

A Visual-Inertial Hybrid Controller Approach to Improving Immersion in 3D Video Games

Alexander Wong

Department of Systems Design Engineering
University of Waterloo
a28wong@engmail.uwaterloo.ca

Abstract

Advances in various areas such as graphics, sound, physics and artificial intelligence have improved the level of player immersion into the gaming environment significantly over the years. However, current game controller systems do not fully facilitate natural bodily motions in an accurate and responsive manner, which may affect player immersion within a gaming environment. This paper presents a novel visual-inertial approach that can potentially improve immersion in 3D video games.

The proposed game controller system utilizes visual sensors (i.e. cameras) and inertial sensors (i.e. accelerometers and gyro-sensors) in a synergistic fashion to provide better 3D spatial positioning information than either of the individual technologies can provide. As a result, the proposed game controller system facilitates highly accurate, responsive, and natural control over 3D environments. This makes it well suited for potentially improving player immersion in future 3D video games. Furthermore, several applications of the proposed game controller system are presented to illustrate its potential for improving immersion in 3D video games.

Introduction

One of the biggest goals a game developer strives for when developing a 3D video game is to create a truly immersive experience for a player. Immersion can be defined as a state of mind where you become unaware of your physical environment and, according to Allen Varney, experience “intense focus, loss of self, distorted sense of time, and effortless action” (Varney, 2006). Tremendous strides have been made by both software developers and hardware vendors to bring video game graphics to a level that approaches photo-realism. This allows for the creation of lush and highly believable worlds that help draw the player into the game experience. Similarly, improvements in artificial intelligence, physics, and sound have contributed greatly to creating a truly involving environment. However, until recently, few advancements have been made to one of the most important components in creating a truly immersive experience: the game controller system.

Alongside visual and audio outputs, the game controller system is one of the key components that connect the player to the game environment. The game controller system allows the player to interact with the environment portrayed by the game. However, even today, the

most believable and photo-realistic 3D game environments are controlled using conventional, general-purpose controller devices such as game controllers and mouse/keyboard combinations. Examples of some current general-purpose controller systems used for 3D video games are shown in *Figure 1*.



Figure 1: Examples of current general-purpose controller systems used for 3D video games.
Top-left: Razer *DeathAdder* Gaming Mouse; Top-right: Razer *Tarantula* Gaming Keyboard;
Bottom: *Xbox 360* wireless controller

While general-purpose game controllers have evolved over time to handle increasingly complex 3D game environments, the fundamental design of such game controllers has changed little over the years. One of the problems with these devices with respect to player immersion is the fact that they provide a very unnatural interface between the player and the game environment, which is not very intuitive in the context of game environment. There is no natural correspondence between the actions of the player and what happens in the game environment. For example, popular fighting games such as the *Virtua Fighter* and *Tekken* series provide photo-realistic 3D characters fighting in photo-realistic game environments. However, the natural punching, kicking, and blocking motions performed by the characters are mapped to non-natural button presses on a game controller, as opposed to natural player motions. This effectively breaks the illusion of hand-to-hand combat and draws the player away from a truly immersive experience within the video game. Therefore, game controller systems that allow for natural and intuitive player motions that correspond to the game environment may improve the *flow* (Csikszentmihalyi, 1990) of the gaming experience, where the player is fully immersed and focused on the gaming environment at hand. This is directly related to *tactical immersion* (Adams, 2003), where the player is completely engaged in a physical and immediate way. This type of immersion will be the focus of this paper.

The main contribution of this paper is the introduction of a visual-inertial controller approach that may improve player immersion in 3D video games. The proposed game controller system utilizes low cost visual and inertial sensors to provide highly accurate, responsive, and intuitive control over 3D environments using natural bodily motions. The proposed approach may help in improving player immersion in future 3D video games. First, a discussion of alternative game controller systems available in the market and currently under research is presented to illustrate their relative strengths and weaknesses in creating an immersive experience for the player within the gaming environment. The proposed visual-inertial game controller system is then described and presented in detail. Furthermore, a discussion on the prototype implementation of the game controller system is presented. A number of different applications for the proposed game controller system are then discussed to illustrate how the system may be used to enhance the level of player immersion within game environments. Finally, conclusions are drawn and future work is presented.

