Orientation Tracking for Outdoor 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 90265
azuma@HRL.com

Abstract

The biggest single obstacle to building effective augmented reality (AR) systems is the lack of accurate wide-area sensors for tracking 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, their drift rates in portable strapdown configurations are too great for practical use. In this paper, we present a hybrid approach to orientation tracking that integrates inertial and vision-based sensing. We exploit the complementary nature of the two technologies to compensate for the weaknesses in each component. Analysis and experimental results demonstrate the effectiveness of this approach.

1. Introduction

1.1 Motivation

The key technological challenge to creating an augmented reality (AR) is to maintain accurate registration between real and computer-generated objects. As AR users move their viewpoints, the graphic virtual elements must remain aligned with the observed positions and orientations of real objects. The perceived alignment depends on accurately tracking the viewing pose, relative to either the environment or the annotated object(s) [1, 6]. The tracked viewing pose defines the virtual camera pose used to project 3D graphics onto the real world image, so tracking accuracy directly determines the visually-perceived accuracy of AR alignment and registration [1].

Several AR tracking technologies have been developed for indoor applications, however none migrate easily to outdoor settings. Indoors, we can often calibrate the environment, add landmarks, control lighting, and limit the operating range to facilitate tracking. To calibrate, control, or modify outdoor environments, however, is unrealistic.

Our effort stems from a program focused on developing tracking technologies for wide-area augmented realities in unprepared outdoor environments[1]. We describe a hybrid orientation tracking system combining inertial sensors and computer vision. We exploit the complementary nature of these two sensing technologies to compensate for their respective weaknesses. Our multiple-sensor fusion is novel in AR tracking system, and the results demonstrate its utility.

1.2 Background

A wealth of research, employing a variety of sensing technologies, deals with motion tracking and registration as required for augmented reality. Each technology has unique strengths and weaknesses. Existing systems can be grouped into two categories: *active*-target, and *passive*-target (Table 1). Active-target systems incorporate powered signal emitters, sensors, and/or landmarks (fiducials) placed in a prepared and calibrated environment. Demonstrated active-target systems use magnetic, optical, radio, and acoustic signals [5]. Passive-target systems are completely self-contained, sensing ambient or naturally occurring signals or physical phenomena. Examples include compasses sensing the Earth's magnetic field, inertial sensors measuring linear acceleration and angular motion, and vision systems sensing natural scene features.

[1] Other participants in the DARPA funded GRIDS program included UNC and Raytheon.

Active-Active	magnetic-vision [8]
Active-Passive	vision-inertial [3] acoustic-inertial [4]
Passive-Passive	compass-inertial [2] vision-inertial* [9-12]

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

Vision is commonly used for AR tracking [1, 6]. Unlike other active and passive technologies, vision methods can estimate camera pose directly from the same imagery observed by the user. The pose estimate 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. The ability to both track pose and measure residual errors is unique to vision, however vision suffers from a notorious lack of robustness and high computational expense. Combining vision with other technologies offers the prospect of overcoming these problems.

All tracking sensors have limitations. The signal-sensing range 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 degrades vision and proximity to ferrous material distorts compass measurements. Inertial sensors measure acceleration or angular rates, so their signals must be integrated to produce position or orientation. Noise, calibration error, and gravity acceleration impart errors on these signals, producing accumulated position and orientation drift. Position is obtained from double integration of linear acceleration, so the accumulation of position drift grows as the square of elapsed time. Orientation is obtained from a single integration of angular rate, accumulating drift linearly with time.

Hybrid systems attempt to compensate for the shortcomings of a single technology by using multiple sensor types to produce robust results. For example, in [8], active-target magnetic and active-target vision are combined. In [3] a hybrid of inertial sensors and active-target vision creates an indoor AR system. Passive-target vision and inertial sensors create a hybrid tracker for mobile robotic navigation and range estimation [11, 12]. These and other examples are presented in Table 1. See [1] for a more complete overview of tracking technologies.

