

The Hedgehog: A Novel Optical Tracking Method for Spatially Immersive Displays

A. Vorozcovs

W. Stuerzlinger

A. Hogue

R. Allison

Department of Computer Science and Engineering and Center for Vision Research

York University, Toronto, Ontario, Canada.

August 2005

Abstract

Existing commercial technologies do not adequately meet the requirements for tracking in fully-enclosed Virtual Reality displays. We present a novel six degree of freedom tracking system, the *Hedgehog*, which overcomes several limitations inherent in existing sensors and tracking technology. The system reliably estimates the pose of the user's head with high resolution and low spatial distortion. Light emitted from an arrangement of lasers projects onto the display walls. An arrangement of cameras images the walls and the two-dimensional centroids of the projections are tracked to estimate the pose of the device. The system is able to handle ambiguous laser projection configurations, static and dynamic occlusions of the lasers, and incorporates an auto-calibration mechanism due to the use of the SCAAT (Single Constraint At A Time) algorithm. A prototype system was evaluated

relative to a state-of-the-art motion tracker and showed comparable positional accuracy (1-2 *mm* RMS) and significantly better absolute angular accuracy (0.1 *deg* RMS).

1 Introduction

Spatially immersive displays (SIDs) have recently become popular for scientific visualization, training, entertainment and Virtual Reality (VR) research (Brooks, 1999). These types of displays provide a way to fully immerse the user into the virtual world allowing them to be more accurate and productive at many tasks (Arms, Cook, & Cruz-Neira, 1999; Dam, Forsberg, Laidlaw, LaViola, & Simpson, 2000). In order to create a compelling visual world, the VR display must produce correct visual cues (perspective, parallax, stereo). These visual cues vary with the position and orientation (pose) of the user's head, hence in a SID it is necessary to have a system that can accurately and robustly track head pose.

In non-fully-enclosed displays, such as the original three-wall CAVETM (Cruz-Neira, Sandin, & DeFanti, 1993), existing commercial head tracking systems can be used since the tracking equipment can be positioned in such a way that it does not interfere with the user's view of the scene (i.e. behind the user). Such an approach is not possible in a fully-enclosed spatially immersive display. Tracking a user within fully-enclosed SIDs such as COSMOS (Yamada, Hirose, & Isda, 1998), HyPi-6 (Rotzer, 2001), PDC VR-CUBE (PDC, 1998), C6 (Graubard et al., 2002), ALICE (Francis, Goudeseune, Kaczmarski, Schaeffer, & Sullivan, 2003), and IVY (Robinson, Laurence, Zacher, et al., 2002) is a more complex task than within typical 1 to 5 wall displays. The user is confined in a fully-enclosed volume and there is no acceptable location for visible tracking equipment as

it interferes with the desired immersive experience, effectively *removing* the user from the virtual world.

[FIGURE 1 HERE]

1.1 Existing solutions for tracking in fully enclosed displays

Most spatially immersive display systems use fabric or glass projection screens that are optically and acoustically opaque or highly diffusing. In a fully enclosed display, the screens surround the user completely and therefore present a line-of-sight obstacle for optical and ultrasonic tracking systems. In large part due to this problem, the most popular solution for tracking existing fully-enclosed displays is wireless electromagnetic technology, such as Ascension Technologies' MotionStar Wireless® tracking system (Ascension Technology Corp., 2002). However, it is known that electromagnetic tracking systems behave poorly in the presence of metallic objects (Kindratenko, 2000, 2001) and that accuracy and signal strength degrade rapidly with distance from the base emitter. This precludes the use of this type of technology in IVY (the display we intend to use, Figure 1) which contains enough metallic material in its frame to render electromagnetic trackers useless.

A novel approach to overcoming the optical line of sight problem was taken by the designers of the *blue-c* immersive display at ETH Zurich (Spagno & Kunz, 2003). *blue-c* is a three-wall non-enclosed spatially immersive display that features walls made of large liquid crystal panels that can be switched between opaque and transparent modes at a high rate (≈ 50 Hz). This capability enables simultaneous projection and image capture using a number of projectors and cameras located outside the display. Due to its non-enclosed nature, *blue-c* currently uses a wired electromagnetic tracker for head tracking, but the

part-time transparency of the walls could potentially be used for optical head tracking.

The limitations of this (hypothetical) solution are the need for the walls to be opaque for a significant fraction of the time to maintain projection brightness and the limits on screen switching time (6 *ms*) that would limit the responsiveness of the tracker.

To deal with the line-of-sight problem, a compromise is to install the required stationary beacons or detectors in the seams of the display. For example, the *I-Space 5* system by Barco and A.R.T. GmbH uses infrared cameras mounted through C-mount sized holes in the corners of the display to track retro-reflective markers (A.R.T. GmbH, 2004). Another solution, recently installed at the six-sided *-C- Room* at Duke University (InterSense Corp., 2005), uses acoustic emitters in the seams between the side walls and the ceiling of the display to allow an acousto-inertial hybrid tracker to operate within the display. Both of these solutions result in a reduction in the perceived immersion because the tracking equipment is visible to the user.

The negative impact of this solution is threefold: pixel loss, distraction and creation of fixed spatial reference. The first problem, pixel loss, refers to the fact that the area covered by the tracking equipment lacks useful image information. This can be minimized by reducing the physical size of the components. The second problem applies to stereoscopic displays and manifests in the form of distracting double images of the components that appear to move when the user focuses on objects in front of or behind the screen. The impact of the third problem, fixed spatial reference, depends on the intended use of the display but can potentially be severe. For example, experiments in human perception designed to measure the interaction between the visual and vestibular systems, such as reported in (Jenkin, Dyde, Jenkin, & Harris, 2004), often critically depend on the ability of

the fully enclosed display to isolate the user from the real world and eliminate any spatial cues about their absolute orientation with respect to the environment. The stable visual cues provided by the stationary sensors in the seams of the display interfere significantly with the goals of the experiment. This expectation seems to be confirmed by recent work in perceptual stability in immersive VR (Tcheang, Gilson, & Glennerster, 2005). One of the results of (Tcheang et al., 2005) is that adding a stable visual reference to an immersive virtual environment modifies the perception of object motion and observer’s own stability.

1.2 Related work in optical tracking

Most optical tracking systems, such as the HiBall®(Welch et al., 2001) developed at UNC-Chapel Hill, require a line of sight between the fixed and mobile parts of the system. For off-the shelf optical trackers to work in a fully enclosed display, LEDs or other markers that they use would need to be placed inside the display, compromising the immersive experience.

The most effective method to date for fully enclosed SIDs is a hybrid optical-inertial tracking system previously developed by one of the co-authors of this paper and discussed in (Hogue, 2003; Hogue, Jenkin, & Allison, 2004). The system is comprised of an inertial tracker (InertiaCube2™, (InterSense Corp., 2000)) that provides the system with fast relative motion information and a secondary “outside-in” optical system that uses a set of cameras outside of the display viewing the screens. An arrangement of laser diodes in a known geometry is coupled to the inertial sensor and is attached to the user. The projections of the lasers are tracked within each image. The absolute pose is directly computed from the laser projections using geometric constraints. The pose estimates

between optical updates (15 Hz) are derived from the data provided by the inertial sub-system.

This optical approach requires that all four laser spots be imaged by the cameras in each frame. Since there are places where the lasers cannot be seen by the cameras (i.e. corners of the display), this is the major limitation of the system. When less than four lasers are visible, the pose cannot be computed and the head-tracker must be re-initialized.

In conclusion, although various commercial and research tracking technologies have been tried in spatially immersive displays, no single tracking device exists that constitutes a completely effective tracking technology for fully-enclosed SIDs.

