Journal of Biomechanics 49 (2016) 1854–1858

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Estimating missing marker positions using low dimensional Kalman smoothing

M. Burke [a,b,*], J. Lasenby [a]

[a] University of Cambridge, Cambridge CB2 1PZ, UK
[b] Mobile Intelligent Autonomous Systems, Council for Scientific and Industrial Research, Pretoria, South Africa

ARTICLE INFO

Article history:
Accepted 20 April 2016

Keywords:
Motion capture
Missing markers
Kalman filter
SVD

ABSTRACT

Motion capture is frequently used for studies in biomechanics, and has proved particularly useful in understanding human motion. Unfortunately, motion capture approaches often fail when markers are occluded or missing and a mechanism by which the position of missing markers can be estimated is highly desirable. Of particular interest is the problem of estimating missing marker positions when no prior knowledge of marker placement is known. Existing approaches to marker completion in this scenario can be broadly divided into tracking approaches using dynamical modelling, and low rank matrix completion. This paper shows that these approaches can be combined to provide a marker completion algorithm that not only outperforms its respective components, but also solves the problem of incremental position error typically associated with tracking approaches.

1. Introduction

Motion capture is a frequently used tool in biomechanics, typically relying on multiple cameras to track markers placed on joints or limbs of interest. Unfortunately, occlusions and marker detection failures often result in a number of missing markers. Missing markers are common in motion capture applications, and typically result in a large amount of time being spent manually correcting marker trajectories before any in-depth analysis can occur.

The traditional approach to solving this problem uses manual marker correction (sometimes using spline or linear interpolation) or relies on skeleton fitting (Herda et al., 2000). The former is typically inaccurate and only suitable for short occlusion durations, while the latter usually requires specific marker placement and limits researchers to standard skeleton rigs. More recently, a number of studies (Xiao et al., 2011; Lai et al., 2011; Tan et al., June 2013) have shown how missing marker positions can be estimated by using matrix factorisation techniques. While these can be effective, they are sometimes slow, and implementations often inaccessible to many biomechanics practitioners.

Tracking approaches that use dynamical motion models and temporal information to fill in missing trajectories have also been proposed previously, but these are often disregarded due to potential difficulties in designing dynamical models and concerns about efficacy. Wu and Boulanger (2011) use a Kalman filter together with a constant velocity motion model to estimate marker positions, but this approach is extremely susceptible to drift. Unfortunately, this latter behaviour has led a number of works (Feng et al., 2014; Federolf, 2013; Xiao et al., 2011; Baumann et al., 2011) to disregard Kalman filtering, under the misconception that this behaviour applies regardless of the dynamical model used.

Various approaches that attempt to learn a dynamical model while estimating marker positions have been proposed (Li et al., 2009), but are typically too slow to be of practical use.

Liu and McMillan (2006) use a family of low dimensional local linear models trained using multiple prior recordings. When a new sequence is provided, this approach selects an appropriate model for each frame using a random forest classifier, before predicting missing marker positions using the appropriate model. Unfortunately, this approach requires a priori training data, and is only applicable if the markers are placed in fixed, predefined positions that match those used in training.

Aristidou et al. (2008) utilise the fact that the distance between markers on a given limb should remain constant to estimate centres of rotation for limbs, and use centres of rotation tracked using a Kalman filter to infer marker position. Unfortunately, this approach requires that multiple markers are present for each limb, assumes limbs are rigid bodies, and is vulnerable to increasing error as the duration for which markers are missing increases. Tracking using an improved variable turn model and an unscented Kalman filter (Aristidou and Lasenby, 2013) lessens this effect, but still assumes the presence of rigid limbs with at least 3 markers placed on each limb.

* Corresponding author at: Council for Scientific and Industrial Research, Pretoria, South Africa.

http://dx.doi.org/10.1016/j.jbiomech.2016.04.016
0021-9290/

Dorfmüller-Ulhaas (2003) track the rotation and translation of rigid bodies using an extended Kalman filter (a linear approximation to a non-linear motion model is used for state predictions and update), assuming a constant angular velocity rotation model and a constant acceleration translation model. Unfortunately, this approach requires specific marker placement for the rigid bodies to be detected, and is also liable to drift. This approach is similar to that described by Cerveri et al. (2003), who used an extended Kalman filter and second order motion model to track joints using 2D image measurements obtained from multiple cameras.

This paper addresses the drift problem in tracking approaches through the introduction of a fast and accurate marker completion algorithm that combines temporal smoothing and matrix factorisation, and which is accessible to the biomechanics community.[1] The primary contributions of this paper are as follows:

- We show that two previously separate families of approaches to marker completion are in fact complimentary and can be combined.
- We show how the most commonly listed flaw in tracking approaches (incremental error when markers are missing for extended periods of time) can be avoided.
- We provide an approach to marker completion that is completely data driven, requiring no prior knowledge of marker placement or number and making no assumptions about the presence of rigid bodies.