1.3 Approach

Our approach combines prior work in natural feature tracking [7, 9] with inertial and compass sensors [2] to produce a hybrid orientation tracking system. By exploiting the complementary nature of these sensors, the hybrid system achieves performance that exceeds any of the components. The two basic tenets of our approach are:

1) Inertial gyroscope data can increase the robustness and computing efficiency of a vision system by providing a relative frame-to-frame estimate 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 perceived motion of scene features is more sensitive to camera rotation than camera translation. The vision system tracks 2D-image motions, and since these are largely due to rotations, the gyroscope sensors provide a good estimate of these motions. Vision tracking, in turn, corrects the error and drift of the inertial estimates.

2. System overview

Figure 1 depicts the system hardware configuration.

- A compass and tilt sensor module (Precision Navigation TCM2) providing user's heading and two tilt angles in local motion frame. The module is specified to achieve approximately +/- 0.5 degree of error in yaw, at a 16Hz update rate.
- Three gyroscopes (System Donner GyroChip II QRS14-500-103) in an orthogonal configuration to sense angular rates of rotation along three perpendicular axes. The maximum sense range is +/-500 degrees/second, sampled at 1kHz.
- A video camera (Sony XC-999 CCD color camera) providing visual streams for vision based tracker and AR display.

The outputs of these sensors are fused to determine a user's orientation. The compass module and gyro sensors are filtered and fused to provide a prediction of angular motion [2]. From a static location under moderate rotation rates, the fusion algorithm achieves about two degrees of peak registration error. Typical errors are less than one degree while operating in real time [2]. For rapid motions or long tracking periods, the errors become larger due to accumulated

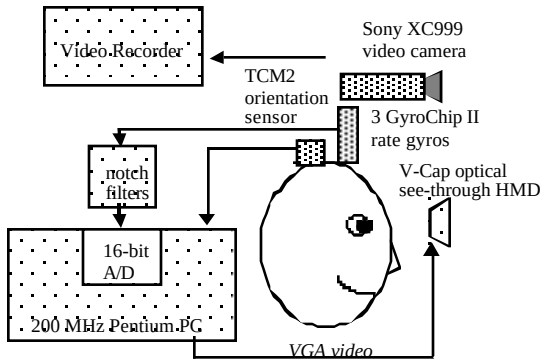


Figure 1- System configuration

gyroscope drift and compass errors. These are corrected by the vision measurements. Our vision tracking does not run in real time, so for our experiments, both the inertial data and video images are recorded for offline processing and fusion.

3. Inertial Tracking

The basic principles behind inertial sensors rest on Newton's laws. We use gyroscopes that sense rotation rate. The gyroscope data are integrated over time to compute relative changes of orientation within the reference frame. The integration of signal and error gives rise to an approximately linear increasing orientation drift.

3.1. Error Sensitivity of Inertial Tracking

We analyzed the error sensitivity of our gyroscope system. We sample the angular rate at 1kHz, and output the integrated orientation at 30Hz to match the imaging frame rate. Integration of the angular rates and a coordinate transformation produces three orientation measurements (Yaw, Pitch, and Roll) of the tracker with respect to the initial orientation.

A vision system can measure the dynamic gyroscope accuracy, so we first determine the relationship between angular rate and image motion. 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 we approximate pixels as sampling the rotation angles uniformly (Yaw and Pitch), the ratio of image pixel motion to the rotation angles (pixel/degree) is

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

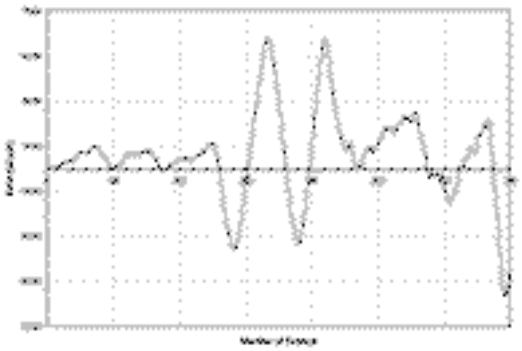


Figure 2 - Average pixel differences between tracked features and features projected by gyro measurements.

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

As a concrete example of this relationship, consider the Sony XC-999 CCD video camera with an 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. Increasing the FOV of the camera with a wide-angle lens reduces the pixel error proportionately, however wide-angle lenses produce significant radial distortions that contribute error [9].