1.3 Contributions

We introduce the **Hedgehog**, a tracking system that utilizes more laser diodes to overcome several limitations inherent in the approach of (Hogue et al., 2004). As in (Hogue et al., 2004), the lasers are tracked visually on each of the display walls. In order to disambiguate and reliably label the laser projections, the diodes' activation state can be changed periodically. This is synchronized with the image capture enabling the algorithm to determine exactly which laser produced the currently tracked laser projection in the image. By modeling the maximum angular velocity and acceleration that a user could potentially induce by moving their head, we are able to maximize the number of lasers that can be reliably and robustly identified. The information provided by all visible lasers contributes to the estimate through the use of a state estimation algorithm. By fully utilizing the Single Constraint At A Time (SCAAT) framework (Welch, 1996), the camera parameters required for the laser projection are optimized online increasing the accuracy of

the tracking system and reducing set-up effort.

[FIGURE 2 HERE]

2 The Hedgehog tracking approach

The basic approach for tracking within a fully-enclosed display (see Figure 2 for an illustration and (Hogue, 2003; Vorozcovs, 2005) for more information) relies on the fact that some of the display surfaces lie outside the view of the user to estimate and track head pose within the environment. A fixed arrangement of low power laser diodes is attached to the back of a helmet worn by the user. The lasers' projections are visible from outside the display, allowing indirect observation of the user's motion. Cameras are positioned outside the SID viewing each projection surface. The projections of the laser beams are tracked by these cameras as they strike the projective surfaces. Using the 2D image measurements of each laser projection we apply each measurement as a constraint that contributes to the estimate of the pose of the device. The redundancy of the lasers allows the Hedgehog to track anywhere within the display. The redundancy of the lasers also allows the system to achieve higher measurement accuracy, since each laser contributes to the overall pose estimate through the use of a state estimation scheme.

2.1 Choice of laser configuration

Various configurations of laser diodes could be used to localize the user. In the previous approach (Hogue, 2003), an arrangement of four orthogonal laser diodes was used. This provides a simple technique for computing the pose of the device, but requires that all laser projections be visible by the camera system at each update. Thus, the system cannot

handle occlusions (when lasers shine into corners or are occluded by an object within the physical space, e.g. the arm of a user).

When more than four lasers are used, it becomes difficult to identify the lasers based on geometrical constraints alone. When active modulation of the lasers is available, the choice of laser configuration is considerably simpler, since no heuristic for identifying the lasers based on geometrical constraints needs to be developed. A configuration based on uniform angular spacing is particularly simple to construct, provides uniform accuracy for all orientations and is easy to characterize due to symmetry.

2.1.1 Choosing the number of lasers

Once the tracker is initialized by turning on each laser in sequence, more than one laser can be used simultaneously. The more lasers are active during regular tracking, the better the tracking accuracy. However, there is a tradeoff between the maximum angular speed supported by the tracker and the number of simultaneously active lasers, because of the need to keep track of laser identities. For a head tracker, peak angular velocities of 1000 *deg/s* need to be supported (Foxlin, 2002). This places a constraint on the minimum angular spacing of the lasers relative to the camera frame rate, because fast rotations of the head can result in some lasers turning into the position of other lasers between two camera exposures. For our 30 Hz camera system and assuming 45 degree spacing of the lasers, the rotational velocity resulting in a laser mismatch between two frames is 675 *deg/s*. Hence, frame rates higher than 30 Hz are needed to support angular rates of 1000 *deg/s*. Further constraints can be introduced by considering the maximum possible acceleration of the head.

2.2 Sources of error

Similarly to most other optical systems, the *resolution* of the Hedgehog tracker scales with the resolution of the cameras used for detecting the laser dots. The *accuracy* depends on how well the cameras are calibrated.

2.2.1 Systematic error

Systematic error can be introduced into the measurements if the cameras are poorly registered to real-world coordinates. This error can be minimized by using auto-calibration, as described below in section 2.10.1.

2.2.2 Noise and Jitter

Two major factors contributing to noise in the pose estimate are image noise in the cameras and wall motion caused by the bending of the flexible screen material (if flexible screens are used, as is the case for IVY). To determine the magnitude and nature of these noise sources, we measured the covariance of the noise in laser dot locations by collecting statistics from the cameras while the tracker was stationary. The standard deviation of the noise in the dot locations was found to be within 0.5 camera pixels, which corresponds to approximately 1 *mm* on the wall. The resulting noise is well described by a zero-mean normal distribution and therefore its effects can be minimized by adjusting the measurement covariance parameter of the Kalman filter (see section 2.6).

[FIGURE 3 HERE]

2.2.3 Uniformity of localization

An important characteristic of any tracking system is how its localization accuracy varies with spatial position. For example, electromagnetic systems are found to have cubic or even quartic falloff with distance from the emitter (Kindratenko, 2000). To determine this parameter for the Hedgehog system, we performed stochastic simulations of the accuracy of the Hedgehog for the case of no systematic error (i.e. ideally calibrated cameras). The simulation uses the same tracking code as the real-life Hedgehog and adds Gaussian noise to laser dot locations.

Figure 3 shows a contour plot of the expected error in translation vs. the location in the display for a 2x2 meter region in the middle of the display, with $\sigma^2 = 1 \text{ mm}$ variance in laser dot localization noise. As expected, localization error is smallest in the middle of the display ($\sigma^2 = 1.4 \text{ mm}$) and increases by about 25% at a distance of 1 m from the middle of the display ($\sigma^2 = 1.8 \text{ mm}$).

Note that the figure is not fully symmetric because the sliding entrance wall of the display was simulated to be in the open position (slid 70 cm away from the display, as can be seen in the 3D rendering in Figure 4). The opening created by the shifted wall, located to the right in Figure 3, creates a local maximum at $Z = +0.3 \text{ m}$ that separates the central region of the display into two local minima (at $Z = -0.4 \text{ m}$ and $Z = +0.8 \text{ m}$).

Other simulations revealed that at least 9 active lasers are required to estimate the pose of the tracking device with better than 1 mm RMS accuracy with our current camera system (Vorozcovs, 2005).

2.3 Tracker Initialization

When more than four lasers are used, it becomes difficult to disambiguate the lasers based on geometrical constraints alone. In that case, active multiplexing is a simple and effective solution to the problem of identifying the lasers. For initialization (“cold start”) of the tracker, the lasers are turned on one at a time to determine their identities. To determine the initial pose of the tracking sensor, we use an adaptation of a linear least-squares exterior orientation algorithm, as presented in (Quan & Lan, 1999), see (Vorzcovs, 2005). This algorithm is, in general, also suitable for continuous tracking of the Hedgehog, but its numerical stability and computational complexity are significantly worse than the Kalman filter algorithm presented in section 2.6. Thus, it is only used for initialization.

2.4 Point tracking based on temporal coherence

After the pose of the tracker is determined using the linear least squares algorithm, laser identification can be performed using a combination of temporal coherence and predictive tracking. If the number of lasers is chosen appropriately (see discussion in section 2.1), all lasers can remain active during normal operation to obtain maximum accuracy. Note that tracking lasers for the Hedgehog based on temporal coherence is a simpler task than tracking a general set of optical markers in 3D, since the paths of the laser dots on the (convex) surface of the display do not intersect and there are minimum spacing guarantees due to the physical size of the device.

2.5 Occlusion handling and eye safety considerations

Both static occlusions (corners of the display, or improper camera placement) and dynamic occlusions (user raising their arm in front of one or several laser diodes) potentially occur in the system. Static occlusions are hard-coded in the description of the environment and can be used to differentiate between anticipated and unanticipated occlusions.

