Fusion of Inertial Sensing to Compensate for Partial Occlusions in Optical Tracking Systems

Changyu He[1,2], H. Tutkun Şen[1], Sungmin Kim[1], Praneeth Sadda[1], and Peter Kazanzides[1]

[1] Johns Hopkins University, Baltimore MD 21218, USA
[2] Beijing Institute of Technology, Beijing 100081, China
hechangyu425@gmail.com, pkaz@jhu.edu

Abstract. Optical tracking is widely used for surgical Augmented Reality systems because it provides relatively high accuracy over a large workspace. But, it requires line-of-sight between the camera and the markers, which can be difficult to maintain. In contrast, inertial sensing does not require line-of-sight but is subject to drift, which causes large cumulative errors, especially for the measurement of position. This paper proposes a sensor fusion approach to handle cases where incomplete optical tracking information, such as just the 3D position of a single marker, is obtained. In this approach, when the optical tracker provides full 6D pose information, it is used to estimate the bias of the inertial sensors. Then, as long as the optical system can track the position of at least one marker, that 3D position can be combined with the orientation estimated from the inertial measurements to recover the full 6D pose information. Experiments are performed with a head-mounted display (HMD) that integrates an optical tracker and inertial measurement unit (IMU). The results show that with the sensor fusion approach we can still estimate the 6D pose of the head with respect to the reference frame, under partial occlusion conditions. The results generalize to a conventional navigation setup, where the inertial sensor would be co-located with the optical markers instead of with the camera.

1 Introduction

We are investigating the use of a head-mounted display (HMD) for presenting navigation information during surgical procedures[1][8]. We implemented our prototype system by mounting an optical see-through HMD (Juxtopia LLC, Baltimore, MD) and a small, commercially-available optical tracking system (Micron Hx40, Claron Technology, Toronto, Canada) on a helmet, as shown in Fig. 1. Ultimately, we envision integrating cameras directly on the HMD to avoid this cumbersome setup. Our goal is not to display high-resolution preoperative images on the HMD, but rather to show simple graphics derived from the navigation information. For example, the navigation information can include models of the patient anatomy that are obtained from preoperative images, such

C.A. Linte et al. (Eds.): AE-CAI 2014, LNCS 8678, pp. 60–69, 2014.

as biopsy target points and tumor outlines. Figure 2 illustrates two possible augmentation strategies: (a) overlaying the preoperative model directly on the patient's anatomy, and (b) displaying a "picture-in-picture" (PiP) virtual view of a model of the surgical instrument with respect to the preoperative model. The first strategy has the advantage that it is not necessary to track the instruments because their positions relative to the target are shown in the augmented view. But, this strategy also introduces technical challenges, especially when using an optical see-through HMD. In particular, it requires accurate calibration of the HMD to the surgeon's eyes (which must be maintained during the procedure, even if the HMD slips) and is subject to "swimming" effects due to the unavoidable delay in rendering overlays. We therefore adopted the second approach, which adds the requirement to track the surgical instrument. In either case, it is necessary to track the surgeon's head so that the the graphics are displayed from an intuitive perspective. This extra tracking requirement would exacerbate the existing concerns with maintaining line-of-sight between the tracking system camera and the markers on the patient, surgical instrument, and surgeon's head. We therefore chose an "inside out" tracking approach, where the camera is mounted on the HMD, as widely done in the augmented reality community [2], [13], [9], [4], [3], [10].

