

Predicting Missing Markers to Drive Real-Time Centre of Rotation Estimation

Andreas Aristidou, Jonathan Cameron, and Joan Lasenby

Department of Engineering
University of Cambridge
Cambridge, UK, CB2 1PZ
{aa462, jic23, j1221}@cam.ac.uk

Abstract. This paper addresses the problem of real-time location of the joints or centres of rotation (CoR) of human skeletons in the presence of missing data. The data is assumed to be 3d marker positions from a motion capture system. We present an integrated framework which predicts the occluded marker positions using a Kalman filter in combination with inferred information from neighbouring markers and thereby maintains a continuous data-flow. The CoR positions can be calculated with high accuracy even in cases where markers are occluded for a long period of time.

Keywords: Kalman Filter, Missing Markers, Joint Localisation, Motion Capture.

1 Introduction

Estimating the location of centres of rotations (CoRs) using marked optical motion capture data is useful within technique analysis for sports training [1]; observation of asymmetries and abnormalities in rehabilitation medicine [2]; and generation of virtual characters for films and computer games [3]. However, even with many cameras, there are instances where occlusion of markers by elements of the scene leads to missing data. In order to unambiguously establish its position, each marker must be visible to at least two cameras in each frame. Although many methods have been developed to handle the missing marker problem, most of them are not applicable in real-time and often require manual intervention.

This paper proposes a real-time approach for estimating the position of occluded markers using previous positions and information inferred from an approximate rigid body assumption. The predicted marker positions are then used to locate the human joints. Without assuming any skeleton model, we take advantage of the fact that for markers on a given limb segment, the inter-marker distance is approximately constant. Thus, the neighbouring markers¹ provide us with useful information about the current positions of any non-visible markers. With a continuous stream of accurate 3d data, we can perform real-time CoR estimation, thereby producing skeletal information for use in visual performance

¹ Neighbouring markers are considered as markers belonging to the same limb segment.

feedback. Experiments demonstrate that the method presented effectively recovers good estimates of the true positions of the missing markers, even if all the markers on a limb are occluded for a long period of time.

2 Related Work

Recent papers have focused on real-time localisation of the CoR. [4] uses the assumption of inter-marker rigidity between neighbouring markers. This allows a closed form sequential algorithm. In this respect it differs from [5,6,7,8]. However, this algorithm neglects frames containing missing markers.

Whilst several methods to estimate the location of missing markers have been proposed, the performance of most is unsatisfactory in the presence of unusual motions or of many contiguous occlusion-effected frames. Methods may interpolate the data using linear or non-linear approaches [9,10,11]; this can produce accurate results, but is useful only in post-processing.

[12] proposes a model-based system for composite interaction devices (CID). The system automatically constructs the geometric skeleton structure and degrees of freedom relations and constraints between segments. The system supports segments with only a single marker, allowing CIDs to be very small. Unfortunately, this is an off-line procedure unsuitable for real-time applications.

[13,14] use an extended Kalman filter to predict the missing marker locations using previous marker positions and a skeletal model. These methods become ineffective when markers are missing for an extended period of time.

[15,16] use post-processing to increase the robustness of motion capture using a sophisticated human model. They predict the 3d location and visibility of markers under the assumption of fixed segment inter-marker distances. However, the skeleton information must be known a priori.

Recently, [17] presented a piecewise linear approach to estimating human motions from pre-selected markers. A pre-trained classifier identifies an appropriate local linear model for each frame. Missing markers are then recovered using available marker positions and the principal components of the associated model. The pre-training session and the classifier limit the approach to off-line applications.

[18] presented a reliable system that could predict the missing markers in real-time under large occlusions. However, this system did not consider the useful information available when the markers are visible by a single camera. The missing markers are usually not entirely occluded and this additional information is used in this paper to produce a more reliable system.

3 Estimating CoR

Locating the CoRs is a crucial step in acquiring a skeleton from raw motion capture data. The data discussed here is from an active marker system, hence no tracking is necessary. To calculate the joints between two sets of markers it is helpful to have the rotation of a limb at any given time. We can estimate the

orientation of a limb at time k relative to a reference frame using the *Procrustes* formulation [19].

