and offers more sophisticated motion-based gameplay mechanics compared to other Wii games available at the time of release. Similarly to other titles in the Zelda series, music contributes considerably to the gaming experience. Particularly, at some point in the game, Link receives a harp that can be used by the player to play simple melodies and strum chords. This is done by pressing a button on the Wii Remote while waving it in the air left and right. These movements correspond to harp strums, and their speed and timing affects the arpeggiated chords produced by the instrument. At some point in the game, a non-playing character (NPC) teaches the player how to use the harp and then challenges them to play at a specific tempo while he sings along. The phases of this interaction are depicted in figure 3. First, the player is instructed by the NPC on how to play the harp. After a brief tutorial, the NPC asks the player to strum the harp following the tempo of his swinging braid. If the player does not strum in time, the NPC will act as a "music teacher", telling the player whether they are strumming too fast or too slow. After this "introductory lesson" (top two screenshots in figure 3), the player has to strum the harp in time with a pulsating circle of light (middle screenshots in figure 3). If the player stays in time, the NPC will start singing along the harp chords. After a few chords at the right tempo, Link will learn the 'Ballad of the Goddess' (bottom screenshots in figure 3) and the player will be able to strum the song again later in the game. According to an interview with the game producer Eiji Aonuma (George, 2011), the design team deliberately chose to allow the player to strum the harp whenever they pleased, not only when playing the instrument is required to solve puzzles and advance in the story. At any point in the game, the player take the harp out and play a few chords in harmony with the background music.

Even though very effective and cleverly designed, the harp in 'The Legend of Zelda: Skyward Sword' gives just a glimpse of the potential and the complexity of musical interactions in video games. In-game musical instruments can be used to advance in the game as well as to play freely whenever the player feels like. Non-playing characters can serve as music teachers as well as performers to play with. Designing sophisticated and rewarding in-game musical interactions is a complex task that could be made easier by making use of recent research findings in the field of virtual reality. Serafin et al. (Serafin et al, 2016) discuss the development of musical instruments in virtual environments. They point out that 'virtual reality musical instruments' (VRMIs) differ from 'virtual musical instruments' (VMIs, i.e. software simulations of existing instruments, mostly focused on sonic emulation) as the former also include a 3-D simulated visual component delivered through virtual reality headsets or other immersive visualisation systems. In addition to presenting an overview of the current state of the art and describing case studies and evaluation methods, they propose various principles for the design of VRMIs. They recommend to take into considerations multiple sensory modalities, suggesting that while sound and visuals remain key elements 'touch and motion should also enter into the equation, as well as full body interaction' (Serafin et al, 2016, p. 26). This not only results in a more immersive experience, it also improves the playability of the instrument, allowing the development of more sophisticated musical skills.

Fig. 3 Phases of an in-game musical interaction with a non-playing character (NPC) using the harp in 'The Legend of Zelda: Skyward Sword.'

Referring to one of Cook's principles for designing computer music controllers stating that 'Copying an instrument is dumb, leveraging expert technique is smart' (Cook, 2001, p. 3), Serafin et al. suggest that VRMIs should make use of existing skills and adopting metaphors from existing real-world interactions. On the other hand, attempting to exactly replicate traditional instruments may not bring about interesting results. The limitations of virtual reality would likely make such interactions difficult and, at the same time, designs that try to faithfully replicate traditional instruments may not take advantage of the distinctive possibilities offered by the medium. In fact, the authors encourage designers 'consider the use of both natural and "magical" interactions and instruments' (Serafin et al, 2016, p. 28), that is, not necessarily constrained by real-world physics, technology, and human anatomy. Additional principles put forward by Serafin et al. focus on the representation and presence of the player's body and on the social dimension of music making, stressing the importance of shared experiences and the sense of presence and agency. Many of these guidelines are informed by the output of research communities active in the fields of computer music, human-computer interaction, and new instruments for musical expression (NIMEs, see (Jensenius and Lyons, 2017) for an overview).

Even though these recommendations are specific for the medium of virtual reality, many principles put forward by Serafin et al. can be applied to the design of in-game musical instruments also for non-VR video games, wether the player interacts with the game world in a first-person view or through an avatar. More examples of how interdisciplinary research outcomes can inform the design of musical interactions in video games is further discussed in section 2.

