Simulation of a Virtual Reality Tracking System

Kapil Patel, Wolfgang Stuerzlinger
Dept. of Computer Science and Engineering
York University
Toronto, Canada

Abstract—**Virtual reality tracking systems are used to detect the position and orientation of a user inside spatially immersive systems. In this paper, we simulate a laser-based tracking system that was originally developed for a six-sided spatially immersive system in environments with one and five walls to evaluate its performance for other installations. As expected, the results show that performance degrades with the number of walls, but they also show that tracking with even one wall is still very feasible.**

Keywords-virtual reality tracking, 6DOF tracking, simulation

I. INTRODUCTION

We focus on the problem of tracking the pose of a person's head or hand moving inside a room spatially immersive Virtual Reality (VR) system as accurately and robustly as possible. A Spatially Immersive Display (SID) is a class of display systems that surround the user with computer-generated imagery replacing all or a large portion of a person's field of view. SIDs have been used for scientific visualization, training, entertainment and VR research. To create a compelling visual world, the SID must produce correct visual cues, such as perspective, parallax, and stereo. These visual cues are highly dependent upon the pose of the user's head, more precisely the position and orientation of the head. Hence, VR tracking systems are an essential technology for SIDs.

Previous work has demonstrated that for a SID with six walls and with a tracking device using 17 lasers whose spots are observed from the outside of the SID by cameras, very high levels of accuracy can be reached [16][17]. Our current work investigates the performance of several variations of this tracking device with different laser configurations and on fewer walls. For this we used software simulator to mimic various user motions within the SID. More precisely, we simulate IVY, the Immersive Virtual environment at York. IVY is a six-sided fully enclosed SID, located at the Dept. of Computer Science and Engineering at York University The display is a cubical 2.4m x 2.4m x 2.4m structure, consisting of fabric projection screens stretched over steel and aluminum frames and a glass floor. The floor and ceiling of IVY are split into two parts to enable it to fit inside limited vertical space. IVY uses eight projectors providing different images for left and right eye using actively switched shutter glasses. A cluster of PCs running Linux provides image generation. IVY has a software framework, the VE library, which makes application development easier. A detailed description of the system is available in [11][12].

A. *Tracking Systems for Spatially Immersive Virtual Reality*

To enable the display of the correct images and also interaction with the user, the image generation system for a SID requires information about the position and orientation of the user's head at any given point in time. Many types of motion tracking systems exist. Magnetic tracking is the oldest tracking technologies used in SIDs. The relative strength of the detected magnetic field provides orientation and the magnitude provides distance information. The main drawbacks are the limited range, as well as sensitivity to environmental distortion as the shape of magnetic field is strongly affected by metallic objects in the environment [7][8].

Marker based optical systems can be categorized by two attributes: 1) active or passive markers and 2) inside-out or outside-in. Active markers are often implemented using infrared light emitting diodes (LEDs). They can be activated and deactivated to enable the tracking system to differentiate between them. Some form of optical sensors, such as linear CCDs, PSDs (position-sensitive detectors) or regular video cameras, is then used as photo-detectors and the pose of the target is determined by triangulation among different cameras. Passive markers use spherical retro-reflective targets, which are illuminated by infrared light from the location of the cameras. Based on asymmetric marker configurations, geometrical constraints are then used to identify markers. There are also other kinds of passive marker systems that use regular video cameras and pattern recognition techniques, see e.g. [1].

For outside-in optical systems markers are placed on the objects that are being tracked and detectors are stationary. Two or more detectors located outside of the tracking volume record the positions of the markers and measurements from several cameras are combined via triangulation to determine the position of each marker in 3D space. With rigidly connected groups of three or more markers 6DOF can be measured for each marker group, see e.g. [10]. For inside-out optical systems markers surround the tracked object and detectors are placed on the tracked objects. The position of the mobile detector is inferred from the known positions of the markers in the infrastructure. Inside-out systems require wide-angle sensors and said markers in the environment. See e.g. the HiBall system [20].

The accuracy of optical systems is determined by characteristics like sensor resolution (for fixed-point sensors, such as CCDs) or position noise for analog systems. Positional accuracy for inside-out and outside-in systems is comparable. However, outside-in systems provide a lower amount of angular resolution because they need to infer orientation from the difference in two position measurements. Inside-out systems can provide better angular resolution because a small change in orientation causes a large shift in the camera image.