Figure 2 illustrates the dynamic gyroscope accuracy we measured experimentally. The 3DOF gyro sensor is rigidly attached to the video camera and continually reports the camera orientation. Rather than attempting to measure the ground-truth absolute orientation of the sensors, we track visual feature motions to evaluate the gyroscope accuracy. We manually select image features (~5) while the camera and gyroscope are at rest. Then during motion, we track these features by our vision method and compare their observed positions to their projected positions that are derived from the 3D orientation changes reported by the gyroscopes. Pixel distances are proportional to the errors accumulated by the inertial system (as described in Eq. 1). Figure 2 plots the average pixel-errors measured for the selected features while rotating the sensors in an outdoor setting. It clearly shows the dynamic variations between the gyroscope data and observed feature motions.

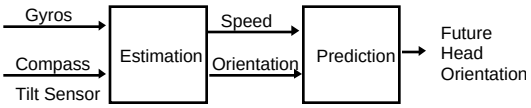


Fig. 3 – Schematic for gyro-compass fusion

3.2. Gyroscope Stabilization by Compass

We can estimate the stabilized angular position and rotation rate of the head from the input of the compass module (TCM2) and the three gyroscopes. This position is extrapolated one frame into the future to estimate the head orientation at the time the image appears on the see-through display (Figure 3). Space limitations prohibit a full explanation of the gyro-compass fusion method; please read [21] for the details. This section will provide an overview of the fusion method and the results.

Sensor calibration is crucial to system performance. The gyroscopes required an estimate of their bias and analog notch filters to remove a high-frequency noise. The compass encountered significant distortions from our environment and the system equipment. The distortions remained relatively constant at a single location over time (30 minutes), so heading calibration was possible with a special non-magnetic turntable (made of Delrin).

The fusion method compensates for the difference in time delays between the two sensors. The gyroscopes are sampled by an A/D at 1.0 KHz, with minimal latency. However, the compass is read at 16 Hz through a serial line. We captured several data runs and determined the average difference in latencies was 92 ms. Therefore, the fusion method incorporates compass measurements by comparing them to gyroscope estimates that are 92 ms old.

Figure 4 shows the dynamic behavior of the filter. The raw compass input (blue line) leads the filter output (red line). The filter compensates for the lagging compass measurements. The filter output retains the smoothness of the gyroscope inputs and is much smoother than the raw compass input. When

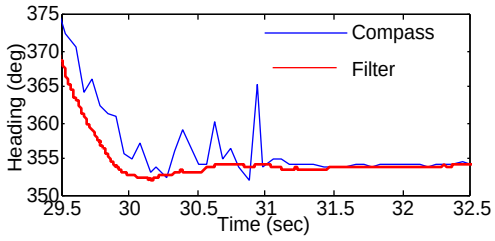


Fig. 4 – Sequence of heading data as system comes to rest

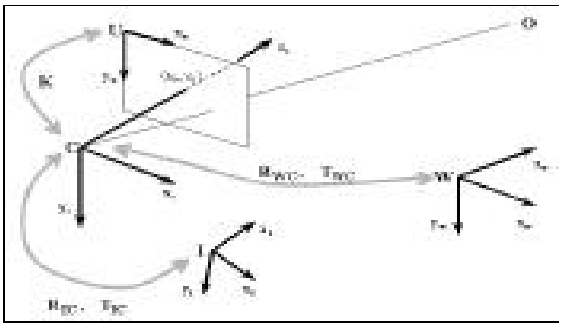


Fig. 5 - Camera model and the related coordinate systems of the hybrid system

the user stops moving, the filter output settles to the compass value, since it provides an absolute heading. Clearly, the registration accuracy of this approach is limited by the absolute heading accuracy. The use of visual input can compensate for compass estimates.

4. Hybrid Inertial-Vision Tracking

The hybrid tracker fuses gyroscope orientation (3D) and vision feature motion (2D) to derive a robust orientation measure. We structure the fusion as predictor-corrector image stabilization. First, an approximate 2D feature-motion is estimated from the inertial data (prediction). Then the vision feature tracking corrects and refines the estimate in the image domain (2D correction). Finally, the estimated 2D-motion residual is converted to a 3D-orientation correction for the gyroscope (3D correction). During this process, an added benefit is realized. The inertial estimate increases the efficiency of the vision tracking by reducing the image search space and providing tolerance to blur and other image distortions.

