

Kinematic model aided inertial motion tracking of human upper limb

Huiyu Zhou and Huosheng Hu

Department of Computer Science, University of Essex, Colchester CO4 3SQ, U.K.

E-mail:zhou@essex.ac.uk, hhu@essex.ac.uk

Abstract— A new motion tracking framework has been developed to estimate the position and orientation of human upper limb. This method fuses data from on-board accelerometers and gyroscopes, which are accommodated in a commercially available inertial sensor MT9. Human upper limb motion can be represented by a kinematic chain in which six joint variables are to be considered: three for the shoulder and three for the elbow. Based on measurements of the inertial sensor placed on the wrist, we then obtain the positions of the wrist and elbow. An extended Kalman filter then fuses the data from these sensors in order to reduce errors and noise in measurements. Preliminary results demonstrate the favorable performance of the proposed strategy.

Index Terms— Motion tracking, upper limb, rehabilitation, inertial sensor, Kalman filter.

I. INTRODUCTION

A. Motivation

It is well recognised that a suitable motor-rehabilitation regimen can facilitate significant functional recovery for post-stroke patients. Normally, this regimen demands ongoing attention and monitoring by either a therapist or an intelligent system. In the last ten years, great interest has been generated toward designing low cost rehabilitation systems, which may work toward avoiding empirical difference of individual therapists. In addition, a successfully self-contained system is applicable to a home-based environment where a rehabilitative course can be conducted straightaway. This can further reduce the expenses of post-stroke rehabilitation, compared to a hospital-based mechanism.

A typical rehabilitation system presumably consists of a motion tracker and a human-computer interface. The former immediately tells the position of an object during trajectory, whilst the latter is built up for undertaking display, data process and communication tasks, etc. In this paper, we will only focus on the former issue. Unfortunately, due to complexity of human movements the problem of achieving accurate and reliable measurements still remains open.

To address this problem, a number of sensing technologies for motion estimation have been developed to improve the performance of motion trackers. These technologies consist of mechanical, inertial, acoustic, magnetic, optical, and radio frequency sensing. However, despite its merits each approach

has inevitable limitations. For example, mechanical components are large in size and not lightweight, electromagnetic sensors cannot reach a target at a distance, and optical sensors easily suffer the occlusion problem. Thus, a successful tracking system requires appropriate selection of potentially used sensors.

B. Previous work

Accelerometers and gyroscopes have been frequently used in navigation and augmented reality modelling [1], [2]. These inertial sensors come up with small size and can provide fairly accurate kinematics measurements. Over the last few years, inertial sensors are also applied into the implementation of human motion tracking [3], [4]. However, position and angle of an engaged inertial sensor cannot be correctly determined due to fluctuation of offsets and measurement noise which leads to integration drift. Therefore, exploring drift-free inertial systems is the ultimate goal in the field of human motion tracking. In this section, existing motion tracking systems will be briefly summarised. In particular, techniques relative to drift reduction will be mainly concerned.

Luinge [5] introduced the design and performance of a Kalman filter to estimate inclination from signals of a tri-axial accelerometer. This design is based on the assumption where the acceleration has a lower frequency bound. Another assumption is that the magnitude of the gravity is 1 *g*. Empirical evidence shows that inclination errors are less than 2 degrees. As a direct extension, an inertial measurement unit (IMU) (a tri-axial accelerometer and a tri-axial gyroscope) was then proposed, combining with a Kalman filter to estimate orientation of human body segments. Unfortunately, the problem of integration drift around the global vertical direction still appears.

Technologies have been developed in an attempt to fit military requirements, i.e. positioning aircraft. Interestingly, these strategies can be directly used in the application of human motion tracking. Foxlin *et al.* [7] revealed the first prototype of the FlightTracker, which is designed to overcome the shortcomings addressed in a hybrid tracking platform that fuses ultrasonic range measurements with inertial tracking. They developed a differential inertial measurement equation

of the pilot's head motion relative to the aircraft, which can then have its drift corrected by periodic optical measurements of the head position relative to the aircraft. Experimental results show that a drift is slower than 1 mm/s or 1 deg/min.