Previous Alternative Controller Systems

Prior to describing the proposed visual-inertial game controller system, it is important to discuss alternative game controller systems that have been introduced over the years. A review of key alternative controller systems helps provide a better sense of both the strengths and weaknesses of such systems in immersing the player into the game environment.

Light-Based Controller Systems

One of the earliest alternative controller systems used for gaming purposes are light-based controller systems. Such systems consist of a light emitter component and a light sensor component. In early light-based controller systems, the visual display acts as the light emitter component while photodiodes are placed in the control device to act as the light sensor component. The most recognizable examples of these early light-based controller systems are the 'light guns' used in early arcade machines and video game consoles such as the *Nintendo Entertainment System*. One of the biggest problems with such devices is that they rely on the characteristics of Cathode Ray Tube (CRT) displays to function properly, making them relatively ineffective for displays based on plasma and LCD technology. More recent light-based controller systems are based on infrared emitters and sensors. Typically, a series of infrared light emitters is placed near the visual display and infrared light sensors are placed into the control device. There are also cases where the emitters are placed on the control device and the sensors are placed near the display. As these systems require additional emitters to be placed near the display, they are independent of the type of display used and therefore work with any type of visual display. Furthermore, these newer systems allow for more accurate control than early systems. Examples of light-based controller systems are shown in *Figure 2*.



Figure 2: Examples of light-based controller systems used for 3D video games.
Top-left: Nintendo *Zapper*; Top-right: Namco *GunCon*; Bottom: Nintendo *Wii* Remote

One of the main contributions of light-based controller systems in the context of player immersion is that they allow the player to interact with certain game environments using more natural and intuitive motions than traditional game controller devices. For example, aiming and shooting a ‘light gun’ requires the same motions as that required to aim a real gun. Similarly, the Nintendo *Wii* remote has been used in several games as a flashlight, where the player uses the remote in the same fashion as they would use a normal flashlight in a real world situation. By corresponding in-game motions with real life motions, the player becomes more immersed into the game experience. Furthermore, light-based systems allow for absolute positioning information to be measured, thus providing a greater level of control over spatial orientation within the game environment.

However, there are several disadvantages to light-based controller systems that reduce their effectiveness in delivering even better immersion into a game. First, the accuracy of current light-based controller systems still leaves much to be desired. This creates a level of frustration that takes the user out of the gaming experience. Second, most light-based controller systems are designed to provide only 2D spatial motion. To work with this restriction, a majority of shooting games that utilize such systems are limited to being what are commonly called ‘rail shooters’, where the player has no direct control over the path taken by their in-game avatar. This lack of full avatar control may further pull the player away from the gaming experience. In many cases, light-based controller systems can be very limiting in providing an immersive experience.

Touch-Screen Controller Systems

Alternative controller systems that have become popular over the last few years are those based on touch-screen technology. These controller devices consist of a touch-screen with which the player interacts either using their fingers or a stylus. The touch-screen senses the touch points and relays that information to software that interprets this information to obtain 2D spatial position and pressure at each touch point. This information can then be used to provide control over the virtual environment. Of particular interest in recent years are multi-touch controller systems, which allow for multiple touch points to be sensed and interpreted in a simultaneous fashion. Popular applications of multi-touch technology include Microsoft *Surface*, Apple *iPhone*, and the research conducted by J. Han (Han, 2005). In terms of game controller systems utilizing touch-screen technology, the most popular example is the touch-screen system used in the Nintendo *DS* portable gaming system. Examples of touch-screen controller systems are shown in *Figure 3*.



*Figure 3: Examples of touch-screen controller systems.
Top: Nintendo DS; Bottom: Microsoft Surface*

One of the biggest advantages of touch-screen controller systems over conventional game controller systems is that they are highly accurate, responsive, and allow for much more natural control over the gaming environment due to their use of flexible finger and stylus motions that can be intuitive to the player. Furthermore, touch and stylus motions are much more natural for

certain games that require writing and material manipulation. For example, puzzle games such as Sudoku and crossword puzzles on the Nintendo *DS* are played using the stylus just as they would be in real life using a pencil. By corresponding in-game motions with real life motions, the player becomes more immersed into the game experience. However, there are several disadvantages to touch-screen controller systems that reduce their effectiveness in improving immersion. First, these systems require a physical touch-screen display to function. While acceptable for small-scale systems such as the Nintendo *DS*, touch-screens are typically expensive and obstructive and not well suited for consumer-level gaming purposes, particularly for situations where large motions are required. Second and more importantly, touch-screen controller systems are designed to provide only 2D spatial control. This limits their effectiveness as a 3D control mechanism and breaks the illusion of interacting within a 3D gaming environment. As such, the lack of true 3D interaction may diminish immersion effects. Therefore, like light-based controller systems, touch-screen controller systems can have limited application in providing an immersive game play experience.