Some laser-based tracking systems perform active scanning with laser beams to determine the angles to the target, such as in Ascension LaserBIRD system. Other systems measure the distance to the target using time-of-flight or interferometric techniques. However, because of the cyclic nature of phase measurements, these systems provide only relative motion information. Acoustic systems usually rely on ultrasonic chirps to determine the pose of the user [14]. Distances between one or more ultrasonic emitters and one or more receivers are determined. Once several distance measurements are obtained, the 3D locations of the receiver can be calculated. Various other approaches have been developed, including time-of-flight, phase-coherent, and spread spectrum based ones [14]. Wireless versions of acoustic systems are also available [6]. Another system used active switching of display walls, where the walls can be made opaque or transparent by applying voltage to the panels [13]. Then cameras can directly track the user inside the display.

Miniature gyroscopes and accelerometers can be used to measure linear and/or angular acceleration. This constitutes another class of tracking systems called inertial trackers. To perform position tracking, acceleration measurements from inertial sensors need to be integrated twice to obtain position measurements, which leads to drift. On the positive side, no external reference is required. With moderately sized equipment it is currently impossible to achieve sufficient accuracy, as the double integration step magnifies the noise and inaccuracies in sensor calibration, causing the position estimate to drift on timescales of seconds. Also, there is orientation drift due to mechanical inaccuracies in gyroscopic technologies. Inertial technologies are often used with hybrid tracking systems, where the other system compensates for the drift.

Due to the complementary strengths and weaknesses of different tracking systems, combinations of tracking technologies often result in better systems, e.g. [6]. A hybrid optical-inertial tracking system was developed by A. Hogue at York University [3][4]. The system combines a commercial inertial tracker, a InertiaCube2 [5], with device that emits four lasers. A custom designed vision system tracks the laser spots on the walls of the SID with a camera outside each projection surface. This system lacks robustness, as it requires that *all* four laser spots be visible to the cameras in each frame.

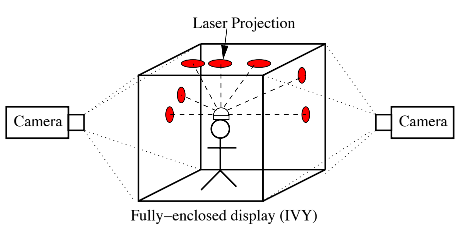

Figure 1. Hedgehog Tracker within fully-enclosed display (IVY)

The Hedgehog [15][16][17] improves on the system presented in [4], overcoming several of its limitations. As in previous work, this system works within IVY, where each (translucent) projection surface is projected on *and* observed by a camera from the outside, see Fig. 1. The tracker is designed to be worn on the head of the user, but could also be held in the hand to track hand motions. In contrast to previous work the Hedgehog uses 17 laser diodes, which gives it higher measurement accuracy and much greater robustness. Due to the algorithms used, any number of visible laser dots can be used to update the location of the sensor. A novel "cold start" methods intelligently controlling the activation state (on or off) of each individual laser diode over time to enable the tracking algorithm to identify which laser dot corresponds to which laser diode. Moreover, the set-up effort is reduced as the system can automatically adjust the camera parameters while the tracker is in operation.

II. SIMULATING THE HEDGEHOG TRACKER

We choose to simulate new versions of the Hedgehog tracker before proceeding with new implementations. Part of the motivation was that IVY recently had to be moved to a new location and the camera infrastructure was not migrated to the new location. Moreover, simulation permits us to investigate many different configurations without having to physically prototype each one. It also permits us to run more simulations in less time. In effect, we can explore the solution space quicker and more comprehensively.

In the simulator the beams emitted by the laser diodes in the Hedgehog are intersected with the SID walls. A simulated camera, with a geometrical setup identical to the real system (Fig. 1) then records each of the resulting spots. The simulation takes also the discretization due to individual camera pixels into account. To derive the spatial pose of the Hedgehog tracker, the system then uses a set of several algorithms: 1) Laser identification 2) Point tracking 3) Pose estimation 4) Motion tracking 5) Start-up and 6) Coordinating the operation of the tracker as described below.

For laser identification during a cold start, the system supports active time-multiplexing of lasers. For simplicity, the system turns on only one given laser at a time in the cold start phase. Unless a laser is invisible to the cameras, e.g. because it hits a seam between screens or due to obstacles, or reflections are present this makes laser identification extremely robust.