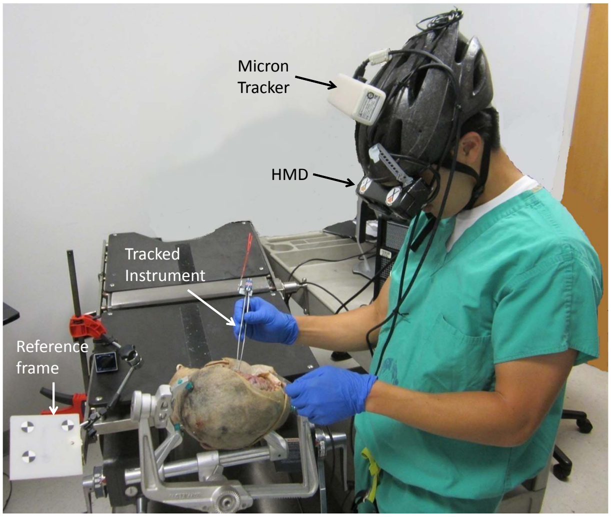
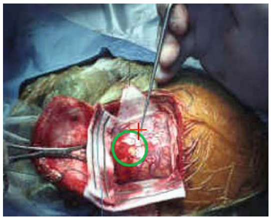
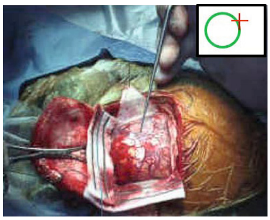


Fig. 1. Cadaver experiment with head-mounted tracking system and display



(a) Overlay on anatomy



(b) PiP virtual view

Fig. 2. Conceptual illustrations of augmented views

We performed phantom experiments [8] and cadaver experiments (unpublished) in procedures emulating resection of convexity (brain surface) tumors. The HMD displayed the pose of the tracked surgical instrument with respect to 2D views of simulated tumor margins (in a real procedure, these would be segmented from preoperative CT or MRI images). As anticipated, mounting the camera on the surgeon's head avoided many line-of-sight issues because the surgeon typically did not block his own view of the patient. But, we discovered that it was difficult to keep the reference frame in the camera field of view because the head-mounted camera is much closer to the scene than an external camera mounted on a tripod or to the ceiling (see Fig. 1). One possible solution is to make a smaller reference frame, but that would reduce the accuracy of the orientation measurement. We therefore pursued a hybrid tracking approach, where we combined optical tracking and inertial sensing. Optical tracking provides drift-free measurement of position and orientation, but subject to a line-of-sight constraint and with slower update rates and higher latency. In contrast, inertial sensing (which includes accelerometers, gyroscopes, and magnetometers) provides low latency, high frequency measurements, but these sensors either provide derivatives of position/orientation and are subject to drift, or provide absolute orientation but are subject to disturbances (e.g., magnetometer). A sensor fusion approach enables us to take advantage of the strengths of each tracking technology [11].

We implement a Kalman Filter to estimate the orientation, but drive the system dynamics by the gyroscope and use the accelerometer and magnetometer to provide measurement updates. The camera always provides the position and, if the full marker frame is visible, the camera orientation is used to estimate the bias of the inertial sensors. Because we use a HMD, we place the IMU near the camera on the helmet and focus only on estimating the transformation between the surgeon's head (camera) and reference frame (i.e., we do not attempt to handle partial occlusion of the tracked instrument). But, the method would apply equally well if the IMU is attached to the optically-tracked markers in a more conventional navigation setup. In this case, it is possible to attach an IMU to each tracked device (reference frame and instrument), with the limitation that these devices cannot be passive and would require a wired or wireless connection to the measurement computer.

Fusion of inertial and optical sensing has been well studied in the literature, as in most of the augmented reality references cited above, but the goals have primarily been to: (1) provide more timely pose estimates than available with just optical tracking, (2) improve the accuracy of the data (especially the orientation), or (3) allow the system to continue to provide pose estimates if optical tracking is occluded for up to a few seconds. Our focus is to handle the situation where we have a partial occlusion of the optically-tracked markers; for example, if up to two of the markers on the Reference Frame in Fig. 1 are blocked. The partial occlusion case was previously addressed, with a different approach, by Tobergte[11], who used individual marker positions to correct the position and orientation estimated by an Extended Kalman Filter (EKF), with the system dynamics driven by the accelerometer and gyroscope feedback. Other researchers have handled partial occlusions by fusing optical and electromagnetic tracking systems [12]. The details of our approach are provided in the following sections.

