

# Performance Analysis of an Outdoor Augmented Reality Tracking System that Relies Upon a Few Mobile Beacons

Ronald Azuma, Howard Neely III, Mike Daily\*  
HRL Laboratories, LLC

Jon Leonard†  
Raytheon Missile Systems

## ABSTRACT

We describe and evaluate a new tracking concept for outdoor Augmented Reality. A few *mobile* beacons added to the environment correct errors in head-worn inertial and GPS sensors. We evaluate the accuracy through detailed simulation of many error sources. The most important parameters are the errors in measuring the beacon and user's head positions, and the geometric configuration of the beacons around the point to augment. Using Monte Carlo simulations, we identify combinations of beacon configurations and error parameters that meet a specified goal of 1m net error at 100m range.

**CR Categories:** H.5.1 [Multimedia Information Systems]: Artificial, augmented and virtual realities. C.4 [Performance of System]: Performance attributes. J.7 [Computers in Other Systems]: Military

**Keywords:** Monte Carlo analysis, simulation, UAV

## 1 MOTIVATION

Soldiers, operating outdoors, could benefit from Augmented Reality (AR) displays that identify landmarks, friendly forces (to prevent fratricide), objectives and targets in the surrounding environment. However, this requires a tracking system that works outdoors in arbitrary, unprepared environments with sufficient accuracy to support AR applications. In a military application, we do not have the luxury of restricting the environment to ones with sufficient visual features of previously known objects to make vision-based tracking approaches feasible. A proposed tracking system for this application must work in virtually any arbitrary environment. Although there has been much progress in outdoor tracking for AR, no existing tracking system can accurately recover 6D pose in all circumstances.

In this paper, we take a step toward this ultimate goal by proposing a new tracking system to meet performance goals defined by military experts, and evaluating this new system in simulation to determine configurations and parameters that provide the desired performance.

While an ideal outdoor AR tracking system would provide perfect 6D pose measurements in all environments and conditions, a less accurate system that focuses only on augmenting a small number of chosen locations is still useful. In our selected application, a small group of soldiers walks outdoors, at night. These soldiers need to see AR augmentations of a small number of objects of interest in the environment: e.g., their own friendly forces, landmarks, targets and objectives. Domain experts stated that these augmentations should appear within 1m of their true

location at 100m range; if that specification is met, this system would be useful for navigation and identification purposes.

This "1 at 100" goal is much larger than ideal (it represents about 0.57 degrees of error, which is ~35 times the typical foveal resolution of 1 arc min.), but it is still difficult to meet in an unprepared environment with limited tracking resources. This is particularly true when GPS errors can be a meter or more. Vision-based tracking approaches will not work without visual features, and we cannot guarantee those will be recognized, so we cannot rely upon those approaches. Instead, we are restricted to the sensors that an infantryman can feasibly carry and any tracking components we can insert into the environment. For example, adding hundreds or thousands of detectable beacons into the environment is not practical, but ten might be feasible.

Therefore, we propose a new hybrid tracking system where a helmet has three rate gyroscopes, two tilt sensors, a GPS sensor and an infrared camera that occasionally observes small numbers (< 10) of *mobile* infrared beacons added to the environment. While some beacons may be fixed at static locations in the environment, we assume many are mounted on unmanned ground vehicles, on miniature unmanned air vehicles (MAVs) (Figure 1), or other soldiers. Occasional beacon observations improve the tracking accuracy. We can move the mobile beacons to where they can best aid the pose recovery when the soldier looks at an object of interest. Since this application occurs at night, high-power beacons will be detectable at ranges of 100 meters or more.

