

A computational method for hip joint centre location from optical markers

A. J. Stoddart, P. Mrázek

Centre for Vision, Speech and Signal Processing, University of Surrey, U.K.

D. Ewins and D. Hynd

Biomedical Engineering Group, University of Surrey, U.K. and
Gait Laboratory, Clinical Biomedical Engineering Centre (CBEC)
Queen Mary's Hospital, London, U.K.

a.stoddart@ee.surrey.ac.uk

Abstract

The problem of hip joint centre location using optical markers located on the skin is addressed. We present a novel computational technique which can recover joint centre location based solely on the marker measurements. In addition a quantitative estimate of the uncertainty in joint centre location is obtained.

We also present some preliminary experimental validation of the technique on a rigid jointed object.

1 Introduction

In a number of biomedical applications it is of interest to compute kinematic and dynamic quantities of human bodies in motion. Commercially available marker-based optical motion capture systems are widely used for this purpose. Many of these systems use retroreflective markers attached to the subject and a number of cameras. The technology for extracting by triangulation accurate 3D coordinates of the markers is well established.

The underlying assumption in this approach is that the human body can be modelled as a collection of rigid segments connected by ball or hinge joints. The skin mounted markers are a relatively non-intrusive indicator of bone position, although soft tissue movement can be a significant source of error. In addition it can be difficult to locate palpable bony landmarks due to the tissue layer between the surface and the bone.

In these systems the number of markers is usually small, between 0 and 5 per 'rigid' segment. Usually the centres of rotation are supplied by an independent method and an inverse kinematics computation is performed to obtain the Euler angles of relative rotations about the joint centre.

There is also much current interest in articulated systems in the field of computer vision. The primary motivation is the rapid progress in techniques for capturing human motion in a variety of applications. In many cases the model of a human as a collection of rigid segments joined by rotational joints is an appropriate approximation or level of abstraction.

In this paper we present a method for extracting the joint centre in the two segment case providing that at least 3 markers are available on each segment. This information is a crucial building block in kinematic analysis of the motion data.

The motion data has a variety of uses in planning and evaluation of surgical and orthotic interventions. However the algorithm will be of general use in the analysis of other systems of articulated segments. The algorithm also quantifies the error in joint centre location.

2 This paper

In this paper we present a novel algorithm for hip joint centre location and uncertainty estimation from marker positions. Firstly, using standard methods we extract the pose of segments from the marker positions. Then we compute the joint centres from the segment poses. An analytic solution is presented followed by a solution to the linearised problem. The linearised solution is preferred since it produces a covariance matrix for the joint centre estimate.

We conclude by presenting some preliminary experimental validation of the technique on a two segment jointed test object.

3 Literature Review

We firstly consider previous work in the field of Biomedical engineering. A highly accurate estimate of hip joint position can be obtained by X-rays from more than one view, but this is not always convenient or justified. At the other extreme there are heuristic methods such as the following method quoted by Bell *et al* [4]. “The hip joint centre lies about 15-20mm directly distal to the midpoint of a line from the pubic symphysis and the anterior superior iliac spines in a frontal plane projection and directly medial from the greater trochanter in the sagittal plane.”

Such methods can be accurate in the region of 20 mm but are generally predicated on specific sets of subjects, e.g. healthy males between 20 and 40 years of age. They are not suitable for subjects widely varying in age or pathological subjects.

For this reason Bell *et al* [4] made a quantitative comparison of several methods of hip joint centre location. They looked at previous heuristic methods based on correlations between palpable anatomical landmarks and a method based on skin mounted optical markers. X-rays provided ground truth. Their method is similar to our own in that they infer the joint centre from optical markers, but they provide no details of their approach and it seems that they did not develop the mathematical framework beyond a rudimentary stage.

Neptune and Hull [13] studied the accuracy of four different techniques for determining the location of the hip joint centre. The reference technique was an invasive technique which involved a screw inserted into the pelvis of a live subject. Mounted on the screw was a triad of optical markers. The true hip joint centre was located by X-rays with the bone mounted marker set in place. The subject then cycled on a stationary bicycle and the hip joint centre was estimated using the reference technique and 2 other non-invasive techniques. The reference technique was believed to be considerably more accurate than the noninvasive techniques for which the quoted errors were up to 13mm rms and 23mm in bias along the worst axis.