2 Future scenarios: perspectives for motion-based gaming from current interdisciplinary research

This section will give an overview of key topics in motion-based interaction research, such as the extraction of useful motion descriptors from motion data and parameter mapping for expressive interaction. We then describe recent research projects involving emotion recognition from full-body motion, and the analysis of body movements indicating different aspects of player engagement.

2.1 Recent research in expressive movement analysis: motion descriptors and mapping

Human motion research is an increasingly active field, contributing to many areas of study such as behavioural sciences, human-computer interaction, health and rehabilitation, systematic musicology , dance studies, ergonomics, user studies and others. Particularly, numerous interdisciplinary studies centred on embodiment and the experience of music as a multimodal medium have been carried out in the past two decades (Godøy and Leman, 2010; Wöllner, 2017; Visi, 2017). Body movement is also a recurring topic among researchers and designers working computer music interaction and new musical instruments (Jensenius, 2014). Analysis of expressive body movement in dance has also been the subject of various studies, some of which employed motion controller such as the Kinect (Camurri et al, 2004; Alaoui et al, 2012).

In these research fields, the way data obtained using motion-sensing technologies are processed to extract useful information is a key research topic. Various computable motion descriptors have been used for movement analysis and interactive applications. The Eyesweb platform (Camurri et al, 2007) contains a set of tools for motion analysis and is compatible with high-end motion capture system as well as low-cost motion controllers. Among motion descriptors, quantity of motion and contraction index are often used for analysing expressivity in applications involving full-body motion (Glowinski et al, 2011; Visi et al, 2014). The former is used to estimate the total amount of displacement over time, while the latter is an indication of the spatial extent of the body. Extracting meaningful features from motion data is a crucial step in the design of advanced interactions with music and sound. Larboulette et al. (Larboulette and Gibet, 2015) recently attempted a thorough review of computable descriptors of human motion, while Piana et al. (Piana et al, 2013) described a set of motion features for emotion recognition (see section 2.2).

The motion controllers described in section 1.1 employ a number of technologies that represent movement in different manners. While camera-like systems often describe body movement through positional data of various points of interests (e.g. the joints in Kinect skeletons) the data returned by the inertial sensors in handheld controllers consists in 3-dimensional vectors describing accelerations and rotational velocities, without a global coordinate system. This has several implications for the computation of motion descriptors useful for sound and music interaction. Visi et al. (Visi et al, 2017b) present motion descriptors dedicated to inertial sensors, and describe their use in music performance and composition.

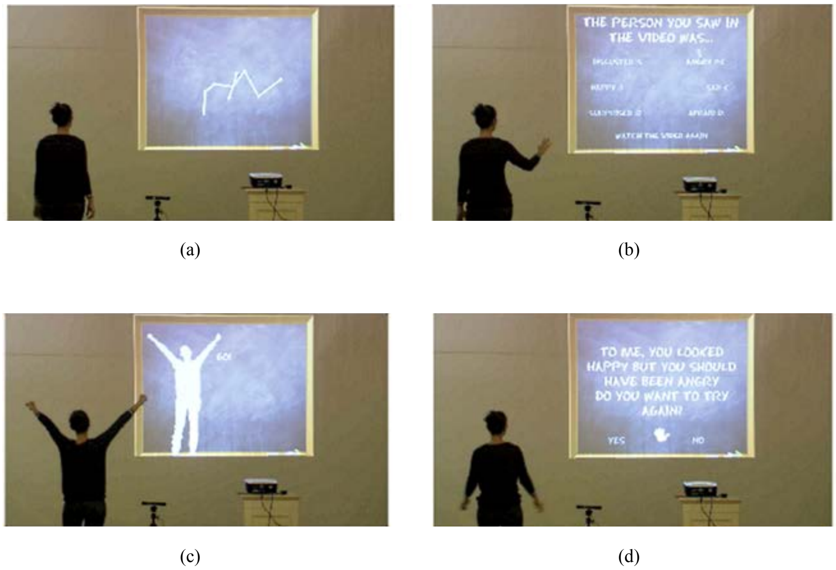
Extracting meaningful descriptors from motion data is not the only crucial step in designing successful motion interactions. Mapping them to parameters for controlling sound and/or visuals is another essential aspect of expressive interaction design. Different approaches to mapping have been the subject of numerous studies (Rovan et al, 1997; Van Nort et al, 2014; Hunt et al, 2002). Recently, machine learning has been increasingly employed for creating gesture-to-sound mappings (Caramiaux et al, 2014) and for creating complex mappings from databases of musical gestures (Visi et al, 2017a).