Unanticipated occlusions are most likely to occur if the user lifts an arm or if there are other obstructions, such as non-tracked observers inside the SID.

Even though the lasers used in the Hedgehog tracker are nominally eye-safe, we reduce the chance of uncomfortable laser-eye encounters using an approach based on a technique for a handheld laser pointer input system, see (Oh & Stuerzlinger, 2002). The lasers that are occluded are time-multiplexed at 1/10th of nominal power. If the laser projection is found again, it signals that the obstruction is gone and the laser can be turned back on continuously.

2.6 State estimation

State estimation for six degree-of-freedom tracking is typically performed using a recursive filtering scheme such as the *Kalman filter* (Kalman, 1960). A Kalman filter is a maximum likelihood stochastic estimator when the process and measurement variables behave as Gaussian Random Variables (GRVs). Even though the motion parameters for head-tracking are not always described by a GRV, Kalman filters are widely used for motion tracking. For more information on the Kalman filter, see (Welch & Bishop, 1995).

2.7 The Single Constraint At A Time approach to state estimation

The SCAAT (Welch, 1996) approach to state and parameter estimation describes using measurements from a *locally unobservable* system to estimate a *globally observable* system. The main observation is that each measurement contains *some* information about the state of the system. In the case of the Hedgehog tracker, each 2D measurement (the coordinate of a laser dot on a display wall) encodes some information about a 6D variable (the pose of the tracker). Using a single measurement at a time approach improves the robustness of the system, increases its temporal responsiveness and eliminates the need for explicit occlusion handling. Only the lasers that are seen in a particular frame are used for updating the estimate.

SCAAT can be used in various types of state estimation filtering schemes, however in this paper we chose to employ the SCAAT algorithm in the form of an Extended Kalman Filter (Welch & Bishop, 1995; Gelb, 1974). The Hedgehog Kalman filter closely resembles the generic vision tracker presented in (Welch, 1996). The estimated state is maintained as a vector that combines a position vector, velocity vector, incremental orientation angles, and their derivatives. Due to the presence of angular quantities, this system model is non-linear. However, it can be linearized by subtracting the absolute orientation and storing it in an external quaternion, as described in (Welch, 1996). Then, only frame-to-frame changes in orientation angles are manipulated by the Kalman filter.

2.8 Data representation

The orientation of the laser diodes used in the tracker is parameterized using unit direction vectors with the origin of the vectors located at the currently estimated position of the

Hedgehog tracker. The measurements (laser dot positions on the walls) are expressed in physical units in the plane of each screen surface relative to one of its corners. To increase numerical stability, all calculations are performed with double precision.

2.9 The Hedgehog Kalman update cycle

Our process model predicts the estimated position and orientation of the tracker using a standard discrete time position-velocity (PV) model. The measurement function has as its input the predicted state, the source parameters (laser direction vector) and the sensor parameters (camera and wall calibration). This function computes what the measurement should be, i.e. projects the state onto the source/sensor pair, predicting the noise-free response of each sensor and source pair given the systems' current state. To achieve this, the measurement function transforms the laser direction vector using the last predicted position and orientation of the tracker and then intersects the resulting beam with the plane of the closest display wall using Plücker matrices (Hartley & Zisserman, 2000). This produces a 2D point in the wall coordinate frame, which is the predicted measurement.

The Jacobian matrix required for transforming the covariance of the estimate is computed numerically by perturbing the state elements by a small ϵ and producing multiple predicted observations. Given the observation predicted from the unperturbed state the residual is computed and the Jacobian is constructed.

The next step in the measurement update is to compute the Kalman gain. The residual error between the predicted observation and the actual measurement from the sensor is used to calculate the Kalman gain matrix. The predicted state and error covariance are then corrected using the calculated Kalman gain matrix.

The final step in the measurement update is to update the orientation quaternion with the incremental orientation angles estimated in the state and then zero the incremental orientation angles. This is a way to maintain local linearity of the system with respect to orientation (see (Welch, 1996)).

2.9.1 Limitations of the Extended Kalman Filter

If the camera frame rate is low relative to the amount of rotational motion, the linearization process as described above might lead to large estimation errors. To combat this, the residuals between the predicted and actual measurements are monitored and the filter is iterated until the average residual error from all available observations falls below a particular empirically determined threshold (i.e. the target accuracy). Observations of the estimation algorithm showed that only two or three iterations are typically required to reach a RMS residual error of 2 mm even if the intra-frame angular motion is on the order of 10 degrees.

2.10 Camera and wall calibration

Prior to operation, the cameras observing the walls are intrinsically calibrated to remove the effects of radial distortion and a homography (projective planar transformation) per camera is computed which transforms image coordinates into the wall coordinate frame. This offline calibration is performed manually by measuring stationary laser points and recording their sub-pixel image coordinates. By measuring at least 4 points, a homography transformation can be computed (Hartley & Zisserman, 2000) transforming the 2D image coordinates into a 2D point that lies in the wall plane. The location of each screen in space

is also physically measured (the wall transform). Since the screens are assumed to be planar, a homography and a wall transformation are sufficient to determine the 3D locations of the laser dots.

2.10.1 Autocalibration

Manual camera calibration provides a good initial estimate of the camera to wall plane transformation. After applying the homography and the wall transformation, we obtain 3D points that lie close to the real wall plane. However, manual measurements are prone to error and the points might not be exactly at the appropriate locations on the real wall. To further increase estimation accuracy, we employ *autocalibration*, a simple and effective approach described in (Welch, 1996). Autocalibration uses the ability of the Kalman filter to estimate both state variables and system parameters simultaneously. To perform autocalibration, the estimated state vector is augmented with the parameters of the source and sensor for each pair of observations, effectively creating a Kalman filter per source/sensor pair. In our system, due to the rigidity of the head-worn device, the source parameters (i.e. the laser direction vectors) do not need to be auto-calibrated and can be simply measured with sufficiently high accuracy. We employ auto-calibration only to compute the parameters of the wall, i.e. the wall normal and translation to the 3D display coordinate frame. Thus, only 6 parameters need to be auto-calibrated.

The vector of wall calibration parameters, is added to state vector to create an augmented state vector. Similarly, the covariance matrix, state transition matrix and process noise matrix are augmented with the appropriate parameters. The amount of process noise specified for the parameters being autocalibrated is set to an empirically

determined constant low value. This ensures that the autocalibrated parameters change very slowly and the noise in the incoming measurements does not corrupt the highly accurate calibration values. The rest of the algorithm proceeds normally using the augmented versions of the matrices, i.e. the Jacobian is numerically computed in the same manner but for the augmented measurement vector. At the end of the measurement update, we collapse the augmented state vector by extracting the corrected sensor parameters and transferring them into the sensor models, discarding the covariance information. Then, the filter is ready for the next iteration. This approach does not maintain the full covariance information between the states of the system, due to the ad-hoc pairing of sources (lasers) with sensors (cameras). However, it works well in practice, reducing the need for a laborious manual calibration.

[FIGURE 4 HERE]

2.11 Monitor Algorithm

To oversee the operation of the tracker, a *monitor* algorithm is used. The monitor checks the residuals between the predicted points and the observed points. The average and maximum values of the residuals are used by an heuristic algorithm to decide if the estimate has become invalid. If the residual error is large or not enough lasers are being updated, the estimate is likely incorrect and the system reinitializes the tracker.