Inertial sensors can provide important cues about the observed scene structure, i.e. vertical and horizontal references. This information can be incorporated alongside a visual modality so as to simplify the 3-D reconstruction of the observed world. Lobo and Dias [8] presented a framework for using inertial sensor data in vision systems. Using the vertical reference provided by the inertial sensors, the image horizon line can be determined. The camera's focal length will be recovered using just a vanishing point. Knowing the geometry of a stereo rig and its pose from the inertial sensor, people can obtain the collineation of level planes, assuming that sufficient constraints have been placed on vertical features and leveled planar patches. The main weakness of this method is that vertical world features may not be available in some circumstances, e.g. flat surfaces, and cluttered scenes, etc. Similar work also has been described in [9], [10].

In this paper, we propose an inertial sensor based motion tracker, incorporating a Kalman filter that fuses measurements from tri-axial accelerometers and tri-axial gyroscopes. This strategy is based on a novel kinematic model and targets at removing undesirable biases or noise. At this stage, the proposed system is developed for tracking movements of a two-joint upper limb. However, we believe that the concept buried in the proposed system can be easily transplanted onto the case of tracking lower limb motion.

The rest of this paper is arranged as follows. Section II introduces the analytical solution to the kinematic problem of human arm movements. To minimise errors or inconsistency, a Kalman filter is adopted in Section III. In Section IV we present a preliminary evaluation on the proposed system against a commercial motion tracking system. Finally, this paper is summarised in Section V.

II. KINEMATIC MODELS

Human arm movements can be represented by kinematic chains. The kinematic chain of our concern consists of six joint variables, i.e. three for the shoulder and the others for the elbow. Note, the proposed system is outstanding from any other existing 7 degree-of-freedom limb model [11], where in the latter the motion of the elbow joint only has one degree-of-freedom. The Denavit-Hartenberg (DH) convention and methodology [12], [13] can be used to derive the arm kinematics.

A. Forward kinematics

The forward kinematics specify the Cartesian position and orientation of the local frame attached to the human arm relative to the base frame which is attached to the still joint (shoulder). They are provided by multiplying a series of matrices parameterised by joint angles. The coordinate

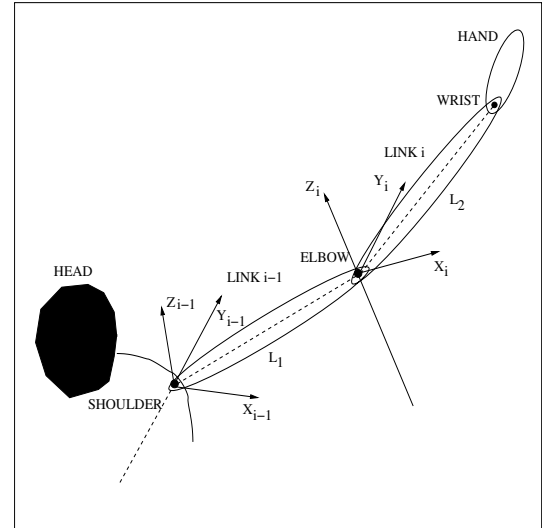


Fig. 1. Kinematics of a two-joint human arm.

frames and their transformation are illustrated in Fig. 1. It is intended to construct the transform that defines frame i to the frame $i-1$, where $i = 1, 2, 3$. In fact, there are two links in this particular case, whereas an inertial sensor (MT9-B, Xsens Dynamics Technologies, Netherlands) is presumably placed upon an area adjacent to the human wrist joint.

Based on the DH convention, we have a general transformation from joint i to joint $i+1$ as

$${}^{i-1}T_i = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & l_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1}d_i \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1}d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where c and s stand for the cosine and sine functions respectively. θ and α are rotating angles in the local frame i .

Therefore, one will have multiplied link transformations as follows

$${}^0T_2 = {}^0T_1 {}^1T_2 = \begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & 0 \\ s\theta_1 c\alpha_0 & c\theta_1 c\alpha_0 & -s\alpha_0 & 0 \\ s\theta_1 s\alpha_0 & c\theta_1 s\alpha_0 & c\alpha_0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c\theta_2 & -s\theta_2 & 0 & L2 \\ s\theta_2 c\alpha_1 & c\theta_2 c\alpha_1 & -s\alpha_1 & -s\alpha_1 L1 \\ s\theta_2 s\alpha_1 & c\theta_2 s\alpha_1 & c\alpha_1 & c\alpha_1 L1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

where L_1 is the distance between the centres of the shoulder and the elbow joints, and L_2 is the distance between the elbow joint and the inertial sensor.