2 Sensor Fusion Approach

The hybrid tracking unit contains one stereo camera (MicronTracker Hx40) and one IMU that is rigidly attached to the camera. The MicronTracker tracks special patterns at approximately 20 fps; the images are transferred to the host computer via a FireWire port. The IMU is a custom design with tri-axial accelerometer, gyroscope and magnetometer feedback. The 9 data values are provided to the host computer via a USB port; for the experiments in this paper, we used a data rate of 160 Hz. The two tracking units are registered by a calibration procedure [5]. The reference frame consists of three markers attached to a plastic board, as shown in Fig. 1. The reference frame is assumed to remain stationary during the procedure (more precisely, all other measurements are made relative to this frame so, without loss of generality, it can be assumed to be stationary). The test setup also includes a surgical instrument ("tool") that contains three tracked markers, though in a smaller physical arrangement (see Fig. 1). Our method does not handle occlusion or partial occlusion of the tool – if the marker frame is not visible, the tool model is not displayed in the augmented reality view. Note that tracking of the surgical instrument would not be required if the preoperative model were overlayed in-situ (on the anatomy), as in Fig. 2(a), which we believe would be more feasible with a video see-through HMD.

2.1 IMU Bias Estimation

The IMU provides three different types of three-axis measurements: (1) the gyroscope measures the angular rate of rotation, (2) the accelerometer measures the linear acceleration and the acceleration due to gravity, and (3) the magnetometer measures the earth's magnetic field (i.e., magnetic North).

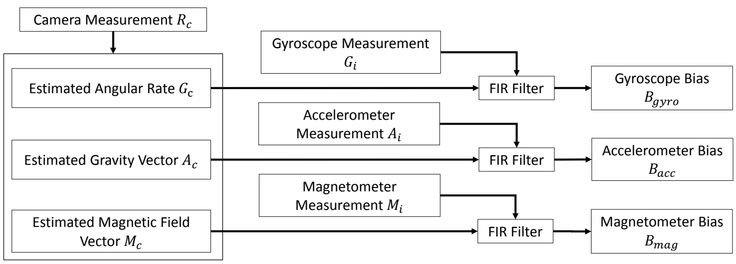


Fig. 3. Diagram of the bias calibration FIR filter

An optical tracking system can accurately track the position and orientation of objects within its field of view (FOV). We consider the optical tracker to be the ground truth and make the reasonable assumption that it does not contain a bias. Thus, we can use the output of the optical tracking unit to estimate the bias of the IMU sensors. Specifically, we used the cosine algorithm to calculate the true values A_{cn} and M_{cn} from the Euler angles obtained from the camera data. The bias can be estimated by subtracting these true values from the sensor feedback, and then using a Finite Impulse Response (FIR) filter to attenuate the noise, as shown in Fig. 3. This is given by the following equation, where the subscripts i and c indicate measurements from the IMU and camera, respectively, and a_n are the coefficients of the FIR filter ($\sum_n a_n = 1$):

$$\begin{bmatrix} B_g \\ B_a \\ B_m \end{bmatrix} \approx \sum_n a_n \begin{bmatrix} G_{in} - G_{cn} \\ A_{in} - A_{cn} \\ M_{in} - M_{cn} \end{bmatrix} \tag{1}$$

The FDAtool was used to design the low pass FIR filter. Sampling frequency was 500, cutoff frequency was set as 20Hz and the order of the FIR filter was 20. This approach makes the following simplifying assumptions: (1) there is negligible acceleration, and (2) the latency between the IMU and optical tracker measurements is negligible. Both of these assumptions are satisfied under quasi-static conditions, where the surgeon's head is not moving very much.