The reference technique was computed using pose transformations estimated by direct computation and heuristic (semi-quantitative) estimation of errors. It is implied that the error in joint centre was much less than 1.6mm.

The main source of errors in the non-invasive techniques are simplistic anatomical assumptions and soft-tissue movement. We believe that more powerful mathematical tools allow the optimal use of redundant data in this type of computation, although it is typical in many fields to use only minimal marker sets and correspondingly simpler mathematical techniques.

In a more recent study Lu and O'Connor [10] presented a method for obtaining bone segment joint angles from marker measurements called the Global Optimisation Method (GOM). They suggest that they improve upon previous work in that the joint constraints are rigorously enforced. (In much previous work on multiple segments this is not true!) However the joint centres are supplied as inputs to their technique. They do consider in some detail the magnitude and nature of the skin marker measurement error. They present results of numerical experiments with a random phase sinusoidal error in marker position which show that their method is better than methods which do not properly enforce joint constraints.

We now consider previous work in the field of computer vision. The area of human motion capture from one or more cameras is currently under active investigation by a number of investigators, for a recent review see Aggarwal and Cai [1].

The goal of much of the work described by Aggarwal and Cai is to capture motion under uncontrolled conditions without markers. The problem is severely underdetermined and many authors contend that model based methods will be useful for a solution to this problem. We consider only those methods that use a 3D articulated volumetric model of rigid segments.

The essence of model-based methods is the assumption that it is desirable to add prior knowledge to constrain and facilitate scene interpretation. Most authors therefore assume detailed kinematic knowledge a priori. This usually includes the dimensions of the subject and the joint positions. Some typical examples of work in this area are listed [18, 15, 16, 12, 17, 5, 6, 8].

Apart from human tracking there is other work worth noting as being relevant to the task of joint centre location. Closely related work is that of Ashbrook *et al* [2, 3] who address the automatic construction of models of articulated objects from range data.

Probably the most closely related work to our own is that of Heap and Hogg [7] who investigated pivot location for the study of articulated objects by point distribution models in polar co-ordinates. They presented a least squares expression similar to the analytic method in this paper. Ours differs in two minor ways. We extend the 2D case to 3D, and by a different choice of coordinate systems reduce by half the number of variables over which it is necessary to solve. They did not consider the possibility of predicting the pivot positional error but instead study it by numerical experiments.

4 The problem

We suppose that there are 2 objects/segments labelled by $\alpha = 0, 1$ and each object has markers labelled by $i = 0..N_\alpha - 1$. The marker position is given in the world coordinate system by $\vec{y}_i^\alpha(t)$ where a set of measurements is made over discrete time instants $t = 0..N_t - 1$.

We define a body fixed coordinate system attached to each segment. In this coordinate system the markers have position \vec{x}_i^α which does not vary with time. As the segment moves over time the coordinate transform required to convert the body centred coordinates \vec{x}_i^α to the world coordinates $\vec{y}_i^\alpha(t)$ is denoted by $f^\alpha(t) = \{R^\alpha(t), \vec{d}^\alpha(t)\}$. The coordinate transform consists of a 3×3 rotation matrix $R^\alpha(t)$ and a translation vector $\vec{d}^\alpha(t)$. We write

$$\vec{y}_i^\alpha(t) = f^\alpha(t) * \vec{x}_i^\alpha = R^\alpha(t)\vec{x}_i^\alpha + \vec{d}^\alpha(t) \quad (1)$$

All 3 coordinate systems are at this point entirely arbitrary and it is helpful to remove some of this degeneracy to simplify the problem. We therefore impose that

the two body fixed coordinate systems coincide with the world coordinate system at $t = 0$, hence that $f^0(0) = f^1(0) = \mathbf{I}$, the identity transform.

We suppose that the two segments are joined by a ball joint which has body fixed coordinate \vec{c}^0 on segment 0 and \vec{c}^1 on segment 1. Because we have chosen that the body fixed coordinate systems coincide at $t = 0$ it follows that $\vec{c}^0 = \vec{c}^1$ and we can discard the superscript.