The location of the joints can be calculated using [4]. This approach takes advantage of the approximation that all markers on a segment are attached to a rigid body. Suppose the markers are placed on two segments (x and y) joined by a CoR. Let the CoR location in frame k be \mathbf{C}^k . The vectors from the CoR to markers in the reference frame are denoted by \mathbf{a}_x^i and \mathbf{a}_y^j for limb x and y respectively, where i and j are marker labels. The position of the markers in frame k is given by:

$$\mathbf{x}_i^k = \mathbf{C}^k + R_x^k \mathbf{a}_x^i \tilde{R}_x^k \quad \mathbf{y}_j^k = \mathbf{C}^k + R_y^k \mathbf{a}_y^j \tilde{R}_y^k \quad (1)$$

where R_x and R_y are the rotors (quaternions) expressing the rotation of the joint limbs x and y respectively. \tilde{R} is the quaternion conjugate of R . Let \mathbf{b}_{ij}^k be the vector from \mathbf{x}_i^k to \mathbf{y}_j^k , that is:

$$\mathbf{b}_{ij}^k = \mathbf{x}_i^k - \mathbf{y}_j^k = R_x^k \mathbf{a}_x^i \tilde{R}_x^k - R_y^k \mathbf{a}_y^j \tilde{R}_y^k \quad (2)$$

Now a cost function S can be constructed that has a global minimum at the correct values of \mathbf{a}_x^i and \mathbf{a}_y^j if the data is noise free, and returns a good estimate in the presence of moderate noise.

$$S = \sum_{k=1}^m \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \left[\mathbf{b}_{ij}^k - \left(R_x^k \mathbf{a}_x^i \tilde{R}_x^k - R_y^k \mathbf{a}_y^j \tilde{R}_y^k \right) \right]^2 \quad (3)$$

where n_x , n_y are the number of markers on limbs x and y respectively, and m is the number of frames used for the calculations. The minimum is given by the solution of the simultaneous linear equations, obtainable by differentiation:

$$\mathbf{a}_x^i = \frac{1}{m} \sum_{k=1}^m \tilde{R}_x^k \bar{\mathbf{b}}^k R_x^k + \frac{1}{m} \sum_{k=1}^m \tilde{R}_x^k R_y^k \bar{\mathbf{a}}_y \tilde{R}_y^k R_x^k \quad (4)$$

$$\mathbf{a}_y^j = \frac{1}{m} \sum_{k=1}^m \tilde{R}_y^k \bar{\mathbf{b}}^k R_y^k + \frac{1}{m} \sum_{k=1}^m \tilde{R}_y^k R_x^k \bar{\mathbf{a}}_x \tilde{R}_x^k R_y^k \quad (5)$$

where

$$\bar{\mathbf{b}}^k = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \mathbf{b}_{ij}^k \quad \bar{\mathbf{a}}_w = \frac{1}{n_w} \sum_{i=1}^{n_w} \mathbf{a}_w^i \quad w = \{x, y\} \quad (6)$$

Having calculated the R_w^k and $\bar{\mathbf{a}}_w$, we can locate the CoR. However, due to occlusions, there are instances where not all marker positions are available. If and only if all markers are available on one limb segment, w , the CoR may be estimated using only R_w^k in the current frame and $\bar{\mathbf{a}}_w$ as estimated in the previous frame, via (1). If markers are occluded on both segments a method such as that in the next section is needed.

4 Missing Marker Position Estimation

The marker position estimates can be predicted using a Kalman filter [20], where constraints from the neighbouring markers are applied for a more reliable system.

4.1 Framework

The process model to update the *state* of the Kalman filter is given by (7), where the state \mathbf{x}_k at frame k is obtained from the state at frame $k - 1$;

$$\mathbf{x}_k = A\mathbf{x}_{k-1} + B\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \quad (7)$$

where A is the state transition model, B is the control input model, \mathbf{u}_{k-1} is the control vector and \mathbf{w}_{k-1} is the process noise. The measured data \mathbf{Z}_k is

$$\mathbf{Z}_k = H\mathbf{x}_k + \mathbf{v}_k \quad (8)$$

where H is the *observation model* and \mathbf{v}_k is the observation noise. \mathbf{w} and \mathbf{v} are assumed to be zero mean multivariate normal with covariance Q and R .