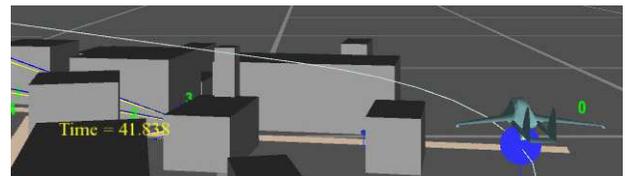


Figure 1: MAV beacon (right), flying over buildings (left)

## 2 PREVIOUS WORK

This is not the first tracking system proposed for outdoor operation; we list some representative previous works here. Azuma et. al. demonstrated a motion-stabilized tracking system outdoors [1], but that relied upon a compass that is not likely to be as accurate or robust as our beacon-based approach. Many approaches rely upon vision-based tracking. In particular, the closest approaches to ours merge inertial sensors with vision-based corrections [3][12][17]. However, our application cannot rely upon vision-based tracking, so we use beacons instead.

The concept of beacon or marker-based corrections is not new, dating back to early work by Bajura and Neumann [2] and a thorough demonstration by State et. al. [13]. Commercial tracking systems made by 3rdTech [16], InterSense [5] and others rely upon large arrays of accurately known markers or beacons. The major difference with our work is that instead of having many *static* beacons at *well-known* locations that support accurate pose recovery, we propose and evaluate a system that has only a few

\* email: { azuma, neely, mjdaily }@HRL.com

† email: jon@raytheon.com

mobile beacons, where there is *significant uncertainty* in the beacon locations. Since we only have a few beacons and they move, we cannot easily reduce beacon location errors through standard photogrammetric or autocalibration techniques. We trade less accurate results (that still fall within a range specified to be useful in this application) for a tracking system that can feasibly operate at almost any outdoor location.

Our approach is related to “cooperative localization” in the field of robotics, where mobile robots observe both stationary and moving objects to aid position recovery [11].

Finally, some previous works conduct simulations to optimize certain tracker parameters, such as marker locations. In one example, Vogt et. al. use Monte Carlo simulations to study the noise that results from different clusters of markers [14]. We also conduct Monte Carlo simulations, but to evaluate the performance of our own particular tracking concept.

### 3 CONTRIBUTION

The contribution of this paper comes in two parts. First, we propose a new tracking approach for an outdoor AR application based upon small numbers of mobile beacons, accepting reduced accuracy for more ubiquitous operation. Second, we analyze the performance of this new tracking concept, determining which parameters are most important and what values and configurations meet the accuracy specification. The variables include the number of beacons, geometric configuration of the beacons, errors in the sensors, etc. To conduct the analysis, we built a detailed simulation framework that models many types of errors and supports Monte Carlo simulations.

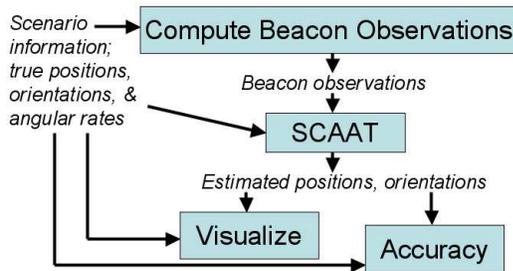


Figure 2: System dataflow

### 4 MOTION DATA

The simulation requires realistic position and orientation motions for the objects that move: the soldier and the mobile beacons. Generating reasonable position data is not a problem, since we do not use accelerations and only need position as a function of time. We can define paths algorithmically (such as “racetrack” curves for the MAV beacons, which also specify the MAV orientation) or apply lowpass filtering to paths traced on a map. However, for soldier orientation we need both the absolute orientation and its first derivative (to drive simulated angular rate gyroscopes). These orientations and rates must be perfectly matched. Synthesizing realistic data is difficult; e.g. we can generate orientations and rates by summing a few sinusoids, but that generally does not resemble realistic human head motions.

Therefore, we generated head orientation data by measuring and modifying actual recorded data. We attached three orthogonal rate gyroscopes (Systron Donner GyroChip II) to a HiBall sensor. As we rotated the sensors, we recorded orientations from the HiBall and angular rates from the gyroscopes (at 1 kHz). Then we applied a non-causal lowpass

filter to the gyro data (that avoids adding any phase shifts) [7]. From the original recorded orientation, we numerically integrate the gyro data to generate future orientations as a function of time. These do not match the HiBall orientations perfectly. So we use Powell’s method [10] to search for bias and scale parameters for each gyroscope such that the integrated rate information perfectly matches the last recorded HiBall orientation at the end of the motion sequence. The end result is a set of perfectly matched angular rates and quaternions, which approximately match the recorded HiBall quaternions.