From Eq. (2), the position and orientation of the end-effector (the inertial sensor) can be determined if all the

joint angles have been a-priori available. In a real application, however, the situation is totally reverse. In other words, the position or orientation of the end-effector is known by some means but the joint angles or the position of the elbow need to be derived. To accomplish this, inverse kinematics of human arm are necessarily studied.

B. Inverse kinematics

Let the coordinates of the elbow joint and the sensor be denoted by $p_1(x_1, y_1, z_1)$ and $p_2(x_2, y_2, z_2)$, where the latter can be obtained by using the following approximation

$$\mathbf{p}_2(x_2, y_2, z_2)_{i+1} = \mathbf{p}_2(x_2, y_2, z_2)_i + \int_{t_i}^{t_{i+1}} \mathbf{a} dt \quad (3)$$

where t_i and t_{i+1} are two time instants, and \mathbf{a} is the combinatorial acceleration represented as

$$\mathbf{a} = R\tilde{\mathbf{a}} + \mathbf{g} \quad (4)$$

where R is the rotation matrix relating the sensor body-fixed frame to the earth bound frame, $\tilde{\mathbf{a}}$ is the tri-axial acceleration vector detected in the sensor body-fixed frame, and \mathbf{g} is the gravity vector [7].

Then, one can have projected coordinates on three orthogonal planes, i.e. x - y plane

$$\begin{cases} x_2 = \sqrt{L_2^2 - (z_2 - z_1)^2} \cos \phi_z + x_1 \\ y_2 = \sqrt{L_2^2 - (z_2 - z_1)^2} \sin \phi_z + y_1 \end{cases} \quad (5)$$

y - z plane

$$\begin{cases} y_2 = \sqrt{L_2^2 - (x_2 - x_1)^2} \cos \phi_x + y_1 \\ z_2 = \sqrt{L_2^2 - (x_2 - x_1)^2} \sin \phi_x + z_1 \end{cases} \quad (6)$$

x - z plane

$$\begin{cases} x_2 = \sqrt{L_2^2 - (y_2 - y_1)^2} \cos \phi_y + x_1 \\ z_2 = \sqrt{L_2^2 - (y_2 - y_1)^2} \sin \phi_y + z_1 \end{cases} \quad (7)$$

where ϕ_x , ϕ_y and ϕ_z are the Euler angles of the sensor around x -, y - and z -axis, respectively.

Our next goal is to estimate the elbow position. Indeed, the solution for the elbow position (x_1, y_1, z_1) is sought using the conditions as follows:

If $\phi_y = \pm \frac{\pi}{2}$, then

$$x_1 = x_2 \quad (8)$$

When $-\frac{\pi}{2} < \phi_y < \frac{\pi}{2}$, or $\frac{3\pi}{2} < \phi_y < 2\pi$

$$x_1 = x_2 - L_2 \sqrt{\frac{\cos^2(\phi_z) + \sin^2(\phi_x) \sin^2(\phi_z)}{\sin^2(\phi_x) \sin^2(\phi_z) + \frac{1}{\cos^2(\phi_y)}}} \quad (9)$$

Otherwise, if $\frac{\pi}{2} < \phi_y < \frac{3\pi}{2}$, or $-\frac{3\pi}{2} < \phi_y < -\frac{\pi}{2}$

$$x_1 = x_2 + L_2 \sqrt{\frac{\cos^2(\phi_z) + \sin^2(\phi_x) \sin^2(\phi_z)}{\sin^2(\phi_x) \sin^2(\phi_z) + \frac{1}{\cos^2(\phi_y)}}} \quad (10)$$

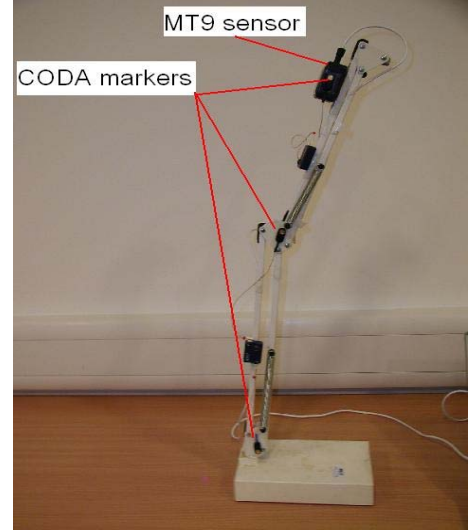


Fig. 2. A desk lamp attached with a MT9 sensor and three CODA markers.