For point tracking, a temporal coherence method is used, which predicts each laser's path. This prediction uses a 3D geometrical model of the environment combined with the most recent Hedgehog pose to predict where a given laser spot should show up next in the 2D camera images. This is then used to identify the nearest laser spot in a camera image, within a given threshold distance.

For pose estimation a linear N-point algorithm is used, which computes the pose of the tracker based on the constraints imposed by the geometric configuration. This algorithm works with 5 or more observed laser points and each additional observed point contributes to the accuracy of the solution. It does not require an initial pose estimate and is hence used during the cold start phase.

For motion tracking a recursive least-squares algorithm is used at run time. More precisely, the system uses a variant of a Kalman filter called SCAAT (Single constraint at a time) [18].

SCAAT keeps track of value of system state variables and also their statistical uncertainty. This algorithm is “recursive” as it requires a prior reasonably close estimate of the system state and incorporates a new measurement to produce a new estimate of the system state in a “predictor-corrector” fashion. SCAAT allows us to estimate the state of a globally observable system using measurements from a locally unobservable system, as it can update the pose estimate using any number of visible laser dots, even a single one. SCAAT is also capable of “dual estimation”, i.e. it is capable of estimating not only the state variables of the system but also the parameters of its subcomponents. This enables auto-calibration of cameras. Due to the recursive nature of this algorithm it is not suitable for initialization of the tracker.

For coordinating the operation of the tracker at run time the Monitor algorithm is used. The Monitor is a finite state machine that has also the purpose of initializing the Hedgehog and restarting it in case of a tracking failure. This can happen e.g. due to prolonged blockage of (nearly) all lasers. Coordination of the tracker takes place as follows: The monitor initializes the tracker by acquiring a number of points using the laser identification algorithm. After a sufficient numbers of lasers are identified, the pose estimation algorithm is used to produce an initial estimate. The system then transitions to the use of SCAAT filter for updates. At this time all remaining lasers can be activated and will contribute to the pose estimate. After continuous tracking has begun, the Monitor keeps track of the measurement residuals by finding the maximum deviation between the predicted and observed laser locations. If the error is larger than an empirically determined threshold, the estimate is likely incorrect and the tracker is restarted.

In practice this system works very well and its first instantiation has been shown to perform equal or better than commercially available systems. In a stationary position, the tracker achieves 0.2 mm and 0.01 degree RMS error. During movements, translations can be tracked with 1.6 mm and rotations with 0.2 degree RMS error, which is significantly better than the real-world results of other systems [2]. The accuracy of the Hedgehog depends on the number of lasers in use at any particular time, the number of pixels in the camera and also the quality of overall calibration.

In our simulations, we set the measurement noise to 1 centimeter. This effectively simulates a camera system worse than that used in the original system. Within this constraint, the results of simulations agree reasonably well with the real-world performance in relative terms (but not in absolute terms). As the 45º configuration of the original Hedgehog was designed for a 6 wall SID, we decided to try different configuration of lasers, i.e. different angles between the lasers. One configuration targeted at a 5 or 6-wall system employs 5 lasers in a 90º angle configuration. However, we also wanted to investigate smaller angles between lasers, such as 30 and 15º configurations. Note that the angle between lasers combined with the camera frame rate and the distance to the screens also puts an implicit constraint on the maximum angular velocity of the device, as otherwise the system could confuse one laser with another during fast rotations. Also, we also wanted to investigate the effect of different numbers of walls, as many SIDs have less than 6 surfaces. In particular, we wanted to explore the case of a single display surface, which is common in large-screen visualization setups. Furthermore, there is a weak interdependency between the angular configuration and the minimum number of walls in the sense that a 90º angle configuration will not work with a single wall system. Finally and mostly for simplicity, we decided to limit ourselves to symmetric, regular laser arrangements.

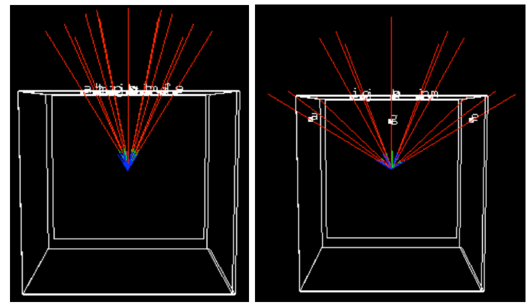

Figure 2. Hedgehog simulator graphics with 15 & 30º configuration.