2.2 Orientation Estimation

We use a Kalman Filter (KF) to estimate orientation from IMU sensor feedback, as illustrated in Fig. 4. Our state consists of a unit quaternion that represents the orientation. We use a discrete-time state equation that describes the evolution of the system state X_k, starting from the system state at the previous step X_{k-1} with the evolution law described by the matrix A and responding to system inputs u_k with the evolution law B.

$$\hat{X}_k^- = A_k \hat{X}_{k-1} + B u_k \tag{2}$$

The discrete-time matrix A_k has the following form, where $Gx(k)$, $Gy(k)$, and $Gz(k)$ are the angular velocities measured by the gyroscope, after removing the bias terms:

$$A_k = \begin{bmatrix} 1 & -Gy(k)*t & -Gx(k)*t & -Gz(k)*t \\ Gy(k)*t & 1 & Gz(k)*t & -Gx(k)*t \\ Gz(k)*t & -Gy(k)*t & 1 & -Gx(k)*t \\ Gx(k)*t & Gy(k)*t & -Gz(k)*t & 1 \end{bmatrix} \tag{3}$$

The Kalman gain matrix is calculated by:

$$\hat{K}_k = P_k^- H^T (H P_k^- H^T + R_k)^{-1} \tag{4}$$

where P_k^- is the "a priori" error covariance, H is the Jacobian matrix that relates the measurement to the system state vector, and R_k is the measurement noise covariance matrix. We compute P_k^- and H_k in the standard manner. The measurement noise covariance R presents the level of trust of the measurement. To enable the Kalman filter to adapt to track both normal-speed motion and abrupt motion, we divide R into two parts, normal measurement noise covariance R_a and adaptive noise covariance R_b:

$$R = R_a + R_b \tag{5}$$

In our approach, we use the accelerometer to sense the gravity vector. In the ideal case, the quadratic sum of the accelerometer measurements, A_x, A_y and A_z, should be g^2, where g is gravity. But, the accelerometer reading is also affected by motion (acceleration). We therefore define our error model as the difference between the quadratic sum and g^2, as follows:

$$R_b = \left| Ax^2 + Ay^2 + Az^2 - g^2 \right| \tag{6}$$

The measurement update is:

$$\hat{X}_k = \hat{X}_k^- + \hat{K}_k(Z_k - H\hat{X}_k^-) \tag{7}$$

where the orientations are converted from quaternions to Euler angles, subtracted, and then converted back to quaternions. The measurement model for the accelerometer and magnetometer readings mentioned in [6] is used to define the system measurement Z_k.

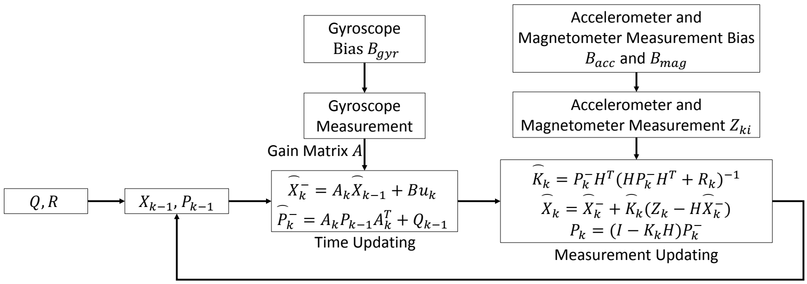


Fig. 4. Block diagram of Kalman Filter for orientation estimation

2.3 Position Estimation

When all the marker points are in the field of view, the camera can capture the position and orientation of the whole marker frame from the spatial position of the marker points. As soon as any of the marker points is blocked, the camera cannot give the orientation of the marker frame. But, it is still possible to obtain the position of any marker that is in the field of view (see Fig. 5). For the Micron Tracker, these stray markers are called `XPoints`. Other tracking systems, such as Polaris (Northern Digital, Inc., Waterloo, Canada) can also provide the positions of stray markers (i.e., those not associated with a defined rigid body).