At all subsequent time instants the world position of the two joint centres must coincide, and this is expressed by the joint constraint,

$$0 = f^0(t) * \vec{c} - f^1(t) * \vec{c}. \quad (2)$$

The problem we would like to solve is to find \vec{c} given a series of measurements $\vec{y}_i^\alpha(t)$. The measurements will be noisy and we approximate the true unknown noise distribution by isotropic Gaussian noise $n_i^\alpha(t)$ with variance $\sigma^2 = Tr(W_y)$.

$$\vec{y}_i^\alpha(t) = f^\alpha(t) * \vec{x}_i^\alpha + n_i^\alpha(t) \quad (3)$$

Finally a note on rotations. We represent rotation where appropriate by either the 3×3 rotation matrix R or the 3×1 vector $r = \theta \hat{n}$ where the rotation is by angle θ in radians about axis \hat{n} . Use of the latter representation is necessary for discussions of rotational covariance.

5 Obtaining segment poses

We propose a two stage solution. Firstly we use an existing method to obtain an estimate of the pose for each segment at each time instant. The least squares estimate for the coordinate transform $\hat{f}^\alpha(t)$ will minimise the objective function

$$E[f^\alpha(t)] = \sum_{i=0}^{N_\alpha-1} [\vec{y}_i^\alpha(t) - f^\alpha(t) * \vec{x}_i^\alpha]^2 \quad (4)$$

Kanatani [9] has proposed an analytic solution based on the singular value decomposition (SVD) of a 3×3 matrix which provides a unique solution provided that neither set of points is collinear and that there are at least 3 points.

The body fixed coordinates are not initially supplied and must be estimated. To start with we set $\vec{x}_i^\alpha \leftarrow \vec{y}_i^\alpha(t=0)$. The $\vec{y}_i^\alpha(t)$ are noisy measurements and we can do better by iterating over the following process a few times. In step 1 we are given \vec{x}_i^α and solve for $\hat{f}^\alpha(t)$ using (4). In step 2 we improve the \vec{x}_i^α using

$$\vec{x}_i^\alpha = \frac{1}{N_t} \sum_{t=0}^{N_t-1} \hat{f}^\alpha(0) [\hat{f}^\alpha(t)]^{-1} \vec{y}_i^\alpha(t) \quad (5)$$

In our experience this procedure always converges to 15 decimals of precision in less than 5 iterations. It also guarantees that $\hat{f}^\alpha(0) = \mathbf{I}$ which is not automatically guaranteed by equation (4).

Once the iterative process has converged the pose estimates $\hat{f}^\alpha(t)$ are optimal in a least squares sense.

Kanatani [9] and later Pennec and Thirion [14] have discussed the estimation of uncertainty in pose estimation. We follow the latter approach and obtain a 6×6 covariance matrix $W_f^\alpha(t)$ for the uncertainty in the pose estimate.

6 Obtaining the centre of rotation

In this section we present an analytic solution for the joint centre. We propose that this problem be formulated as a second least squares estimation problem that uses the output of the first. The joint centre is estimated by minimising the objective function

$$E[\vec{c}] = \sum_{t=0}^{N_t-1} [f^0(t)\vec{c} - f^1(t)\vec{c}]^2 \quad (6)$$

Before continuing we change slightly the way that the coordinate transform is expressed. We define $\vec{p}^\alpha(t)$ by

$$f^\alpha(t) * \vec{c} = R^\alpha(t)\vec{c} + \vec{d}^\alpha(t) = R^\alpha(t) [\vec{c} - \vec{p}^\alpha(t)] \quad (7)$$

The objective may be expanded as

$$E[\vec{c}] = \sum_{t=0}^{N_t-1} [(\vec{c} - \vec{p}^0(t))^2 + (\vec{c} - \vec{p}^1(t))^2 - 2(\vec{c} - \vec{p}^0(t))^\top R^0(t)^\top R^1(t) (\vec{c} - \vec{p}^1(t))] \quad (8)$$

By applying $\partial E / \partial \vec{c} = 0$ we obtain the least squares estimate of the joint centre as

$$\vec{c} = \left[2N_t I_3 - \sum_t R^0(t)^\top R^1(t) - \sum_t R^1(t)^\top R^0(t) \right]^{-1} \sum_t [\vec{p}^0(t) + \vec{p}^1(t) - R^0(t)^\top R^1(t) \vec{p}^1(t) - R^1(t)^\top R^0(t) \vec{p}^0(t)] \quad (9)$$

We note that the matrix inversion in this equation may not be possible, signifying that the joint centre may not be determined. A necessary condition to determine a joint centre is that at least 3 time instants are considered and that the rotations relative to $t = 0$ must be non-null and around different axes. This will ensure that the matrix is invertible.