III. Simulation Results

We ran the simulator with chosen configurations – laser diodes at 15, 30, 45 & 90º as well as from 1 to 6 walls to analyze how well the system can track the Hedgehog pose in these environments. As IVY was already non-operational at the beginning of this work, we were unable to record real user motions. Hence, we decided to use Lissajous curves to simulate the motions of the user both in terms of movement and rotation. These curves are reasonably similar to a human walking and rotating his head inside a SID in real life as they include both straight segments and curved paths. We used a Lissajous curve with coefficients 2 and 3.

The process of simulation and logging of RMS error and calculating the average RMS error with different number of lasers, walls, angles, rotation and movements was automated with the help of Perl script. Each simulation of a stationary tracker lasted 20 seconds, and for movements, rotations, or their combination 170 seconds.

The script removed one display wall at a time in the following order: Floor, Back, Right, Left, Front and Ceiling. Note that the “ceiling only” case is also equivalent to a user pointing a hand-held device at a large screen, so we expect our results to generalize to the scenario of pointing a tracking device at a large screen. Each display configuration is simulated with a tracker configuration between 5 to 17 lasers (only 5 lasers for the 90º configuration). Moreover, we also varied the maximum number of “active” lasers, i.e. lasers that are actively used to track the pose, from 5 up to the maximum in the current tracker configuration. Not all data can be shown due to space constraints, but we selected representative results. In the graphs the vertical y-axis shows the average RMS error and the horizontal x-axis shows the number of lasers in the configuration as well as the number of active lasers simulated. The vertical axis is shown in meters. All simulations with an even (maximum) number of lasers failed to initialize due to some unidentifiable software problem. Finally, we removed all values above 20 cm RMS error as outliers to make the plots more informative.

For the simulation of the Hedgehog with no movement (H), i.e. a stationary tracker, and for any number of walls, all the configurations show approximately the same amount of error. See Fig. 3. The 15º configuration has no outliers and 45 and 30º have about the same amount.

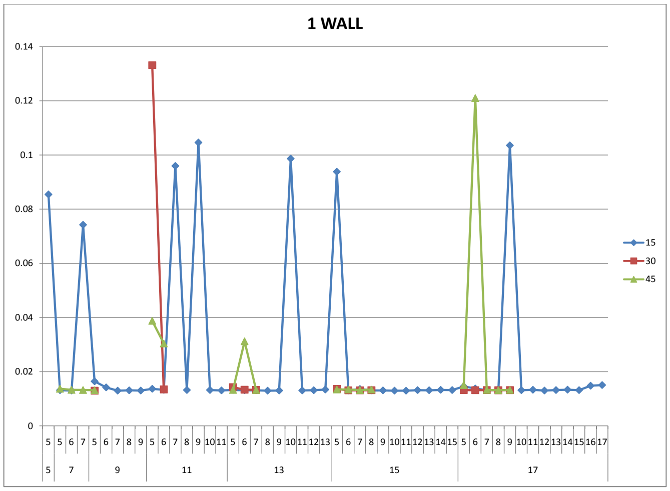

Figure 3. Results for stationary tracker with 1 wall, (similar for 2-6 walls).

For the simulation of the Hedgehog with only movements (HM) and for SIDs between 3 to 6 walls, the 45º configuration is better than the 30º one and the 15º one in turn. With 1 or 2 walls the 45º case has the most number of outliers, while the 15º configuration shows the least. With 6 walls, and only 5 lasers, 90º shows the least amount of error, with 15º being second and 45º third. See Fig 4. and 5.

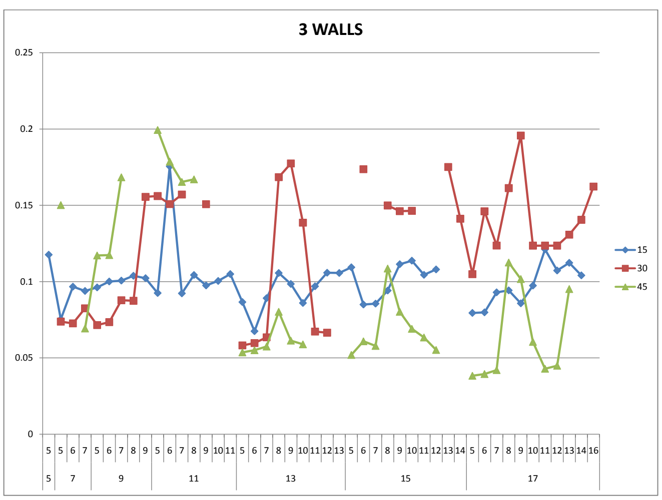

Figure 4. Results for movements with 3 walls (4-6 walls similar).