Vision-Based Controller Systems

A particularly interesting type of alternative controller system in the context of player immersion within a virtual environment is based on vision sensing technology. In vision-based controller systems, various extremities of an object (or markers placed on the object) are recorded using one or more cameras. The acquired visual data is processed (in real-time for video games) using computer vision techniques to determine absolute position and/or orientation information about the object being tracked. This position and/or orientation information can then be used to control and interact with the game environment. Traditionally, vision-based controller systems require very expensive equipment and therefore have been limited to studio environments for professional motion capture purposes. However, advances in manufacturing technology have resulted in low-cost digital camera systems and powerful microprocessors that support consumer-level motion tracking purposes. The best known example of vision-based controller system in the gaming industry is the Sony *EyeToy*, which consists of an inexpensive digital camera and games that make use of the visual data provided by the camera to control the game environment. Another recent vision-based controller system is the Kick Ass *Kung-Fu* platform, which creates a large cushioned playfield using custom computer vision technology that can be used for martial arts training.

There are a number of important advantages to vision-based controller systems when they are compared to the other controller systems in terms of improving player immersion. First and foremost, vision-based systems are capable of providing accurate absolute positional and/or orientation information in 3D space. By interpreting the motion of the avatar in 3D space, the player can use natural body motions to truly interact with the game environment as they would with real life environments. By using corresponding motions in both the real world and the virtual world, player immersion in the game environment may be improved since there are fewer interruptions caused by unnatural motion. Furthermore, vision-based systems are much more flexible than other controller systems as they are less constrained by physical limitations and allow for greater freedom of motion.

Despite the aforementioned benefits of using vision-based controller systems to improve game immersion, there are several important limitations to using such systems alone. First, vision-based controller systems are highly susceptible to visual occlusion. The object being tracked must be visible at all times. More importantly, most computer vision algorithms used to extract 3D position and orientation information from visual data are computationally very expensive. For example, to track the 3D position and orientation of a human body, it is necessary to first automatically identify tracking features on the body using feature extraction methods such as SIFT (Scale Invariant Feature Transform) (Lowe, 2004). Then you must identify the semantic association between these identified features and human body parts, and infer the 3D position and motion of these associated human body segments by estimating the 3D model based on the observed location and prior motion information using statistical state estimation methods such as Kalman filtering (Kalman, 1960) and particle filtering (Kitagawa, 1987). As such, current vision-based game controller systems use basic computer vision algorithms that limit player interaction to very simple 2D motions. Vision-based systems that allow for complex 3D motions would be too sluggish to allow for a true sense of immersion in a 3D game environment. In these respects, current consumer-level game controller systems that rely solely on vision technology may be too limiting to provide player immersion within a 3D game environment.

Inertial-Based Controller Systems

Another promising type of alternative controller system that has been used to improve immersion in game environments are those based on inertial sensing technology. Inertial sensors such as accelerometers and gyro-sensors have traditionally been used extensively for vehicle navigation and tracking (Dias et al., 1995; Powell et al., 2003) due to the size and cost of early inertial sensors. However, advances in Micro-Electro-Mechanical Systems (MEMS) technology have resulted in the development of miniature, low-cost inertial devices for human motion tracking (Luinge, 2002). Inertial-based controller systems utilize devices such as accelerometers and gyro-sensors to measure linear acceleration and angular velocity, respectively. This information can then be used to interpret the 3D motion performed by the player and used to interact with the game environment accordingly. The most recent uses of inertial sensors for the purpose of game control are the Nintendo *Wii* Remote and the Sony *SIXAXIS* wireless controller. An inertial-based controller system designed for tracking full body motion during game play has also been proposed using a network of accelerometers (Whitehead, 2007).