4.1 Camera Model and Coordinates

Our system includes a CCD video camera with a rigidly mounted 3DOF inertial sensor. The four principal coordinate systems are shown in Figure 5: world $\mathbf{W}:(x_w, y_w, z_z)$, camera-centered $\mathbf{C}:(x_c, y_c, z_c)$, inertial-centered $\mathbf{I}:(x_I, y_I, z_I)$, and 2D image coordinates $\mathbf{U}:(x_u, y_u)$.

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

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

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

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

where the matrix **K**

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

represents 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.

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

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

Since we only measure 3D-orientation motion, only the rotation transformation needs to be determined.

4.2 Static Calibration

4.2.1 Camera Parameters

Camera calibration determines the intrinsic parameters **K** and the lens distortion parameters. We use the method described in [9]. 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 are assumed constant for our experiments.

4.2.2 Transformation Between Inertial and Camera Frames

The transformation between the inertial and the camera coordinate systems relates the measured inertial motion to camera motion and image feature motion. Measuring this transformation is difficult, especially with optical see-through display systems [1]. We describe a motion-based calibration, as opposed to the boresight methods presented in [3].

Equation (5) relates the inertial tracker frame and the camera coordinate frame. The rotation relationship between the two coordinates is

$$\omega_C = \left[\mathbf{R}_{Ic} \right] \omega_I \qquad (6)$$

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 data. We need to compute the camera's angular velocity ω_C in some way, in order to determine the transformation matrix $\mathbf{R}_{Ic}$ from equation (6).

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

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

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 the camera. Eliminating the translation term and substituting from equation (6), we have

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

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 $\left[\dot{x}_u, \dot{y}_u \right]$ 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 between position tracking sensors.

* For simplicity we omitted the lens distortion parameters from the equation. A complete form can be found in [9] for the method we used.

4.3 Dynamic Registration

The static registration procedure described above establishes a good initial calibration, however the gyroscope accumulates drift over time and produces errors with motion. The distribution of drift and error is difficult to model for analytic correction. Our strategy for dynamic registration is to minimize the tracking error in the 2D plane of the perceived image.

4.3.1 Tracking Prediction

Suppose N features are detected in the scene. Our goal is to automatically track these features as the camera moves in the following frames. 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 as

$$\mathbf{x}_{it} = \mathbf{x}_{it-1} + \Delta\mathbf{x}_{it}$$

$$\Delta\mathbf{x}_{it} = \Lambda\,\omega_c \qquad (9)$$

where Λ is given by equation (8).

4.3.2 2D Tracking Correction

Inertial data predicts the motion of image features. The correction refines these predicted positions by local image searches for the true features. Our robust motion tracking approach integrates three motion analysis functions, feature selection, tracking, and verification, in a closed-loop cooperative manner to cope with complex imaging conditions [7]. Firstly, in the feature selection module, 0D (points) and 2D (regions) tracking features are selected for their suitability for tracking and motion estimation. The selection process also uses data from a tracking evaluation function that measures the confidence of the prior tracking estimations.

Once selected, features are ranked according to their evaluations and fed into the tracking module. A differential-based local optical-flow calculation utilizes normal-motions in local neighborhoods to perform a least-squares minimization to find the best affine motion estimate. 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 low, the result is refined iteratively until the estimation error converges. For details, see [7].

4.3.3 3D Tracking Correction

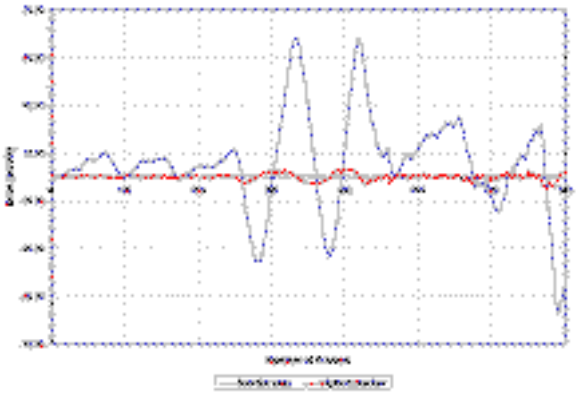
Let $\omega_I = \omega_c + \Delta\omega$ be the orientation from the inertial sensor, in which ω_C is the *real* camera motion, and $\Delta\omega$ is the gyroscope *drift* that we want to estimate and correct. From equations (7) and (8), we derive the relationship between the gyro error $\Delta\omega$ and the resulting 2D error of image velocity as

$$\dot{\mathbf{x}}^I_u - \dot{\mathbf{x}}^C_u = \Lambda\,\Delta\omega \qquad (10)$$

