Hybrid Inertial and Vision Tracking for Augmented Reality Registration

Suya You* Ulrich Neumann* Ronald Azuma §

*Integrated Media Systems Center
University of Southern California
Los Angeles, CA 90089-0781
{suyay|uneumann}@graphics.usc.edu

§HRL Laboratories
3011 Malibu Canyon Rd
Malibu, CA 90256
azuma@HRL.com

Abstract

The biggest single obstacle to building effective augmented reality (AR) systems is the lack of accurate wide-area sensors for trackers that report the locations and orientations of objects in an environment. Active (sensor-emitter) tracking technologies require powered-device installation, limiting their use to prepared areas that are relatively free of natural or man-made interference sources. Vision-based systems can use passive landmarks, but they are more computationally demanding and often exhibit erroneous behavior due to occlusion or numerical instability. Inertial sensors are completely passive, requiring no external devices or targets, however, the drift rates in portable strapdown configurations are too great for practical use. In this paper, we present a hybrid approach to AR tracking that integrates inertial and vision-based technologies. We exploit the complementary nature of the two technologies to compensate for the weaknesses in each component. Analysis and experimental results demonstrate this system's effectiveness.

1. Introduction

One of the key technological challenges for creating an augmented reality (AR) is to maintain accurate registration and tracking between real and computer-generated objects. As users move their viewpoints, the graphic virtual elements must remain aligned with the observed 3D positions and orientations of real objects. The alignment depends on accurately tracking the viewing pose, relative to either the environment or the annotated object(s) [15]. The tracked viewing pose defines the virtual camera pose used to project 3D graphics onto the real world image, so the tracking accuracy directly determines the visually-perceived accuracy of AR alignment and registration [1, 3].

A wealth of research, employing a variety of sensing technologies, deals with motion tracking and registration as required for augmented reality applications. Each technology has unique strengths and weaknesses. Tracking technologies may be grouped into three categories: *active*-target, *passive*-target, and *inertial*. Active-target systems incorporate powered signal emitters and sensors placed in a prepared and calibrated environment. Examples of such systems use magnetic, optical, radio, and acoustic signals. Passive-target systems use ambient or naturally occurring signals. Examples include compasses sensing the Earth's field and vision systems sensing intentionally placed fiducials (*e.g.*, circles, squares) or natural features. Inertial systems are completely self contained, sensing physical phenomena created by linear acceleration and angular motion. See [1, 12] for more complete overviews of tracking technologies.

Each tracking approach has limitations. The signal sensing range of as well as man-made and natural sources of interference limit active-target systems. Passive-target systems are also subject to signal degradation, for example poor lighting or proximity to steel in buildings can defeat vision and compass systems. Inertial sensors measure acceleration or motion rates, so their signals must be integrated to produce position or orientation. Noise, calibration error, and the gravity field impart errors on these signals, producing accumulated position and orientation drift. Position requires double integration of linear acceleration, so the accumulation of position drift grows as the square of elapsed time. Orientation only requires a single integration of rotation rate, so the drift accumulates linearly with elapsed time.

Hybrid systems attempt to compensate for the shortcomings of each technology by using multiple measurements to produce robust results. Active-target magnetic and passive-target vision are combined in [18]. Inertial sensors and active-target vision are combined in [2]. These and other examples are presented in Table 1.

Hybrid Approach	
Active-Active	vision-magnetic [3]
Active-Passive	magnetic-vision [18]
Active-Inertial	vision-inertial [2] acoustic-inertial [8]
Passive-Passive	
Passive-Inertial	compass-inertial [7][21] vision-inertial*
Inertial-Inertial	

Table 1 – Examples of hybrid tracking approaches (including *this work)

Vision is commonly used for AR tracking and registration [11, 13, 17, 20]. Unlike other active and passive technologies, vision methods track camera pose directly from the same imagery observed by the user. The tracked pose (position and orientation) is often relative to the object(s) of interest, not a sensor or emitter attached to the environment. This has several advantages: a) tracking may occur relative to moving objects; b) tracking measurements made from the viewing position often minimize the visual alignment error; and c) tracking accuracy varies in proportion to the visual size (or range) of the object(s) in the image [13]. The ability to both track pose and manage residual errors is unique to vision, however vision suffers from a notorious lack of robustness and high computational expense. Combining vision and inertial technologies offers one approach to overcoming these problems.