For debugging and demonstration purposes, a graphical depiction of tracker operation is generated on the fly using real-time computer graphics techniques (OpenGL). Figure 4 shows one of the diagnostic screens of the software. The dots denote laser hits with identity information indicated by using the first 17 letters of the alphabet. Search regions for

frame-to-frame temporal matching of points are shown as circles. The radius of each circle is equal to the search distance value used by the temporal coherence tracking algorithm and is derived using the maximum head velocity assumption (see section 2.1.1).

[FIGURE 5 HERE]

3 Hardware Implementation Details

Our implementation of the Hedgehog hardware consists of 17 laser diodes arranged in a spherically symmetric manner with 45 degree spacing (see Figure 5). The prototype is 9 cm in diameter, 6 cm in height and weighs approximately 200 grams. Each laser diode is a visible red laser diode in the 645nm wavelength range and is individually controlled by a PIC microcontroller through a serial interface. The lasers are embedded in a custom housing machined from a single piece of Delrin® plastic. The orientation of the diodes was measured to within an accuracy of 0.2 degrees. A low-bandwidth wireless serial link (9600 bps at 433 MHz with FM modulation) allows the tracker to operate without a tether.

The display, IVY at York University (Robinson, Laurence, Hogue, et al., 2002), is equipped with 8 digital IEEE-1394 cameras connected to a single-CPU Linux PC through 3 IEEE-1394 buses. Two cameras are required for both the floor and ceiling due to space restrictions in the design of the display. Each camera provides gray-scale images at 640x480 resolution at 30 fps. We set the exposure time of the cameras to 5 ms. This ensures that the laser spot is the only bright point in the image and is thus easily identified by thresholding, even in the presence of motion blur. We verified that the point detection algorithm works reliably at angular velocities over 1000 *deg/s*.

4 Experimental Results

In order to determine the accuracy of the Hedgehog tracking system, we performed several comparisons of the tracker with ground truth and the IS-900 tracker from InterSenseTM (Foxlin, Harrington, & Pfeifer, 1998). The IS-900 provides a way to determine Hedgehog's performance relative to a state-of-the-art commercial system. The IS-900 has a nominal positional resolution of 0.75 *mm* and an angular resolution of 0.05 degrees. The static and dynamic accuracy of the IS-900 has been independently compared to other kinds of trackers and was found to be better than magnetic trackers (Kindratenko, 2001) and worse than optical methods (Gilson, Fitzgibbon, & Glennerster, 2003, 2005). Therefore, comparing the Hedgehog with the IS-900 is a reasonable choice.

To use the IS-900 in IVY, we placed the IS-900 base transmitter (with two SoniStripsTM) on a rigid frame standing vertically at the entrance of IVY (aligned with the X-axis of the display) and one of its sensors was rigidly attached to the Hedgehog tracking sensor. The measurements reported by both systems were appropriately transformed into display coordinates (relative to the center of IVY). In order to accommodate the IS-900, the rear wall of the display was left open approximately 70cm and the camera parameters for this wall were adjusted accordingly (this can be seen in the simulator screenshot, Figure 4).

[FIGURE 6 HERE]

4.1 Positional Experiments

Several straight lines were defined on the floor of IVY using stretched string and the trackers were placed on moveable posts of various heights. We initialized both tracking systems and recorded data from both the Hedgehog tracker and the IS-900 while moving

the connected trackers along the straight lines defined by the strings. The results for a test run covering the perimeter of the display are shown in Figure 6.

A quantitative estimate of tracking accuracy can be obtained by fitting straight lines to the reported position data and finding the RMS (Root Mean Square) deviation from each line. The RMS error for the Hedgehog tracker was found to be 1.6 *mm* versus 1.4 *mm* for the IS-900. Other tests performed closer to the center of the display showed similar results, confirming the simulation results described in section 2.2.3.

Examination of the raw pose data for a stationary tracker with 17 lasers (see section 4) shows that the jitter in the raw data is 0.2 *mm* (RMS) for position and 0.01 *deg* (RMS) for orientation. These values are consistent with the 1 *mm* laser dot localization error, because of the effect of averaging 17 lasers.

[FIGURE 7 HERE]

4.1.1 Dynamic behavior and latency

Figure 7 shows a comparison of the dynamic motion behavior of the two trackers. This plot allows us to estimate the latency of the Hedgehog tracker relative to the IS-900 by finding the best temporal alignment of the two motion curves. In this case, the best fit value was close to 100 *ms*. Since the specified latency of the IS-900 tracker is 4 *ms*, we can assume a rough estimate of 100 *ms* for the latency of this particular implementation of the Hedgehog system.

4.2 Angular Experiments

[FIGURE 8 HERE]

4.2.1 Angular linearity

The Hedgehog tracker was placed on a rotational stage with a Vernier scale permitting angular measurements accurate to 0.5 degrees. We initialized the tracking system and recorded data for every 0.5 degrees over an interval of 7 degrees on the rotational stage in the first experiment and every 5 degrees over an interval of 40 degrees in a second experiment. The combined results of these measurements are shown in Figure 8. The deviation from linearity was 0.08 degrees (RMS) for the 7 degree interval and 0.20 degrees (RMS) for the 40 degree interval. The increased error for the larger interval can be explained by the variation in calibration accuracy of the different cameras. Since the Hedgehog tracker is symmetric with respect to 45 degree rotations, the results for other orientations are similar and are not given.

[FIGURE 9 HERE]

4.2.2 Angular repeatability

A dynamic angular repeatability test was performed with both trackers. In this experiment, the rigidly connected trackers were rotated back and forth several times through an angle of 10 degrees using the rotation stage and then returned to the original orientation. The resulting plot of the reported angles (Figure 9) shows that after the rotation stage was brought back to the original orientation, the Hedgehog estimate returned to within 0.4 degrees of the start location, but the acousto-inertial estimate has drifted by more than 3 degrees. This makes it impossible to directly determine the angular accuracy of the IS-900 tracker. In particular, a plot similar to Figure 8 cannot be easily constructed for the IS-900 because of the drift in reported positions for this experiment.

Since the IS-900 sensor was well within its operating range (distance to the base station was 1.8 meters, with the quoted operating range of 3 meters) and the orientation of the ultrasonic sensor was nearly optimal (parallel to the base station), these motion artifacts can be potentially attributed to imperfect sensor fusion. An evaluation of the IS-900 using a vision-based system (Gilson et al., 2003) also found motion artifacts with the IS-900.

4.3 Summary of the experiments

Using the data presented above we can conclude that the positional accuracy of the Hedgehog tracking system is comparable to the IS-900 tracker but the angular accuracy is much higher, especially if the tracker is rotated in small increments. The angular repeatability test revealed a limitation of the acousto-inertial tracker with respect to small motions and reinforced one of the original justifications for developing the Hedgehog system: perception experiments performed in IVY, such as described in (Jenkin et al., 2004) require subtle rotations and translations of the head to be tracked in a reliable way. The Hedgehog tracker was designed to provide a reliable spatial reference for these kinds of experiments.

5 Summary

In this paper we introduced the Hedgehog optical tracking system for fully enclosed display environments. Our approach to tracking in enclosed displays has many advantages over existing technologies:

1. The geometry of the Hedgehog system provides highly accurate tracking in fully

enclosed displays.

2. The user is untethered and there is no interference with the metallic frame of the display.
3. The sources of noise and systematic error are few and well-defined.
4. The system is robust to multiple occlusions with only a minor loss of accuracy.
5. The hardware is constructed from inexpensive and readily available materials.
6. The system is adaptable to different display types made from planar surfaces (polygonal, single-wall, etc.).
7. The system performs automatic adjustment of the calibration parameters increasing accuracy and reducing set-up effort.

A prototype of the Hedgehog tracker was also shown to compare favorably with a state-of-the-art commercial tracking system.

