RESEARCH PAPER
International Journal of Recent Trends in Engineering, Issue. 1, Vol. 1, May 2009

Low-Cost Multiple Degrees-of-Freedom Optical Tracking for 3D Interaction in Head-Mounted Display Virtual Reality

Yang-Wai Chow
School of Computer Science and Software Engineering, University of Wollongong, Australia
Email: caseyc@uow.edu.au

Abstract— **Interacting with the 3D content present in games and virtual environments generally involves some form of 3D interaction. As such, the design and development of 3D spatial interaction devices are important in creating more realistic and immersive user experiences in 3D virtual environment applications through natural and intuitive human expression. In general, current commercially available 3D input devices for virtual reality applications like data gloves, multiple DOF sensors and trackers, etc. typically come with a heavy price tag. The objective of this research is to investigate an approach to setting up an inexpensive 6-DOF optical tracking system using Wii Remotes, which is adequate for 3D interaction in an interactive Head-Mounted Display (HMD) virtual reality system. For the purpose of HMD virtual reality, a user should ideally be able to use a 3D interaction device in a space surrounding the user. This cannot be achieved when using this game controller in the conventional manner. Also, normal usage of the controller only allows for relative positioning and cannot reliably track 6-DOF. This paper outlines a method of using Wii Remotes for 3D spatial interaction in an area surrounding the user. This paper also presents experimental results conducted in order to benchmark the accuracy of the system, by comparing the system's position and orientation estimates with the readings obtained from a commercial 6-DOF magnetic tracker.**

Index Terms—**3D input device, optical tracking, virtual reality, Wii Remote**

I. INTRODUCTION

Interactive virtual reality and 3D virtual environment applications are becoming increasingly common nowadays. These applications typically require two key technologies; namely, 3D displays and multiple Degrees-of-Freedom (DOF) input devices [1]. While much advancement has been made toward improving the technology required by interactive 3D virtual environment applications, a lot of this progress has focused on technology for the generation of real-time 3D computer graphics. This can be seen in the increasingly powerful, yet affordable, Graphics Processing Units (GPUs) available in the market these days. The driving force behind the rapid progress in this area has mainly been attributed to the demand for better video game technology. However, the development of 3D input devices has evolved relatively slowly. To this end, the desktop is still very much dominated by the mouse, and overall only a small variety of input devices is commercially available at present [2].

Interacting with the 3D content present in games and virtual environments generally involves some form of 3D interaction. 3D interaction has been defined in [3] as human-computer interaction in which the user's tasks are performed directly in a 3D spatial context. This however does not necessarily involve 3D input devices. Conventional 2D input devices like the keyboard and mouse, or the standard gamepads that have numerous buttons and analog controllers are probably not ideal, and certainly not intuitive, for interaction in a 3D spatial context [4]. Furthermore, the unnatural mapping between these 2D devices with 3D content to some extent reduces the user's immersive experience [5]. Thus, this highlights the impact that the development of 3D spatial interaction devices and techniques can offer to 3D virtual environment interaction, via more natural and intuitive human expression.

Typical 3D input devices for virtual reality applications like data gloves, 3D mouse, pointing or gesture devices as well as other multiple DOF sensors or trackers that are available in the market at this point in time often come with a heavy price tag. A number of these input devices and commercially available tracking systems have been documented in [1] and [6]. In fact, the cost of high-end virtual reality systems has long been a factor that has inhibited the development of high-quality virtual environment applications and hindered its use in mainstream society [7]. Consequently, there is much interest in developing affordable tracking systems for virtual reality [8].

However, since with the release of the Nintendo Wii in November 2006, there has been a strong push in the direction of 3D interaction devices for games and virtual environments. In particular the Wii's video game controller, the Wii Remote (informally known as the 'Wiimote'), presents players with an innovative way of interacting with 2D/3D game content. The somewhat simple yet effective optical tracking and motion sensing technology provided by the Wii Remote has given rise to many interesting interaction possibilities. This has in turn revolutionized the way in which certain video games are being developed and played.