The predicted state \mathbf{y}_k and its error E_k can be written as

$$\mathbf{y}_k = A\hat{\mathbf{x}}_{k-1} + B\mathbf{u}_{k-1} \quad E_k = AP_{k-1}A^T + Q \quad (9)$$

where $\hat{\mathbf{x}}$ refers to the *estimate* and P is the covariance of the state estimate.

The *Kalman gain* between actual and predicted observations is:

$$K_k = E_k H^T (H E_k H^T + R)^{-1} \quad (10)$$

Thus given an estimate $\hat{\mathbf{x}}_{k-1}$ at $k - 1$, the time update predicts the state value at frame k . The measurement update then adjusts this prediction based on the new \mathbf{y}_k . The estimate of the new state is

$$\hat{\mathbf{x}}_k = \mathbf{y}_k + K_k (\mathbf{Z}_k - H\mathbf{y}_k) \quad (11)$$

K_k is chosen to minimise the steady-state covariance of the estimation error. Finally, the error covariance matrix of the updated estimate is;

$$P_k = (I - K_k H) E_k \quad (12)$$

In this work, a constant velocity model is used. Hence,

$$\mathbf{y}_k = \mathbf{x}_{k-1} + \dot{\mathbf{x}}_{k-1} dk \quad (13)$$

where \mathbf{x}_i and $\dot{\mathbf{x}}_i$ are the position and velocity in frame i , and dk is the time step.

4.2 Applying Constraints

The observation vector, \mathbf{Z}_k , gives the observed position of the tracked marker when available, otherwise it represents estimated position. The state vector represents true position and velocity as given above. To cope with cases where markers are missing for long periods of time, we implement a tracker that uses information from both the previous frames and the current positions of neighbouring visible markers. We assume three markers on each limb. In the presence of noise the observation vector is updated as given below for 4 different scenarios;

All markers are visible on a given limb

$$\mathbf{Z}_k = H\mathbf{x}_k + \mathbf{v}_k \quad (14)$$

where \mathbf{x}_k is the current state of a tracked marker on the limb. In this case H is the identity and R is determined empirically. Many factors contribute to marker noise, and hence R , including optical measurement noise miscalibration of the optical systems, reflection, motion of markers relative to the skin and motion of the skin relative to the rigid body (underlying bone).

One missing marker on a limb segment

$$\mathbf{Z}_k = H\hat{\mathbf{x}}_1^k + \mathbf{v}_k \quad (15)$$

where $\hat{\mathbf{x}}_1^k$ is the estimated position of the occluded marker m_1 in frame k . $\hat{\mathbf{x}}_1^k$ can be calculated as given below. Firstly we calculate $\mathbf{D}_{1,2}^{k-1}$ and $\mathbf{D}_{1,3}^{k-1}$ which correspond to the vectors between marker m_1 and markers m_2, m_3 in frame $k-1$ respectively. These vectors are given by $\mathbf{D}_{i,j}^{k-1} = \mathbf{x}_j^{k-1} - \mathbf{x}_i^{k-1}$. Thereafter, these vectors are rotated as $\hat{\mathbf{D}}_{i,j}^k = R^{k-2,k-1}\mathbf{D}_{i,j}^{k-1}\tilde{R}^{k-2,k-1}$ where $R^{p,q}$ is the rotor expressing the rotation between frames p to q , assuming that the rotation of the markers between two consecutive frames remains constant. One obvious way to proceed is to calculate the point $\tilde{\mathbf{x}}_1^k$ which is an average of the estimated positions in frame k using the $\hat{\mathbf{D}}$ vectors;

$$\tilde{\mathbf{x}}_1^k = \frac{(\mathbf{x}_2^k - \hat{\mathbf{D}}_{1,2}^k) + (\mathbf{x}_3^k - \hat{\mathbf{D}}_{1,3}^k)}{2} \quad (16)$$

where \mathbf{x}_i^k is the position of marker i in frame k . We now improve on this estimate by finding the solution of the intersection of the two spheres in frame k with centres $\mathbf{x}_2^k, \mathbf{x}_3^k$ and radii $|\hat{\mathbf{D}}_{1,2}^k|$ and $|\hat{\mathbf{D}}_{1,3}^k|$ respectively. $\hat{\mathbf{x}}_1^k$ is assigned as the closest point on the circle of intersection to $\tilde{\mathbf{x}}_1^k$. Figure 1 illustrates this process.