2. Missing marker position estimation

We briefly describe the two primary families of approaches to missing marker completion, before introducing our approach, which combines elements of both.

2.1. Marker smoothing

Knowledge of the expected motion of markers in the form of a dynamical model can be used together with detected markers to estimate marker positions. Let $\mathbf{x}_t$ denote the position vector of markers at time t, and $\mathbf{z}_t$ the set of measured marker positions. Assuming N markers, the state vector is constructed as $\mathbf{x}_t = [x_1, y_1, z_1 \ldots x_N, y_N, z_N]^T$. Our goal is to estimate $\mathbf{x}_t$ using measurements $\mathbf{z}_t$.

The Kalman filter (Kalman, 1960) is frequently used for problems like these, as it provides an optimal solution to tracking problems when states are governed by linear Gaussian motion and observation models. Let us assume that a state $\mathbf{x}_t$ evolves as

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \tag{1}$$

with process noise $\mathbf{w}_t$ drawn from a zero-mean Gaussian distribution with process covariance $\mathbf{Q}$, and $\mathbf{F}_t$ a linear transition matrix that describes how states are likely to evolve over consecutive time steps. Further, let us assume that we can obtain measurements $\mathbf{z}_t$, which are related to the state at time t by the equation

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t, \tag{2}$$

with measurement or observation noise $\mathbf{v}_t$ drawn from a zero-mean Gaussian distribution with measurement covariance $\mathbf{R}$, and $\mathbf{H}_t$ a linear measurement matrix that maps measurements to states. The Kalman filter provides an optimal estimate of the state $\mathbf{x}_t$ given a history of measurements up to time t for models of this form. In post-processing applications, the Kalman filter can be extended to the Rauch–Tung–Striebel (RTS) smoother (Rauch et al., 1965), which provides an optimal estimate of the state $\mathbf{x}_t$ given all measurements in the sequence (see Appendix A for the smoothing recursions).

Unfortunately, it can be hard to design the dynamical model $\mathbf{F}_t$, particularly without prior knowledge of marker placement, expected motion and the relationship between markers. A naive approach would be to model each marker's motion independently, but this failure to account for marker correlations limits the achievable accuracy and results in large errors if markers are missing for prolonged periods of time.

2.2. Low rank matrix completion

Motion capture sequences are typically recorded at an extremely high frame-rate, and there is often little change in motion over consecutive frames. As a result, the sequences can usually be described in a low dimensional space. This property has led to a number of approaches that try and find missing marker positions by using a low dimensional representation of motion capture sequences to reconstruct the original data.

We briefly illustrate these techniques using a representative approach termed mSVD (Srebro et al., 2003), but there are multiple decompositions that could be used. Let $\mathbf{X}$ denote a $T \times d$ training set, formed by stacking all marker position vectors in a motion capture sequence horizontally. Here, T is the length of the motion capture sequence, while d denotes the dimensionality of the state vector $\mathbf{x}_t$, typically $3N$, where N is the number of markers.

A low dimensional representation of this matrix can be obtained by performing singular value decomposition (SVD), a factorisation of a matrix into the form

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*, \tag{3}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, $*$ denotes a conjugate transpose, and $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values of $\mathbf{X}$, all positive and listed in decreasing order (Stewart, 1993). The magnitude of the singular values can be viewed as a measure of a mode's (columns of $\mathbf{U}$) contribution to the matrix $\mathbf{X}$. A low rank approximation of the matrix $\mathbf{X}$ can be obtained by discarding the modes and basis functions (rows of $\mathbf{V}$) of $\mathbf{X}$, which correspond to singular values of smaller magnitude.

mSVD is an iterative approach that decomposes motion capture data, discards a portion of the basis functions, reconstructs the original data, replaces missing values in the sequence with reconstructed ones, and repeats until convergence. Essentially, this approach uses the low dimensional representation of sequences to find correlations between markers, and uses markers that are present to provide information about those that are missing. Unfortunately, this can be slow and memory intensive, as it relies on multiple decompositions.

3. Low dimensional Kalman smoothing

In previous sections, we introduced smoothing and low rank matrix completion approaches to missing marker problems. The former requires the design of a complex motion model typically utilising knowledge of marker placement, while the latter can be slow and memory intensive due to its iterative nature. Our approach combines the two by projecting markers into a lower dimensional space learned from the sequence, performing Kalman smoothing in this space using a random walk motion model and then returning to the original space, using correlated markers to reduce the average error in each marker position estimate.

In our formulation we let $\mathbf{X}$ denote an $M \times d$ training set, formed by taking M position vectors in a motion capture sequence

[1] Matlab and Python implementations are available at https://github.com/mgb45/MoGapFill, together with a Python plugin for Vicon Nexus.