For simulation of the Hedgehog with rotation (HR) and between 4 and 6 walls, the 15º configuration is better than both the 30 and 45º ones, which perform about the same. Between 1 and 3 walls, clearly 15º is better and more stable than 30º and this is better and more stable than 45º in turn. With 6 walls, and only 5 lasers, the 15º configuration works best, the 30 and 90º ones show about the same amount of error and 45º performs worst. See Fig. 6 and 7.

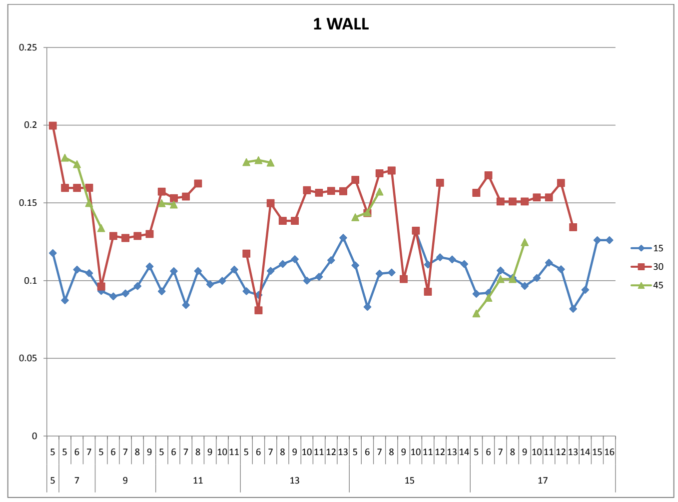

Figure 5. Results for movements with 1 wall (2 walls similar)

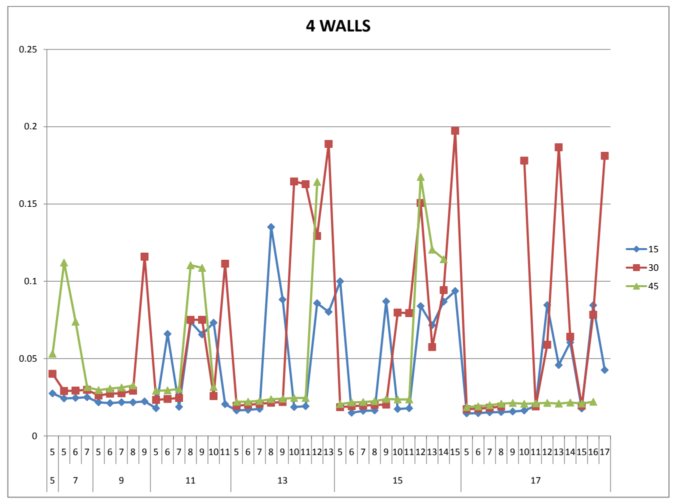

Figure 6. Results for rotation with 4 walls (5-6 walls similar).

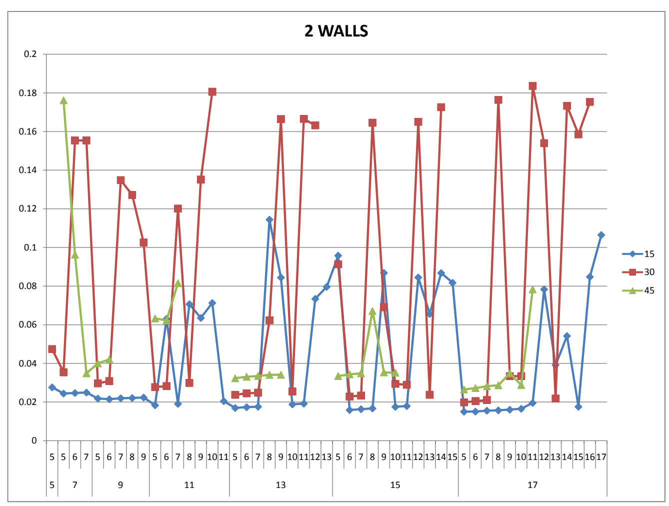

Figure 7. Results for rotation for 2 walls (1 & 3 walls similar).