### 5 SIMULATION

The simulation code consists of several modules, written in C++. Figure 2 shows the main dataflow. The “Compute Beacon Observations” module determines which beacons can be seen at each moment in time, computing the (x, y) image coordinates of each observed beacon. These image coordinates may be inaccurate due to modeled errors applied in the observation computation. Then the SCAAT algorithm [15] is used to compute the estimated pose (position and orientation), given the beacon observations and corrupted measurements from the gyros and GPS sensor on the soldier. The Visualize module, which uses the Open Scene Graph library [9], renders the results from both external and first-person simulated AR viewpoints (Figure 3). The Accuracy module computes various error metrics that compare the computed pose against the true pose. The scenario information controls the environment (e.g., buildings, beacons) and error parameters. We use a configuration of buildings based upon the McKenna urban training facility at Ft. Benning.

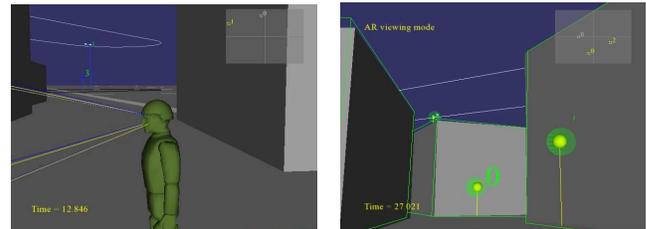


Figure 3: (Left) External view of soldier in Visualize module. (Right) First-person simulated AR view in Visualize module

The SCAAT filter incorporates three types of measurements: beacon observations, and gyro and GPS measurements. Our state vector holds linear position and velocities, and incremental orientation and velocities (with absolute orientation maintained and updated outside of the state). The SCAAT filter measurement functions require computing what the measurement should be, given the current state. This is trivial for gyro and GPS measurements, as they correspond one-to-one to the incremental orientation velocities and the linear position states, respectively. Beacon measurements are more complicated, and we incorporate those in the same way the HiBall does [16]. The SCAAT filter updates at 1 kHz, with a gyro measurement available at each iteration. Beacon and GPS updates occur much less frequently, at about 30 Hz and 2 Hz, respectively.

The SCAAT algorithm requires an initial estimate of the soldier’s position and orientation. The first GPS measurement provides the initial position estimate, while we get the initial pitch and roll orientation components from two tilt sensors. We compute the initial heading by observing one beacon. Given our estimated position, pitch and tilt, and approximately known beacon position, we can derive a heading estimate.

The Visualize module simulates rendering AR wireframes from the soldier’s first-person view by first rendering the buildings and environment using the true head position and orientation, then rendering the green AR wireframe augmentations based on the estimated pose. Hidden surface removal is implemented when rendering the green wireframes by using the stencil buffer [8].

We performed several checks to verify that the simulation operated correctly. We verified that perfect inputs yielded perfect outputs: when all simulated errors were removed, the estimated poses matched the true positions and orientations. We use Welch and Bishop’s original SCAAT code, modified to use our sensors. The Visualize module simultaneously displays the apparent beacon screen space coordinates as computed by the “Compute Beacon Observations” module and as rendered from the simulated AR viewpoint, so we can visually verify that the two match. Since it is easy to make mistakes in simulation, such careful implementation and verification were necessary to give us confidence in the simulation results.