But, although we can obtain the position, P_n of a stray marker, it is necessary to compute the position of the frame origin P_o. This requires three pieces of information: (1) identification of which marker point (n) was measured, (2) the distance vector (d_n) from this point to the frame origin, and (3) the orientation, R, of the reference frame, as illustrated in the following equation:

$$P_o = P_n - Rd_n \tag{8}$$

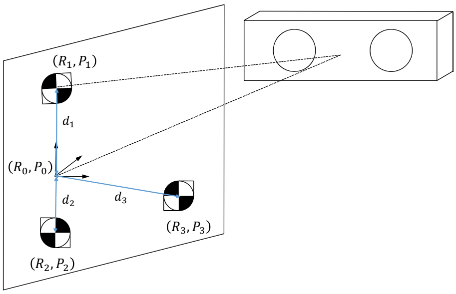


Fig. 5. When all markers are visible, the 6D pose of the reference frame, (R_0, P_0), can be calculated; if any marker is blocked, only the 3D positions, P_i, of the visible markers (i = 1, 2, or 3) are available

where d_n is obtained from the marker definition file and R is obtained from the IMU measurements, as described in Section 2.2. The identification of the marker point, n, is done using a nearest neighbor approach where the position of the marker is compared to the prior estimated positions of all markers, and the closest marker selected.

3 Experiments and Results

The hybrid tracking system (IMU and camera) was mounted on a helmet and moved in random 6-D motions. All sensor data was captured by the host PC. The raw data contained valid data for all markers (i.e., there were no actual occlusions). To test the tracking performance under partial occlusion conditions, we temporarily invalidated some of the recorded marker positions. This stopped the bias estimation process and relied on the orientation estimated from the IMU measurements (Section 2.2) to recover the frame origin, as described in Section 2.3. Note that in our system, the orientation is always estimated by the IMU, so the only effect of marker occlusion is to stop estimation of the bias terms.

We sequentially rotated the tracking unit by about 40 degrees around each of the three axes. Figure 6-left shows the estimated orientation under partial occlusion conditions. As expected, there is no noticeable impact on the estimated orientation when the markers are blocked because the orientation is computed from the inertial sensor data. Thus, the only effect of marker occlusion is that sensor biases are not compensated by the optical tracking data. To determine the typical drift rates for the inertial sensor biases, we performed an experiment where we computed the orientation from the inertial data over a 24 minute time interval, with a stationary sensor. We expressed the orientation as pitch, roll, and yaw angles, and computed the RMS errors by first subtracting the mean value from each set of angles. The resulting RMS orientation errors, expressed as pitch, roll, and yaw, were 0.0821, 0.0495, and 0.0917 degrees, respectively. This data characterized the orientation error due to both sensor bias drift and noise (Fig. 7).