Part of the success and attention that the Wii Remote has attracted can be attributed to the fact that it is a cost effective 3D interaction device. In fact, its low-cost has been the main driving factor behind much research effort [9]–[11]. As noted in [12], the potential and future development of input devices which can track 6-DOF that are in the price range as that of the Wii Remote, will go a long way toward improving 3D spatial interaction and providing much more realistic and immersive virtual environment experiences for the general public.

This study aims to develop an inexpensive 6-DOF optical tracking system using Wii Remotes as a 3D input device that is suitable for use in immersive Head-Mounted Display (HMD) virtual reality. Previous work focused on various system design issues and 3D user interaction approaches using the system [13]. This paper presents experimental results carried out in order to evaluate and assess the accuracy of the system in comparison with a commercially available 6-DOF magnetic tracking system.

II. Background

This section presents a brief overview of the Wii Remote, as well as the way in which this game controller has been incorporated into virtual environment applications developed by other people. It also gives some background as to the purpose and motivation behind this research, and outlines issues that had to be considered when attempting to use this game controller for 3D spatial interaction in HMD virtual reality.

A. The Wii Remote

The Wii Remote is a cost effective wireless device that provides basic optical tracking and motion sensing technology. The device combines an infrared sensor with 3-axis accelerometers, vibration feedback, a speaker and a variety of buttons all within a single device [14]. Wireless communication is made available by means of Bluetooth connectivity. Furthermore, the controller can also be connected to a number of other low-cost extensions, like the 'Nunchuk' extension, which in addition to 2 buttons and a control stick also has similar motion-sensing technology [15].

Optical tracking can be achieved via the infrared optical sensor mounted in front of the controller that can simultaneously detect up to four infrared light sources, as long as they are within the sensor's limited field-of-view. A number of different sources have reported slightly different field-of-view measurements [14], [16], [17]. The reason for these unofficial estimates is that to date Nintendo has not released any official technical specifications about the technology contained within their game controller. The detected image positions of these infrared light sources are reported over a 2D resolution of around 1024×768 pixels [14].

The optical sensor is typically used in conjunction with the 'sensor bar'. Nintendo's official sensor bar basically consists of two groups of infrared LEDs, with wavelengths of around 900 nm without modulation, located at either end of the bar [16]. When used for optical tracking, normal usage requires the Wii Remote to be pointed in the direction of the sensor bar, which is typically placed at a fixed location either above or below the TV or monitor. This configuration allows for relative positioning of the device with respect to the sensor bar, which effectively acts as two infrared light sources.

Tracking in virtual reality refers to the process of tracing the scene coordinates of moving objects in real-time. Tracking often requires the position and orientation of an object to be recovered. In the 3D domain, this can be represented by 6 independent variables; 3 translational coordinates (x, y, z) and 3 rotational angles (yaw, pitch, roll). These systems are called 6-DOF tracking systems. Real-time tracking is essential in interactive virtual reality because these applications demand the generation of real-time 3D graphics based on the tracked information [18]. When used in the conventional manner, the Wii Remote can only detect relative positioning and cannot reliably track 6-DOF.

B. Related Work

The Wii Remote has been used for a variety of different applications. Examples include applications for motion capture [11], gesture based applications [19], [20], robot control interface [21], for conducting a virtual orchestra [22], and many others.

A number of people have demonstrated head-tracking for desktop virtual reality by mounting infrared light sources on a user's head and detecting these using a fixed location controller, this is sufficient to give the impression of parallax in a display [9], [14]. Others have used the Wii Remote as a pointing and manipulation device in conjunction with large screen virtual environment displays [23], [24] and for virtual environment navigation [10]. In addition, attempts have also been made to place multiple sensor bars at a number of locations within the display area, in order to increase the interaction space for a virtual reality theatre [17].

C. 3D Spatial Interaction in HMD Virtual Reality

The virtual environment applications described in the previous section were only concerned with fixed screen displays. The requirements for immersive HMD virtual reality fundamentally differ from such systems, primarily because the location of the display screen is not fixed. In HMD virtual reality, the user views the virtual environment through display screens that are in front of the user's eyes. When the user moves his/her head, these screens move as well and the display has to be updated in real-time based on the user's head position and orientation, generally resulting in a 360 degree field of regard. This means that unlike fixed location displays, the user should be allowed to interact with the virtual environment by physically turning around. Therefore, in order to adequately interact with the surrounding virtual environment, the user should ideally be able to use the 3D input device in a 360 degree space in the horizontal plane around the user.

