Implementation of an Extended Kalman Filter for optical motion capture with real–time 3D visualization

Urbano Lugrís[1], Rubén Vilela[1], Emilio Sanjurjo[1], Francisco Mouzo[1] and Florian Michaud[1]

[1] *Department of Naval and Industrial Engineering, University of A Coruña, ulugris@udc.es*

The technology for tracking reflective markers in 3D using IR cameras is currently mature and widely available, which makes optical motion capture the most popular approach for analyzing human motion. The method consists of recording the trajectories of several markers attached to anatomical landmarks, and then using these trajectories to reconstruct the motion of a rigid multibody model that represents the underlying skeletal system.

This method has multiple sources of error, due to the inaccuracy inherent to biomechanical systems. Unless CT scan or similar data is available, the exact geometry of the skeletal system is unknown. In addition, the ideal joints commonly used in multibody models do not always represent the actual kinematics with accuracy. Furthermore, the marker trajectories are perturbed by noise from the capture system and, more importantly, they undergo significant displacements relative to the skeleton, known as skin motion artifact.

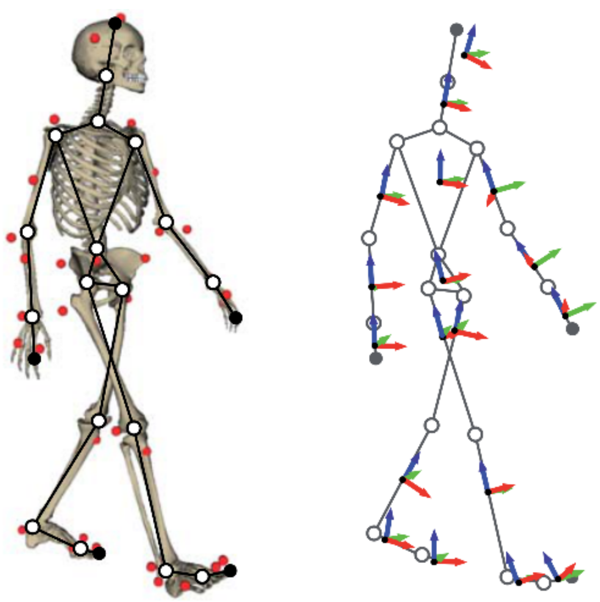

Fig. 1: Marker configuration and multibody model

Obtaining the best estimate of the skeletal motion requires the solution of two problems, which are commonly tackled using optimization techniques: first, the geometry of the multibody model must be fit to the actual skeleton, and secondly, its motion must be adjusted to track the recorded marker trajectories as closely as possible. There exist several methods in the literature for addressing these tasks [1, 2, 3], although in general they carry out the motion reconstruction as a post process, so that real–time applications such as biofeedback–based rehabilitation are not possible. Moreover, the task of obtaining said marker trajectories poses its own problems. Markers must be identified or *labeled* at every time step, and this is a nontrivial problem, since motion capture systems do not always provide the marker coordinates in a specific order. In addition, marker occlusions can occur during the capture, so there is no guarantee that all markers will be available at every time step.

In order to overcome these problems and attain real–time motion reconstruction, an efficient Extended Kalman Filter is proposed in this work. In the literature, most applications of Kalman filter to motion capture are designed for inertial sensors, although there exist some implementations based on optical markers [4, 5]. However, they use computationally–expensive observation functions, thus reducing their applicability to real–time applications.

The 52–DOF multibody model used in this work, depicted in Fig. 1, is formed by 18 bodies: pelvis, trunk, neck, head, arms, forearms, hands, thighs, calves, feet and toes. Most joints are considered spherical; however, in order to reduce the number of markers, the connection between neck and trunk is modeled as a universal joint, and the toe segments are attached to the feet through revolute joints. Therefore, the state vector $\mathbf{x}$ comprises 156 variables: the absolute position of the lumbar joint, three absolute Euler angles per body (except relative angles for the neck and toe segments), and all their associated velocities and accelerations. The filter uses a simple second–order kinematic model [6] to propagate the states of each degree of freedom i from instant t to $t+h$:

$$\begin{Bmatrix} x_i \\ \dot{x}_i \\ \ddot{x}_i \end{Bmatrix}_{t+h} = \begin{bmatrix} 1 & h & \frac{1}{2}h^2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} x_i \\ \dot{x}_i \\ \ddot{x}_i \end{Bmatrix}_{t} \qquad \mathbf{Q}_i = \begin{bmatrix} \frac{h^5}{20} & \frac{h^4}{8} & \frac{h^3}{6} \\ \frac{h^4}{8} & \frac{h^3}{3} & \frac{h^2}{2} \\ \frac{h^3}{6} & \frac{h^2}{2} & h \end{bmatrix} \sigma_i^2 \tag{1}$$

