

HUMAN BODY MODELLING USING QUADRATIC DEFORMATIONS

João K. Fayad¹, Alessio Del Bue², Lourdes Agapito¹, and Pedro M. Q. Aguiar²

¹Queen Mary, University of London
Mile End Road, E1 4NS London, UK
e-mail: {jrkf,lourdes}@dcs.qmul.ac.uk

²Instituto de Sistemas e Robótica - Instituto Superior Técnico
Av. Rovisco Pais, 1049-001 Lisbon, Portugal
e-mail: {adb,aguiar}@isr.ist.utl.pt

Keywords: Human Body Models, Biomechanics, Factorization, Non-linear Modelling.

Abstract. *The inference of the skeleton structure and joint locations of the human body is of focal interest in many applications such as clinical analysis and sport performance optimization. One of the major assumptions when dealing with this problem is that the body parts are rigid or quasi-rigid. However, soft tissue artifacts may appear with muscular contraction when observing the motion of the body. In this paper, we tackle this problem by proposing a more general shape model that accounts for quadratic deformations. Our approach takes Motion Capture (MOCAP) data as input and enables the extraction of more accurate estimates for the rigid component of the different body segments using a factorization framework. The parameters of the model are computed using a Levenberg-Marquardt non-linear optimization scheme. The proposed method can later be combined with other approaches that allow the construction of an articulated skeleton based on the rigid segments to create a model of the human body. The algorithm was extensively tested using synthetic data, and its performance assessed by comparison with ground truth. We also present a qualitative analysis of the results obtained with MOCAP sequences.*

1 INTRODUCTION

Human motion analysis is an essential tool to investigate the dynamic evolution of body articulations and its anthropometry. In this context, one of the major assumptions to date has been to consider the body as a union of rigid articulated shapes. However, this assumption is far from being realistic. The soft-tissues (muscles and skin) that surround the bone structure can flex and stretch causing non-rigid deformations. Therefore, approximating the body as rigid may lead to consistent errors in the measurements.

In this work we focus on human motion analysis using optical MOCAP systems where the 3D coordinates of some reflective markers placed on the body are recovered from their 2D coordinates viewed by a set of synchronised cameras. These systems are nowadays very popular as they are less invasive than other solutions. Optical MOCAP systems are not restricted in their use to the field of medicine where their main application is to gait analysis [1]. Athletes use MOCAP data to optimise their performance given the analysis of their natural pose and motion [2, 3] and in computer graphics, MOCAP data is employed to animate 3D models of synthetic actors [5]. The fields of identification, security and surveillance [?] have also benefited from the use of 3D motion data.

This paper deals with the applications to medicine and sports, where an accurate description of the properties of the articulations is required. This contrasts with other fields where realistic appearance suffices [6]. In medicine, MOCAP systems are used for *gait analysis* since the study of alterations to normal gait patterns are of great importance both for diagnosis and treatment in conditions such as cerebral palsy, prosthetic limbs or orthoses and total joint replacements [1, 7, 8]. In this scenario it is crucial to obtain accurate estimates of the exact location of the joints and their axes of rotation which are not biased by the deformations caused by muscles and skin. While the relevant information used to build articulated human models is in fact given by the bone structure, the reflective markers used by MOCAP systems are inevitably placed on the skin. When examining a subject, there is an inherent relative non-rigid motion between the soft-tissues surrounding the bone and the bone itself. These deformations create artifacts on the data that degrade the accuracy of the recovered model. These *soft-tissue artifacts* [6, 9, 10] are one of the major sources of error in current algorithms.

Our proposed solution attempts to model the *soft-tissue artifacts* explicitly with a quadratic model which is able to describe such deformations and at the same time be subject-specific. This latter aspect is important since *a priori* models may not fully encode the variability of the human body. For instance, different medical conditions such as obesity may drastically change the shape and the deformations of the studied subject. Besides, methods that do not rely on positioning of the markers in specific locations are preferable since they have better chances of not introducing further errors in the estimation and the specific landmarks might not be easily found on all patients [1].