Under normal usage, in order to obtain relative positioning of the Wii Remote or to use it as a pointing device, the device has to be pointed at the sensor bar. While this is adequate for fixed screen displays, it severely restricts the interaction scope in the case of HMD virtual reality. Even though it is possible to increase the number of sensor bars to enlarge the interaction space, as in [17], it would be rather impractical to completely surround the user with infrared light sources.

Furthermore, unlike head-tracking for desktop virtual reality applications, head-tracking accuracy with low latency is absolutely vital in immersive HMD virtual reality systems. Otherwise the user may suffer from a variety of adverse side effects collectively known as simulator sickness or cybersickness. In this respect, the noisy readings supplied by the Wii Remote makes it ill suited for user head-tracking in HMD virtual reality.

Nevertheless, the game controller can be used as a hand-held 3D interaction device for applications where perfect accuracy is not essential. Moreover a human user cannot really hold a hand-held device perfectly stationary, so slight inaccuracies or lag in the data outputs may not impact user satisfaction. For example, from the user's perspective in an application like a virtual reality game, slight inaccuracies and delays might be tolerable as long as it does not impede user task performance in the virtual environment.

III. System Design

This section outlines the design of a system using Wii Remotes whereby the game controller can be used as a 3D interaction device in a space around the user. The system design uses at least two Wii Remotes; one that the user holds and moves around as the 3D input device, and an overhead Wii Remote used to track infrared light sources. Fig. 1 gives an overall depiction of the system setup.

940nm wavelength infrared emitters with a 160 degree viewing angle were used. Four clusters of these infrared light sources were attached around the handheld Wii Remote in a non-coplanar configuration. When used in the space within the overhead Wii Remote's field-of-view, this controller would be able to detect these infrared light sources. Based on the relative 2D image positions of the infrared light sources as seen and reported by the overhead infrared sensor, calculations can then be performed to estimate the position and orientation of the handheld Wii Remote.

Figure 1. System setup.

Estimating the position and orientation of a rigid-body object with known geometric feature points using images with corresponding image points, has been the topic of much research. This system used the POSIT algorithm as described in [25]. This algorithm requires a minimum of 4 feature points arranged in a non-coplanar configuration. This coincides with the maximum number of infrared light sources that the Wii Remote's infrared sensor can handle at any one time. A Kalman filter was used to smoothen jitters in the position and orientation estimates. This setup allows for 6-DOF; 3D positioning, 360 degrees yaw and approximately +/- 75 degrees pitch and roll. A number of 3D user interaction approaches that can be implemented with this system has previous been outlined in [13].

For 6-DOF camera tracking, the overhead Wii Remote can actually be substituted with other video cameras. For example, Woods et al. [26] demonstrate a 6-DOF mouse using a USB video camera, computer vision software and fiducial markers. However the attraction for using the Wii Remote for this purpose is that the device's optical infrared sensor generally outperforms comparably priced webcams in terms of refresh rate and resolution [14]. In addition, high-performance video cameras are typically a lot more expensive.

IV. Results

To evaluate the accuracy of the system, a Polhemus Patriot was used. This is a commercially available real-time 6-DOF magnetic tracking system. Its specifications document a 60Hz update rate, a static accuracy of 0.1 inch RMS for x, y, z position and a 0.75 degree RMS for orientation [27]. The Patriot's sensor was attached to the handheld Wii Remote, which was moved to different positions and rotated to different orientations in the interaction space below the overhead Wii Remote. The magnetic tracker's readings were then compared to the system's calculated position and orientation estimates in order to benchmark the system. Hence, the magnetic tracker's readings were regarded as the expected values. Differences between the expected values and the estimated values were taken over 10000 frames.