Two missing markers on a limb segment

$$\mathbf{Z}_k = H\hat{\mathbf{x}}_j^k + \mathbf{v}_k \quad (17)$$

where $\hat{\mathbf{x}}_j^k$ is the estimated position of the occluded marker m_j ($j = 1, 3$) in frame k . $\hat{\mathbf{x}}_j^k$ is given by:

$$\hat{\mathbf{x}}_j^k = \mathbf{x}_2^k - \hat{\mathbf{D}}_{j,2}^k \quad (18)$$

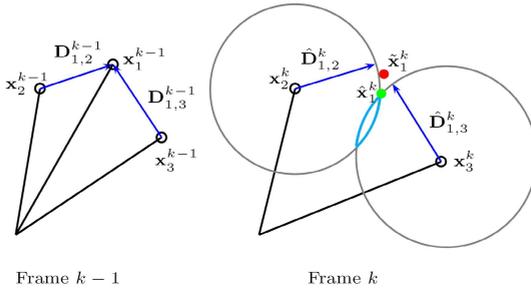


Fig. 1. The observation vector in the case of 2 visible markers. The red dot, $\hat{\mathbf{x}}_1^k$, represents the average value as given in equation 16. The green dot, $\tilde{\mathbf{x}}_1^k$, is the point on the intersection of the 2 spheres which is closest to $\hat{\mathbf{x}}_1^k$.

where \mathbf{x}_2^k is the position of the visible marker m_2 on the limb in the current frame and $\hat{\mathbf{D}}_{j,2}^k$ is as described above. In this case, we are using the constant velocity assumption as we cannot estimate the rotation.

All markers on a limb segment are missing: Here we consider two possible subcases; the case where the other limb segment has some markers visible and the case where both limb segments have all of their markers occluded. If some markers on the other limb segment are visible, the missing marker positions can be calculated using the CoR estimate, $\hat{\mathbf{C}}_k$ as calculated in Sect. 3. In that case the observation vector of the Kalman filter is updated as:

$$\mathbf{Z}_k = H\hat{\mathbf{x}}_j^k + \mathbf{v}_k \tag{19}$$

where $\hat{\mathbf{x}}_j^k$ is the estimated position of the occluded marker m_j ($j = 1, 2, 3$) in frame k . $\hat{\mathbf{x}}_j^k$ is given by;

$$\hat{\mathbf{x}}_j^k = \hat{\mathbf{C}}_k + \hat{\mathbf{D}}_{j,c}^k \tag{20}$$

where $\hat{\mathbf{D}}_{j,c}^k$ is an estimate of the distance between marker m_i and the CoR. This approach takes advantage of the fact that the distance between markers and the CoR is constant. This distance is equal to $\hat{\mathbf{D}}_{j,c}^k = R^{k-2,k-1}\mathbf{D}_{j,c}^{k-1}\tilde{R}^{k-2,k-1}$ where $\mathbf{D}_{j,c}^{k-1} = \mathbf{x}_j^{k-1} - \mathbf{C}_{k-1}$. This assumes that the rotation of the markers between two consecutive frames remains constant.

If both limb segments have all markers occluded, only information from previous frames is used. The observation vector, \mathbf{Z}_k , in this instance is calculated using a quaternion based method. This method also assumes that the segment rotation between two consecutive frames is constant. The observation vector is now equal to

$$\mathbf{Z}_k = H\hat{\mathbf{x}}^k + \mathbf{v}_k \tag{21}$$

where $\hat{\mathbf{x}}^k$ is equal to $\hat{\mathbf{x}}^k = R^{k-2,k-1}\mathbf{x}^{k-1}\tilde{R}^{k-2,k-1}$.