There are a number of key advantages to using inertial-based controller systems for improving immersion within a game environment. First, inertial-based controller systems can be used to provide positional and/or orientation information in 3D space. By interpreting the motion of the avatar in 3D space, the player can use the same motions to interact with the game environment as they would with real life environments. Second, information acquired by inertial sensors can be computationally processed into 3D position and orientation in a very efficient manner. This allows for a large number of data samples to be acquired and processed in real-time, thus resulting in highly responsive control over the game environment. Additionally, such systems are not affected by visual occlusion. Finally, inertial-based systems are more flexible than other controller systems (with the exception of vision-based systems) as they allow the player to perform a wide range of motions. All of these factors help improve player

immersion into the game environment by allowing for responsive interaction in 3D space using natural bodily motions.

Inertial-based controller systems do have several important limitations that may prevent them from delivering an immersive experience when used as the sole game controller source. First, inertial-based controller systems are not capable of measuring absolute 3D positioning information. This limits the ability of such systems to provide truly accurate 3D spatial interaction within a game environment, given the lack of an absolute frame of reference. Also, inertial-based systems do not measure 3D position and orientation information explicitly. As such, 3D position and orientation information must be derived based on measured 3D accelerations and angular velocities. This derivation process results in rapid error accumulation and thus prevents such systems from providing reliable position and orientation information over a long period of time. To work around these issues, game developers have relied on using only acceleration and angular velocities to infer motion gestures for controlling the game environment. However, the use of gesture-based control is inaccurate and often leads to player frustration that diminishes a player's sense of immersion. This limits the ability of current consumer-level game controller systems that rely solely on inertial sensing technology to provide 'flawless' player immersion within a 3D game environment.

Proposed Visual-Inertial Hybrid Approach

Of the aforementioned alternative game controller systems, the two most promising systems for improving immersion in 3D video games are those based on vision sensing technology and inertial sensing technology. Each of these types of systems has its own strengths and weaknesses such that neither will suffice on its own for interacting with the game environment. Inertial-based controller systems are capable of providing highly responsive control that is not susceptible to visual occlusion, but are unable to maintain accuracy over a long period of time. Furthermore, they are unable to provide absolute 3D positioning information. On the other hand, vision-based controller systems are capable of providing highly accurate 3D position information over a long period of time. However, due to the computational complexity of computer vision algorithms used to extract position information from visual data, vision-based controller systems provide sluggish control over the game environment. Additionally, vision-based systems are susceptible to visual occlusion. Based on this information, it can be observed that inertial and vision-based controller technologies are complementary in nature: each has strengths that compensate for the weaknesses of the other. Combining vision-based and inertial-based technologies can create a more responsive, accurate, and reliable game controller system that may improve player immersion in the game environment.

Inspired by the synergistic nature of these two technologies, the proposed game controller system takes a hybrid vision-inertial approach to providing more natural and intuitive interaction within a 3D game environment. The proposed game controller system utilizes the theoretical concepts presented by Wong, which focuses on integrating low-cost sensors to produce improved sensor accuracy and reliability (Wong, 2006). The proposed game controller system is comprised of both hardware and software components designed to interpret the 3D motion of the player.

Hardware Architecture

An overview of the hardware components of this vision-inertial system is shown in *Figure 4*. The proposed system consists of three main types of hardware components. The first type of hardware component is what will be referred to as *motion nodes*. Each motion node contains a series of three miniature accelerators and three miniature gyro-sensors that acquire 3D motion data with a full six degrees of freedom. This is illustrated in *Figure 5*. Furthermore, each motion node is uniquely identifiable by the controller system. The motion nodes can be mounted on wearable equipment (such as a glove, shirt, cap, ankle bracelets, and so forth) and static objects (a gun or a racquet, etc.) to allow for natural player movement. The second hardware component is a digital video camera, which tracks the 3D position of the motion nodes relative to the player's surrounding environment. The third hardware component is a processing unit, which computes the data received from the digital camera and the motion nodes to produce the absolute 3D position and orientation coordinates of each motion node. The processing unit used in the proposed system is essentially the video game console or personal computer upon which the game is played. The 3D position and orientation information produced by the processing unit is then interpreted and used by the 3D video game to provide natural motion control over the game environment. By allowing the player to fully interact with the 3D game environment in real-time using natural bodily motions in a highly accurate and responsive manner, the *flow* of the game experience may be improved by allowing the player to focus on the game environment.