A histogram which depicts the distribution of the system's positional differences as compared with the Polhemus Patriot tracker's readings is shown in fig. 2. The histogram in fig. 3 in turn shows the distribution of orientation errors. These histograms give an overall depiction of the frequency at which differences between the expected values and the estimated values occurred. It can be seen from the figure that the majority of the differences in position are mainly concentrated between -0.5 and 0.5 cm. Differences in rotation, depicted in fig. 2, are more spread out and are generally within -1 and 1 degrees.

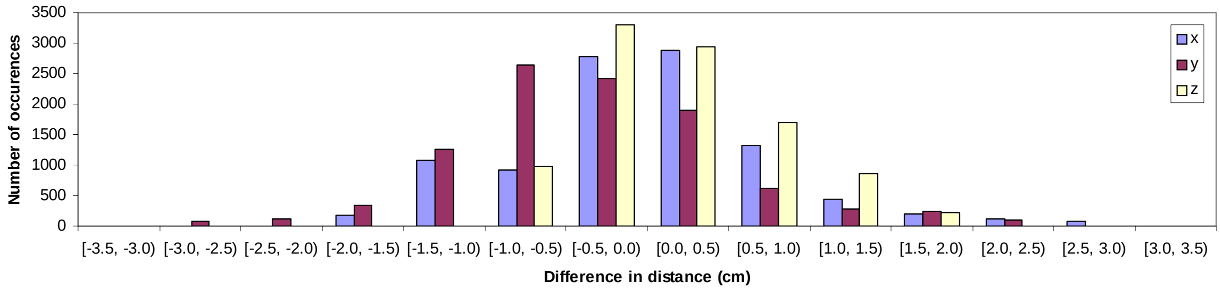

Figure 2. Accuracy of position estimates.

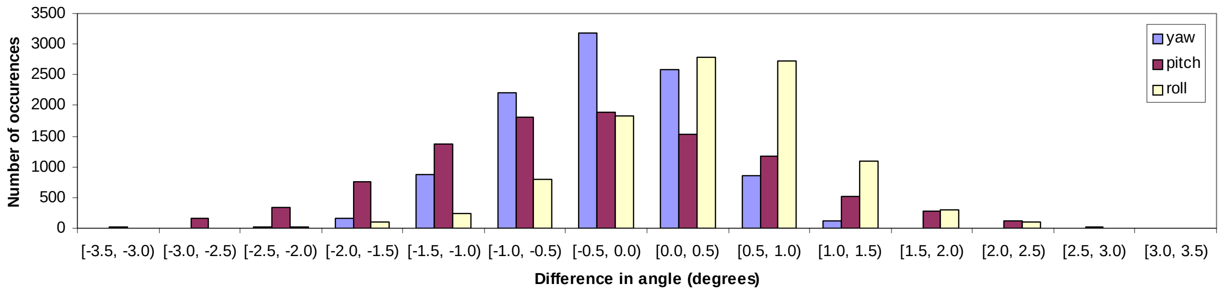

Figure 3. Accuracy of orientation estimates.

TABLE I.
COMPARISONS BETWEEN ESTIMATED POSITION AND POSITION REPORTED BY THE TRACKER

	x	y	z
Max difference (cm)	3.0	2.4	2.2
Min difference (cm)	-2.1	-3.1	-1.2
Average difference (cm)	0.6	0.7	0.5
Standard Deviation	0.78	0.82	0.58
RMS error	0.78	0.88	0.61

TABLE II.
COMPARISONS BETWEEN ESTIMATED ORIENTATION AND ORIENTATION REPORTED BY THE TRACKER

	yaw	pitch	roll
Max difference (degrees)	1.9	2.8	2.7
Min difference (degrees)	-2.4	-3.3	-2.9
Average difference (degrees)	0.5	0.8	0.6
Standard Deviation	0.58	1.02	0.70
RMS error	0.63	1.08	0.78

Table 1 and 2 show more detailed comparisons between the estimated position and orientation values and the values as reported by the tracker. It shows the maximum, minimum and average differences between the expected and estimated distance and rotation values, as well as the standard deviation and the resulting RMS error for each degree for freedom.

While it can be seen from the results that the system is not completely accurate, it can be used for applications that do not require perfect accuracy. An example of such an application is a virtual reality game or entertainment system. Slight inaccuracies in such systems are usually tolerable and will not greatly influence user satisfaction.