Our long-term goal is to develop stable, accurate and robust tracking methods for wide-area augmented realities, especially in unprepared indoor or outdoor environments. To achieve this, our laboratory explores a range of related issues, including robust natural feature detection and tracking methods [16], extendible vision tracking with natural features and new-point estimation techniques [14], and Kalman filters for pose estimation. This work combines our methods for fiducial and natural feature tracking with inertial gyroscope sensors to produce a hybrid tracking system. The two basic tenets of this work are:

1) Inertial gyro data can increase the robustness and computing efficiency of a vision system by providing a frame to frame prediction of camera orientation.
2) A vision system can correct for the accumulated drift of an inertial system.

We consider the case when the scene range is many multiples of the camera focal length. Under this condition, the 2D motion of image features is more sensitive to camera rotation than camera translation. People can rotate their heads very quickly, so in the case of a head-mounted camera, the 2D image motions are often mainly due to head rotation. Vision pose tracking methods often compute 2D-image motion. Since these motions are often due to rotation, inertial gyro sensors can aid the vision system in tracking these motions. Vision can in turn correct the long term drift of the inertial sensors.

The remainder of the paper describes our approach and method for camera and gyro calibration. We also present the results of our analysis and experiments.

2. Problem Statement

2.1 Inertial Tracking

The basic principles behind inertial sensors for determining orientation and position rest on Newton's laws [19, 4]. Two devices, gyroscopes and accelerometers, are contained in an inertial sensor, affixed to the three perpendicular axes of a body. Accelerometers measure linear acceleration vectors with respect to the inertial reference frame. In order to subtract the acceleration component due to gravity, the orientation of the linear accelerometers must be accurately known at all times. We focus on gyro devices that measure rotation rate. The gyro outputs are integrated over time to compute relative changes of orientation within the reference frame. The integration of signal and error gives rise to a linearly increasing orientation drift. Correction techniques may include magnetic compass measurements [7, 19]. However, compass signals are also noisy and especially subject to errors induced by ferrous materials. Indoor or urban compass signals are consequently unreliable. We attempt vision-based corrections in the hope that this approach will generalize to a wide range of environments.

2.2 Error Sensitivity of Inertial AR Tracking System

In this section, we analyze the error sensitivity of inertial tracker to an augmented reality tracking system. The inertial device we used for experiment is a three-degree of freedom (3DOF) orientation tracker produced by InterSense (Model IS-300). This device incorporates three orthogonal gyroscopes to sense angular rates of rotation along its three perpendicular axes. It also has sensors for the gravity vector and a compass [7] to compensate for gyro drift. The measured angular rates are integrated to obtain the three orientation measurements (Yaw, Pitch, and Roll). This system is specified as achieving approximately 1° RMS static orientation accuracy and 3° RMS dynamic accuracy, with 150Hz maximum update rate. Although adequate for interactive applications in virtual reality, this accuracy is inadequate for AR tracking. To demonstrate this, we map the specified error into the 2D image domain.

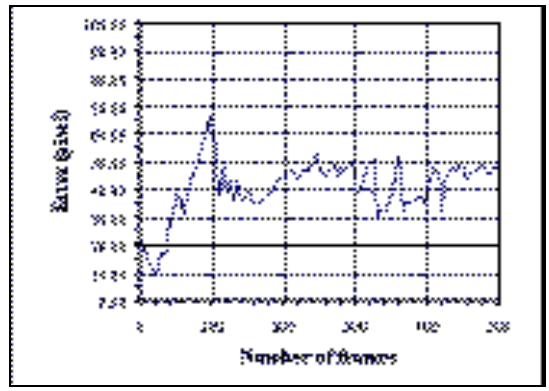
Let (f_x , f_y) be the effective horizontal and vertical focal lengths of a video camera (in pixels), (L_x , L_y) represent the horizontal and vertical image resolutions, and (θ_x,θ_y) be the field-of-view (FOV) of the camera, respectively. If pixels sample the rotation angles uniformly (Yaw and Pitch), the ratio of image pixel motion to the rotation angles (pixel/degree) is