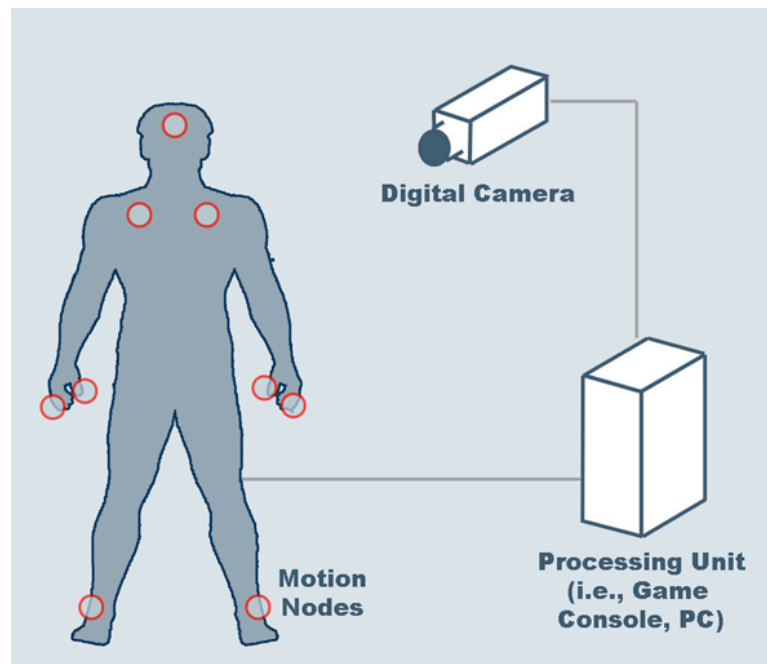


Figure 4: Hardware architecture of the proposed game controller system

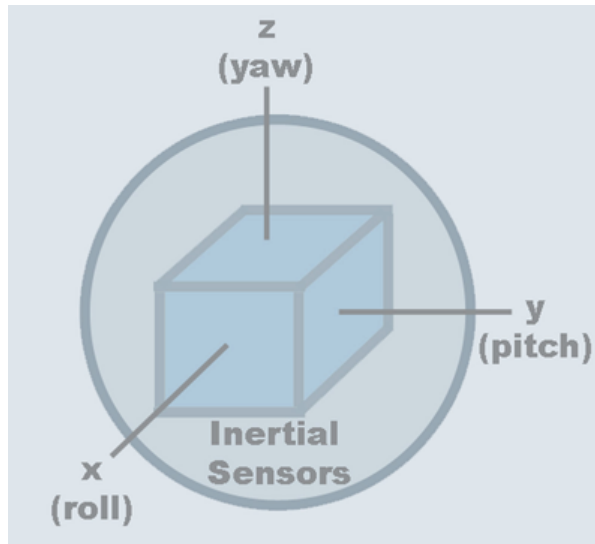


Figure 5: Overview of a motion node

Software Architecture

The heart of the proposed system lies in the software residing on the video game console or personal computer, which takes the raw inertial and vision data received from the motion nodes and digital camera, respectively, and produces highly accurate and responsive absolute 3D position and orientation information. This information is then used to provide natural control over the game environment in 3D space. An overview of the software architecture is shown in Figure 6.