7 Uncertainty Estimation

We wish to obtain an estimate of the uncertainty in the position of the joint centre. To do so requires that we have some understanding of the errors in the inputs to the computation, in this case the pose estimates. We have already seen that these errors may be estimated as part of the segment pose computation. The computation of the pose uncertainty is in turn based on knowledge of the uncertainty in the marker positions.

Since we expect the pose errors to have non-trivial structure it is important to take these errors into account during the joint estimation step. This provides some difficulty in that rotations do not form a linear vector space, so that a covariance may not be defined unless one linearises about some point in pose space and works with small rotational errors only. Pennec [14] gives a detailed and rigorous framework in which to handle such uncertainties. In particular he draws attention to the fact that the (small) errors must carefully be separated out from the (usually large) rotations under investigation before the usual statistical averaging steps are performed.

In this section we simply sketch the computation in outline. We use the joint constraint that states that (in the absence of noise)

$$0 = f^0(t) * \vec{c} - f^1(t) * \vec{c} \quad (10)$$

Our measurements are the poses $f^\alpha(t)$ which have covariance $W_f^\alpha(t)$. We wish to estimate the joint centre \vec{c} and its uncertainty W_c .

BMVC99

The joint constraint may be viewed as an implicit measurement equation

$$h(x_t, a) = 0 \quad (11)$$

where we wish to solve for state vector a based on measurements x_t . When discussing the non-linear least squares problem we introduce new symbols x_t and y_t for the real and pseudo generalised measurements. These are distinct from \bar{x}_i^α and $\bar{y}_i^\alpha(t)$ as used earlier in the paper.

In the case of our problem we see that the following substitutions must be made. Our measurement vector \hat{x}_t is the 12×1 column vector consisting of two 6 element poses $f^0(t), f^1(t)$. The measurement vector has a 12×12 covariance matrix

$$W_{x,t} = \begin{bmatrix} W_f^0(t) & 0 \\ 0 & W_f^1(t) \end{bmatrix} \quad (12)$$

This problem may be linearised as shown in the appendix and the solution for the joint centre and its covariance is given by equations (22) and (23).

The following identifications must be made: The Jacobian M_t is given by a 3×3 matrix

$$M_t = \frac{\partial h}{\partial a} = R^0(t) - R^1(t) \quad (13)$$

The other Jacobian that is needed is the 3×12 matrix

$$\frac{\partial h}{\partial x_t} = \left[\frac{\partial R^0}{\partial r^0} \vec{c}, I_3, -\frac{\partial R^1}{\partial r^1} \vec{c}, -I_3 \right] \quad (14)$$

Finally the 3×1 pseudo-measurement is given by

$$\hat{y}_t = -\bar{d}^0 + \bar{d}^1 \quad (15)$$

It is necessary to iterate as the result depends on the linearisation point. We initialise the iteration using the analytic computation of section 6.

8 Results

An experiment was performed with a 5 camera optical system. One of the authors held by hand a two segment mechanical test rig on which the markers were mounted and swung it like a pendulum making sure that over the 20 second time interval the second segment moved around at least 2 of the axes.

The experiment was performed with a mechanical rig as the main focus of our work is to determine what can be estimated for rigid objects. A subsequent investigation is necessary to look more closely at the effects of soft tissue movement on the rigidity assumption.

A 20 second (1200 frame) data set of a 2 segment kinematic chain with 3 markers per segment was acquired. In order to predict the uncertainty of the rotation centre we need to supply an estimate of the uncertainty in the marker positions.

