Chapter 13

3D Arm Motion Tracking for Home-based Rehabilitation

Y. Tao and H. Hu

13.1 Introduction

This paper presents a real-time hybrid solution to articulated 3D arm motion tracking for home-based rehabilitation by combining visual and inertial sensors. The Extended Kalman Filter (EKF) is used to fuse the different data modalities from two sensors and exploit complementary sensor characteristics. Due to the non-linear property of the arm motion tracking, upper limb geometry information and the pin-hole camera model are used to improve the tracking performance. The experimental results show the real-time performance and reasonable accuracy of our system.

Recent developments in visual tracking tend to track 3D arm motion using a single camera (Goncalves *et al.,* 1995). Most visual tracking algorithms of arm motion are however computationally expensive and have low tracking accuracy. They normally need manual initialisation, and their motion model or structure model is difficult to generalise. In contrast, inertial tracking has attracted much attention recently in human motion analysis (Bachmann *et al*, 1999). Most of the work focused on using only accelerometers (or gyros) attached to a human body to detect and analyse the human motion. Zhou and Hu (Zhou and Hu, 2005) used both acceleration and rate of turn to track 3D human arm motion in real time. But inertial tracking suffers from the drift problem.

Integrating visual and inertial sensors into a motion tracking system proves to be of particular value to achieve robust and applicable data (Foxlin *et al.,* 2004). But existing work has mainly been done on tracking the pose of a rigid object. Our interest in this paper is to integrate visual and inertial sensors for tracking the 3D motion of articulated objects, *e.g.,* human upper limbs. The purpose is to develop a 3D motion tracking model for home-based rehabilitation projects, which should be cheap, accurate and in real time. Traditionally stroke patients rely on the help of physiotherapists or well-trained carers to diagnose their rehabilitation activities during physiotherapy. We propose to develop an intelligent system to support the stroke patients in doing rehabilitation at home so that the burden on hospitals and physiotherapists can be reduced. The motion tracking method proposed in this paper will serve as one of the major models for the home based intelligent system.

The rest of the paper is organised as follows: Section 13.2 presents the system overview. The integration of vision and inertial sensors for tracking arm movements is presented in Section 13.3. Our experimental results on the performance test of the proposed hybrid motion tracking method are shown in Section 13.4. Finally, conclusions and future work are presented in Section 13.5.

13.2 System Overview

Inspired by Johansson's psychology experiments (Johansson, 1973) using moving light displays (MLDs), we focus on tracking three arm joints: shoulder, elbow and wrist. It is assumed in our method that the shoulder joint is fixed during the motion and its position is known *a priori*. This has been widely used in many arm motion tracking systems (Goncalves *et al.,* 1995). It is based on the fact that the shoulder joint moves little with the arm motion. We further assume that the length of forearm L_2 and upper arm L_1 are known *a priori*. The tracking task is now to track the elbow and wrist joints in 3D.

The system configuration and related coordinate systems for the arm motion tracking using the proposed hybrid method are shown in Figure 13.1, including three coordinate frames: a camera frame (C), an inertial sensor frame (B), and a world reference system (W).

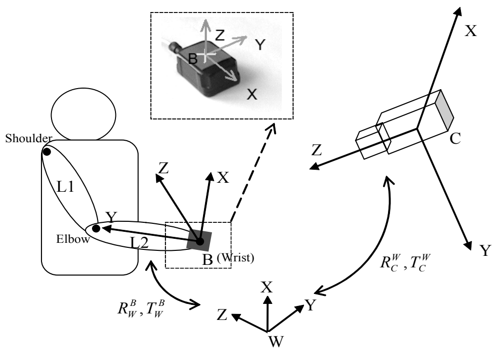


Figure 13.1 System configuration

13.3 Data Fusion & Accuracy Improvement

13.3.1 System State Vector

Normally the arm pose can be represented by a state vector with six degrees of freedom, *i.e.,* three coordinates of elbow joint position $PE = (PE_x, PE_y, PE_z)$ and three coordinates of wrist joint position $PW = (PW_x, PW_y, PW_z)$. The size of the state vector can be reduced from six to three, represented by Equation (13.1):

$$X(t) = [PW(t)] = \lfloor PW_x(t) \quad PW_y(t) \quad PW_z(t) \rfloor \tag{13.1}$$

The main idea is to convert the elbow and wrist joint tracking problem to the position tracking of the wrist joint only. The elbow joint is then inferred from wrist joint tracking and arm geometry information. The detailed procedures are as follows: (1) The elbow joint in local frame B is constant and can be represented as $PE_B = (0, L_2, 0)$; (2) Given the pose R_W^B, T_W^B of the wrist joint with respect to W, the elbow joint position can be represented in W as follows: $PE_W = R_W^B * PE_B + T_W^B$.