Research on motion descriptors and mapping strategies for musical interaction can inform the design of gestural interactions in gaming. As an example, the collection of game apps developed for the Mogees Play[4] sensor stemmed from the development of a device for musical applications, which itself employed various gesture recognition technologies resulting from academic research (Zamborlin et al, 2014).

2.2 Body motion, emotion, engagement

Motion-based gaming has received attention from researchers for different purposes. There are many instances of serious games for healthcare (Wattanasoontorn et al, 2013). Among these, Piana et al. (Piana et al, 2016) developed two games designed to help children affected by autism recognise and express emotions through full-body movement. In ‘Guess the Emotion’, the player is presented with a short video showing the silhouette or a stick figure of a person expressing a basic emotion (e.g. happiness or anger) through body movement (Fig. 4a). Then, the player has to guess which emotion was expressed by the silhouette by picking one from a list of possible answers (Fig. 4b). If the answer is correct, the player gains points and the game asks them to express the same emotion through body movement alone (Fig. 4c). If the emotion recognised by the system is correct, the player gains additional points, otherwise they will be asked if they want to try to perform that emotion again (Fig. 4d).

[4] `http://www.mogees.co.uk/play`



(a) (b) (c) (d)

Fig. 4 ‘Guess the Emotion’ game by Piana et al. (Piana et al, 2016).

To recognise the emotions expressed by the player, the system uses an RGB-D camera like the Kinect to extract a set of motion descriptors relevant for online emotion recognition. Classification is then performed by a classifier based on Support Vector Machines (SVM) (Burges, 1998).

A similar approach is adopted in the second game, ‘Emotional Charades’. The game involves two players with different roles: the ‘Actor’ chooses an emotion and expresses it by moving in front of the sensor. The second player (the ‘Observer’) will have to guess which emotion was chosen by the Actor by looking at the on-screen silhouette recorded by the sensor. The computer will try to classify the emotion as well, and the guesses of the Observer and the classifier will be shown on screen. The Actor then reveals what the correct answer is, and will gain more points if both the computer and the Observer have guessed right. Similarly, the Observer will gain points if they guessed the emotion expressed by the Actor right. Then the players switch roles.

Bianchi-Berthouze (Bianchi-Berthouze, 2013) studies the relationship between body movement and player engagement in motion-based video games. She argues that body movement has a strong influence on the sense of presence of the player and on the overall engagement experienced while playing. By examining the movement patterns players adopt when playing games using both desktop and motion controllers, she proposes a taxonomy of body movements that will be the basis of a hypothetical model of the relationship between player body movement, controller, and player engagement. This taxonomy includes five classes: ‘Task-Control Body Movements’ (defined by the controller and necessary to control the game),

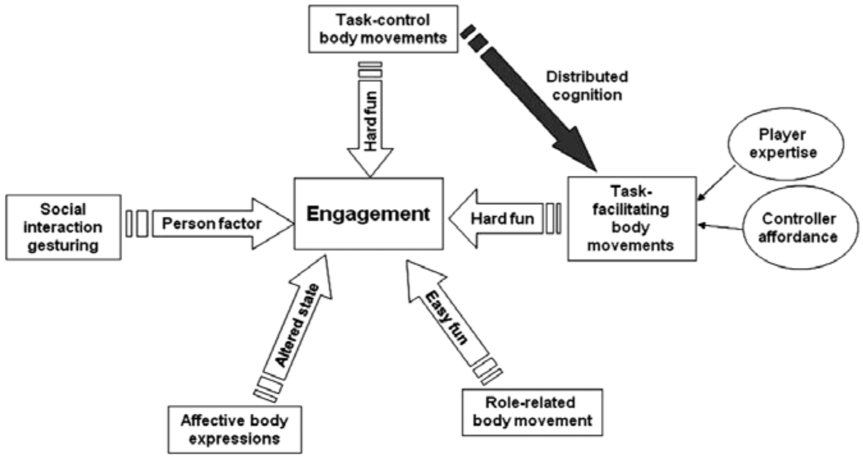


Fig. 5 Bianchi-Berthouze's movement-based engagement model for the design of full-body interactions in video games (Bianchi-Berthouze, 2013).