In summary, the two main current challenges of human motion analysis systems based on MOCAP are dealing with soft-tissue artifacts and achieving subject specific modelling. In this paper we address both these problems by proposing a new algorithm to model the non-rigid motion caused by soft tissue artifacts using a quadratic deformation model. Our technique does not assume a specific configuration of the markers, requiring only that they cover the whole body part. Since no prior models are used, all the results will be specific to the sequence being analysed

1.1 Previous Work

Systems used to capture human motion are generally optical MOCAP systems, based on markers placed on the skin of the subject performing the motion. Most of the existent clinical systems for gait analysis use markers placed at specific anatomic landmarks. After acquiring the 3D trajectories, *a priori* gait models (such as the *Helen Hayes* model) are fitted to the observations using standard regression techniques. In general, the hip, knee and ankle articulations are modelled as joints with 3 degrees of freedom [1]. This setup has the disadvantage of requiring accurate marker placement on the exact anatomical landmarks. Evidence has been brought up that most common equations used for regression provide unsatisfying results [11].

Other approaches have instead used anatomical calibration to fit the motion of the subject to an existent model of the joints. In this method the subject performs *predetermined* motions designed to facilitate the estimation of the joint model parameters [1], which are then captured using MOCAP systems. The disadvantage of this method is that in some medical conditions the predetermined motions are impossible to perform.

As mentioned in Section 1, one of the major sources of error in the 3D measurements are the soft-tissue artifacts. Several attempts have been made to cope with this problem including the use of optimization techniques to fit the motion capture data to a model. Some attempts have focused on describing the bone displacement by observing the soft tissue deformation over the entire range of motions [12, 13]. The main issue here is the difficulty in defining which is the true bone motion since there is no clear consensus for this kind of measurements [1].

As MOCAP systems evolved, a higher number of points were possible to track. This allowed the development of point cluster techniques (PCTs) based on the redundancy of the information to reduce the effect of soft-tissue artifacts [14]. This technique was later extended to deal directly with soft-tissue artifacts by reformulating the problem as a chi-squared estimation. However, the functional form which models the relative displacement between bone and the skin has to be explicitly provided [12] and it is subject to the action performed by the patient.

In a different context, human body modelling from video sequences has recently seen significant advances in the field of Computer Vision. Due to the high variability in image data, subject specific modelling is the natural choice for this type of inference. In particular, Structure from Motion (SfM) methods can obtain 3D models and joint localization exclusively from the observed 2D image data.

The original factorization algorithm for SfM was able to recover the shape and motion of a rigid object moving in a scene [15, 16] viewed by an orthographic camera. Later, the approach was extended to recover structure and motion of various objects moving independently [17], and to deal with non-rigid objects with small deformations, using linear combinations of basis shapes [18, 19, 20, 21]. Finally, factorization methods were also extended to deal with articulated objects. While the first approaches were based on fitting MOCAP data to previously existent models of the objects [22, 23], recent developments resulted on approaches that are able to infer articulated structure based solely on the 2D motion data [24, 25].

These factorization approaches have inspired recent work that use 3D data provided by MOCAP systems in order to build articulated models of the human body, first considering the body segments as rigid [26]. Later, a weighted factorization method was proposed in order to deal with the non-rigidity of the human body segments [?].

Although marker-based MOCAP systems are the most popular systems in clinical research, alternative approaches have also been considered. Among them, there are methods such as stereo radiography, bone pins, external fixation devices, and single or double plane fluoroscopy.

However these methods are either invasive, they limit the range of motions, or they require the exposure of the subject to radiation [1]. Models based on magnetic resonance imaging (MRI) have also been used [28, 29]. Although they can provide a detailed image data of bones and muscles, this technique only works within small volumes and so their application is limited.

1.2 Objectives

In this paper we present a method for modelling non-rigid deformations of human body parts. This method is based on the Structure from Motion (SfM) approaches described in Section 1.1 but for the case of 3D data and non-rigid bodies. While previous approaches have attempted to minimize the contribution of non-rigid points to the overall motion, by modelling the deformations of body segments explicitly we are able to model the non-rigid nature of soft-tissues. This allows the underlying rigid component of the body segment to be recovered more reliably. The rigid component can later be linked with the approach described in [26] in order to create more accurate articulated models for the human body.

2 FACTORIZATION METHODS FOR STRUCTURE FROM MOTION

Factorization methods for SfM decompose the measurement matrix (which contains the coordinates of the tracked points in each frame) into the product of two factors: *motion* and *shape*. The shape parameters represent the 3D geometric properties of the object while the motion parameters represent the time-varying parameters of the motion (*e.g.* rotations and translations of the rigid body) that the shape performs in a metric space.