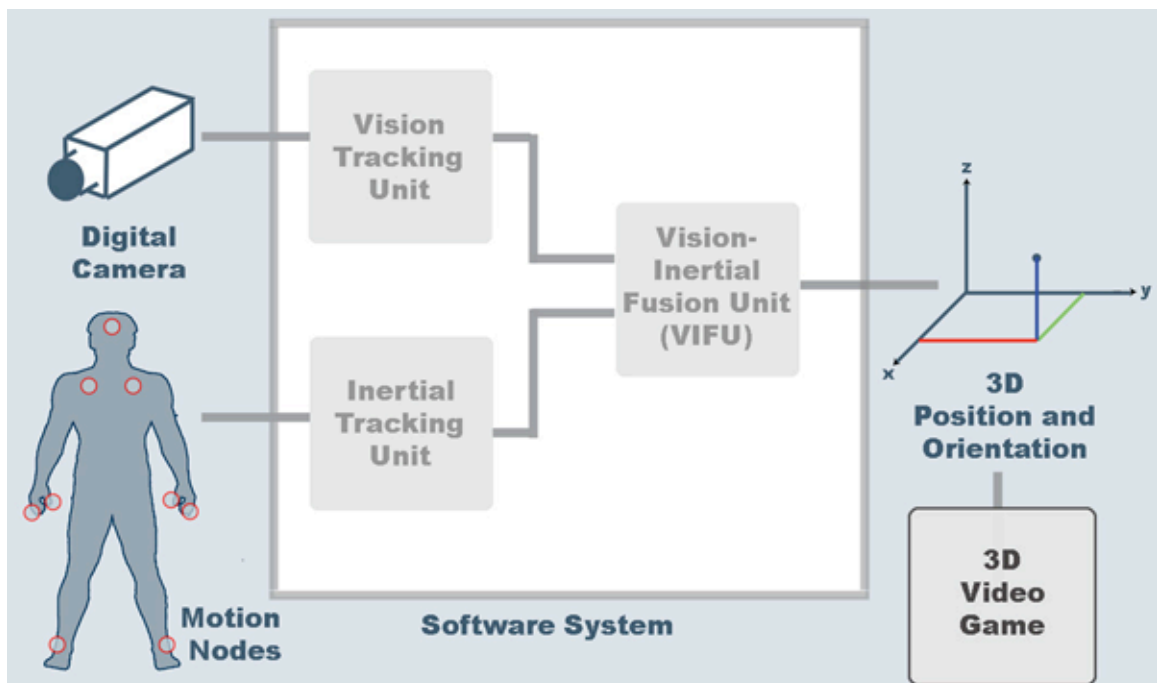


Figure 6: Software architecture of the proposed game controller system

The proposed software system consists of three main software units:

1. vision tracking unit,
2. inertial tracking unit, and
3. vision-inertial fusion unit.

The vision tracking unit takes the raw vision data from the digital camera and extracts absolute 3D positioning information about the individual motion nodes in real-time. Similarly, the inertial tracking unit takes the raw inertial data from the motion nodes and derives relative 3D positioning and orientation information about the individual motion nodes in real-time, except at a much high sampling frequency. Finally, the vision-inertial fusion unit takes the 3D motion information from the vision tracking unit and the inertial tracking unit and combines them to produce absolute 3D position and orientation information that is more reliable, accurate, and responsive than either of the individual technologies can provide. The final set of 3D motion information is then translated into a standardized coordinate frame of reference such that it can be used directly by the 3D video game to provide natural motion control over the game environment.

Vision Tracking Unit

The vision tracking unit is responsible for extracting 3D motion information from the raw streaming video data in real-time. The vision tracking unit consists of two main processes: motion node detection and position estimation. To allow for vision-based tracking of the individual motion nodes, it is necessary to detect and isolate the motion nodes from video data obtained from the digital camera. The motion node detection process must be both highly accurate and performed in real-time to provide responsive control. Fortunately, several characteristics of the proposed system allow for the use of simple and efficient computer vision algorithms while maintaining a high level of accuracy. The individual motion nodes are designed to be uniquely identifiable. As such, they can be isolated relatively efficiently without the need for very complex computer vision algorithms. Most importantly, since the proposed system makes use of inertial data that can be processed in real-time at a very high sampling frequency, the video data acquired from the digital camera can be processed at a much lower sampling frequency without noticeably affecting the accuracy of the system. Therefore, this allows the vision tracking unit to perform motion node detection in real-time without creating a large computational overhead on the system.

Once the individual motion nodes have been detected, position estimation is performed to determine the 3D position of the detected motion nodes relative to the camera's frame of reference. This is accomplished by utilizing the characteristics of the digital camera as well as an inverted pinhole camera model to estimate both the motion of the nodes along the image plane as well as the approximate distance of the nodes from the camera. The extracted 3D motion is then sent to the vision-inertial fusion unit for further processing.