For the simulation of the Hedgehog with movements and rotation (HMR), with 5 and 6 walls the 45º configuration is slightly better than the 30º one. Below 5 walls, 30º is better than the 45º configuration. As the number of walls decreases 15º has the least amount of outliers compared to the 30 and 45º configurations. If we limit the system to only 5 lasers and in a 6 wall SID, 90º achieves the least amount of error, with 15º being second and followed by 30º and 45º. In general, and with 5 and 6 walls, the results match the measured data from the Hedgehog prototype reasonably close, if slightly optimistic. We see this as validation that the simulator produces practically relevant results. With motions and rotations, the error increases, which is to be expected and also matches the experience with the original Hedgehog. See Fig. 8 and 9.

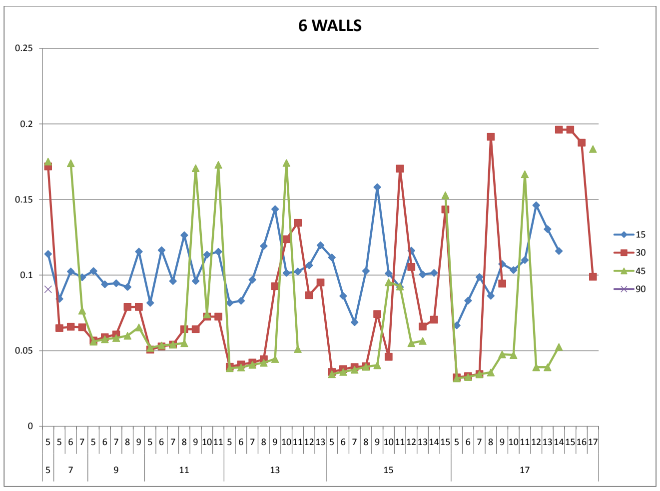

Figure 8. Results for movement & rotation with 6 walls (5 walls similar). Vertical axis in meters.

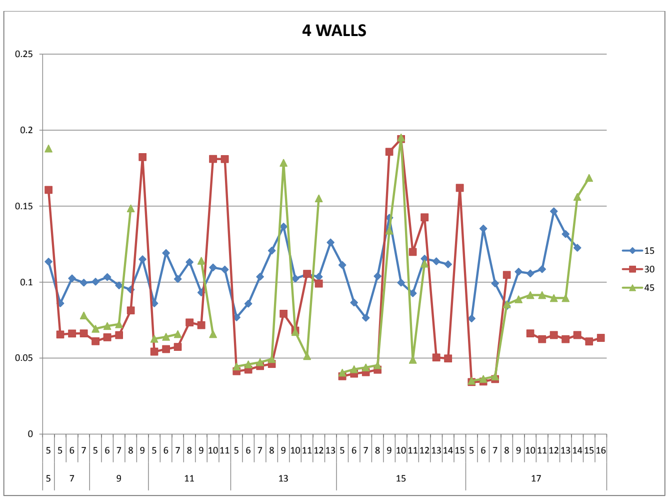

Figure 9. Results for movement & rotation with 4 walls (1-3 walls similar).

In all the graphs, as the number of lasers increases the error decreases, which is to be expected as more constraints can be used to refine the pose. However and with an increasing number of active lasers, the error decreases slightly, which is opposite to expectations. We were unable to determine the exact cause, but can point out that this increase is usually not very strong. Potentially it is a consequence of the chosen setting of the measurement uncertainty.

If all the lasers are projected onto a single display surface (e.g. the ceiling, or a large display wall) using the 15º configuration the number of outliers is smallest and the stability of the system increases. However, the amount of errors is almost an order of magnitude higher, which is to be expected, as the pose cannot be constrained as much due to the geometry. Nevertheless, this is a very interesting application scenario, as it shows that one can use this to build a cheap single-camera tracking system that is very competitive with other systems in this class.

If there are only movements involved, the 45º configuration gives the least amount of error, but the number of outliers increases as the number of walls is reduced. In scenarios that include only rotations, 15º is the best choice for any number of walls. However when both movements and rotations are involved, the 45º configuration yields the least error with 5 and 6 walls. However, a 30º configuration is better for 4 walls and less. In a sense, this is a retroactive validation of the design decisions behind the original Hedgehog system, which was created for a 5 or 6 wall configuration.

In summary, we can recommend a 15º configuration for a single display surface, a 30º configuration for systems with 2 to 4 walls, and a 45º configuration for 5 or 6 wall SIDs.