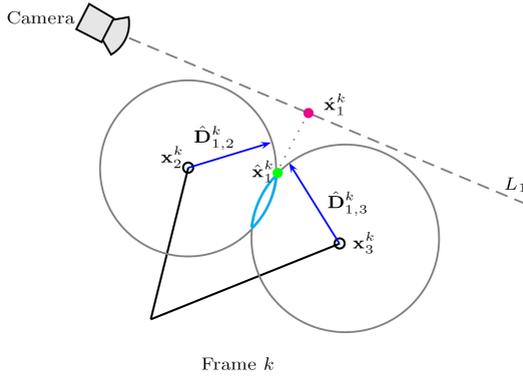


Fig. 2. The observation vector in the case of 2 visible markers and one marker visible only by one camera. The magenta dot, $\hat{\mathbf{x}}_1^k$, is now used for the calculation of the observation vector, $\mathbf{Z}_k = H\hat{\mathbf{x}}^k + \mathbf{v}_k$.

However, the motion capture system also provides us with additional information which could be used for prediction of missing marker locations. Each marker can be reconstructed by the motion capture system if it is visible in at least two cameras. It is often the case that some markers are visible in only one camera. This information identifies a line, L_1 , starting from the camera and passing through the position of the missing marker. By relaxing the constraints that the inter-marker distance is constant and accepting that the real position of the marker is on the line L_1 , we can obtain a more accurate estimate of the position of the relevant marker. This position, $\hat{\mathbf{x}}_1^k$, corresponds to the projection from the point $\hat{\mathbf{x}}_1^k$ onto the line L_1 , as in Fig. 2. This is applicable for the cases in which the motion capture system fully reconstructs one or two markers locations and another marker is visible in one camera. If a limb segment has only one known and one partially visible marker, the system is more reliable when it first predicts the partially visible marker and then the entirely occluded marker.

5 Results and Discussion

Experiments were carried out using a 16 camera Phasespace motion capture system capable of capturing data at 480Hz [21]. The algorithm was implemented

Table 1. Average results on real data with occlusions generated by deletions. Case of one missing marker on each limb segment for more than 1500 frames.

The error when the missing markers are entirely occluded		The error when the missing markers are visible by one camera	
	Error (mm)		Error (mm)
Marker position	3.348151	Marker position	0.585351
CoR (when $\bar{\mathbf{a}}_w$ is updated using the predicted data.)	5.905546	CoR (using the predicted markers positions)	1.281958