$$L_x/\theta_x = \frac{L_x}{2\tan^{-1}(L_x/2f_x)}$$

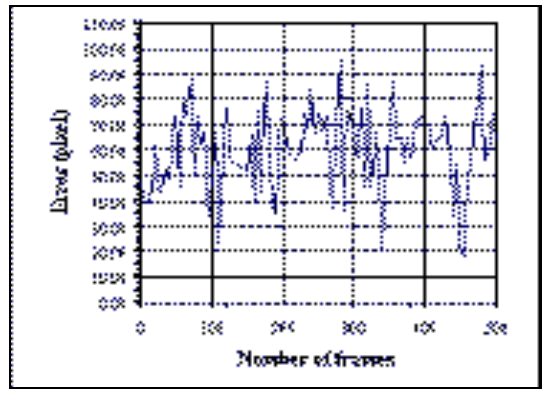
$$L_y/\theta_y = \frac{L_y}{2\tan^{-1}(L_y/2f_y)} \qquad (1)$$

To illustrate a concrete example of this relationship, consider the Sony XC-999 CCD video camera with F 1:1.4, 6 mm lens. Through calibration, we determine the effective horizontal and vertical focal lengths as f_x =614.059 pixels, and f_y =608.094 pixels, with a 640×480 image resolution. The ratios are L_x/θ_x =11.625 pixel/degree, and L_y/θ_y =11.143 pixel/degree. That is, each degree of orientation angle error results in about 11-pixels of alignment error in the image plane. In our actual use experience, the error of the inertial tracker may become larger than the specified one-degree. Increasing the FOV of the camera by using a wide-angle lens reduces the pixel error proportionately, however wide-angle lenses produce significant radial distortions that also contribute to pixel error [3].

Figure 1 illustrates the dynamic accuracy we measured experimentally with the inertial tracker. In our experiment, the 3DOF inertial gyro sensor is attached to a video camera to continually report the camera orientation. We do not attempt to measure a ground-truth absolute pose of the sensor/camera, rather we track visual feature motions to evaluate the gyro sensor accuracy relative to the image. By back-projecting the 3D orientation changes reported by the inertial sensor, we compare the gyro motion estimates with the observed feature motions in the image plane. Changes in the image-space distances are proportional to the errors accumulated by the inertial system. We believe this method simulates an AR system annotating visual features. The experiment allows us to evaluate the tracking of orientation-only inertial sensors. The error measure is appropriate since the ultimate metric of any augmented reality is the perceived image. Two kind sequences, a far-view (>100 feet) and near-view*



(a) Far-view scene



(b) Near-view scene

Fig. 1 – Average pixel differences between tracked features and backprojected features for Fig.3 (a) distant and Fig. 3 (b) near scenes.

(Figure 3 (a), (b)), each of 500 frames, are used for the test. Figure 1 illustrates the average error distributions for the two scenes. It clearly shows the dynamic drifts between the gyro data and tracked features.

3. Hybrid Inertial-Vision Tracking

Our prototype hybrid tracker fuses inertial orientation (3DOF) data with vision feature tracking to stabilize performance and correct inertial drift. We treat the fusion as an image stabilization problem. Approximate 2D feature-motion is derived from the inertial data, and vision feature tracking corrects and refines these estimates in the image domain. Furthermore, the inertial data also serves as an aid to the vision tracking by reducing the search space and providing tolerance to interruptions.

While our current experiments focus on a hybrid of 3DOF inertial and vision-based technologies, the methods are useful for 6FOF systems incorporating Gyros as well as other sensors such as accelerometers, GPS, compass, and pedometer measurements.

* We only consider pure rotation of the camera. Although we carefully pan the camera to avoid translations, minor translation is injected by the offset between the rotation axis and the optical center of the camera. For completeness, we consider both a far-view (campus) scene with feature ranges of over 100 feet and a near-view (office) scene that is more sensitive to minor translation.