It is anticipated that the use of a second overhead Wii Remote will very likely improve the position and orientation estimates. With proper camera calibration the actual positions of the feature points can be obtained through triangulation [9]. In addition, using more overhead sensors at strategically placed positions will also increase the user's interaction space. This will be the subject of future work.

As with other optical tracking systems, this system suffers from the line-of-sight limitation. In other words, tracking can only be performed as long as there is line-of-sight between the sensor and the infrared light sources. Also from the Wii Remote infrared sensor's point of view, if the light sources are too close together, they will be reported a single point. Nonetheless, if line-of-sight is lost, pitch and roll estimations can still be obtained through tilt-sensing (i.e. estimating pitch and roll orientation of the controller with respect to gravity) by using the handheld controller's accelerometer readings. An attachment to the Wii Remote, known as Wii MotionPlus which purports to use 3 orthogonally aligned gyroscopes will likely improve orientation sensing, however this attachment has yet to be released [28].

V. CONCLUSION

This paper investigates a low-cost 6-DOF optical tracking approach to 3D interaction for immersive HMD virtual reality using Nintendo Wii Remotes. Other than for immersive HMD virtual reality, this approach can also be used in other virtual reality systems that require a large area of interaction around the user, such as for virtual reality CAVE systems or systems that use wide-screen displays. The system was evaluated in terms of accuracy by comparing the position and orientation estimates with readings from a commercial 6-DOF magnetic tracking system. While the system is not perfectly accurate, it can be used in virtual reality applications where slight inaccuracies will not impede user task performance. Future work will focus on obtaining user feedback by setting up a virtual reality game type application and conducting usability studies.

ACKNOWLEDGMENT

The author would like to acknowledge the support of the UOW URC Small Grant used for this research.

REFERENCES

[1] E. Moritz, and T. Wischgoll, "Usability of multiple degree-of-freedom input devices and virtual reality displays for interactive visual data analysis," in *Proc. of the 2007 ACM Symposium on Virtual Reality Software and Technology (VRST)*, New Beach, California, 2007, pp. 243–244.

[2] B. Fröhlich, J. Hochstrate, A. Kulik, and A. Huckauf, "On 3D input devices," *IEEE Computer Graphics and Applications*, vol. 26, no. 2, pp. 15–19, 2006.

[3] D.A. Bowman, J. Chen, C.A. Wingrave, J. Lucas, A. Ray, N.F. Polys, Q. Li, Y. Haciahmetoglu, J.S. Kim, S. Kim, S. Boehringer and T. Ni, "New directions in 3D user interfaces," *The International Journal of Virtual Reality*, vol. 5, no. 2, pp. 3–14, 2006.

[4] M. Katzourin, D. Ignatoff, L. Quirk, J.J. LaViola Jr. and O.C. Jenkins, "Swordplay: Innovating game development through VR," *IEEE Computer Graphics and Applications*, vol. 26, no. 6, 2006, pp. 15–18.

[5] S. Seedharam, E.S. Zurita and B. Plimmer, "3D input for 3D worlds," in *Proc. of Australasian Computer-Human Interaction Conference (OzCHI 2007)*, Adelaide, Australia, 2007, pp. 227–230.

[6] M. Ribo, "State of the art report on optical tracking," Technical Report VRVis 2001-25, TU Wien, 2001.

[7] H. Godbersen, "Virtual environments for anyone," *IEEE Multimedia*, vol. 15, no. 3, 2008, pp. 90–95.

[8] T. Pintaric, and H. Kaufmann, "Affordable infrared-optical pose-tracking for virtual and augmented reality," in *Proc. Trends and Issues in Tracking for Virtual Environments Workshop (IEEE Virtual Reality)*, Charlotte, North Carolina, pp. 44–51.

[9] S.Hay, J. Newman and R. Harle, "Optical tracking using commodity hardware," in *IEEE International Symposium on Mixed and Augmented Reality 2008*, Cambridge, UK, 2008, pp. 159–160.

[10] L. Deligiannidis and J. Larkin, "Navigating inexpensively and wirelessly," in *Proc. of Human System Interaction (HSI'08)*, Krakow, Poland, 2008, pp. 165–169.