$\mathbf{Q}_i$ is the covariance of the process noise which, in turn, depends on the variance of the acceleration, σ_i^2.

In the proposed filter, 36 optical markers rigidly attached to anatomical landmarks act as absolute position sensors. The nonlinear observation function $\mathbf{z} = \mathbf{h}(\mathbf{x})$ is formed by 36 blocks of the following form:

$$\mathbf{z}_i = \mathbf{r}_b + \mathbf{R}_b \bar{\mathbf{m}}_i = \mathbf{r}_1 + \sum_{j=2}^{b} (\mathbf{R}_{j-1} \bar{\mathbf{r}}_j) + \mathbf{R}_b \bar{\mathbf{m}}_i \tag{2}$$

Assuming that marker i is rigidly attached to the anatomical segment b, its absolute position, $\mathbf{z}_i$, can be obtained as the sum of two components: the absolute position of the proximal joint, $\mathbf{r}_b$, and the relative position of the marker, $\mathbf{R}_b\bar{\mathbf{m}}_i$, which is, in turn, the result of applying the rotation matrix of the segment, $\mathbf{R}_b$, to the local coordinates of the marker, $\bar{\mathbf{m}}_i$. The absolute rotation matrices $\mathbf{R}_b$ depend on the Euler angles from the state vector, and the absolute joint positions $\mathbf{r}_b$ are obtained recursively, starting at the lumbar joint, being $\bar{\mathbf{r}}_j$ the coordinates of the proximal joint of each segment j within the local frame of its predecessor in the kinematic chain, $j-1$.

The use of absolute rotations for describing the orientation of most segments greatly simplifies the derivation of the Jacobian of the observation function $\mathbf{H}(\mathbf{x})$, due to the avoidance of cumulative rotation matrix products. This, along with the fact that all the matrices involved in the filter are very sparse, leads to a highly efficient algorithm: it performs several times faster than real time with 100 Hz IR cameras, on a regular desktop PC.

Moreover, this filter helps overcoming the main technical problems associated to marker–based motion capture. Since the states are propagated in time using a second–order model, the filter is robust to brief marker occlusions. And the problem of marker labeling is easily addressed by relating, at every time step, the current observed and estimated marker positions, using a nearest neighbor approach.

The performance and robustness of the filter will be verified by implementing a C++ application with real–time 3D visualization.

References

[1] J. A. Reinbolt, J. F. Schutte, B. J. Fregly, B. I. Koh, R. T. Haftka, A. D. George, and K. H. Mitchell, “Determination of patient–specific multi–joint kinematic models through two–level optimization,” *Journal of Biomechanics*, vol. 38, no. 3, pp. 621–626, 2005.

[2] S. S. U. Gamage and J. Lasenby, “New least squares solutions for estimating the average center of rotation and the axis of rotation,” *Journal of Biomechanics*, vol. 35, pp. 87–93, jan 2002.

[3] C. L. Vaughan, B. L. Davis, and J. C. O’Connor, *Dynamics of Human Gait, 2nd Edition*. Cape Town: Kiboho Publishers, 1999.

[4] F. De Groote, T. De Laet, I. Jonkers, and J. De Schutter, “Kalman smoothing improves the estimation of joint kinematics and kinetics in marker–based human gait analysis,” *Journal of Biomechanics*, vol. 41, pp. 3390–3398, 2008.

[5] P. Cerveri, A. Pedotti, and G. Ferrigno, “Robust recovery of human motion from video using Kalman filters and virtual humans,” *Human Movement Science*, vol. 22, no. 3, pp. 377–404, 2003.

[6] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons Inc, 2001.