2.1 Rigid objects

A factorization method for 3D MOCAP data was proposed in [26] as an extension of the original method first proposed by Tomasi and Kanade [15] for 2D measurements. The approach takes as input the 3D coordinates of set of P feature points tracked over F frames by a MOCAP system. These trajectories are arranged into a $3F \times P$ measurement matrix W as:

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_F \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_F \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \cdots & x_P \\ y_1 & y_2 & \cdots & y_P \\ z_1 & z_2 & \cdots & z_P \end{bmatrix} + \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_F \end{bmatrix} = MS + T, \quad (1)$$

where R_i and T_i are respectively a 3×3 rotation matrix and a $3 \times P$ translation matrix of the rigid motion at frame i . The vector $[x_j y_j z_j]^T$ represents the 3D coordinates of point j in the local reference system (see Figure 1). The translational component T_i encodes the coordinates of the centroid of the point cloud at each frame W_i . If the shape centroid at frame i is expressed as \mathbf{t}_i the matrix translation T_i is given by $T_i = \mathbf{t}_i \mathbf{1}_P^T$ where $\mathbf{1}_P^T$ is a vector of P ones. Thus, it frequently occurs that, instead of W , we consider a registered form of this matrix *i.e.* we use a matrix \tilde{W} such that:

$$\tilde{W} = W - T = MS. \quad (2)$$

It is easy to see that W is rank deficient, with $rank(W) \leq 4$ in the noiseless case. In the registered case $rank(\tilde{W}) \leq 3$. However, when performing real experiments there will always be some noise which will increase the rank of \tilde{W} . Noise can be caused by the uncertainty in the position of the tracked feature points in the MOCAP system or some non-rigidity of the tracked objects.

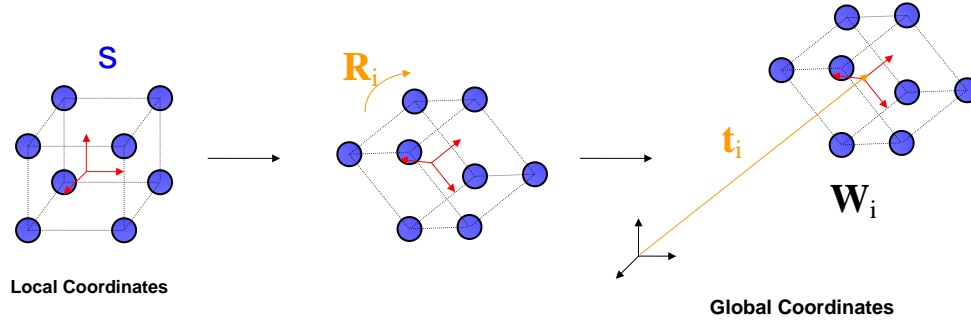


Figure 1: Graphical representation of the physical meaning of the motion and shape factors. The shape matrix S contains the 3D coordinates of the (blue) points that define the body, on the local (red) referential. The rotation matrix R_i and translation vector t_i represent the coordinate transformations that describe S on the global (black) referential, resulting in W_i .

Consider the truncated singular value decomposition (SVD) of matrix \tilde{W} defined by:

$$\hat{W} = U_k \Sigma_k V_k^T, \quad (3)$$

where Σ_k $k \times k$ is a diagonal matrix with the largest k singular values of \hat{W} , U_k $3F \times k$ is a matrix with the correspondent k left singular vectors and V_k $P \times k$ a matrix with the correspondent k right singular vectors. This result is the best approximation of \tilde{W} to a $rank - k$ matrix in the Frobenius norm sense. Since we know the ideal value for $rank(\tilde{W})$ to be 3, we can use a $rank - 3$ truncated SVD as a first global optimal fit to the measurements. This factorization not only reduces noise but also provides a way to separate the data into a motion factor and a shape factor:

$$\hat{M} = U_3 \Sigma_3^{1/2}; \quad (4)$$

$$\hat{S} = \Sigma_3^{1/2} V_3^T. \quad (5)$$

Note that $\hat{M} \neq M$, as this factorization provides no guaranty that \hat{M} is a collection of rotation matrices. Similarly the shape \hat{S} is not guaranteed to lie in Euclidean space. There exists an ambiguity in this factorization as any invertible 3×3 matrix A will satisfy the equality:

$$\hat{M} \hat{S} = \hat{M} A A^{-1} \hat{S}. \quad (6)$$

This ambiguity is solved by choosing the invertible transformation A such that $M = \hat{M}A$. Similarly, $S = A^{-1}\hat{S}$. Details on the computation of A can be found in [26].