'Task-Facilitating Body Movements' (performed to facilitate game control but not recognised by the controller), 'Role-Related Body Movements' (typical of the role adopted by the player in the game scenario but not recognised by the controller, e.g. head-banging while playing Guitar Hero), 'Affective Expressions' (gesturing expressing the affective state the player during game play, generally not recognised by the controller), and 'Expressions of Social Behaviour' (movements that facilitate and support interaction between players, not currently recognised by the controller) (Bianchi-Berthouze, 2013, pp. 49-51). This movement classes are then related to the four engagement factors previously identified by Lazzaro (Lazzaro, 2004), resulting in a model useful for a more systematic approach to the design of full-body interactions in video games (Fig. 5). To show how motion control can facilitate engagement and result in a broader set of emotions in the player, Bianchi-Berthouze (Bianchi-Berthouze, 2013) reports a series of empirical studies focused on a group of participants playing the music-themed rhythm game Guitar Hero with different controllers. Overall, the experimental results showed that, when using a motion-enabled controller, players tend to shift between the engagement types related to the different movement classes described above, showing movements that relate to role-play and enjoyment. On the other hand, when using a controller that 'do not require and do not afford natural body movements' (Bianchi-Berthouze, 2013, p. 60), she observed a general lack of overt movements beyond those necessary to control the game. She then concludes that a more complete and compelling gaming experience involves body movements from all the five classes of the taxonomy, which can also be used as mood and attitude induction mechanisms.

Bianchi-Berthouze's findings (Bianchi-Berthouze, 2013) as well as the emotion recognition techniques employed by Piana et al. in their serious games (Piana et al, 2016) constitute useful resources for game designers. Their work on affective states and emotional engagement shows that there is still considerable potential in motion control that has not been fully explored in mainstream video game design.

3 Discussion and Conclusion

By looking at recent instances of motion control in commercial video games and at the findings of interdisciplinary research involving video games and body movement, one can clearly see that the field is steadily evolving.

Albeit showing increasing sophistication, motion-based interactions in commercial games is still somewhat limited, especially when involving music and sound. Research in neighbouring fields such as computer music and human-computer interaction can lead to useful insight for making musical interaction in games more profound an effective. As seen in the dance games mentioned in section 1.2, recent motion-based games are starting to include features and play modes that allow for skill-learning and training. This is a departure from more casual gaming modalities that characterised earlier motion-based games, which may allow for more sophisticated gaming experiences. Research in the field of new musical instruments design can provide useful guidelines for obtaining more profound musical interactions in video games. One of the most cited tenets of digital musical instrument design is ‘low entry fee with no ceiling to virtuosity’ (Wessel and Wright, 2001), meaning that getting started with a digital musical instrument should be relatively easy but this should not hinder the development of higher degrees of expressivity. The same principle can be applied to musical interactions in video games, wether these are based on dance or on in-game musical instruments. The learning curve should allow for casual gaming and, at the same time, for the development of virtuosity. This way, progressive skill-learning is possible, increasing the longevity of the game and the depth the musical interactions it affords. Moreover, appropriation and the emergence of personal styles contribute to define the identity of a musical instrument and characterise its use (Zappi and McPherson, 2014). Similarly, musical interactions in video games would benefit from the emergence of personal styles, which make the experience more intimate and distinctive for each player.

With the increasing development and availability of virtual reality and augmented reality HMDs, video games are likely to progressively untether from fixed screens. This process is making room for more sophisticated motion-based interactions that also take greater consideration of body movements related to affective states, role-playing, and social interactions with other players or non-playing characters. If the players’ gaze will be eventually freed from having to stare at a fixed screen, the perception of one’s own presence in the game will increasingly shift towards other sensory modalities, enabling players to perceive and interact with each other in a more natural and less mediated manner, thus allowing more intimate musical interactions. This would contribute to making video games an even more immersive and emotionally compelling experience: a medium increasingly useful for healthcare, education, and other fields beyond entertainment.