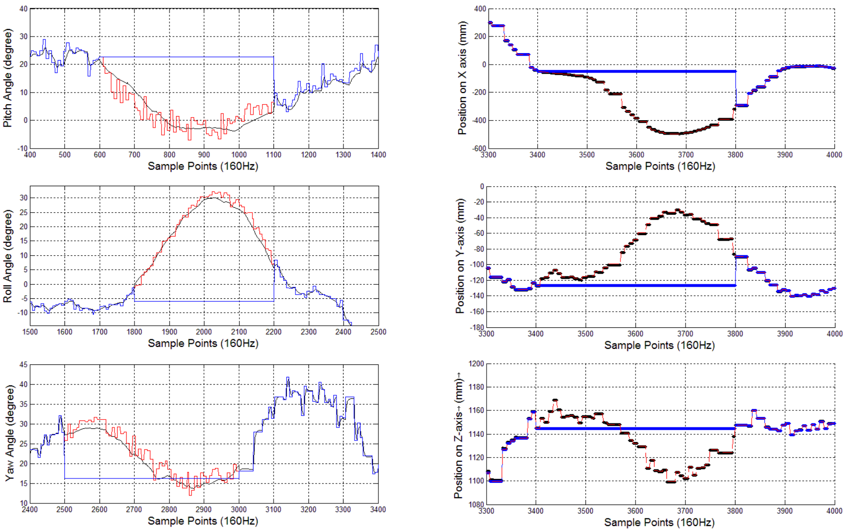


Fig. 6. Left: Orientation estimated by optical tracker (blue line) and our proposed method (black line), when we simulate partial occlusion for sample points 600-1000, 1800-2200 and 2500-3000. Red line shows the ground-truth, which is provided by the optical tracker. Right: Position estimated by optical tracker (blue line) and our proposed method (black line), when we simulate partial occlusion for sample points 3400-3800. Red line shows the ground-truth position from the optical tracker.

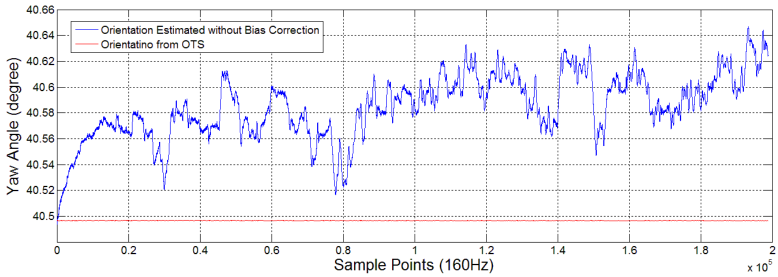


Fig. 7. Yaw angle measured by optical tracker (red) and IMU without bias correction (blue) over a 24 minute interval

We then performed experiments to demonstrate the estimation of the position under partial occlusion conditions. We manually moved the hybrid tracking system in the X, Y, and Z directions, as shown in Fig. 6-right. The position error due to partial occlusion is relatively small, even though it is affected by inaccuracies in both the position and orientation measurements. This is illustrated in Fig. 8, which shows the difference between the frame origin reported by the Micron Tracker (since it was never really blocked) and the position estimated by our method, with simulated blocking of all but one marker.

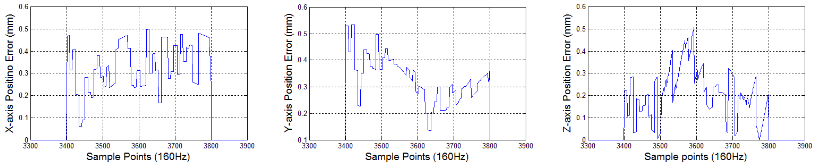


Fig. 8. Position errors for the three axes during simulated occlusion of all but one marker

4 Conclusions and Future Work

This paper presented a sensor fusion approach that uses inertial measurements to compensate for partial occlusion of optically-tracked markers. In contrast to other approaches, the orientation is always obtained from the inertial measurements, via a Kalman Filter where the system dynamics are driven by the gyroscope and the accelerometer and magnetometer provide measurement updates. Rather than discarding the orientation measured by the camera, however, we use it to estimate the bias of the inertial sensors (gyroscope, accelerometer, and magnetometer). The position is always obtained from the optical tracker, because we assume that at least one marker will be visible. If some markers are occluded, however, the optical tracker can only provide the positions of the visible markers and therefore we use the marker design geometry, in conjunction with the IMU-estimated orientation, to compute the frame position.

In practice, we expect that this sensor fusion approach will provide good accuracy over relatively long periods of partial marker occlusion. The determining factors include the stability of the estimated bias terms. If the sensor biases drift, it will be necessary to restore full line-of-sight so that the biases can be re-estimated. This is particularly important for the magnetometer bias, which can have large variations due to magnetic field disturbances.