When N rigid objects are moving independently, this same approach can be taken considering $rank(\tilde{W}) \leq 3N$ or $rank(W) \leq 4N$, and the motion and shape factors are upgraded accordingly:

$$\bar{W} = \begin{bmatrix} M_1 & M_2 & \dots & M_N \end{bmatrix} \begin{bmatrix} S_1 & & & \\ & S_2 & & \\ & & \dots & \\ & & & S_N \end{bmatrix}. \quad (7)$$

2.2 Rigid articulated objects

In the case of two rigid bodies linked by a joint there is a loss in the degrees of freedom of the motion when compared with two independently moving rigid objects. This causes a decrease in the rank of W [24, 25]. For the sake of simplicity only two bodies will be considered. Still, the formulation is analogous for longer chains.

2.2.1 Universal joint

We define a universal joint as a joint in which each of the two bodies is at a fixed distance from a common joint centre, with their rotations independent (see Figure 2). This is also known as a *spherical joint* in the mechanics literature.

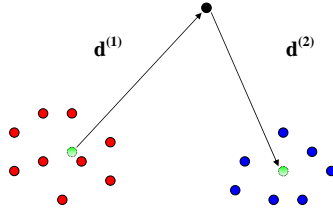


Figure 2: Universal joint. The first body is represented by the red points, while the second body is represented by the blue points. The joint centre is shown as a black point. The 3-vector $\mathbf{d}^{(1)}$ stands for the 3D coordinates of the joint centre, in the local coordinate system of the first body. The 3-vector $\mathbf{d}^{(2)}$ stands for the 3D coordinates of the joint centre, in the local coordinate system of the second body.

Let $\mathbf{d}^{(1)} = [u, v, w]^T$ be the 3D coordinates of the joint centre, in the local coordinate system of the first body; $-\mathbf{d}^{(2)} = [u', v', w']^T$ be the 3D coordinates of the joint centre, in the local coordinate system of the second body; $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$ be the $3F \times 3$ matrices corresponding to a collection of 3×3 global rotation matrices over F frames, for the first and second bodies, respectively; and $\mathbf{t}^{(1)}$ and $\mathbf{t}^{(2)}$ be the $3F$ -vectors corresponding to the global translation vectors of the first and second bodies, respectively.

The geometric constraint of this joint is equivalent to stating that the joint centre satisfies the motion of each of the two bodies that form the joint. This implies W can be factorized as:

$$W = [W^{(1)} \mid W^{(2)}] = [\mathbf{R}^{(1)} \quad \mathbf{R}^{(2)} \quad \mathbf{t}^{(1)}] \begin{bmatrix} \mathbf{S}^{(1)} & \mathbf{D}^{(2)} \\ \mathbf{0}_{3 \times P_1} & \mathbf{S}^{(2)} + \mathbf{D}^{(2)} \\ \mathbf{1}_{P_1}^T & \mathbf{1}_{P_2}^T \end{bmatrix}, \quad (8)$$

where $W^{(1)}$ and $W^{(2)}$ are, respectively, the measurement matrices for the first and second body; $\mathbf{D}^{(2)} = \mathbf{d}^{(2)} \mathbf{1}_{P_2}^T$, $\mathbf{1}_{P_1}$ a P_1 -vector with all entries equal to 1 and $\mathbf{1}_{P_2}$ a P_2 -vector with all entries equal to 1, where P_1 and P_2 are the number of points belonging to the first and second bodies, respectively; and $\mathbf{0}_{3 \times P_1}$ is a $3 \times P_1$ zero matrix. In this case, we have that $\text{rank}(W) \leq 7$. Naturally, in order to separate W into $W^{(1)}$ and $W^{(2)}$, we must assume the body segmentation is known. Details on how to factorize W can be found in [26].

2.2.2 Hinge joint

We define a hinge joint as a joint in which two bodies can rotate around an axis, while their respective distances to the axis remain constant (see Figure 3). In this case, their rotations are not independent.