Our simulator models many types of errors. Most of these are modeled as Gaussian error distributions, where we specify the one-sigma value, and the errors change with time. The errors are:

- *Uncompensated gyro bias errors:* We typically assume a 100 degree/hour drift rate.
- *Static beacon, mobile beacon and soldier position errors:* These are the position errors from the GPS sensors.
- *Limited 2D sensor resolution:* Beacon observation coordinates are rounded to the nearest 0.5 pixels (to simulate subpixel resolution).
- *2D image noise:* White noise is added to the beacon observations, based upon the distance to the beacon.
- *2D image smearing:* This simulates the effects of the sensor having non-zero integration time.
- *Latency:* Uncompensated delays in reporting the positions of mobile beacons cause beacon position errors.

## 6 ANALYSIS

Our military domain experts stated that the accuracy should be tested when the soldier kept his head still and focused upon the object of interest, 100m away, with at least one beacon visible. Under such circumstances, gyro bias errors are not a significant factor, as long as the net drift is small enough so the orientation state can be corrected by the occasional beacon observations.

We measure compliance with the “1 at 100” goal by computing a scatterplot error metric. Imagine that a real object was 100m in front of the user, along the gaze direction, no matter how the head orientation changed. If the estimated pose were perfect, a rendered virtual annotation at the position of that hypothetical object would always be at the dead center of the image. Errors in the estimated pose cause the virtual annotation to be rendered away from the center. For our 1280 by 1024 infrared image sensor with a field of view of 40 by 32 degrees, the scatterplot coordinates must be within 17.58 pixels in X and 17.85 pixels in Y from the center to meet the “1 at 100” goal. The scatterplot error metric captures the net combined effect of position and orientation errors, as seen in screen space.

We can guarantee the “1 at 100” goal is met if a beacon is at the point of interest and the sum of all position errors (for the beacons and the soldier) is under 1 meter. This requirement is too stringent, however, since the errors usually will not add up in a worst-case fashion, so we can often tolerate larger errors. We built a version of the simulator that contained the “Compute Beacon Observations,” SCAAT, and Accuracy modules to

support Monte Carlo analyses. Given a configuration of beacons and parameters specifying the distribution of each type of error, each Monte Carlo run randomly selected the values for each error source from those distributions, and then computed estimated poses for a brief sequence where the soldier head remained still. For each combination of configuration and error distribution, we ran 20,000 Monte Carlo simulations to determine the percentage of times the “1 at 100” goal was met (via the scatterplot metric).

In all tested configurations, the soldier’s head is at (0, 0, 2), where X is to the right, Y is forward along the direction of gaze, and Z points straight up. The point of interest is at (0, 100, 2). All positions are in meters. Latency was set to zero in these tests.

Space limitations prevent us from listing the exact details of the Monte Carlo simulations. We ran 24 simulations on 13 different beacon configurations, using 1-4 different error distributions on each configuration. The beacon configurations varied from two to nine beacons and explored different geometries of placing beacons around the point of interest (close to the point, far from the point, in front of and behind the point, all beacons to the left of the point, etc.) Based on these Monte Carlo simulations, we derived “rules of thumb” that identify the types of beacon configurations and error distributions that meet the goal.

## 7 DISCUSSION

The most important parameters for meeting the “1 at 100” goal are the geometric placement of the beacons around the point of interest (as seen from the soldier’s viewpoint), and the errors in measuring the positions of the beacons and the soldier’s head. In comparison, other errors, such as the gyro bias, image noise and image resolution, are not significant factors.

Good geometric placement consists of beacons at different heights, equidistant to the left and the right of the point of interest, and preferably at the same range as the point. If the ranges differ, beacons should be both in front of and behind the point. It is bad to have all the beacons in front of the point, or all behind the point, or all to the left or all to the right of the point. Having widely different distances to the left and right (e.g., 5m left and 30m right) also reduces the success rate. Simulations do not meet the accuracy goal with such bad beacon placements.

Note that we test registration accuracy only at the point of interest. Previous works [2][13] have noted that with one beacon whose location is accurately known, registration errors are minimized at that beacon’s location, because if the estimated position is incorrect, a beacon-based correction rotates the orientation to fixate upon that beacon. This minimizes registration error at the beacon’s 3D location but can cause much larger registration errors at other 3D points away from that beacon. Our studies suggest that with *multiple* beacons where there is *significant uncertainty* in knowing the beacon positions, it is best to distribute them in space at the same range and equidistantly left and right with respect to the soldier’s viewpoint. This dictates where to steer the mobile beacons.