The tri-axial gyros in the inertial sensor supply the turning rates, which can be integrated to be angular variations using the updating equation as

$$\dot{\Phi} = E_b^n \omega_b \quad (11)$$

where $\dot{\Phi}(\phi_x, \phi_y, \phi_z)$ is the derivative of the Euler angles in terms of time, ω_b is the reading of the gyros, and the rotation rate transformation matrix E_b^n from the body to world frame is given as

$$E_b^n = \begin{bmatrix} 1 & s\phi_x \tan \phi_y & c\phi_x \tan \phi_y \\ 0 & c\phi_x & -s\phi_x \\ 0 & s\phi_x \sec \phi_y & c\phi_x \tan \phi_y \end{bmatrix} \quad (12)$$

Finally, the solutions for y and z will be explicitly available, where

$$\begin{cases} y_1 = y_2 - \sqrt{L_2^2 - (x_2 - x_1)^2} \cos(\phi_x) \\ z_1 = z_2 - \sqrt{L_2^2 - (x_2 - x_1)^2} \sin(\phi_x) \end{cases} \quad (13)$$

This technique allows the signs of x_1 , y_1 and z_1 to be correctly determined whilst removing the redundancy of the non-linear system.

To avoid the soft tissue effect that may mislead the analysis, in this paper we intend to test the proposed strategy using a desk lamp that contains a two-joint arm moving up-down and rotating around the base (Fig. 2), where the inertial sensor is placed at a point as the ‘‘wrist’’, and a nearer joint is regarded as the ‘‘elbow’’. A part of the results is shown in Fig. 3, where the solid lines are the ground truth, and the dotted lines are from the proposed kinematic models. Clearly, errors and noise exist throughout the whole data set. Therefore, a dynamic filtering strategy needs to be explored in order to remove errors or noise. Kalman filtering is considered in this system due to its plausible concepts and consistent performance.

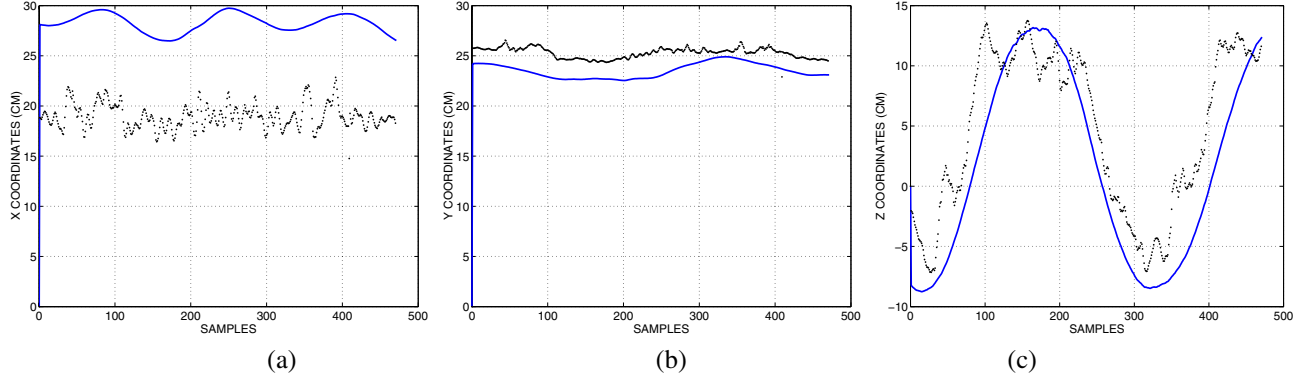


Fig. 3. Position estimates of the “elbow” on the lamp arm using the proposed kinematic models (dotted lines) against the ground truth (solid lines).