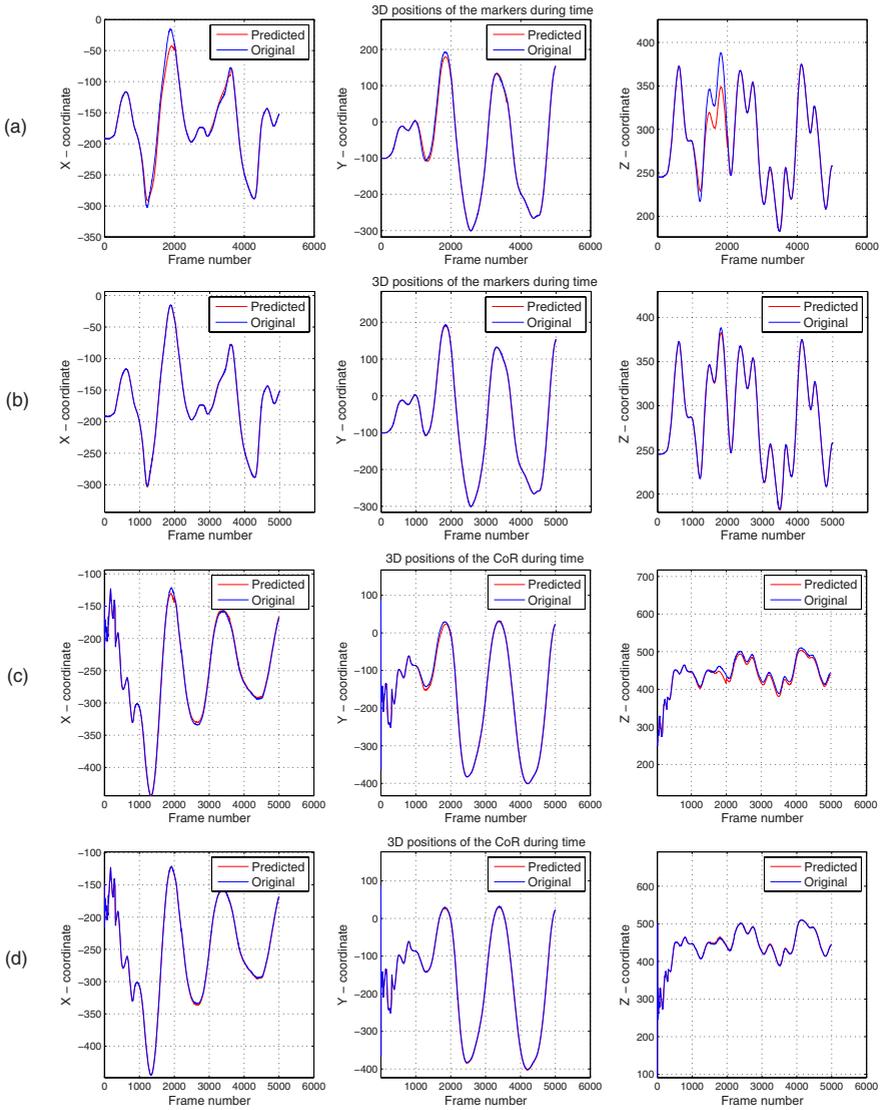


Fig. 3. An example of the 3D positions of predicted and the true coordinates of the *Markers* (a), (b) in the case of one missing marker on a limb segment and a missing marker visible by one camera respectively, and the *CoR* (c), (d) under the same conditions. The occlusion periods are between frames 1000-2000 and 3000-3600.

in MATLAB and run on a Pentium IV PC. The system can process up to 350 *frames per second* (using MATLAB). Our datasets comprise both simulated and real data (i.e. captured data with natural occlusions or occlusions generated by artificial deletion) with more than 5000 *frames* in each. There are two categories: one with 7 segment leg datasets and the other with 5 segment arm datasets. The

3d location of the markers can be reliably reconstructed even when we have single marker occlusion for more than 1000 *frames* at a time, returning mean position errors of less than 3.5mm from the true value. The position of the CoR using the predicted marker positions can be calculated with a mean error of approximately 6.35mm in cases where one marker on each limb segment is entirely occluded, this increases to 11.8mm in cases where 2 out of 3 markers on a limb are not visible. However, in the case where one of the limb segments has all its markers available, the CoR can be calculated with higher accuracy using information only from that limb segment, as in Sect. 3, where $\bar{\mathbf{a}}_w$ is now updated using the predicted positions of the markers. The error between the true and the CoR estimate for that instance is 5.9mm when one marker is occluded and 9.5mm when two markers are missing. This error is significantly decreased to 0.6mm for marker position estimations and 1.3mm for the CoR positions estimation in the case where the missing markers are visible by one camera. Table 1 presents the average results (over 30 runs) in the case of one missing marker on each limb segment. Figure 3 shows an example of the 3D position variation over time between the true and the predicted position of the markers and the CoR for two particular cases of occlusion.

6 Conclusion and Future Work

This paper describes an algorithm related to the problem of using marker-based optical motion capture data to automatically establish a skeleton model to which the markers are attached. It presents a prediction method that estimates the missing markers and reconstructs the skeletal motion. These positions are calculated in real-time using a Kalman filter updated via information from neighbouring visible markers. Also, the system takes advantage of the information returned by each single camera regarding the position of the missing markers. The predicted data is then used for real-time joint localisation. This approach works reliably even if large sequences with occluded data exist, in which more than 1 marker is occluded on each limb, and also when the limb rapidly changes direction. Future work will introduce biomechanical constraints to restrict motions to those from a feasible set and a reliable model for predicting the rotors expressing the rotation of each limb segment.