5.1 Possible improvements and future work

Inexpensive and fast ($120Hz$) dedicated point tracking cameras have recently been made commercially available (Naturalpoint Inc., 2005). Using this kind of camera with the Hedgehog system would significantly lower tracking latency, eliminate the need for image processing on the CPU and further reduce the cost of the system. On the other hand, if very high accuracy is desired, linear CCDs could be used instead of video cameras to increase both the measurement rate and accuracy, albeit at a higher cost.

If a further reduction in the average power of the lasers is desired, the lasers can be modulated *within* a single camera frame and only turned on for the exposure time of the cameras. This approach is shown to work in (Pavlovych & Stuerzlinger, 2004) and can reduce the duty cycle of the lasers by a factor of 6 (5 ms / 33 ms).

By employing a network of cameras, a wide-area tracking system could be built using the Hedgehog principle, similar to the HiBall (Welch et al., 2001), Constellation (Foxlin et al., 1998) and Vistracker (Foxlin & Naimark, 2003) wide-area trackers. An application that would benefit from the high angular accuracy of the Hedgehog system is tracking of Head-Mounted Displays for Mixed and Augmented Reality.

Acknowledgments: We would like to acknowledge Jeff Laurence from the Center for Vision Research and Andriy Pavlovich for helping us build the hardware and Dr. Michael Jenkin for guidance and financial support. The financial support of NSERC Canada and the Canadian Foundation for Innovation (CFI) is gratefully acknowledged.

References

- Arms, L., Cook, D., & Cruz-Neira, C. (1999). The benefits of statistical visualization in an immersive environment. In *Virtual Reality, 1999. Proceedings* (p. 88-95). IEEE.
- A.R.T. GmbH. (2004). *The ARTTrack 2/C camera system*. Retrieved from <http://www.ar-tracking.de/view.php?page=97>, on June 25, 2005.
- Ascension Technology Corp. (2002). *MotionStar Wireless 2*. Retrieved from <http://www.ascension-tech.com/products/motionstarwireless.php>, July 2005.
- Brooks, F. J. (1999). What's real about virtual reality? In *Computer Graphics and Applications* (Vol. 19, p. 16-27). IEEE.

- Cruz-Neira, C., Sandin, D., & DeFanti, T. (1993). Surround-screen projection based virtual reality: The design and implementation of the CAVE. In *Proc. SIGGRAPH '93* (p. 135-142). New York: ACM Press.
- Dam, A. van, Forsberg, A., Laidlaw, D., LaViola, J. J., & Simpson, R. (2000). Immersive VR for scientific visualization: a progress report. In *Computer Graphics and Applications* (Vol. 20, p. 26-52). IEEE.
- Foxlin, E. (2002). Motion tracking requirements and technologies. In *Handbook of virtual environment technology* (chap. 7). Lawrence Erlbaum Associates. (ed. Stanney, K. M. Chapter 8)
- Foxlin, E., Harrington, M., & Pfeifer, G. (1998). Constellation: a wide-range wireless motion-tracking system for Augmented Reality and virtual set applications. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (pp. 371-378). New York, NY, USA: ACM Press.
- Foxlin, E., & Naimark, L. (2003). VIS-Tracker: A Wearable Vision-Inertial Self-Tracker. In *Proc. of IEEE Virtual Reality 2003 (VR2003)*. (March 22-26, 2003, Los Angeles, CA)
- Francis, G., Goudeseune, C., Kaczmarek, H., Schaeffer, B., & Sullivan, M. (2003). ALICE on the Eightfold Way: Exploring Curved Spaces in an Enclosed Virtual Reality Theatre. In H.-C. Hege & K. Polthier (Eds.), *Visualization and Mathematics III*. Springer.
- Gelb, A. (1974). *Applied optimal estimation*. MIT Press, Cambridge, MA.
- Gilson, S. J., Fitzgibbon, A. W., & Glennerster, A. (2003). Dynamic performance of a tracking system used for Virtual Reality displays.
- Gilson, S. J., Fitzgibbon, A. W., & Glennerster, A. (2005). Quantitative analysis of accuracy of an inertial/acoustic 6DOF tracking system.

- Graubard, B., Chen, F., Min, Z., Lwakabamba, B., Weber, R., Rover, D., et al. (2002).
Lessons Learned: Installing a Wireless System in the C6 Virtual Reality Environment.
In *Proc. SIGGRAPH 2002*.
- Hartley, R., & Zisserman, A. (2000). *Multiple view geometry*. Cambridge University Press;
ISBN:0521623049.
- Hogue, A. (2003). *MARVIN: a Mobile Automatic Realtime Visual and INertial tracking system*. Unpublished master's thesis, York University, Toronto, Canada.
- Hogue, A., Jenkin, M. R., & Allison, R. S. (2004). An optical-inertial tracking system for fully-enclosed VR displays. In *Proc. of IEEE 1st Canadian Conference on Computer and Robot Vision (CRV'2004)*. (May 17-19, London, Ontario, Canada)
- InterSense Corp. (2000). *InertiaCube2*. Retrieved from <http://www.isense.com/products/prec/ic2/index.htm>, July 2005.
- InterSense Corp. (2005). InterSense installs Wireless IS-900 Tracking in Dukes six-sided immersive display. Retrieved from <http://www.intersense.com/news/InterSenseTrackerIssue7.pdf>, July 2005.
- Jenkin, H. L., Dyde, R. T., Jenkin, M. R., & Harris, L. R. (2004). Pitching up in IVY. In *Proc. ICAT 2004, Korea*.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D), 35-45.
- Kindratenko, V. (2000). A survey of electromagnetic position tracker calibration techniques. In *Virtual reality: Research, development, and applications*. (vol.5, no.3, pp. 169-182)
- Kindratenko, V. (2001). A comparison of the accuracy of an electromagnetic and hybrid ultrasound-inertia position tracking system. *Presence: Teleoperators and Virtual Envi-*

ronments, 10(6), 657-663.

- Naturalpoint Inc. (2005). *OptiTrack Camera System*. Retrieved from <http://www.naturalpoint.com/optitrack/>, July 2005.
- Oh, J.-Y., & Stuerzlinger, W. (2002). Laser pointers as collaborative pointing devices. In *Graphics interface*.
- Pavlovyh, A., & Stuerzlinger, W. (2004). Laser Pointers as Interaction Devices for Collaborative Pervasive Computing. In *Advances in pervasive computing* (p. 315-320). OCG. (ISBN 3-85403-176-9)
- PDC. (1998). *Primeur: Advancing European Technology Frontiers, World's first fully immersive VR-CUBE installed at PDC in Sweden, 1998*. (Center for Parallel Computing)
- Quan, L., & Lan, Z. (1999, Aug). Linear n-point camera pose determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21, 774-780.
- Robinson, M., Laurence, J., Hogue, A., Zacher, J., German, A., & Jenkin, M. (2002). IVY: Basic design and construction details. In *Proc. ICAT*. Tokyo, Japan.
- Robinson, M., Laurence, J., Zacher, J., Hogue, A., Allison, R., Harris, L. R., et al. (2002). IVY: The Immersive Visual environment at York. In *6th International Immersive Projection Technology Symposium, March 24-25, 2002, Orlando, Fl.* .
- Rotzer, I. (2001). *Synthetic worlds within six walls*. (Fraunhofer Magazine, 2:2001)
- Spagno, C. P., & Kunz, A. M. (2003). Construction of a Three-sided Immersive Telecollaboration System. In *Proc. of IEEE Virtual Reality 2003 (VR2003)* (p. 37-44). (March 22-26, 2003, Los Angeles, CA)
- Tcheang, L., Gilson, S. J., & Glennerster, A. (2005, July). Systematic distortions of perceptual stability investigated using immersive virtual reality. *45(16)*, 2177-2189.