III. EXTENDED KALMAN FILTER

The Kalman filter estimates a process by taking a form of feedback control: the filter updates the process state recursively and obtains feedback in the form of measurements. In the process, the Kalman filter can be divided into two steps: *time update* and *measurement update*. The former provides a prediction, and the latter acts as a corrector, incorporating a new measurement into the *a priori* estimate to conduct an improved *a posteriori* estimate. Since the relation between measurements and the process is non-linear, an extended Kalman filter (EKF) is actually used.

We represent arm motion with a state vector

$$\mathbf{x} = [\mathbf{a}, \Phi, \mathbf{p}_w, \mathbf{p}_e]^T \quad (14)$$

where \mathbf{p}_w and \mathbf{p}_e the Cartesian coordinates of the wrist and elbow respectively. The arm’s dynamic motion can be expressed as

$$\mathbf{x}_{i+1} = \mathbf{A}_i \mathbf{x}_i + \mathbf{w}_i \quad (15)$$

where \mathbf{w}_i is the process noise, and \mathbf{A}_i is the state transition matrix that includes the relationships as follows

$$\begin{cases} \mathbf{a}_{i+1} = \mathbf{a}_i \\ \Phi_{i+1} = \Phi_i \end{cases} \quad (16)$$

The transition matrix also consists of the kinematic models, which have been discovered in Section II B, e.g. Eqs. (3)-(4), (8)-(10), and (13).

A measurement model is required to predict the ideal noise-free response of the sensors, given the filter’s current estimate of the target state. Therefore, the measurement vector \mathbf{z}_i and its corresponding measurement function \mathbf{h} are

$$\mathbf{z}_i = \mathbf{h}(\mathbf{x}_i) + \mathbf{v}_i \quad (17)$$

where \mathbf{v}_i is the measurement noise vector, and \mathbf{z}_i is described as

$$\mathbf{z}_i = \tilde{x}_{1_i}^2 + \tilde{y}_{1_i}^2 + \tilde{z}_{1_i}^2 - L_1^2 \quad (18)$$

where

$$\begin{cases} \tilde{x}_{1_i} = \frac{x_1 L_1}{\sqrt{x_{11}^2 + y_{11}^2 + z_{11}^2}} \\ \tilde{y}_{1_i} = \frac{y_1 L_1}{\sqrt{x_{11}^2 + y_{11}^2 + z_{11}^2}} \\ \tilde{z}_{1_i} = \frac{z_1 L_1}{\sqrt{x_{11}^2 + y_{11}^2 + z_{11}^2}} \end{cases} \quad (19)$$

where (x_{11}, y_{11}, z_{11}) are the coordinates of the first sample in the sequence. Ideally, \mathbf{z}_i must be a null vector. Due to noise and errors, we consider the mean of the history of \mathbf{z}_i (from 0 to $i - 1$) as *synthetic* measurements.

The corresponding Jacobian function of the measurement function \mathbf{h} is

$$\mathbf{H}_{i,j} = \frac{\partial \mathbf{h}(\mathbf{x}_i)}{\partial \mathbf{x}_j} \quad (20)$$

The process and measurement noise covariance matrices \mathbf{Q} and \mathbf{R} have been presumably *stationary* although they are actually varied. Doing this allows to dramatically reduce uncertainty in computation. They are given by

$$\begin{cases} \mathbf{Q} = E(\mathbf{w}_i \mathbf{w}_i^T) \\ \mathbf{R} = E(\mathbf{v}_i \mathbf{v}_i^T) \end{cases} \quad (21)$$

Given an initial state estimate \mathbf{x}_0 and error covariance estimate \mathbf{P}_0 , the extended Kalman filter updates the state estimate $\tilde{\mathbf{z}}_k$ and the associated state covariance matrix \mathbf{P}_k as follows [14]:

The time update equations are given by

$$\begin{cases} \mathbf{x}_{i+1}^- = \mathbf{A}_i \mathbf{x}_i \\ \mathbf{P}_{i+1}^- = \mathbf{A}_i \mathbf{P}_i \mathbf{A}_i + \mathbf{Q} \end{cases} \quad (22)$$