3.1 Camera Model and Coordinates

The configuration of our system includes a CCD video camera with a rigidly mounted 3DOF inertial sensor. There are four principal coordinate systems, as illustrated in Figure 2: the world coordinate system $\mathbf{W}:(x_w, y_w, z_z)$, the camera-centered coordinate system $\mathbf{C}:(x_c, y_c, z_c)$, the inertial-centered coordinate system $\mathbf{I}:(x_I, y_I, z_I)$, and the 2D image coordinate system $\mathbf{U}:(x_u, y_u)$.

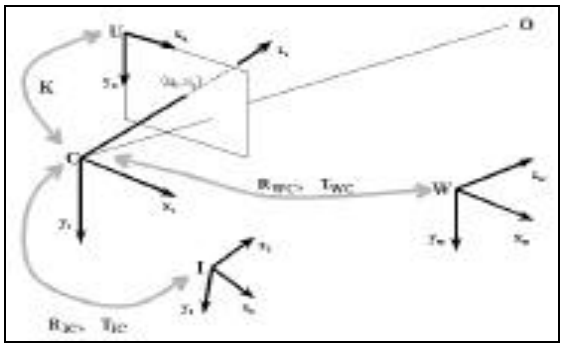


Fig. 2 - Camera model and the related coordinate systems of the hybrid system.

A pinhole camera models the imaging process. The origin of **C** is at the projection center of camera. The transformation from **W** to **C** is

$$\mathbf{W} :\rightarrow \mathbf{C} : \quad \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \left[\mathbf{R}_{wc} \,\middle|\, -\mathbf{R}_{wc}\,\mathbf{T}_{wc}\right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{2}$$

where the rotation matrix $\mathbf{R}_{wc}$ and the translation vector $\mathbf{T}_{wc}$ characterize the orientation and position of the camera with respect to the world coordinate frame. Under perspective projection, the transformation from **W** to **U** is

$$\mathbf{W} :\rightarrow \mathbf{U} : \quad \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = [\mathbf{K}] \left[\mathbf{R}_{wc} \,\middle|\, -\mathbf{R}_{wc}\,\mathbf{T}_{wc}\right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{3}$$

where the matrix **K**

$$\mathbf{K} = \begin{bmatrix} \alpha_x f & 0 & u_0 \\ 0 & \alpha_y f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

contains the *intrinsic* parameters of the camera*, f is the focal length of camera, α_x, α_y are the horizontal and vertical pixel sizes on the imaging plane, and (u_0, v_0) is the projection of camera center (principal point) on the image plane. The intrinsic parameters are calibrated offline.

Camera orientation changes are reported by the inertial tracker, so the transformation between the **C** and **I** is needed to relate inertial and camera motion. For rotation $\mathbf{R}_{Ic}$ and translation $\mathbf{T}_{Ic}$ the transformation is

$$\mathbf{I} :\rightarrow \mathbf{C} : \quad \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \left[\mathbf{R}_{Ic}\right] \begin{bmatrix} x_I \\ y_I \\ z_I \end{bmatrix} + \left[\mathbf{T}_{Ic}\right] \tag{5}$$

Since we only use the 3DOF orientation motion of the inertial tracker, only the rotation transformation needs to be determined. Our automatic calibration method is detailed below.

3.2 Static Calibration

3.2.1 Camera Parameters

Camera calibration determines the intrinsic parameters **K** and the lens distortion parameters. We use the method described in [13]. A planar target with a known grid pattern is imaged at measured offsets along the viewing direction. The intrinsic parameters and coefficients of radial lens distortion are computed by an iterative least-squares estimation. These parameters remain constant during our tracking experiments.

3.2.2 Transformation Between Inertial Frame and Camera Frame

The transformation between the inertial and the camera coordinate systems relates the inertial data to the camera motion, and hence to the image feature motions. Measuring this transformation is difficult, especially with optical see-through display systems [1]. We describe a motion-based calibration, as opposed to the boresight techniques presented in [2, 3]. For previously stated reasons, only the rotation component of the transformation needs to be determined.

Equation (5) relates the position transformation between the inertial tracker frame and the camera coordinate frame. The rotation motion relationship between the two coordinates can be derived

$$\omega_{\mathrm{C}} = \left[\mathbf{R}_{Ic}\right] \omega_{\mathrm{I}} \tag{6}$$

* For simplicity we omitted the lens distortion parameters from the equation. A complete form can be found in [13] for the method we used.

where, ω_C and ω_I denote the angular velocity of scene points, relative to the camera coordinate frame and the inertial coordinate frame, respectively.