- Vorozcovs, A. (2005). *The Hedgehog: A Novel Optical Tracking System for Fully Enclosed Spatially Immersive Displays*. Unpublished master's thesis, York University, Toronto, Canada.
- Welch, G. (1996). *SCAAT: Incremental Tracking with Incomplete Information*. Unpublished doctoral dissertation, Department of Computer Science, UNC-Chapel Hill, Chapel Hill, NC.
- Welch, G., & Bishop, G. (1995). *An introduction to the Kalman Filter* (Tech. Rep. No. TR95-041). University of North Carolina at Chapel Hill.
- Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K., & Colucci, D. (2001). High-Performance Wide-Area Optical Tracking — The HiBall Tracking System. In *Presence: Teleoperators and virtual environments*. (Vol 10, No. 1, Pp. 1-21)
- Yamada, T., Hirose, M., & Isda, Y. (1998). Development of a complete immersive display: COSMOS. In *Proc. VSMM'98* (p. 522-527).

List of Figures

1	IVY: The Immersive Visual environment at York. IVY is shown here with the rear (entry) wall removed in order to show the structure of the device more clearly.	30
2	Optical Tracking Approach. The user wears many low power laser diodes whose projections on each screen surface are tracked via cameras outside of the display. The head pose is determined from these projections alone. . . .	31
3	Stochastic simulation results for localization error vs. position in the display. The sliding wall is located to the right of the figure.	32
4	The Hedgehog software simulation environment.	33
5	The Hedgehog hardware. 17 laser diodes are arranged in a symmetrical hemispherical arrangement. As shown in (a) the IS-900 tracker is rigidly attached for experimental purposes.	34
6	Positional Linearity Experiment. Data for both trackers is shown.	35
7	Dynamic behavior of the Hedgehog vs the IS-900. Hedgehog data is advanced by ≈ 100 ms, a value that can be used as an estimate of relative latency between the trackers.	36
8	Angular Measurements of the Hedgehog	37
9	Comparison of the dynamic behavior of the two trackers for small angular displacements. Dashed lines represent the initial position (at approximately 5 deg) and the extremes of the rotational motion.	38

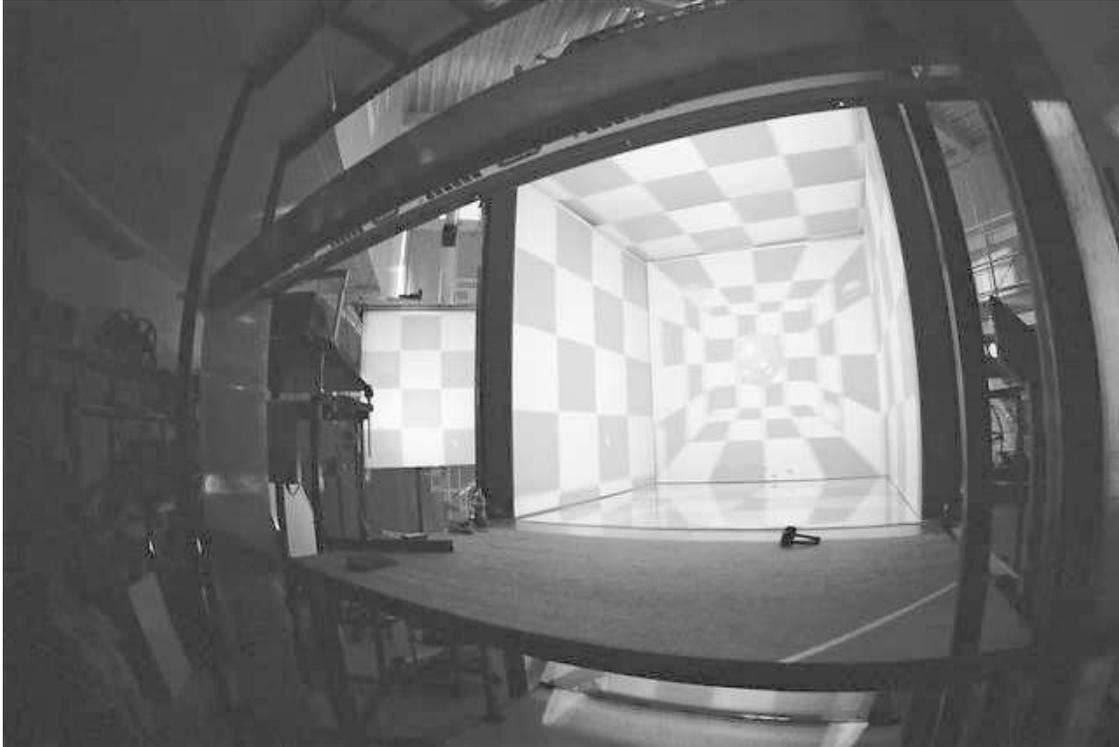


Figure 1: IVY: The Immersive Visual environment at York. IVY is shown here with the rear (entry) wall removed in order to show the structure of the device more clearly.

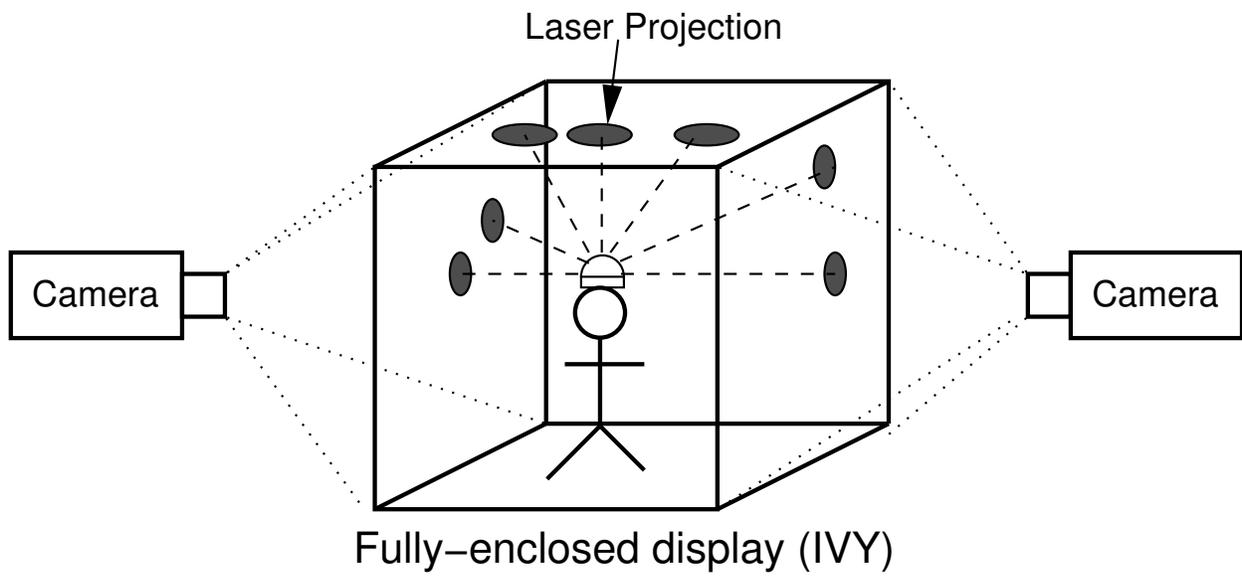


Figure 2: Optical Tracking Approach. The user wears many low power laser diodes whose projections on each screen surface are tracked via cameras outside of the display. The head pose is determined from these projections alone.