[11] S. Attygalle, M. Duff, T. Rikakis and J. He, "Low-cost, at-home assessment system with Wii Remote based motion capture," in *Proc. of Virtual Rehabilitation 2008*, Vancouver, Canada, 2008, pp. 168–174.

[12] J.J. LaViola Jr., "Bringing VR and spatial 3D interaction to the masses through video games," *IEEE Computer Graphics and Applications*, vol. 28, no. 5, 2008, pp. 10–15.

[13] Y.W. Chow, "The Wii Remote as an input device for 3D interaction in immersive head-mounted display virtual reality," in *Proc. IADIS International Conference Gaming 2008: Design for Engaging Experience and Social Interaction*, Amsterdam, The Netherlands, 2008, pp. 85–92.

[14] J.C. Lee, "Hacking the Nintendo Wii Remote," *Pervasive Computing*, vol. 7, no. 3, 2008, pp. 39–45.

[15] Nintendo, http://www.nintendo.com/wii/what/controllers

[16] A. Shirai, E. Geslin, and S. Richir, "WiiMedia: Motion analysis methods and applications using a consumer video game controller," in *Proc. of ACM SIGGRAPH Sandbox Symposium 2007*, San Diego, California, 2007, pp. 133–140.

[17] T. Schou and H.J. Gardner, "A Wii Remote, a game engine, five sensor bars and a virtual reality theatre," in *Proc. of Australasian Computer-Human Interaction Conference (OzCHI 2007)*, Adelaide, Australia, 2007, pp. 231–234.

[18] M. Ribo, A. Pinz, and A.L. Fuhrmann, "A new optical tracking system for virtual and augmented reality applications," in *Proc. of IEEE Instrumentation and Measurement Technology Conference*, vol. 3, Budapest, Hungary, 2001, pp. 1932–1936.

[19] T. Schlömer, B. Poppinga, N. Henze and S. Boll, "Gesture recognition with a Wii controller," in *Proc. of the 2nd International Conference on Tangible and Embedded Interaction (TEI' 08)*, Bonn, Germany, 2008, pp. 11–14.

[20] S.J. Castellucci, I.S. MacKenzie, "UniGest: Text entry using three degrees of motion," in *Proc. of ACM CHI 2008*, Florence, Italy, 2008, pp. 3549–3554.

[21] M. Lapping-Carr, O.C. Jenkins, D.H. Grollman, J.N. Schowertfeger and T.R. Hinkle, "Wiimote interfaces for lifelong robot learning," in *Proc. of AAAI Symposium on using AI to Motivate Greater Participation in Computer Science*, Palo Alto, California, 2008.

[22] B. Bruegge, C. Teschner, P. Lachenmaier, E. Fenzl, D. Schmidt and S. Bierbaum, "Pinocchio: Conducting a virtual symphony orchestra," in *Proc. of the International Conference of Advances in Computer Entertainment Technology (ACE)*, Salzburg, Austria, 2007, pp. 294–295.

[23] L.Gallo, G.D. Pietro and I. Marra, "3D interaction with volumetric medical data: experiencing the Wiimote." in *Proc. of the 1st International Conference on Ambient Media and Systems*, Quebec, Canada, 2008, article no. 14.

[24] H.J. Lee, H. Kim, G. Gupta and A. Mazalek, "WiiArts: Creating collaborative art experience with WiiRemote Interaction," in *Proc. of the 2nd International Conference on Tangible and Embedded Interaction (TEI'08)*, Bonn, Germany, 2008, pp. 33–36.

[25] D.F. DeMenthon, and L. S. Davis, "Model-based object pose in 25 lines of code," *International Journal of Computer Vision*, vol. 15, pp. 123–141, 1995.

[26] E. Woods, P. Mason, and M. Billinghurst, "MagicMouse: An inexpensive 6-degrees-of-freedom mouse," in *Proc. of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia (GRAPHITE)*, Melbourne, Australia, 2003, pp. 285–286.

[27] Polhemus, http://www.polhemus.com

[28] Nintendo, http://e3.nintendo.com/pressrelease/