References

Alaoui SF, Caramiaux B, Serrano M, Bevilacqua F (2012) Movement qualities as interaction modality. In: Proceedings of the Designing Interactive Systems Conference, ACM, pp 761–769

Bernardo F, Arner N, Batchelor P (2017) O Soli Mio : Exploring Millimeter Wave Radar for Musical Interaction. In: Erkut C (ed) NIME'17 International Conference on New Interfaces for Musical Expression, Copenhagen, Denmark, May, pp 283–286, URL http://homes.create.aau.dk/dano/nime17/papers/0054/paper0054.pdf

Bianchi-Berthouze N (2013) Understanding the Role of Body Movement in Player Engagement. HumanComputer Interaction 28(1):40–75, DOI 10.1080/07370024.2012.688468, URL http://www.tandfonline.com/doi/abs/10.1080/07370024.2012.688468

Brown D, Renney N, Stark A, Nash C, Mitchell T (2016) Leimu: Gloveless Music Interaction Using a Wrist Mounted Leap Motion. Proceedings of the International Conference on New Interfaces for Musical Expression 16:300–304, URL http://www.nime.org/proceedings/2016/nime2016{_}paper0059.pdf

Burges CJC (1998) A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2:1–43

Camurri A, Mazzarino B, Ricchetti M, Timmers R, Volpe G (2004) Multimodal Analysis of Expressive Gesture in Music and Dance Performances. In: Camurri A, Volpe G (eds) Gesture-based Communication in Human-Computer Interaction, 2915, Springer, Berlin, pp 20–39, DOI 10.1007/978-3-540-24598-8_3, URL http://link.springer.com/10.1007/978-3-540-24598-8{_}3

Camurri A, Coletta P, Demurtas M, Peri M, Ricci A, Sagoleo R, Simonetti M, Varni G, Volpe G (2007) A Platform for Real-Time Multimodal Processing. In: Proceedings International Conference Sound and Music Computing, July, pp 11–13, URL http://www.smc-conference.org/smc07/SMC07Proceedings/SMC07Paper59.pdf

Caramiaux B, Françoise J, Schnell N, Bevilacqua F (2014) Mapping Through Listening. Computer Music Journal 38(3):34–48, DOI 10.1162/COMJ_a_00255, URL http://www.mitpressjournals.org/doi/abs/10.1162/COMJ{_}a{_}00255

CBS Interactive (2017a) Dance Central for Xbox 360 Reviews - Metacritic. URL http://www.metacritic.com/game/xbox-360/dance-central

CBS Interactive (2017b) Just Dance for Wii Reviews - Metacritic. URL http://www.metacritic.com/game/wii/just-dance

Champy A (2007) Elements of Motion: 3D Sensors in Intuitive Game Design. Analog Dialogue 41(2):11–14, URL http://www.analog.com/en/analog-dialogue/articles/3d-sensors-in-intuitive-game-design.html

Collins K (2008) Game sound : an introduction to the history, theory, and practice of video game music and sound design. MIT Press, URL https://mitpress.mit.edu/books/game-sound

Cook P (2001) Principles for designing computer music controllers. Proceedings of the International Conference on New Interfaces for Musical Expression 2001 (NIME 2001) pp 3–6, URL http://www.nime.org/proceedings/2001/nime2001{_}003.pdfhttp://dl.acm.org/citation.cfm?id=1085152.1085154

Die Gute Fabrik (2014) Presskit — Johann Sebastian Joust. URL http://www.jsjoust.com/presskit/

George R (2011) Behind the Scenes of Zelda: Skyward Sword. URL http://uk.ign.com/articles/2011/11/11/behind-the-scenes-of-zelda-skyward-sword

Glowinski D, Dael N, Camurri A, Volpe G, Mortillaro M, Scherer K (2011) Toward a Minimal Representation of Affective Gestures. IEEE Transactions on Affective Computing 2(2):106–118, DOI 10.1109/T-AFFC.2011.7, URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5740837http://ieeexplore.ieee.org/document/5740837/

Godøy RI, Leman M (eds) (2010) Musical Gestures: Sound, Movement, and Meaning. Routledge

Google ATAP (2017) Project Soli. URL https://atap.google.com/soli/

Halston T (2016) New PS4 camera vs old PS4 camera comparison - what's the difference? URL https://www.psu.com/news/31081/New-PS4-camera-vs-old-PS4-camera-comparison---whats-the-difference