The angular motion ω_I, relative to the inertial coordinate system, is obtained from the inertial tracker output. We need to compute the camera's angular velocity ω_C in some way, in order to determine the transformation matrix $\mathbf{R}_{Ic}$ based on equation (6).

General camera motion can be decomposed into a linear translation $\mathbf{V}_C = [V_{Cx}, V_{Cy}, V_{Cz}]^T$ and an angular motion $\omega_C = [\omega_{Cx}, \omega_{Cy}, \omega_{Cz}]^T$. Under perspective projection, the 2D-image motion resulting from camera motion can be written as

$$\begin{aligned} \dot{x}_u &= \left[\frac{-f V_{Cx} + x_u V_{Cz}}{z_C} + \frac{x_u y_u}{f}\omega_{Cx} - f\left(1+\frac{x_u^2}{f^2}\right)\omega_{Cy} + y_u \omega_{Cz}\right] \\ \dot{y}_u &= \left[\frac{-f V_{Cy} + y_u V_{Cz}}{z_C} + f\left(1+\frac{y_u^2}{f^2}\right)\omega_{Cx} - \frac{x_u y_u}{f}\omega_{Cy} + x_u \omega_{Cz}\right] \end{aligned} \tag{7}$$

where $(\dot{x}_u, \dot{y}_u)$ denotes the image velocity of point (x_u, y_u) in the image plane, z_C is the range to that point, and f is the focal length of camera. Eliminating the translation term and substituting from equation (6), we have

$$\dot{\mathbf{x}}_u = \Lambda \left[\mathbf{R}_{Ic}\right] \omega_I \tag{8}$$

where

$$\Lambda = \begin{bmatrix} \frac{x_u y_u}{f} & -f\left(1+\frac{x_u^2}{f^2}\right) & y_u \\ f\left(1+\frac{y_u^2}{f^2}\right) & -\frac{x_u y_u}{f} & -x_u \end{bmatrix}$$

In words, given knowledge of the internal camera parameters, the inertial tracking data ω_I, and the related 2D motions $[\dot{x}_u, \dot{y}_y]$ of a set of image features, the transformation $\mathbf{R}_{Ic}$ between the camera and the inertial coordinate systems can be determined from equation (8). This approach can also be used to calibrate the translation component of position tracking devices.

3.3 Dynamic Registration

The static registration procedure described above establishes a good initial calibration, however the inertial tracker accumulates drift over time and errors with motion. The distribution of drift and error is difficult to model for analytic correction. Our strategy of dynamic registration is to minimize the tracking error in the image plane, relative to the visually-perceived image.

Suppose N points are annotated in the scene. Their projections in the image are (x_i, y_i), $i = 1,2\cdots N$. Our goal is to automatically track these features as the camera moves in the following frames. Our method computes a tracking prediction from the inertial data, followed by a tracking correction with vision.

3.3.1 Tracking Prediction

Let ω_C be the camera rotation from frame $I(\mathbf{x}, t-1)$ to frame $I(\mathbf{x}, t)$. For the scene points O_i, their 2D positions in the image frame $t-1$ are $\mathbf{x}_{it-1} = [x_{it-1}, y_{it-1}]^T$. The positions of these points in the frame t, due to the related motion (rotation) between the camera and the scene, can be estimated

$$\begin{aligned} \mathbf{x}_{it} &= \mathbf{x}_{it-1} + \Delta \mathbf{x}_{it} \\ \Delta \mathbf{x}_{it} &= \Lambda\, \omega_C \end{aligned} \tag{9}$$

where, Λ is determined by equation (8).

3.3.2 Tracking Correction

Inertial data predicts the motion of image features. The correction refines these predicted image positions by doing local searches for the true features. A robust motion tracking approach is used for the correction strategy. The novel part of the approach [16] is it integrates three motion analysis functions, feature selection, tracking, and verification, in a closed-loop cooperative manner to copy with complicated imaging conditions. Firstly, in the feature selection module, 0D and 2D tracking features are selected for their reliable tracking and motion estimation suitability. The selection and evaluation processes also use data from a tracking evaluation function that measures the confidence of the last tracking estimation.