The measurement update equations are

$$\begin{cases} \mathbf{K} = \mathbf{P}_{i+1}^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_{i+1}^- \mathbf{H}^T + \mathbf{R})^{-1} \\ \mathbf{x}_{i+1} = \mathbf{x}_{i+1}^- + \mathbf{K} (\mathbf{z}_{i+1} - \tilde{\mathbf{z}}_{i+1}) \\ \mathbf{P}_{i+1} = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}_{i+1}^- \end{cases} \quad (23)$$

where $\tilde{\mathbf{z}}_i$ is the predicted measurement using Eq. (22). \mathbf{K} is the Kalman gain that weights the measurement difference in order to minimise errors on state estimation.

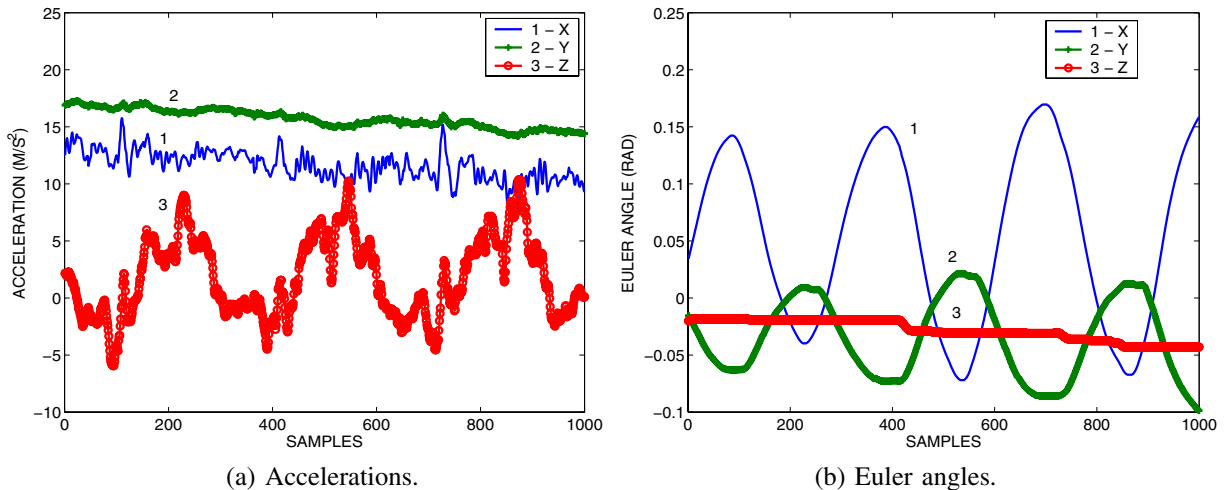


Fig. 4. Acceleration and Euler angle measurements of the MT9 sensor attached at the “wrist”.

In the implementation, to avoid complicated configuration we have undertaken $\sin(\gamma) = \gamma$ and $\cos(\gamma) = 1$ when $\gamma \rightarrow 0$. This guarantees \mathbf{H} to be further simplified for easy manipulation.

IV. RESULTS AND DISCUSSION

In this section, the performance of the proposed kinematic models (Section II) and the Kalman filtering is in comparison to that of the commercially available motion tracking system CODA CX1 using the lamp arm shown in Fig. 2. Before the evaluation starts, an experimental environment needs to be set up. The plane that the lamp’s arm lies on is perpendicular to the CODA system. The viewing direction of the CODA system is aligned with the Z-axis of the world coordinate system, and the X-axis is vertical to the floor. Both systems run at a sampling rate of 100 Hz. In addition, a proper alignment must be achieved between these involved systems so as to start numerical comparison.

We have obtained the measurements of accelerations and Euler angles, illustrated in Fig. 4. It is evident that the readings of accelerations are encoded with significant noise. We have found that this noise is mainly due to inconsistent motion of the lamp joints. Therefore, motion estimates along the three orthogonal axes are believed to be inevitably noisy. In Fig. 5, we observe a few interesting points, which can be summarised as follows.