Holz D (2014) Leap Motion Sets a Course for VR. URL http://blog.leapmotion.com/leap-motion-sets-a-course-for-vr/

Hunt AD, Wanderley MM, Paradis M (2002) The importance of Parameter Mapping in Electronic Instrument Design. In: Casey C, Schneider K, Hammond E (eds) Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, pp 88–93, URL http://www.nime.org/proceedings/2002/nime2002{_}088.pdf

Jensenius AR (2014) To gesture or Not? An Analysis of Terminology in NIME Proceedings 2001-2013. In: Proceedings of the International Conference on New Interfaces for Musical Expression, pp 217–220, URL http://www.nime.org/proceedings/2014/nime2014{_}351.pdf

Jensenius AR, Lyons MJ (2017) A NIME Reader, Current Research in Systematic Musicology, vol 3. Springer International Publishing, Cham, DOI 10.1007/978-3-319-47214-0, URL http://link.springer.com/10.1007/978-3-319-47214-0

Kuchera B (2017) Nintendo Switch's World of Goo shows off system's Wii-style pointer controls. URL https://www.polygon.com/2017/3/16/14947750/nintendo-switch-world-of-goo-controls-wii

Larboulette C, Gibet S (2015) A review of computable expressive descriptors of human motion. In: Proceedings of the 2nd International Workshop on Movement and Computing - MOCO '15, ACM Press, New York, New York, USA, pp 21–28, DOI 10.1145/2790994.2790998, URL http://dl.acm.org/citation.cfm?id=2790994.2790998http://dl.acm.org/citation.cfm?doid=2790994.2790998

Lazzaro N (2004) Why We Play Games: Four Keys to More Emotion Without Story. Tech. rep., XEODesign Inc., Oakland, CA

Nintendo (2017) How to Calibrate the Controllers. URL http://en-americas-support.nintendo.com/app/answers/detail/a{_}id/22340/{~}/how-to-calibrate-the-controllers

Piana S, Staglianò A, Camurri A, Odone F (2013) A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. In: IDGEI International Workshop, Crete, Greece

Piana S, Staglianò A, Odone F, Camurri A (2016) Adaptive Body Gesture Representation for Automatic Emotion Recognition. ACM Transactions on Interactive Intelligent Systems 6(1):1–31, DOI 10.1145/2818740, URL http://dl.acm.org/citation.cfm?id=2896319.2818740

Rovan JB, Wanderley MM, Dubnov S, Depalle P (1997) Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Kansei-The Technology of Emotion Workshop Proceedings of the AIMI International Workshop pp 68–73

Serafin S, Erkut C, Kojs J, Nilsson NC, Nordahl R (2016) Virtual Reality Musical Instruments: State of the Art, Design Principles, and Future Directions. Computer Music Journal 40(3):22–40, DOI 10.1162/COMJ_a_00372, URL http://www.mitpressjournals.org/doi/10.1162/COMJ{_}a{_}00372

Sixense Entertainment (2011) Razer Hydra. URL http://sixense.com/razerhydra-3

STMicroelectronics (2017) Semiconductor Solutions from STMicroelectronics Selected by Nintendo for Nintendo Switch. URL http://www.st.com/content/st{_}com/en/about/media-center/press-item.html/t3934.html

Takahashi D (2017) Nintendo Switch has high-tech Joy-Con controllers with motion detection camera. URL https://venturebeat.com/2017/01/12/nintendo-switch-has-high-tech-joy-con-controllers-with-motion-detection/

Trojaniello D, Cereatti A, Pelosin E, Avanzino L, Mirelman A, Hausdorff JM, Della Croce U (2014) Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. Journal of NeuroEngineering and Rehabilitation 11(1):152, DOI 10.1186/1743-0003-11-152, URL http://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-11-152http://www.jneuroengrehab.com/content/11/1/152

Van Nort D, Wanderley MM, Depalle P (2014) Mapping Control Structures for Sound Synthesis: Functional and Topological Perspectives. Computer Music Journal 38(3):6–22, DOI 10.1162/COMJ_a_00253, URL http://www.mitpressjournals.org/doi/abs/10.1162/COMJ{_}a{_}00253