Inertial Tracking Unit

The inertial tracking unit is responsible for extracting 3D motion information from the raw streaming inertial data obtained from the motion nodes in real-time. The inertial tracking unit consists of two main processes: 3D orientation estimation and 3D position estimation. The 3D orientation estimation process performs several important tasks to obtain the relative 3D orientation of the individual motion nodes. First, error reduction techniques are used to filter some of the device noise caused by the miniature gyro-sensors in the motion nodes. The filtered angular velocity data from the gyro-sensors are then transformed from the coordinate frame of reference of the associated motion node to coincide with the camera coordinate frame of reference. Finally, the relative 3D orientation is obtained by performing a single integration on the transformed 3D angular velocity information.

Similar to the 3D orientation estimation process, several important tasks are performed to obtain the relative 3D position of the individual motion nodes. First, error reduction is performed to filter device noise from the accelerometers within the motion nodes. Next, the filtered data is transformed to align with the camera coordinate frame of reference. Finally, the transformed 3D linear acceleration information is processed using a double integration to determine the relative 3D position information of the motion nodes. The extracted relative 3D motion is then sent to the vision-inertial fusion unit for further processing.

Vision-Inertial Fusion Unit

The vision-inertial fusion unit is responsible for combining the relative 3D motion from the inertial tracking unit and the absolute 3D motion from the vision tracking unit to produce the final set of 3D position and orientation information used to provide natural control over the game environment. The sensor fusion performed in this unit is crucial to providing 3D motion control that is more accurate, responsive, and reliable than either of the individual controller technologies can provide on its own. The proposed system takes advantage of the high-frequency nature of the inertial data (which allows for highly responsive control) and combines it with the long-term stability of the vision data (which allows for highly accurate and reliable control). The vision-inertial fusion unit utilizes an error-state complementary Kalman filter configuration (Maybeck, 1979; Cooper et al., 1994), which continuously estimates the time-varying error of the high-frequency inertial data using an error model and the low-frequency vision data. The error estimate is then used to correct the motion data in real-time. This approach has been shown to produce significantly improved 3D motion data that is both accurate, reliable, and responsive (Wong, 2006). The final set of 3D motion information is then translated into a standardized coordinate frame of reference and sent to the 3D video game to be interpreted into 3D motion control over the game environment.

Prototype Implementation

Based on the visual-inertial hybrid game controller system design outlined above, a prototype system was implemented for the Microsoft *Windows* environment, written in C++ in Microsoft *Visual Studio*. For the digital camera system, a low-cost PC camera capable of recording 30 frames per second at a 640 x 480 resolution was used. Each motion node contained

Analog Devices ADXL202 low-cost MEMS accelerometers and low-cost gyro-sensor devices, which were capable of performing inertial motion measurements at a high sampling frequency of 2048 Hz. Both the camera and inertial sensing devices used in the prototype system were chosen because these components are readily available and can be manufactured for very low costs. Several motion tracking tests were performed over extended periods of time to test the long-term reliability and accuracy of the prototype system. It is important to note that there are currently no game controller systems available that are capable of providing accurate absolute 3D motion information, making direct comparisons with existing systems difficult. Based on experimental results, it was observed that the prototype system was capable of providing accurate motion tracking, with an accumulated position error of under 0.01 m and an accumulated orientation error of less than 7 degrees for over 45 minutes without the need for recalibration. This level of motion tracking accuracy demonstrates that the proposed game controller system can be implemented using low-cost components and therefore may be well suited for controlling character movement within a 3D game environment on consumer-level entertainment systems.

Applications in 3D Video Games

With the proposed game controller system explained, it is important to provide a number of different game applications to illustrate how the system may be used to enhance the level of player immersion within game environments. As such, this section addresses possible implementation scenarios within two popular genres of video games: fighting games, and role-playing games.

Fighting Games

As mentioned previously, while the graphics and fight mechanics of popular fighting games such as the *Virtua Fighter* and *Tekken* series have improved greatly over the years, they are limited by conventional game controllers that do not allow for natural player motions such as punching, kicking, and blocking. Games like *Wii Sports* (boxing) allow for more natural punching action, but their reliance on motion gestures instead of true 3D position and orientation make certain motions frustrating and render player movement around the environment impossible to track. This disconnection between the player's actions and the in-game character's actions may prevent the player from becoming fully immersed in the gaming experience. Using the proposed game controller system, motion nodes can be attached to the player's hands, head, and ankles so that the 3D motion of the player can be tracked in real-time. This allows the player to not only perform punches and kicks in an accurate and responsive manner, but would also allow the player to move around the 'ring' to evade an opponent's attacks. This may create a better sense of immersion within the game environment for the player, facilitating the delivery of a more complete *kinetic game* (Parker, 2006).