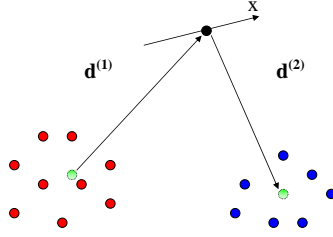


Figure 3: Hinge joint. The first body is represented by the red points, while the second body is represented by the blue points. The joint centre is shown as a black point. The x -axis represents the rotation axis. The 3-vector $\mathbf{d}^{(1)}$ stands for the 3D coordinates of the joint centre, in the local coordinate system of the first body. The 3-vector $\mathbf{d}^{(2)}$ stands for the 3D coordinates of the joint centre, in the local coordinate system of the second body.

By analysing the geometry of the joint, we see that every vector belonging to any of the two bodies that is parallel to the joint axis, must remain so throughout the movement. Without loss of generality, let us choose an appropriate local coordinate system, where the axis of rotation of the joint is coincident with the x -axis. Let us also keep the same notation used for the universal joint in Section 2.2.1. To comply with the joint constraints, the first column of $\mathbf{R}^{(1)}$ must be equal to the first column of $\mathbf{R}^{(2)}$. We can now define the rotation matrices as $\mathbf{R}^{(1)} = [\mathbf{c}_1 \ \mathbf{c}_2 \ \mathbf{c}_3]$ and $\mathbf{R}^{(2)} = [\mathbf{c}_1 \ \mathbf{c}_4 \ \mathbf{c}_5]$. In this case, the measurement matrix \mathbf{W} is given by

$$\mathbf{W} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \mathbf{c}_4 & \mathbf{c}_5 & \mathbf{t}^{(1)} \end{bmatrix} \begin{bmatrix} x_1^{(1)} & \cdots & x_{P_1}^{(1)} & x_1^{(2)} & \cdots & x_{P_2}^{(2)} \\ y_1^{(1)} & \cdots & y_{P_1}^{(1)} & 0 & \cdots & 0 \\ z_1^{(1)} & \cdots & z_{P_1}^{(1)} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & y_1^{(2)} & \cdots & y_{P_2}^{(2)} \\ 0 & \cdots & 0 & z_1^{(2)} & \cdots & z_{P_2}^{(2)} \\ & & \mathbf{1}_{P_1}^T & & & \mathbf{1}_{P_2}^T \end{bmatrix}, \quad (9)$$

Again, details about the factorization of \mathbf{W} can be found in [26].

2.3 Quasi-rigid objects

To account for the non-rigidities present on the human body, a weighted factorization approach was proposed in [27]. This method estimates the best rigid shape and motion factors of a non-rigid body by giving less weight in the computations to highly deforming points.

Let us first rearrange the data matrix \mathbf{W} defined in Eq. (1) as:

$$\tilde{\mathbf{W}} = \begin{bmatrix} \bar{\mathbf{w}}_{11}^T & \bar{\mathbf{w}}_{12}^T & \cdots & \bar{\mathbf{w}}_{1P}^T \\ \bar{\mathbf{w}}_{21}^T & \bar{\mathbf{w}}_{22}^T & \cdots & \bar{\mathbf{w}}_{2P}^T \\ \vdots & \vdots & & \vdots \\ \bar{\mathbf{w}}_{F1}^T & \bar{\mathbf{w}}_{F2}^T & \cdots & \bar{\mathbf{w}}_{FP}^T \end{bmatrix}, \quad (10)$$

where the matrices $\tilde{\mathbf{W}}_j = [\bar{\mathbf{w}}_{1j}^T \cdots \bar{\mathbf{w}}_{Fj}^T]^T$ contain the 3D coordinates of point j along the F frames. Let us also assume the rearranged data matrix corresponding to the motion of the best rigid description of the body is known and denoted by $\bar{\mathbf{w}}^{(r)}$. Given $\bar{\mathbf{w}}^{(r)}$, the 3D distance of the measured position and the ideal rigid position for point j can be defined as:

$$\mathbf{E}_j = \tilde{\mathbf{W}}_j^{(r)} - \tilde{\mathbf{W}}_j, \quad (11)$$

with E being an $F \times 3$ matrix. Given this definition, a point that presents a non-rigid motion will have a higher value for $\|E_j\|$ when compared to close to rigid motions. Thus, a natural definition for the weight matrix can be given by:

$$C_j = \text{cov}(E_j)^{-1}, \quad (12)$$

as higher errors will give higher covariance values.