Visi F (2017) Methods and Technologies for the Analysis and Interactive Use of Body Movements in Instrumental Music Performance (March), URL https://pearl.plymouth.ac.uk/handle/10026.1/8805?show=full

Visi F, Schramm R, Miranda E (2014) Gesture in performance with traditional musical instruments and electronics. In: Proceedings of the 2014 International Workshop on Movement and Computing - MOCO '14, ACM Press, New York, NY, USA, pp 100–105, DOI 10.1145/2617995.2618013, URL http://dl.acm.org/citation.cfm?id=2618013http://dl.acm.org/citation.cfm?doid=2617995.2618013

Visi F, Caramiaux B, Mcloughlin M, Miranda E (2017a) A Knowledge-based, Data-driven Method for Action-sound Mapping. In: NIME'17 International Conference on New Interfaces for Musical Expression

Visi F, Coorevits E, Schramm R, Miranda ER (2017b) Musical Instruments, Body Movement, Space, and Motion Data: Music as an Emergent Multimodal Choreography. Human Technology 13(1):58–81, DOI 10.17011/ht/urn.201705272518, URL http://humantechnology.jyu.fi/archive/vol-13/issue-1/musical-instruments-body-movement-space-and-motion-data/@@display-file/fullPaper/Visi{_}Coorevits{_}Schramm{_}Reck-Miranda.pdf

Visi F, Georgiou T, Holland S, Pinzone O, Donaldson G, Tetley J (2017c) Assessing the Accuracy of an Algorithm for the Estimation of Spatial Gait Parameters Using Inertial Measurement Units: Application to Healthy Subject and Hemiparetic Stroke Survivor. 4th International Conference on Movement Computing

Wasenmüller O, Stricker D (2017) Comparison of kinect v1 and v2 depth images in terms of accuracy and precision. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer, Cham, vol 10117 LNCS, pp 34–45, DOI 10.1007/978-3-319-54427-4_3, URL http://link.springer.com/10.1007/978-3-319-54427-4{_}3

Wattanasoontorn V, Boada I, García R, Sbert M (2013) Serious games for health. Entertainment Computing 4(4):231–247, DOI 10.1016/j.entcom.2013.09.002, URL http://linkinghub.elsevier.com/retrieve/pii/S1875952113000153

Weichert F, Bachmann D, Rudak B, Fisseler D (2013) Analysis of the Accuracy and Robustness of the Leap Motion Controller. Sensors 13(5):6380–6393, DOI 10.3390/s130506380, URL http://www.ncbi.nlm.nih.gov/pubmed/23673678http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3690061http://www.mdpi.com/1424-8220/13/5/6380/

Wessel D, Wright M (2001) Problems and Prospects for Intimate Musical Control of Computers. In: Poupyrev I, Lyons MJ, Fels SS, Blaine T (eds) Proceedings of the 2001 International Conference on New Interfaces for Musical Expression

(NIME-01), Seattle, WA, United States, pp 11–14, URL http://www.nime.org/proceedings/2001/nime2001{_}011.pdf

Wöllner C (2017) Body, Sound and Space in Music and Beyond: Multimodal Explorations. Routledge, DOI 10.4324/9781315569628

Yang L (2017a) Knuckles Cap Sense Overview. URL http://steamcommunity.com/sharedfiles/filedetails/?id=943495896

Yang L (2017b) Knuckles Quick Start. URL http://steamcommunity.com/sharedfiles/filedetails/?id=943406651

Zamborlin B, Bevilacqua F, Gillies M, D'inverno M (2014) Fluid gesture interaction design. ACM Transactions on Interactive Intelligent Systems 3(4):1–30, DOI 10.1145/2543921, URL http://dl.acm.org/citation.cfm?id=2543921http://research.gold.ac.uk/9619/1/TiiS2013.pdfhttp://dl.acm.org/citation.cfm?doid=2567808.2543921

Zappi V, McPherson A (2014) Dimensionality and Appropriation in Digital Musical Instrument Design. In: Caramiaux B, Tahiroglu K, Fiebrink R, Tanaka A (eds) Proceedings of the International Conference on New Interfaces for Musical Expression, Goldsmiths, University of London, London, United Kingdom, pp 455–460, URL http://www.nime.org/proceedings/2014/nime2014{_}409.pdf