IV. Conclusion

We presented the simulation of a laser-based tracking system in spatially immersive display environments that use between one and six walls to show images and track the device at the same time. As expected, the accuracy results show that performance degrades with decreasing number of walls, but they also show that tracking with one wall is still feasible.

For future work we are planning to investigate asymmetric configurations where the lasers are arranged in irregular intervals and without having a common origin. This might provide a compromise between the different trade-offs and may yield a device that performs well regardless of the number of display walls.

References

[1] E. Foxlin, L. Naimark, VIS-Tracker: A Wearable Vision-Inertial Self-Tracker, in Proc. of IEEE Virtual Reality 2003, 199-206.

[2] S. J. Gilson, A. W. Fitzgibbon and A. Glennerster, Quantitative analysis of accuracy of an inertia/acoustic 6DOF tracking system in motion. Journal of Neuroscience Methods, 154(1-2), pp 175-182, 2006.

[3] A. Hogue, MARVIN: a Mobile Automatic Realtime Visual and INertial tracking system. Master's thesis, York University, Toronto, Canada, 2003

[4] A. Hogue, M. R. Jenkin and R. S. Allison, An optical-inertial tracking system for fully-enclosed VR displays. In Proc. of IEEE 1st Canadian Conference on Computer and Robot Vision 2004, 22-29.

[5] InterSense Corp., InertiaCube2, 2000. Retrieved from http://www.isense.com/products/prec/ic2/index.htm, July 2005.

[6] InterSense Corp., IS-900 and IS-1200 tracking system, Retrieved from http://www.isense.com/products, Dec. 2010.

[7] V. Kindratenko, A survey of electromagnetic position tracker calibration techniques. In Virtual reality: Research, development, and applications, 5(3), pp 169-182, 2000.

[8] V. Kindratenko, A comparison of the accuracy of an electromagnetic and hybrid ultrasound-inertia position tracking system. Presence: Teloperators and Virtual Environments, 10(6), pp 657-663, 2001.

[9] Lissajous curve, retrieved from http://en.wikipedia.org/wiki/Lissajous_curve, Dec. 2010.

[10] Northern Digital Inc. OptoTak Certus, retrieved from http://www.ndigital.com/certus.php, 2005.

[11] M. Robinson, J. Laurence, J. Zacher, A. Hogue, R. Allison, L. Harris, M. Jenkin, W. Stuerzlinger, Growing IVY: Building the Immersive Visual Environment at York, In Proc. ICAT 2001, 85-89.

[12] M. Robinson, J. Laurence, J. Zacher, A. Hogue, R. Allison, L. Harris, M. Jenkin, and W. Stuerzlinger, IVY: The Immersive Visual environment at York. In 6th International Immersive Projection Technology Symposium 2002.

[13] C. P. Spagno, and A. M. Kunz, Construction of a Three-sided Immersive Telecollaboration System, In Proc. of IEEE Virtual Reality 2003, 37-44.

[14] N. M. Vallidis, WHISPER: A Spread Spectrum Approach to Occlusion in Acoustic Tracking, PhD thesis, University of North Carolina at Chapel Hill, 2002.

[15] A. Vorozcovs, The Hedgehog: A Novel Optical Tracking Method for Fully Enclosed Spatially Immersive Displays, MSc Thesis, York University 2005.

[16] A. Vorozcovs, A. Hogue, W. Stuerzlinger, The Hedgehog: A Novel Optical Tracking Method for Spatially Immersive Displays, In Proc. IEEE Virtual Reality 2005, 83-89.

[17] A. Vorozcovs, W. Stuerzlinger, A. Hogue, R. Allison, The Hedgehog: A Novel Optical Tracking Method for Spatially Immersive Displays, Presence: Teleoperators & Virtual Environment, 15(1), pp 108-121, 2006.

[18] G. Welch, G. Bishop, SCAAT: Incremental Tracking with Incomplete Information, In Proc. SIGGRAPH 1997, 333-344.

[19] G. Welch, G. Bishop, An Introduction to the Kalman Filter, University of North Carolina at Chapel Hill, Technical Report 95-041, 1995.

[20] G. Welch, G. Bishop, L. Vicci, S. Brumback, K. Keller, D. Colucci, High- Performance Wide-Area Optical Tracking - The HiBall Tracking System Presence: Teleoperators and virtual environments 10(1), pp 1-21, 2001.