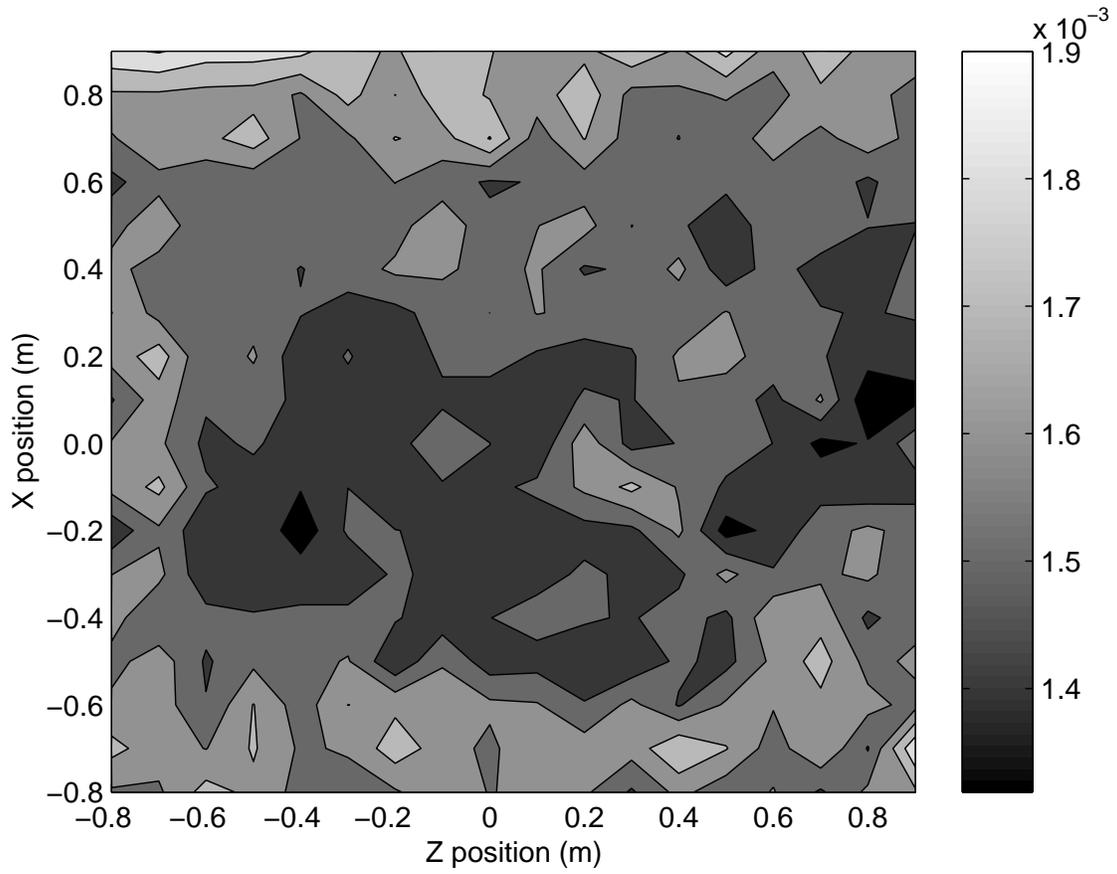


Figure 3: Stochastic simulation results for localization error vs. position in the display. The sliding wall is located to the right of the figure.

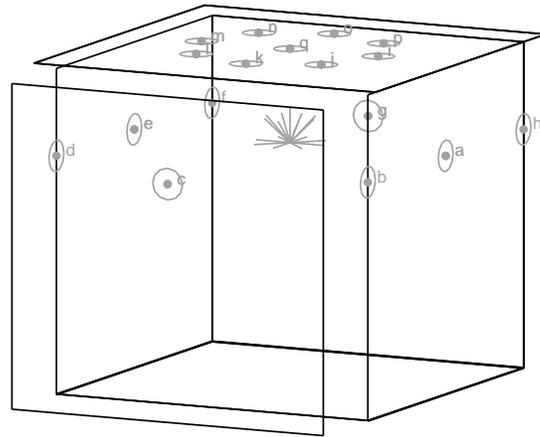
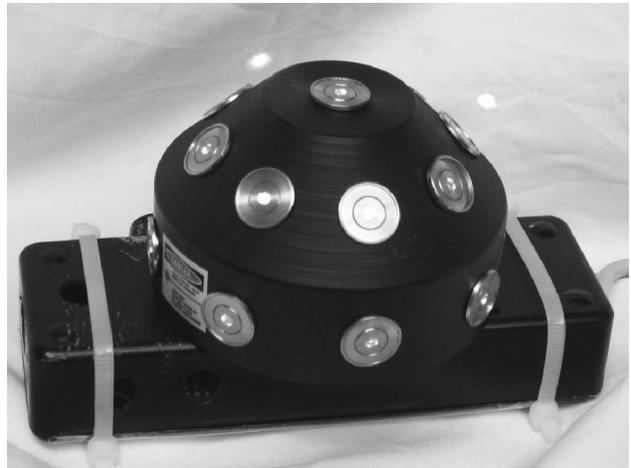


Figure 4: The Hedgehog software simulation environment.



(a)



(b)

Figure 5: The Hedgehog hardware. 17 laser diodes are arranged in a symmetrical hemispherical arrangement. As shown in (a) the IS-900 tracker is rigidly attached for experimental purposes.