The weighted factorization algorithm is an iterative approach that estimates the motion and shape factors in an alternated fashion. Since the best rigid description is not known, the observed data matrix W is used instead. The error defined by Eq. (11) is then computed against the reconstruction provided by the weighted factorization at the current iteration, while the initial estimates for the motion and shape factors are given by the rigid factorization approach described in Section 2.1. With this considerations, the motion and shape factors are computed using the following equations:

$$\mathbf{s}_j = \left(\sum_{i=1}^F \mathbf{M}_i^T C_j \mathbf{M}_i \right)^{-1} \sum_{i=1}^F \mathbf{M}_i \tilde{\mathbf{w}}_{ij}; \quad (13)$$

$$\mathbf{m}_i = \left(\sum_{j=1}^P \mathbf{S}_j^T C_j \mathbf{S}_j \right)^{-1} \sum_{j=1}^P \mathbf{S}_j \mathbf{w}_{ij}; \quad (14)$$

where the 12-vector \mathbf{m}_i and the 3×12 matrix \mathbf{S}_j are defined by:

$$\mathbf{m}_i = \begin{bmatrix} \mathbf{r}_{1i} \\ t_{xi} \\ \mathbf{r}_{2i} \\ t_{yi} \\ \mathbf{r}_{3i} \\ t_{zi} \end{bmatrix}; \quad (15)$$

$$\mathbf{S}_j = \begin{bmatrix} \mathbf{s}_j^T & 1 & & & & \\ & & \mathbf{s}_j^T & 1 & & \\ & & & & \mathbf{s}_j^T & 1 \end{bmatrix}. \quad (16)$$

Details about these equations can be found in [27].

Finally the algorithm can be summarised into the following steps:

1. Initialization: Compute the initial estimations for M , S and \mathbf{t} using the rigid body factorization method described in Section 2.1
2. Use the current estimates of \mathbf{t}_i for $i = 1 \cdots F$ to register the data matrix W .
3. With the current estimate of M and the registered matrix \hat{W} , compute a new estimation for S using Eq. (13).
4. Based on the estimate of S from Step 3, compute a new M and \mathbf{t}_i based on Eq. (14).
5. Repeat Steps 2 to 4 until convergence of the Frobenius norm of E is achieved.

3 QUADRATIC MODEL FOR NON-RIGID BODIES

The idea behind the weighted factorization approach was that an improved rigid representation of the human body segments could be achieved by giving less weight to the information given by non-rigid motion to the computations. Contrastingly, the new quadratic model for non-rigid bodies actually aims at modeling non-rigid motion, and thus it is able to recover the underlying rigid shape that gives origin to the motions in the first place. This rigid shape can again be used with the methods described in [26] to estimate more accurate joint parameters, as the rigid shape was estimated more accurately.

One of the most popular approaches when modelling deformable bodies is to approximate them as a linear combination of different rigid basis shapes [30, 18]. However, deformations occurring on the human body due to soft-tissue tend to have quadratic behaviour (*e.g.* muscle contractions), increasing considerably the number of basis shapes required to accurately approximate the deformations. This has a negative impact on computational costs, allowing us to use only a limited number of basis shapes, and thus compromising accuracy. Moreover, due to the high number of parameters to estimate, it is common to obtain various local minima when applying minimisation schemes to solve this problem.

If we are to deal with quadratic deformations, the most logical approach leads to use a quadratic model for deformable bodies. Inspired by previous works in the field of computer graphics [5, 31], we present a new quadratic model for non-rigid bodies using geometric constraints, built as an extension to the factorization-based rigid body model described in Section 2.1.