It is clear that the orientation and position R_W^B, T_W^B of the wrist joint with respect to world frame W are the variables to be estimated. The MT9 inertial sensor can measure the orientation R_W^B directly and our previous work (Tao *et al.,* 2005) shows it is quite accurate. We use the unit quaternion $q = (q_w, q_1, q_2, q_3)$ to represent the orientation information R_W^B in this paper. The remaining unknown variable is T_W^B, which has the same meaning as *PW,* and needs to be estimated.

13.3.2 Sensor Fusion

We assume a constant acceleration motion model for the wrist position tracking. Inertial data is used as input signals in the dynamic model of the system:

$$X(t+1) = \begin{bmatrix} P(t+1) \\ v(t+1) \end{bmatrix} = \begin{bmatrix} P(t) + v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 \\ v(t) + a(t)\Delta t \end{bmatrix} \tag{13.2}$$

where P is used to represent PW and v represents $\dot{PW}$ for simplicity. *a(t)* is the acceleration of the moving object represented in the world frame W. It is converted from the inertial output $a^B = \{a_x^B, a_y^B, a_z^B\}$ by the equation: $a(t) = q(t) * a^B(t) * q(t)' - g$, where $*$ denotes the quaternion multiplication; g is the gravity; and *q(t)* is the unit quaternion representing the orientation of the frame B relative to the reference frame W.

Figure 13.2 shows a colour marker attached to a MT9 sensor and used for visual tracking of the sensor position. The centre of the colour patch is regarded as the origin of the inertial local frame B, and it is also the 2D image projection of the wrist joint. The reason for using colour is to achieve real-time performance and robustness to occlusion.

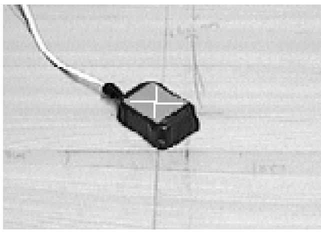


Figure 13.2 Colour tracking marker

The visual measurements $z = (I_x, I_y)$ of the colour marker are related to the state vector of the system, which is obtained by the measurement equation (13.3).

$$Z(t) = \begin{bmatrix} I_x(t) \\ I_y(t) \\ 1 \end{bmatrix} = KI * \begin{bmatrix} \dfrac{P_{C,x}(t)}{P_{C,z}(t)} \\ \dfrac{P_{C,y}(t)}{P_{C,z}(t)} \\ 1 \end{bmatrix} = KI * \left(R_C^W * \begin{bmatrix} P_x(t) \\ P_y(t) \\ P_z(t) \end{bmatrix} + T_C^W \right) \qquad (13.3)$$

where *KI* is the camera intrinsic parameter and is calibrated offline; R_C^W, T_C^W is the pose of the reference frame *W* represented in the camera frame *C*; $P_C = (P_{C,x}, P_{C,y}, P_{C,z})$ is the wrist joint represented in the camera frame *C* and *P* is the state vector to be estimated.

The sensor fusion is implemented with the EKF that operates in a predictor-corrector model, as shown in Figure 13.3. Note that $\hat{X}(t)$ is the state estimate of the system, $\hat{X}(t)^-$ is the predicted state vector, and *Z(t)* is the visual measurements at time *t*. It is assumed in our experiments that the accelerometer measurements are Gaussian noise and isotropic with $\sigma = 0.1 m/s^2$ and the visual measurement noise is also assumed to be Gaussian and isotropic with $\sigma = 2 pixels$.

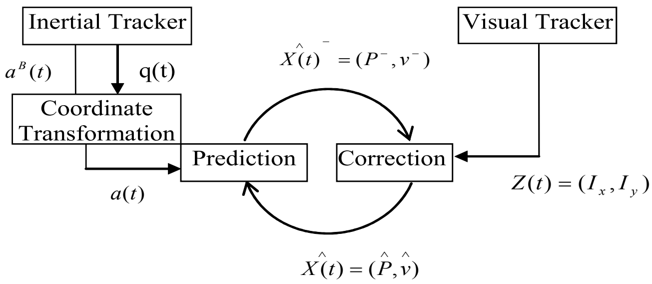


Figure 13.3 Sensor fusion framework

13.3.3 Accuracy Improvement

The EKF fusion method gives reasonable results, but the accuracy is not enough when inputs are noisy. Therefore, an optimisation method is adopted here to improve the tracking accuracy:

$$f(X(t)) = \frac{1}{2}((PS - PE(X(t))^2 - l^2)^2 + (I_x - proj(X(t))_x)^2 + (I_y - proj(X(t))_y)^2) \quad (13.4)$$

where the first term means the elbow joint lies on a sphere surface, whose centre is located in the shoulder joint and the radius is the upper arm length. The second and third terms are to restrict the wrist joint so it lies on a back-projected line through the camera projection centre and the 2D image position of colour tracking.