The “1 at 100” goal is met under two classes of error values. First, we can achieve this with small position errors and a few (e.g., 2) well-placed beacons. For example, with two beacons at (-10, 100, 6) and (10, 100, 2), beacon  $\sigma=0.5m$ , soldier GPS horizontal  $\sigma=1m$  and vertical  $\sigma=2m$ , 98.9% of the trials succeeded. Second, we can achieve the goal with a large number (e.g., 9) of well-placed beacons with larger position errors. We simulated 9 beacons where beacon  $\sigma=1m$ , soldier GPS horizontal  $\sigma=1.8m$  and vertical  $\sigma=3.2m$ , and achieved success in 97.1% of the runs. Large position errors generally do not meet the goal. For example, with three beacons at (-20, 100, 15), (20, 100, 15)

and (0, 100, 30), both beacon and soldier GPS horizontal  $\sigma=1.8\text{m}$  and vertical  $\sigma=3.2\text{m}$ , only 15.2% of trials met the accuracy goal.

## 8 LIMITATIONS

We have not simulated the problem of beacon identification: given that many beacons exist, and our image sensor detects a beacon, which one did we detect? We did not simulate this error because incorrect identification of which beacon we are observing can cause huge errors in pose recovery; therefore it is crucial to prevent such errors from occurring in the first place.

There are at least three approaches for identifying beacons. In the ID CAM system [4], beacon ID's are encoded in rapidly amplitude-modulated signals that are detected by special image sensors where each pixel detector is "smart" enough to recognize these identification signals. The amplitude changes are encoded and detected at kilohertz rates, which is fast enough to usually avoid a major problem in beacon detection: when the beacon's apparent location in the image sensor changes with time, it is easy to lose track of the beacon if it also vanishes occasionally due to amplitude modulation. However, the ID CAM approach requires special sensors that we are not likely to be able to use in our system. Naimark and Foxlin [6] address this problem by modulating the amplitude but never letting the brightness drop below the detection threshold. This approach requires a thorough understanding of the strength of the beacon signal at different ranges and viewing angles. Finally, the HiBall tracker [16] avoids ambiguity by temporally multiplexing the beacons. Only one beacon is flashed at any instant of time, so the beacon ID at that time is unambiguous. Different beacons are flashed at different times. This approach requires accurate temporal synchronization between the beacons and the image sensors.

The Monte Carlo analyses assume that all error sources are uncorrelated. This is a reasonable assumption for most of the error sources, but it may not be valid for the GPS sensors. Since many GPS errors are due to atmospheric effects, a group of GPS sensors that are geographically close to each other are likely to experience similar errors. Even after differential or carrier-phase corrections are applied, the remaining errors may be correlated. Such correlations would actually improve our results, since what is crucial is the relative error amongst the beacons and soldier positions. If they are all offset by the same amount, then the position errors do not degrade the accuracy of the pose recovery.

We did not simulate systematic optical distortions (e.g. radial distortion). Calibration should compensate for such errors.

Our analysis so far has relied upon simulations. We have not yet actually built this system and demonstrated it in reality.

## 9 CONCLUSIONS AND FUTURE WORK

This performance analysis provides evidence that our system concept can work and can meet the desired "1 at 100" accuracy goal, if certain "rules of thumb" are followed in geometric placement of the beacons and expected errors of key system components fall within stated limits. We have demonstrated multiple configurations that meet the accuracy requirement in 95-100% of the Monte Carlo trials.

The next logical step is to build a prototype system that validates this system concept and verifies that the simulated performance matches reality. We propose to construct this in two phases. The first will use only ground-based mobile beacons, while the second phase will add flying MAV beacons. We ultimately intend to use head-worn infrared sensors being developed by the DARPA MANTIS program.