Role-playing Games

Another genre that could greatly benefit from the proposed visual-inertial system in the context of game immersion is role-playing games. While the graphics, physics, and sound of popular role-playing games such as *Elder Scrolls: Oblivion* create a very realistic virtual environment, the control mechanics leave much to be desired from a player interaction

perspective. The way the player interacts with this game world may feel unrealistic, in stark contrast with the photo-realistic game environment being presented visually. This is particularly evident during combat, where different weapons and magic spells are performed using the same generic combination of button presses. This can disconnect the player from their game avatar, the linkage which is essential to the success of a role-playing game. Using the proposed game controller system, motion nodes can be attached to the player's hands to allow the player to manipulate in-game objects (such as jars and doors) and weapons (such as swords and bows) differently within the game environment in a responsive and realistic manner using natural hand motions. Similarly, the player can use natural hand motions to conjure different magical spells during combat. Finally, attaching motion nodes to the body of the player allows the player to block and evade attacks through natural body movements. These features would allow the player to better relate to their in-game avatar and may create a greater sense of immersion and interaction with the game environment for a more enjoyable experience.

Conclusions and Future Research

This paper proposes a novel visual-inertial game controller system that utilizes vision and inertial sensing technologies that may improve player immersion within a 3D game environment. The proposed system allows the player to interact with the game environment using fully natural bodily motions in an accurate, reliable, and responsive fashion. It is believed that the proposed alternative game controller system may be used to deliver improved player immersion in future 3D video games. Further work to be done includes investigating the integration of other sensing technologies into the existing game controller framework proposed here. Furthermore, the visual-inertial system will be tested on real games to investigate whether it may have a noticeable impact on player immersion.

Acknowledgements

Special thanks go to the Natural Sciences and Engineering Research Council of Canada for funding and support.

References

- Adams, E. (2003). *Postmodernism and the Three Types of Immersion*. Gamasutra, http://designersnotebook.com/Columns/063_Postmodernism/063_postmodernism.htm
- Cooper, S., and Durrant-Whyte, H. (1994). *A frequency response method for multi-sensor high-speed navigation systems*. Proceedings from the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 1–8.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- Dias, J., Lobo, J., Lucas, P., and Almeida, A. (1995). *Inertial navigation system for mobile land vehicles*. Proceedings from ISIE'95: IEEE International Symposium on Industrial Electronics, 843–848.
- Han, J. (2005). *Low-cost multi-touch sensing through frustrated total internal reflection*. Proceedings from the 18th Annual ACM Symposium on User Interface Software and Technology, 115-118.
- Kalman, R. (1960). A new approach to linear filtering and prediction problems. *ASME-Journal of Basic Engineering*, 35-45.
- Kitagawa, G. (1987). Non-gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82, 1032-1063.
- Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- Luinge, H. (2002). *Inertial sensing of human movement*. Enschede: Twente University Press.
- Maybeck, P.S. (1979). *Stochastic models, estimation and control: Volume I*. New York: Academic Press.
- Parker, J. (2006). *Human motion as input and control in kinetic games*. Proceedings from Futureplay 2006: The International Academic Conference on the Future of Game Design and Technology.
- Powell, J., Gebre Egziabher, D., Lee Boyce Jr., C., & Enge, P. (2003). An inexpensive DME-aided dead reckoning navigator. *Navigation*, 50(4), 247–264.
- Whitehead, A. (2007). *Sensor networks as video game input devices*. Proceedings from ACM Futureplay.
- Varney, A. (2006). Immersion Unexplained. *The Escapist*.

http://www.escapistmagazine.com/articles/view/issues/issue_57/341-Immersion-Unexplained

Wong, A. (2006). *Low-cost visual/inertial hybrid motion capture system for wireless 3D controllers*. Unpublished master's thesis, Waterloo: University of Waterloo.