We assume that the marker errors are isotropic which is reasonable since the cameras are separated by angles of the order of 90 degrees. To estimate the errors we consider a pair of markers on a rigid segment well separated in distance d . We define the rms error as $\sigma_d = E[(d - \bar{d})^2]$. Because this is a combination of two measurements it is an estimate of twice the diagonal element of the marker covariance matrix, i.e. $3\sigma_d^2 = 2Tr(W_y)$. We combined many pairs of measurements to obtain an estimate of the marker covariance matrix from a value of $\sigma_d = 1.4mm$ and hence $\sqrt{tr W_y} = 1.7mm$.

We show the results in table 1. We repeat the computation for varying number of frames. For each case we list the trace of the covariance of the joint position.

BMVC99

Table 1: Tests of the uncertainty prediction for varying t

N_t	$\sqrt{tr(W_c)}$ (mm)	v
4	1872.02	0.46
8	96.04	2.62
16	76.60	3.18
32	35.93	2.97
75	3.04	1.82
150	1.10	3.05
300	0.50	2.99
600	0.30	1.38
1200	0.20	-

For the full 20 second sequence it is estimated that we can predict the joint center location to 0.20mm. As we decrease the number of frames two effects cause the uncertainty in joint location to increase. Firstly there are fewer measurements being used. Secondly the angular range begins to decrease. In the case of $N_t = 4$ only 0.07 seconds have elapsed and the angular range is very small causing an uncertainty of 1900mm.

Using the 1200 frame value for the joint centre as ground truth we have computed a validation index for the smaller N_t cases. The validation index is based on the Mahalanobis distance e and is defined as

$$v = \sqrt{\frac{e}{3}}, \quad e = (c - c_{true})^T W_c^{-1} (c - c_{true}) \quad (16)$$

The Mahalanobis distance is computed from the ground truth c_{true} , the predicted covariance W_c and the actual estimate c . The expected value of the validation index for a number of trials is 1 if the covariance estimate is reliable.

In table 1 we show results for the validation index. We have used the 1200 frame value for the joint centre as “ground truth” and computed the validation index for shorter sequences. We justify this by arguing that the long frame value is clearly very much more accurate than the short sequences.

The rows in the table are only partially statistically independent. The validation index is a little high and future trials will investigate this. Overall it seems that the covariance has very much the expected behaviour.

The final row suggests that a 20 second data sequence with markers accurate to 1.7mm rms can yield a joint centre accurate to 0.20mm rms, which is impressive in comparison with other methods.

The individual sets of 3 markers were about 100-200mm apart and the two sets were about 1m apart with the joint halfway between.

9 Conclusion

The problem of hip joint centre location using optical markers located on the skin has been addressed. We have presented a novel computational technique which can recover joint centre location and uncertainty based solely on the marker measurements. We have presented preliminary results that suggest our accuracy estimates are reasonable.

There are several issues that remain to be addressed. The algorithm described is very close to optimal for Gaussian noise and no significant improvements may reasonably be expected in this framework. However when considering the problem

of hip joint location soft tissue movement may be modelled by better noise models that exploit likely systematic correlations. Work is in progress to extend the scheme to address problems with larger numbers of segments.

A Solving non-linear least squares problems

In this appendix we present details of the approach to linearisation of the measurement equations following the work of Pennek and Thirion [14].

We suppose that we have $i = 0..N_c - 1$ non-linear constraint equations

$$h_t(x_t, a) = 0, \quad (17)$$

which depend on measurement vectors x_t and a state vector a to be determined. In the absence of noise the constraint is a strict equality, we use a hat ($\hat{\cdot}$) to denote noisy or estimated quantities.

The Taylor expansion of the constraint equations to first order about some estimate of the state vector \hat{a} and a noisy measurement \hat{x}_t is given by

$$-h(\hat{x}_t, \hat{a}) + \frac{\partial h}{\partial a} \hat{a} = \frac{\partial h}{\partial a} a + \frac{\partial h}{\partial x_t} (x_t - \hat{x}_t) \quad (18)$$

$$y_t = M_t a + w_t \quad (19)$$

In equation (19) we see that the first order equation may be converted to a conventional linear estimation problem by appropriate definitions of the ‘‘pseudo-measurement’’ y_t and the Jacobian matrix M_t . The covariance of the noise on the pseudo measurement y_t is denoted $W_{y,t}$ and is computed using the covariance of the noise on the real noisy measurement \hat{x}_t from

$$W_{y,t} = \frac{\partial h}{\partial x_t} W_{x,t} \frac{\partial h}{\partial x_t}^\top \quad (20)$$

We seek the unbiased minimum variance estimate \hat{a} . This is defined as the value of a which has expectation value equal to the true value of a

$$E[\hat{a}] = a \quad (21)$$

for which the expectation value of the variance $E[\hat{a}^\top \hat{a}]$ is a minimum.