Now, the task is to find *X (t)*, which minimises (13.5):

$$X(t)^+ = \arg\min_{X(t)} \{f(X(t))\} \quad (13.5)$$

This is a non-linear least square problem. The famous optimisation algorithm Levernberg-Marquardt (LM) method is employed here to find the system's state vector *X(t)* at time *t* that minimises the cost function. We combine the sensor fusion method in the last section and the LM algorithm to achieve the good tracking accuracy. The procedures are as follows:

Fuse the data from inertial and visual sensors using the EKF algorithm.

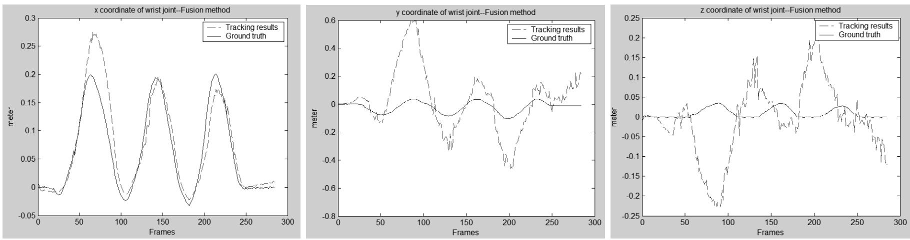
The result from the EKF method is used as the initial value into (13.4) to find the minimum $X(t)^+$ that satisfies (13.5).

13.4 Experimental Results

In this experiment, the camera is fixed and the inertial sensor is attached to the wrist joint. As shown in Figure 13.4, the subject performs a circular motion with a radius of 10 cm.

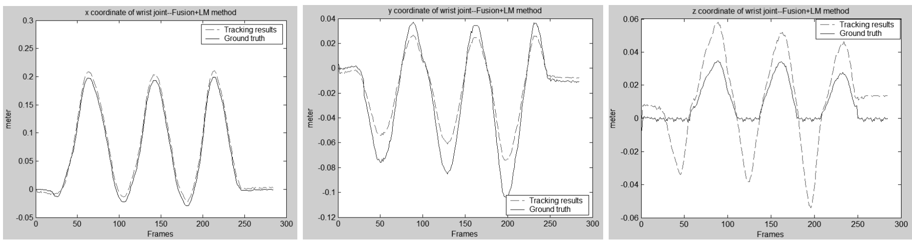
Figure 13.4 A circular motion

Figure 13.5 and Figure 13.6 show the results of wrist joint position from the fusion method and the fusion plus optimisation method respectively. More specifically, Figure 13.5 shows the results from the EKF estimator that can successfully capture arm motion by fusing inertial data with visual data. Fusion results don't have the drift problem. However, the results are still too noisy to be used for rehabilitation applications. After introducing the arm geometry and system configuration constraints, fusion results can be improved. Figure 13.6 illustrates results of the fusion plus optimisation method. The tracking results match the ground truth quite well. Tracking accuracy is within ±5cm for all three coordinates.



(a) x coordinates (b) y coordinates (c) z coordinates

Figure 13.5 Fusion result – circular motion of the wrist joint



(a) x coordinates (b) y coordinates (c) z coordinates

Figure 13.6 Fusion+LM results – circular motion of the wrist joint

13.5 Conclusions and Future Work

This paper presents a new arm motion tracking system based on the integration of vision and inertial sensors. Our method does not impose any special requirements on the environment, the clothes that the subject wears or arm motion. Initial experimental results demonstrate that the proposed method can track arm movements accurately in 3D. It has great potential to be used for home-based rehabilitation.

In the future, we will further test and refine our tracking system by comparing the results with that from a commercial marker based system, CODA or Qualisys, which offers more accurate ground truth.

13.6 References

Bachmann ER, Duman I, Usta UY, McGhee RB, Yun XP, Zyda MJ (1999) Orientation tracking for humans and robots using inertial sensors. In: Proceedings of International Symposium on Computational Intelligence in Robotics and Automation (CIRA’99), Monterey, CA, USA, pp 187-194

Foxlin E, Altshuler Y, Naimark L, Harrington M (2004) FlightTracker: a novel optical/inertial tracker for cockpit enhanced vision. In: Proceedings of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), Washington, DC, USA, pp 212-221

Goncalves L, Bernardo ED, Ursella E, Perona P (1995) Monocular tracking of the human arm in 3D. In: Proceedings of ICCV95 (International Conference on Computer Vision 1995), MIT, Cambridge, MA, USA, pp 764-770

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Perception and Psychophysics 14(2): 210-211

Tao Y, Hu H, Zhou H (2005) Integration of vision and inertial sensors for home-based rehabilitation. In: Proceedings of the 2nd Workshop on Integration of Vision & Inertial Sensors (InerVis 2005), Barcelona, Spain

Zhou H, Hu H (2005) Inertial motion tracking of human arm movements in home-based rehabilitation. In: Proceedings of IEEE International Conference on Mechatronics and Automation, Ontario, Canada