The left-hand of (10), $\dot{\mathbf{x}}^I_u - \dot{\mathbf{x}}^C_u$ is the image velocity difference between the inertial sensor and the real camera motion (or 2D-motion *residual*). The problem of 3D correction is reduced to finding the inertial drift $\Delta\omega$ that minimizes the motion residual $\|\dot{\mathbf{x}}^I_u - \dot{\mathbf{x}}^C_u\| \rightarrow \min$. Then the inertial drift to be corrected is

$$\Delta\omega = \Lambda^{-1}\,(\dot{\mathbf{x}}^I_u - \dot{\mathbf{x}}^C_u) \qquad (11)$$

(a) – Virtual labels annotated over landmarks for video sequences showing vision-corrected (red labels), and inertial only (blue labels) tracking results.



(b) – Hybrid alignment errors for scene (a) sequence showing inertial only (blue line), and vision corrected (red line) errors

Figure 6 – Tracking result of outdoor natural scene

5. Results and Evaluation

We experimentally tested our approach. Figure 6(a) shows a sample frame from a 30 Hz video sequence captured at an outdoor location with moderate rotation rates. In this frame, black dots identify the feature targets that we want to track and annotate. The yellow labels are positioned only by inertial data (fused gyro and compass data), while the red labels show the vision-corrected positions. The resolution of the images is 640×480.

Figure 6(b) illustrates the average pixel errors for inertial-only tracking (blue line), and hybrid inertial-vision tracking (red line), respectively. To obtain these quantitative results, ten distinct features are manually selected in initial frames to establish visual reference points. The selected features are back-projected in each frame based on the camera orientation reported by the tracking system. The average differences between the back-projected image positions and the observed (vision-tracked) feature positions are the measure of tracking accuracy in each frame. The inertial tracking errors are effectively corrected, reducing the average registration error over the image sequence to 4.27 pixels (corresponding to ~0.4 degree of rotation). These results illustrate the value of hybrid tracking.

To obtain these results, our hybrid system ran at about 2-4 frames/sec on and SGI O2. Our current version runs over 10 frames/second on an SGI Onyx2 and multiprocessor PC. Since the 2D-vision correction operates on each feature, the system speed depends on the number of tracked features.

As mentioned before, we assume that scene objects are distant to minimize the effect of position errors. Although this condition is often met in outdoor applications, orientation tracking is insufficient when tracking and annotation features are close to the tracker. In this case, the translation term can not be ignored in the motion model. Additional data is needed to provide position information. Accelerometers and GPS sensors are important data sources that we will investigate in our future work.

Acknowledgments

This work was largely supported by the Defense Advanced Research Project Agency (DARPA) “Geospatial Registration of Information for Dismounted Soldiers.” We also thank Intel, SGI, and the Integrated Media Systems Center for support.

References

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

[2] R. Azuma, B. Hoff, H. Neely III and R. Sarfaty, "A Motion-stabilized Outdoor Augmented Reality System", *Proc. IEEE VR'99,* Houston, TX, 1999, pp. 13-17.

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

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

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

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

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

[8] 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*, 1996, pp. 429-438.

[9] S. You, U. Neumann and R. Azuma, "Hybrid Inertial and Vision Tracking for Augmented Reality Registration", *Proc. IEEE VR'99,* Houston, TX, 1999, pp. 13-17.

[10] Welch Gregory, "Hybrid Self-Tracker: An Inertial / Optical Hybrid Three-Dimensional Tracking System", UNC Chapel Hill, Dept. of Computer Science Technical Report TR95-048, 1995.

[11] R. S. Suorsa and B. Sridhar, "A Parallel Implementation of a Multisensor Feature-Based Range-Estimation Model", *IEEE Trans. On Robotics and Automation*, Vol. 10, No. 6, 1994.

[12] J. Lobo, P. Lucas, J. Dias, and A. T. Almeida, "Inertial Navigation System for Mobile Land Vehicles", *Proc. ISIE’95*, 1995, pp. 843-848.