The solution is given by the Gauss-Markov theorem [11]

$$\hat{a} = \left[\sum_t M_t^\top W_{y,i}^{-1} M_t \right]^{-1} \left[\sum_t M_t^\top W_{y,i}^{-1} y_t \right] \quad (22)$$

with covariance

$$W_a = \left[\sum_t M_t^\top W_{y,i}^{-1} M_t \right]^{-1} \quad (23)$$

References

- [1] J. K. Aggarwal and Q. Cai. Human motion analysis: a review. *Computer Vision and Image Understanding*, 73(3):428–440, 1999.
- [2] A.P.Ashbrook and R. B. Fisher. Segmentation of range data for the automatic construction of articulated. In *IEEE Nonrigid and articulated motion workshop*, pages 62–69, Puerto Rico, 1997.

- [3] A.P.Ashbrook, R. B. Fisher, C. Robertson, and N. Werghi. Segmentation of range data into rigid subsets using surface patches. In *6th Int. Conference on Computer Vision*, pages 201–206, Bombay, India, 1998.
- [4] A. L. Bell, D. R. Pederson, and R. A. Brand. A comparison of the accuracy of several hip center location prediction methods. *Journal of Biomechanics*, 23(6):617–621, 1990.
- [5] C. Bregler. Learning and recognising human dynamics in video sequences. In *IEEE Conf. on Vision and Pattern Recognition*, pages 568–574, Puerto Rico, 1997.
- [6] C. Bregler and J. Malik. Tracking people with twists and exponential maps. In *IEEE Conf. on Vision and Pattern Recognition*, pages 8–15, Santa Barbara, California, 1998.
- [7] T. Heap and D. Hogg. Automated pivot location for the cartesian-polar hybrid point distribution model. In *British Machine Vision Conference*, pages 97–106, Birmingham, England, 1995.
- [8] E. Hunter, P. H. Kelly, and R. C. Jain. Estimation of articulated motion using kinematically constrained mixture densities. In *IEEE Nonrigid and articulated motion workshop*, pages 10–17, Puerto Rico, 1997.
- [9] K. Kanatani. Analysis of 3D rotation fitting. *IEEE Trans. Pattern Analysis and Machine Intell.*, 16(5):543–549, 1994.
- [10] T.-W. Lu and J. J. O’Connor. Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints. *Journal of Biomechanics*, 32:129–134, 1999.
- [11] D. G. Luenberger. *Optimisation by vector-space methods*. John Wiley and Sons, New York, U.S.A., 1969.
- [12] D. D. Morris and J. Rehg. Singularity analysis of articulated object tracking. In *IEEE Conf. on Vision and Pattern Recognition*, pages 289–296, Santa Barbara, California, 1998.
- [13] R. R. Neptune and M. L. Hull. Accuracy assessment of methods for determining hip movement in seated cycling. *Journal of Biomechanics*, 28(4):423–437, 1995.
- [14] X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, 1997.
- [15] J. Rehg and T. Kanade. Visual tracking of high dof articulated structures: An application to human hand tracking. In *Third European Conference on Computer Vision*, pages 35–46, Stockholm, Sweden, 1994.
- [16] J. Rehg and T. Kanade. Model-based tracking of self-occluding articulated objects. In *5th Int. Conference on Computer Vision*, pages 612–617, Cambridge, Massachusetts, 1995.
- [17] S. Wachter and H.-H. Nagel. Tracking of persons in monocular image sequences. In *IEEE Nonrigid and articulated motion workshop*, pages 2–9, Puerto Rico, 1997.
- [18] M. Yamamoto, A. Sato, and S. Kawada. Incremental tracking of human actions from multiple views. In *IEEE Conf. on Vision and Pattern Recognition*, pages 2–7, Santa Barbara, California, 1998.