References

1. Hashiguchi, J., Nivomiya, H., Tanaka, H., Nakamura, M., Nobuhara, K.: Biomechanical analysis of a golf swing using motion capture system. In: Proc. of the Japanese Society for Orthopaedic Biomechanics, vol. 27, pp. 325–330 (2006)
2. Broeren, J., Sunnerhagen, K.S., Rydmark, M.: A kinematic analysis of a haptic handheld stylus in a virtual environment: a study in healthy subjects. *Journal of NeuroEngineering and Rehabilitation* 4,13 (2007)
3. Menache, A.: *Understanding Motion Capture for Computer Animation and Video Games*. Morgan Kaufmann Publishers Inc, USA (1999)

4. Cameron, J., Lasenby, J.: A real-time sequential algorithm for human joint localization. In: ACM SIGGRAPH Posters, USA, p. 107 (2005)
5. Gamage, S., Lasenby, J.: New least squares solutions for estimating the average centre of rotation and the axis of rotation. *J. of Biomechanics* 35(1), 87–93 (2002)
6. Kirk, A.G., O'Brien, J.F., Forsyth, D.A.: Skeletal parameter estimation from optical motion capture data. In: Proc. of the IEEE CVPR, pp. 782–788 (2005)
7. O'Brien, J.F., Bodenheimer, R.E., Brostow, G.J., Hodgins, J.K.: Automatic joint parameter estimation from magnetic motion capture data. In: Proceedings of Graphic Interface, pp. 53–60 (2000)
8. Ehrig, R.M., Taylor, W.R., Duda, G.N., Heller, M.O.: A survey of formal methods for determining the centre of rotation of ball joints. *Journal of Biomechanics* 39(15), 2798–2809 (2006)
9. Wiley, D.J., Hahn, J.K.: Interpolation synthesis of articulated figure motion. *IEEE Comp. Graphics and Applications* 17(6), 39–45 (1997)
10. Rose, C., Cohen, M., Bodenheimer, B.: Verbs and adverbs: Multidimensional motion interpolation. *IEEE Comp. Graphics and Applications* 18(5), 32–40 (1998)
11. Nebel, J.: Keyframe interpolation with self-collision avoidance. In: Proc. of the Workshop on Comp. Animation and Simulation, pp. 77–86. Springer, Heidelberg (2006)
12. Van Rhijn, A., Mulder, J.D.: Optical tracking and automatic model estimation of composite interaction devices. In: IEEE VR Conference, pp. 135–142 (2006)
13. Dorfmüller-Ulhaas, K.: Robust optical user motion tracking using a kalman filter. Technical Report TR-2003-6, Institut fuer Informatik 2, 86159 Augsburg (2003)
14. Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K., Colucci, D.: The HiBall tracker: High-performance wide-area tracking for virtual and augmented environments. In: VRST, December 20–22, 1999, pp. 1–10. ACM, New York (1999)
15. Herda, L., Fua, P., Plänkner, R., Boulic, R., Thalmann, D.: Using skeleton-based tracking to increase the reliability of optical motion capture. *Human Movement Science Journal* 20(3), 313–341 (2001)
16. Hornung, A., Sar-Dessai, S.: Self-calibrating optical motion tracking for articulated bodies. In: Proceedings of the IEEE Conference on VR, Washington, DC, USA, pp. 75–82 (2005)
17. Liu, G., McMillan, L.: Estimation of missing markers in human motion capture. *The Visual Computer* 22(9–11), 721–728 (2006)
18. Aristidou, A., Cameron, J., Lasenby, J.: Real-time estimation of missing markers in human motion capture. In: Proceedings of the iCBBE 2008 (2008)
19. Horn, B.: Closed-form solution of absolute orientation using unit quaternions. *Journal of the Opt. Society of America* 4, 629–642 (1987)
20. Kalman, R.E.: A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 35–45 (1960)
21. PhaseSpace Inc.: Optical motion capture systems, <http://www.phasespace.com>