Once selected, features are ranked according to their evaluation values and fed into the tracking module. The tracking method is a differential-based local optical-flow calculation that utilizes normal-motion information in local neighborhoods to perform a least-squares minimization to find the best fit to motion vectors. Unlike traditional single-stage implementations, the approach adopts a multi-stage robust estimation strategy. For every estimated result, a verification and evaluation metric assesses the confidence of the estimation. If the estimation confidence is poor, the result is refined iteratively until the estimation error converges.

To achieve robust tracking, a novel motion verification and feedback strategy is proposed in a closed-loop tracking architecture. Two different verification strategies are used for the two kinds of tracking features and motion models. Basically, in both case, they depend on the estimated motion field to generation an evaluation frame that

measures the estimation residual. The difference between the evaluation frame and the true target frame measures the error of the estimate. This error information is fed back to the tracking module for motion correction and to the feature detection module for feature re-evaluation. The closed-loop control of the tracking system is inspired by the use of feedback for stabilizing errors in non-linear control system. The process acts as "selection-hypothesis-verification-correction" strategy that make it possible to discriminate between good and poor estimation features, which maximizes the quality of the final motion estimation.

4. Results

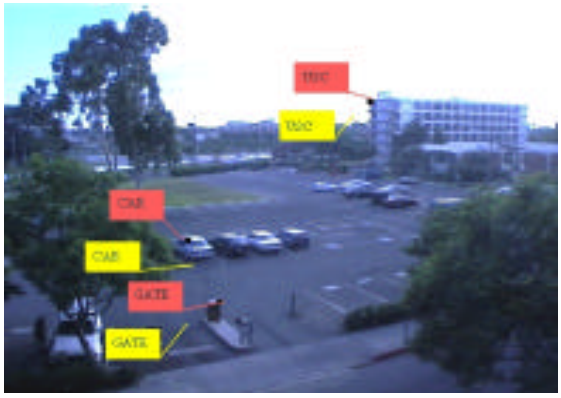
We have conducted extensive experiments to test the proposed fusion approach. Two prototype systems have been built, one is based on the InterSense's 3DOF inertial tracker (Model IS-300), and another is based on a hybrid 3DOF sensors system developed by HRL Laboratories [21]. The current fusion systems achieve about 9 frames/second on a SGI O2 workstation. Figure 3 shows three frames from video sequences captured from three different geographical locations. In these frames, black dots identify the feature points that we want to track and annotate. The yellow boxes are annotation text banners positioned only with inertial data (fused output of each tracker), while the red boxes denote the vision-corrected positions. The resolution of the images is 640x480.

4.1. Inertial-Only Tracking

In this test, only inertial data is used for tracking. Ten distinct features are manually selected in initial frames to establish visual reference points. The selected features are backprojected in each frame based on the camera orientation reported by the inertial tracker. The average differences between the backprojected image positions and the observed (vison-tracked) feature positions are the measure of tracking accuracy in each frame. Figures 4 illustrate the average error distributions for the three scenes confirming that substantial errors occur.

4.2. Hybrid Inertial-Vision: case 1

This test performs inertial tracking with vision correction of the integrated gyro error. As described in section 3.3, the predication of 2D image motion is based on the motion equation (9). This test corrects a feature's motion based on its integrated inertial predicted position. This approach has the disadvantage that inertial drift accumulates, however the drift is unaffected by any errors in the correction process, and this simulates the effect of prolonged occlusion of the vision system. This test shows how well the method corrects the accumulated gyro drifts



(a) Campus scene

(b) Office scene

(c) Pepperdine University scene

Fig. 3 – Virtual labels annotated over landmarks for three video sequences showing vision-corrected (red labels), and inertial only (yellow labels) tracking results. Note: (a) and (b) are based on InterSense's IS-300 inertial tracker, while (c) uses HRL's hybrid tracker.

over long periods of time. Figure 4 illustrates the results for the test scenes.