- The average deviations in the Z estimates of the wrist and elbow by the Kalman filtering are 1.6 and 1.2 cm respectively against -5.4 and -3.8 cm by the kinematic models (Fig. 5 (c) and (f)).
- Both methods have very similar accuracy in the Y and Z estimates for both positions, whose average biases span from 1.0 to 5.0 cm (see Fig. 5 (a), (b), (d) and (e)). In the overall results, no drift turns up significantly.

This indicates that the Kalman filtering readily achieves error and noise minimisation in the case of significantly periodic and regular motion. However, this filtering technique fails to identify noise and errors from a fairly noisy background. Theoretically, to improve the performance of the Kalman filter, we take into account the issues as follows: (1) The process and measurement noise vectors have to be instantly updated. Thus, an adaptive noise estimation will be suggested. (2) The Kalman filtering is based on the assumption of Gaussian distribution. Hence, non-Gaussian noise will not be justified in this context. (3) Noise frequency spreads out in a wide bandwidth, so the Kalman filter is reluctant to follow variations. A band-pass filter can be applied before data goes to the Kalman filtering. (4) Simplification of \mathbf{H} may deteriorate motion estimates in some circumstances.

V. CONCLUSIONS AND FUTURE WORK

We have presented a kinematic model aided inertial tracking system that integrates kinematics of human arm motion and a dynamic filtering strategy. To justify the proposed fusion method, the 3-D positions of a desk lamp joints are estimated immediately after measurements of accelerations and Euler angles. Compared to the commercial tracking system CODA that uses markers, our system has the advantage that it is dramatically cheaper and can obtain reasonable accuracy in motion estimation with a simple setting.

The future work will be addressed to apply the ideas presented here in tracking movements of real human arms. The two-joint arm movements have more complex motion patterns and higher degrees of freedom. To reduce potentially accumulated integration errors (drifts), we intend to fuse the inertial sensing data and visual information. This fusion idea most likely improves the accuracy of localisation of human arm movements.