A major challenge in implementing this system is to outfit a MAV with GPS, IMU and other sensors and communications equipment of the desired performance while meeting stringent size, weight and power requirements. For example carrier-phase GPS sensors can meet the stated error limits but they are often large and heavy and will be difficult to emplace in a MAV.

## ACKNOWLEDGMENTS

Raytheon Missile Systems funded this research. We thank Greg Welch for providing code that implements the SCAAT algorithm.

## REFERENCES

- [1] Azuma, R., B. Hoff, H. Neely III, and R. Sarfaty. A Motion-Stabilized Outdoor Augmented Reality System. *Proc. IEEE Virtual Reality 1999* (Houston, TX, 13-17 March 1999), pp. 252-259.
- [2] Bajura, M. and U. Neumann. Dynamic Registration Correction in Augmented-Reality Systems. *Proc. 1995 Virtual Reality Annual Int'l Symp.* (Research Triangle Park, NC, 10-15 March 1995), pp. 189-196.
- [3] Jiang, B., U. Neumann and S. You. A Robust Hybrid Tracking System for Outdoor Augmented Reality. *Proc. IEEE Virtual Reality 2004* (Chicago, 27-31 March 2004), pp. 3-10.
- [4] Matsushita, N., D. Hihara, T. Ushiro, S. Yoshimura, J. Rekimoto, and Y. Yamamoto. ID CAM: A Smart Camera for Scene Capturing and ID Recognition. *Proc. ISMAR 2003* (Tokyo, 8-10 Oct. 2003), pp. 227-236.
- [5] Naimark, L. and E. Foxlin. Circular Data Matrix Fiducial System and Robust Image Processing for a Wearable Vision-Inertial Self-Tracker. *Proc. ISMAR 2002* (Darmstadt, Germany, 30 Sept. – 1 Oct. 2002), pp. 27-36.
- [6] Naimark, L. and E. Foxlin. Encoded LED System for Optical Trackers. *Proc. ISMAR 2005* (Vienna, Austria, 5-8 Oct. 2005), pp. 150-153.
- [7] NASA Tech Briefs KSC-11471. Digital Low-Pass Filter Without Phase Shift. John F. Kennedy Space Center.
- [8] Neider, J., T. Davis, and M. Woo. *OpenGL Programming Guide*. Addison-Wesley (1993).
- [9] Open Scene Graph. <http://www.openscenegraph.org>
- [10] Press, W., B. Flannery, S. Teukolsky, W. Vetterling. *Numerical Recipes in C*. Cambridge University Press (1988).
- [11] Rekleitis, I. Cooperative Localization and Multi-Robot Exploration. Ph.D. thesis, McGill University (2003).
- [12] Satoh, K., M. Anabuki, H. Yamamoto, and H. Tamura. A Hybrid Registration Method for Outdoor Augmented Reality. *Proc. ISAR 2001* (New York City, 29-30 Oct. 2001), pp. 67-76.
- [13] State, A., G. Hirota, D. Chen, W. Garrett, M. Livingston. Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking. *Proc. SIGGRAPH 1996* (New Orleans, LA, 4-9 Aug. 1996), pp. 429-438.
- [14] Vogt, S., A. Khamene, F. Sauer, and H. Niemann. Single Camera Tracking of Marker Clusters: Multiparameter Cluster Optimization and Experimental Verification. *Proc. ISMAR 2002* (Darmstadt, Germany, 30 Sept. – 1 Oct. 2002), pp. 127-136.
- [15] Welch, G. and G. Bishop. SCAAT: Incremental Tracking with Incomplete Information. *Proc. SIGGRAPH 1997* (Los Angeles, 3-8 August 1997), pp. 333-344.
- [16] Welch, G., 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 vol. 10, #1* (2001), pp. 1-21.
- [17] You, S., U. Neumann, and R. Azuma. Hybrid Inertial and Vision Tracking for Augmented Reality Registration. *Proc. IEEE Virtual Reality 1999* (Houston, TX, 13-17 March 1999), pp. 260-267.