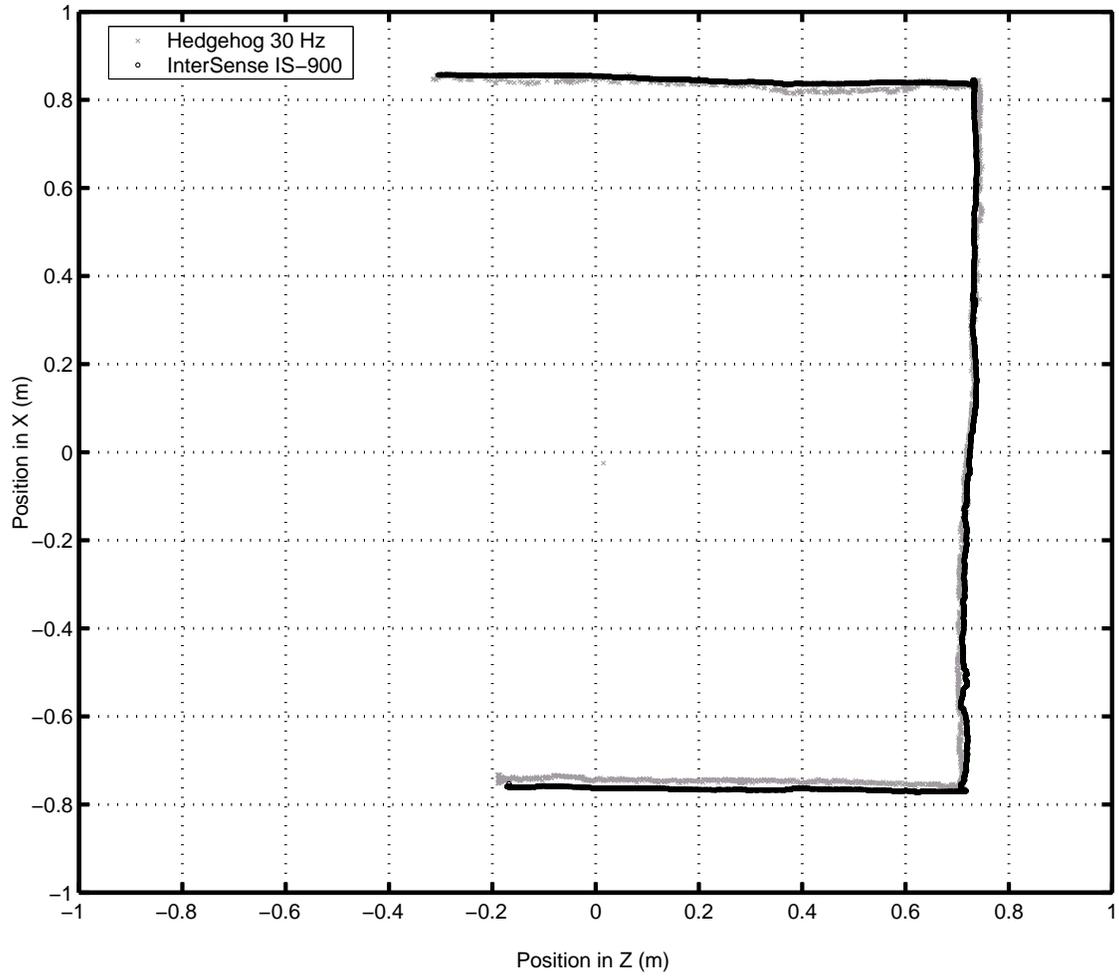


Figure 6: Positional Linearity Experiment. Data for both trackers is shown.

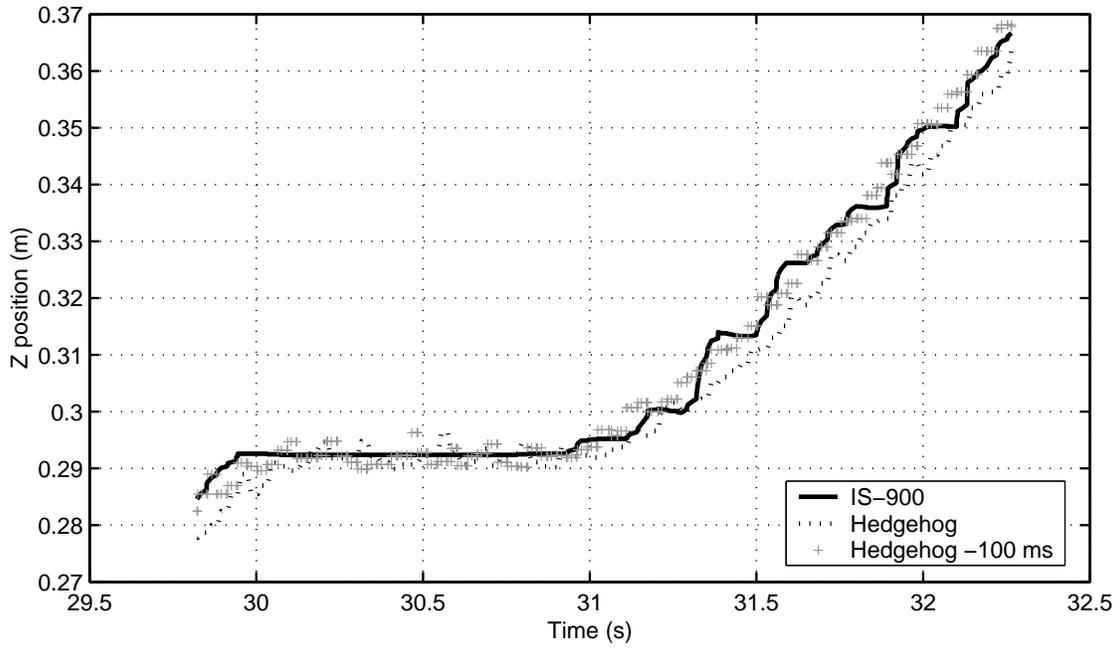


Figure 7: Dynamic behavior of the Hedgehog vs the IS-900. Hedgehog data is advanced by ≈ 100 ms, a value that can be used as an estimate of relative latency between the trackers.

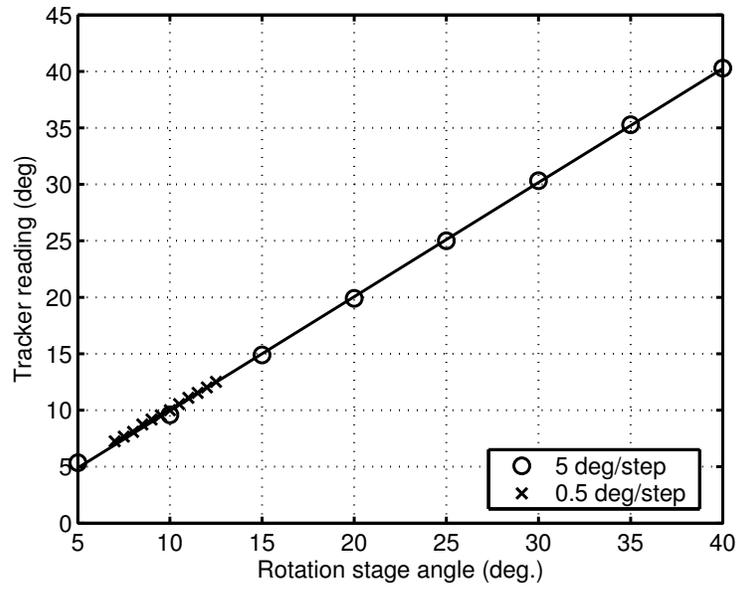


Figure 8: Angular Measurements of the Hedgehog

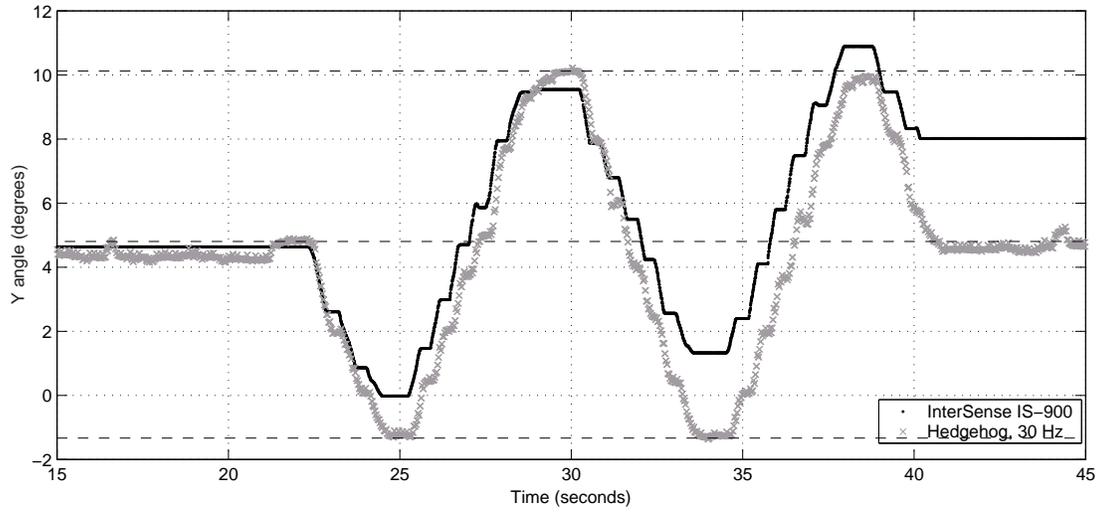


Figure 9: Comparison of the dynamic behavior of the two trackers for small angular displacements. Dashed lines represent the initial position (at approximately 5 *deg*) and the extremes of the rotational motion.