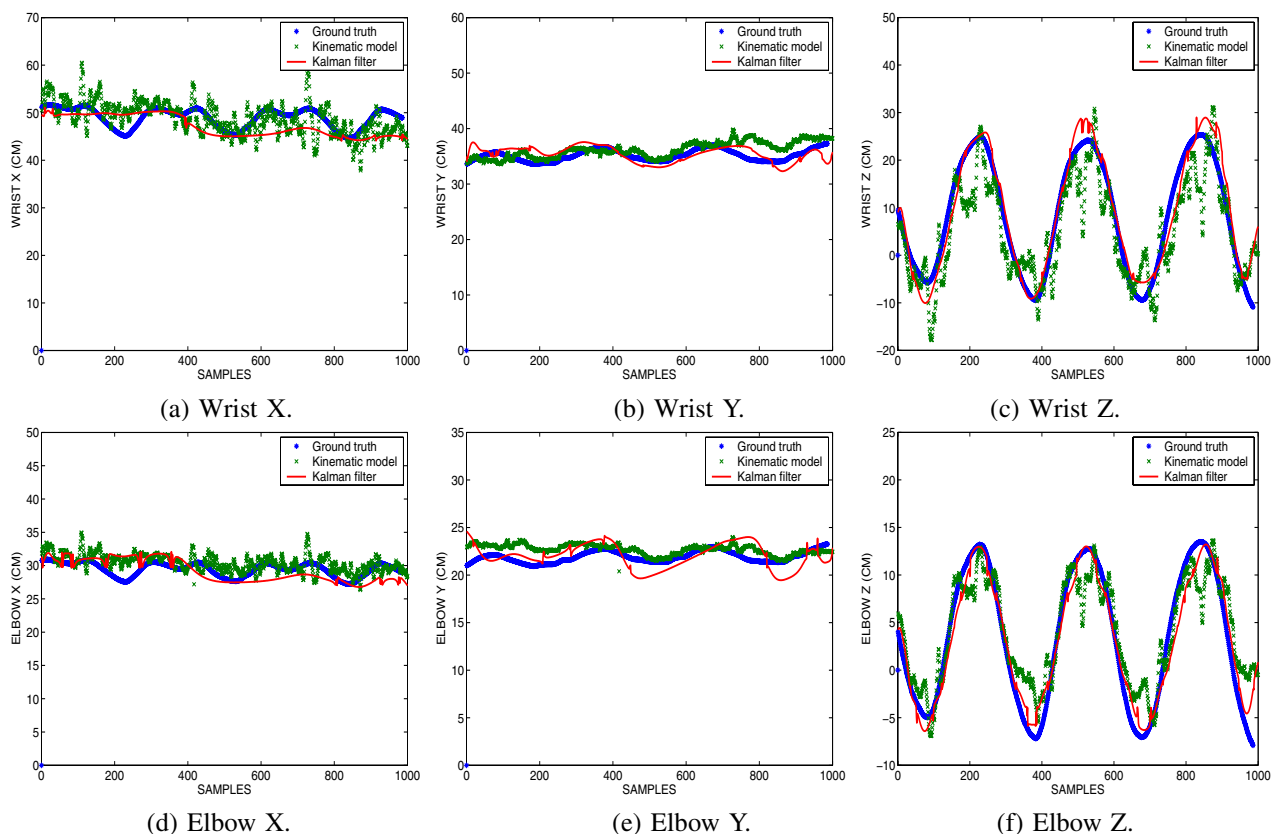


Fig. 5. Position estimates of lamp arm movements.

ACKNOWLEDGEMENTS

This SMART Rehabilitation Project is supported by the UK EPSRC under Grant GR/S29089/01. The authors would like to thank the Charnwood Dynamics Ltd. that kindly provided the CODA CX1 for the evaluation. The other members of EPSRC EQUAL SMART Rehabilitation Consortium are also acknowledged for valuable discussion.

REFERENCES

- [1] E. Nebot and H. Durrant-Whyte, "Inertial calibration and alignment of low cost inertial navigation units for land vehicle applications," *J. of Robotics Systems*, vol. 16, no. 2, pp. 81–92, 1999.
- [2] S. You, U. Neumann, and R. Azuma, "Hybrid inertial and vision tracking for augmented reality registration," in *Proc. of IEEE VR'99, Houston, TX*, Mar. 1999, pp. 260–267.
- [3] J. Zhao and N.I. Badler, "Inverse kinematics positioning using non-linear programming for highly articulated figures," *ACM Transactions on Graphics*, vol. 13, no. 4, pp. 313–336, 1994.
- [4] Y. Uno, M. Kawato, and R. Suzuki, "Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model," *Biological Cybernetics*, vol. 61, pp. 89–101, 1989.
- [5] H.J. Luinge, *Inertial sensing of human movement*, PhD Thesis, Twente University Press, Netherlands, 2002.
- [6] B. Delaney, *The market for visual simulation & virtual reality systems, fifth edition*, CyberEdge Information Services, New York, 2002.
- [7] E. Foxlin, Y. Altshuler, L. Naimark, and M. Harrington, "Flighttracker: a novel optical/inertial tracker for cockpit enhanced vision," in *Proc. of ISMAR*, Nov. 2-5, 2004.
- [8] J. Lobo and J. Dias, "Vision and inertial sensor cooperation using gravity as a vertical reference," *IEEE Trans. on Pattern. Analy. and Mach. Intel.*, vol. 25, no. 12, pp. 1597–1608, 2003.
- [9] F. Panerai, G. Metta, and G. Sandini, "Visuo-inertial stabilization in pace-variant binocular systems," *Robotics and Autonomous Systems*, vol. 30, no. 1-2, pp. 195–214, 2000.
- [10] T. Mukai and N. Ohnishi, "The recovery of object shape and camera motion using a sensing system with a video camera and a gyro sensor," in *Proc. Seventh Int'l. Conf. Computer Vision*, Sep. 1999, pp. 411–417.
- [11] D. Tolani, A. Goswami, and N.I. Badler, "Real-time inverse kinematics techniques for anthropomorphic limbs," *Graphical Models*, vol. 62, pp. 353–388, 2000.
- [12] J.J. Craig, *Introduction to robotics: mechanics and control*, Addison-Wesley, MA, 1989.
- [13] S.B. Niku, *Introduction to robotics: analysis, systems, applications*, Prentice Hall, 2001.
- [14] R.G. Brown and P.Y.C. Hwang, *Introduction to random signals and applied kalman filtering*, John Wiley & Sons, New York, 1992.