4.3. Hybrid Inertial-Vision: case 2

The alternative error correction is incremental correction. In this case, each correction results in an adjustment of the gyro state, consequently, the gyro error accumulation (for perfect corrections) is limited to periods between corrections. The reduced period of drift integration often results in lower accumulated error and better registration as illustrated in figure 4. A drawback of this approach is the possibility that a spurious correction error produces a lingering bias in the result.

5. Conclusions

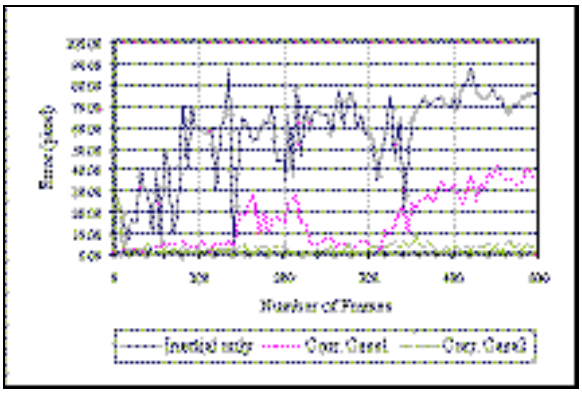
We presented a hybrid approach for AR registration with integrated inertial and vision tracking technologies. Inertial tracking has advantages of robustness, range, and a system that is passive and self-contained. Its major disadvantage is its lack of accuracy and drift over time. Vision tracking is accurate over long periods, but it suffers from occlusion and computation expense. We exploit the complementary nature of these two tracking technologies to compensate for the weakness in each separate component.

We quantitatively analyzed the sensitivities of orientation tracking error. To integrate the inertial and vision subsystems, accurate calibration of the two coordinate systems is critical, and we presented a motion based registration method that automatically computes the orientation transformation.

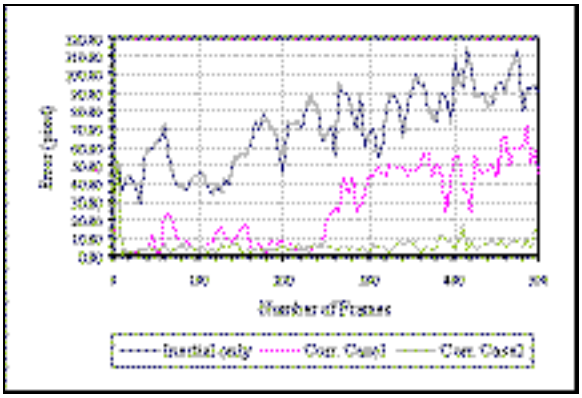
We applied vision corrections to both the accumulated and the incremental gyro error, and we presented our test results for two image sequences.

Acknowledgments

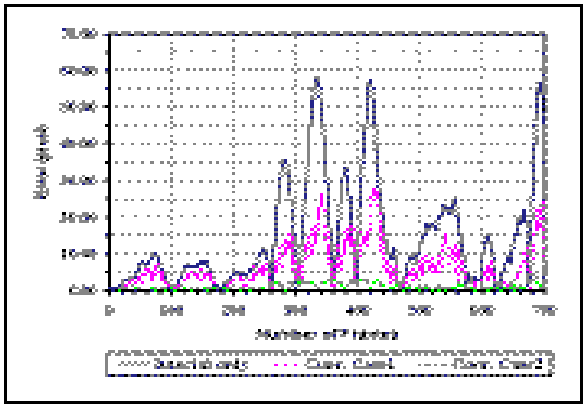
This work was supported by the Defense Advanced Research Project Agency (DARPA) "Geospatial Registration of Information for Dismounted Soldiers." We thank the Integrated Media Systems Center for their support and facilities. We acknowledge the research members of the AR Tracking Group at the University of Southern California for their help and suggestions. We also thank the reviewers for their valuable comments and suggestions.



(a) Campus sequence



(b) Office sequence



(c) Pepperdine University sequence

Fig. 4 – Hybrid alignment errors for different scenes shown in Fig. 3, with inertial only (blue line), hybrid case1 (pink line), and hybrid case2 (green line) methods.
Note: (a) and (b) are based on InterSense's IS-300 inertial tracker, while (c) uses HRL's hybrid tracker.