Our experimental results showed good accuracy during the partial occlusion cases, but further improvements are possible. First, we can compensate for the small time difference between the optical tracker and IMU measurements. Second, we can better estimate the gravity vector in the presence of body accelerations, as done in our prior work that used a Kalman Filter to estimate the gravity from the "noisy" accelerometer measurements [7]. Another possibility is to estimate the acceleration from the optical tracker positions. Third, we currently compute the gyroscope bias by numerically differentiating the orientation angles measured by the camera; it would likely be better to use a Kalman Filter. We could also estimate the gyroscope bias in the orientation EKF, as in [11][7]. Fourth, our current implementation uses just a single visible marker in the partial occlusion case, but the method can be extended to use all visible markers. Finally, the results reported here relied on offline processing of the sensor data. We are currently implementing a real-time version in C++ to support online testing of the hybrid tracking system with augmented reality overlays.

Acknowledgements. We thank Jayfus Doswell and Juxtopia LLC for providing the optical see-through goggles and Dr. George Jallo for his clinical guidance. Dr. Kaisorn Chaichana performed the cadaver experiments. Changyu He received support from the Chinese Scholarship Council. Praneeth Sadda was supported by a JHU Provost Undergraduate Research Award (PURA).

References

1. Azimi, E., Doswell, J., Kazanzides, P.: Augmented reality goggles with an integrated tracking system for navigation in neurosurgery. In: Proc. IEEE Virtual Reality, Orange County, CA, pp. 123–124 (March 2012)
2. Azuma, R., Bishop, G.: Improving static and dynamic registration in an optical see-through HMD. In: SIGGRAPH, pp. 197–204 (July 1994)
3. Chai, L., Hoff, W.A., Vincent, T.: Three-dimensional motion and structure estimation using inertial sensors and computer vision for augmented reality. Presence 11(5), 474–492 (2002)
4. Hoff, W., Vincent, T.: Analysis of head pose accuracy in augmented reality. IEEE Trans. Visualization and Computer Graphics 6(4), 319–334 (2000)
5. Park, F.C., Martin, B.J.: Robot sensor calibration: Solving AX = XB on the Euclidean group. IEEE Trans. on Robotics and Auto. 10(5), 717–721 (1994)
6. Ren, H., Kazanzides, P.: Hybrid attitude estimation for laparoscopic surgical tools: A preliminary study. In: Intl. Conf. of IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5583–5586 (September 2009)
7. Ren, H., Kazanzides, P.: Investigation of attitude tracking using an integrated inertial and magnetic navigation system for hand-held surgical instruments. IEEE/ASME Trans. on Mechatronics 17(2), 210–217 (2012)
8. Sadda, P., Azimi, E., Jallo, G., Doswell, J., Kazanzides, P.: Surgical navigation with a head-mounted tracking system and display. Medicine Meets Virtual Reality (MMVR) 20, 363–369 (2013)
9. Sauer, F., Wenzel, F., Vogt, S., Tao, Y., Genc, Y., Bani-Hashemi, A.: Augmented workspace: designing an AR testbed. In: Proc. IEEE Intl. Symp. on Augmented Reality (ISAR), pp. 47–53 (October 2000)
10. Tao, Y., Hu, H., Zhou, H.: Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation. Intl. J. of Robotics Research 26(6), 607–624 (2007)
11. Tobergte, A., Pomarlan, M., Hirzinger, G.: Robust multi sensor pose estimation for medical applications. In: Proc. IEEE/RSJ Intl. Conf. on Intell, St. Louis, MO, pp. 492–497 (October 2009)
12. Vaccarella, A., De Momi, E., Enquobahrie, A., Ferrigno, G.: Unscented Kalman filter based sensor fusion for robust optical and electromagnetic tracking in surgical navigation. IEEE/ASME Trans. on Instrumentation and Measurement 62(7), 2067–2081 (2013)
13. You, S., Neumann, U., Azuma, R.: Hybrid inertial and vision tracking for augmented reality registration. In: Proc. IEEE Conf. on Virtual Reality, Houston, TX, pp. 260–267 (March 1999)