References

[1] R. Azuma. A Survey of Augmented Reality. *Presence:Teleoperators and Virtual Environments.* Vol. 6, No.4, pp. 355-385, 1997

[2] R. Azuma and G. Bishop. Improving Static and Dynamic Registration in an Optical See-through HMD. *Proc. of SIGGRAPH 95*, 1995.

[3] M. Bajura and U. Neumann. Dynamic Registration Correction in Augmented Reality Systems. *Proc. of IEEE Virtual Reality Annual International Symposium*, pp. 189-196, 1995.

[4] K. Britting. Inertial Navigation System Analysis. Wiley Interscience, New York, 1971.

[5] T. P. Caudell and D. M. Mizell, Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes. *Proc. of the Hawaii International Conference on Systems Sciences*, pp. 659-669, 1992.

[6] S. Feiner, B. MacIntyre and D. Seligmann. Knowledge-Based Augmented Reality. *Communications of the ACM*, Vol. 36, No. 7, pp. 52-62, July 1993.

[7] E. Foxlin. Inertial Head-Tracker Sensor Fusion by a Complementary Separate-Bias Kalman Filter. *Proc. of IEEE Virtual Reality Annual International Symposium*, pp. 184-194, 1996.

[8] E. Foxlin, M. Harrington and G. Pfeifer. Constellation: A Wide-Range Wireless Motion-Tracking System for Augmented Reality and Virtual Set Applications. *Prof. of GRAPHICS 98*, 1998.

[9] M. Ghazisadedy, D. Adamczyk, D. J. Sandlin, R. V. Kenyon and T. A. DeFanti. Ultrasonic Calibration of a Magnetic Tracker in a Virtual Reality Space. *Proc. of IEEE Virtual Reality Annual International Symposium* pp. 179-188, 1995.

[10] D. Kim, S. W. Richards and T. P. Caudell. An Optical Tracker for Augmented Reality and Wearable Computers. *Proc. of IEEE Virtual Reality Annual International Symposium*, pp. 146-150, 1997.

[11] K. Kutulakos and J. Vallino. Affine Object Representations for Calibration-Free Augmented Reality. *Proc. of IEEE Virtual Reality Annual International Symposium*, pp. 25-36, 1996.

[12] K. Meyer, H. L. Applewhite and F. A. Biocca. A Survey of Position Trackers. *Presence: Teleoperators and Virtual Environments*, Vol. 1, No. 2, pp. 173-200, 1992.

[13] U. Neumann and Y. Cho, A Self-Tracking Augmented Reality System. *Proc. of ACM Virtual Reality Software and Technology*, pp. 109-115, 1996.

[14] U. Neumann and J. Park. Extendible Object-Centric Tracking for Augmented Reality. *Proc. of IEEE Virtual Reality Annual International Symposium*, pp.148-155, 1998.

[15] U. Neumann and A. Majoros. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. Proc. of IEEE Virtual Reality Annual International Symposium, pp. 4-11, 1998.

[16] U. Neumann and S. You. Integration of Region Tracking and Optical Flow for Image Motion Estimation. *Proc. of IEEE International Conference on Image Processing*, 1998.

[17] R. Sharma and J. Molineros. Computer Vision-Based Augmented Reality for Guiding Manual Assembly. *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 3, pp. 292-317, June 1997.

[18] A. State, G. Hirota and D. T. Chen, B. Garrett, M. Livingston. Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking. *Proc. of SIGGRAPH'96*, pp. 429-438, 1996.

[19] D. H. Tittertion and J. L. Weston. Strapdown inertial navigation technology. *IEE Radar, Sonar, Navigation and Avionics Series 5*, Peter Peregrinus Ltd. UK 1997.

[20] M. Uenohara and T. Kanade. Vision-Based Object Registration for Real-Time Image Overlay. *Proc. of Computer Vision, Virtual Reality, and Robotics in Medicine*, pp. 13-22, 1995.

[21] R. Azuma, B. Hoff, H. Neely III and R. Sarfaty. A Motion-stabilized Outdoor Augmented Reality System. *Proceedings of IEEE VR'99* (Houston, TX, 13-17 March 1999).