3.1 Quadratic model formulation

Our model for deformable bodies expands the rigid body formulation defined in Eq. (1), to a formulation that uses linear, quadratic and crossed-terms of the previous rigid shape matrix. Let us define the new shape matrix as:

$$\mathbf{S} = \begin{bmatrix} x_1 & x_2 & \dots & x_P \\ y_1 & y_2 & \dots & y_P \\ z_1 & z_2 & \dots & z_P \\ \hline x_1^2 & x_2^2 & \dots & x_P^2 \\ y_1^2 & y_2^2 & \dots & y_P^2 \\ z_1^2 & z_2^2 & \dots & z_P^2 \\ \hline x_1y_1 & x_2y_2 & \dots & x_Py_P \\ y_1z_1 & y_2z_2 & \dots & y_Pz_P \\ z_1x_1 & z_2x_2 & \dots & z_Px_P \end{bmatrix} = \begin{bmatrix} \mathbf{S}^{(\Gamma)} \\ \mathbf{S}^{(\Omega)} \\ \mathbf{S}^{(\Lambda)} \end{bmatrix}, \quad (17)$$

where $\mathbf{S}^{(\Gamma)}$ is the $3 \times P$ linear shape matrix, $\mathbf{S}^{(\Omega)}$ the $3 \times P$ quadratic shape matrix and $\mathbf{S}^{(\Lambda)}$ is the $3 \times P$ cross-terms shape matrix. Given this new structure of \mathbf{S} , we introduce the motion matrix \mathbf{M}_i defined as:

$$\mathbf{M}_i = \mathbf{R}_i \begin{bmatrix} \Gamma_i & \Omega_i & \Lambda_i \end{bmatrix}, \quad (18)$$

where \mathbf{R}_i is a 3×3 rotation matrix, and Γ_i is a 3×3 transformation matrix associated with linear deformations, Ω_i is a 3×3 transformation matrix associated with quadratic transformations and Λ_i is a 3×3 transformation associated with cross-terms deformations. By modeling the rotations in a separate matrix \mathbf{R}_i , we are defining the deformation matrices in the local referential of the body. Using the same formulation as in the model for rigid bodies, and stacking the equation

for all the F frames, we can now define:

$$\tilde{W} = \begin{bmatrix} R_1 & & & \\ & R_2 & & \\ & & \ddots & \\ & & & R_F \end{bmatrix} \begin{bmatrix} \Gamma_1 & \Omega_1 & \Lambda_1 \\ \Gamma_2 & \Omega_2 & \Lambda_2 \\ \vdots & \vdots & \vdots \\ \Gamma_F & \Omega_F & \Lambda_F \end{bmatrix} \begin{bmatrix} S^{(\Gamma)} \\ S^{(\Omega)} \\ S^{(\Lambda)} \end{bmatrix} = MS, \quad (19)$$

where \tilde{W}_i is the data matrix containing the 3D coordinates of the feature points, registered to the origin of the global referential. Note that this model is now characterized by a $rank(\tilde{W}) \leq 9$ constraint. This quadratic model is in fact an extension of the linear rigid body model defined in Section 2.1 to deal with quadratic and cross-value terms, while keeping the same factorization into *motion* and *shape* factors. Note that a rigid body is still easily expressed by this model if we make $\Gamma_i = I$, $\Omega_i = 0$ and $\Lambda_i = 0$ for every frame i . Accordingly, the rank constraints in this case will be still satisfied, giving $rank(\tilde{W}) \leq 3$. By combining the new shape matrices with the associated transformation matrices, we are able to model characteristic soft-tissue motions such as bending, bulging, jiggling or stretching. Detailed description about the role of the different matrices will be addressed in the following sections. For the sake of notation simplicity, we will only consider one frame of the motion, and the i index will be dropped. However, results can be easily extended to a general case.

3.1.1 Linear deformation and shape matrices

Before we can analyse the role of the linear deformation matrix Γ , a few considerations about the model must be done. Every full-rank 3×3 matrix can be expressed with an RQ decomposition, from which results a rotation matrix R and an upper triangular matrix Q . Since in our model rotations are fully given by R_i , a RQ decomposition of $[\Gamma \ \Omega \ \Lambda]$ should not incorporate a rotation component *i.e* the rotation matrix of that decomposition should be the 3×3 identify matrix I , otherwise we would have an ambiguity in the optimization of these components. This implies that Γ must be an upper triangular matrix. Thus, we can now define Γ as:

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \Gamma_{13} \\ 0 & \Gamma_{22} & \Gamma_{23} \\ 0 & 0 & \Gamma_{33} \end{bmatrix}. \quad (20)$$

In order to fully understand the role of Γ in the model, we applied different transformation matrices to a previously built synthetic cubic object (see Figure 4(a)). Due to the particular symmetry of the cube, results drawn from analysing the effects of the transformation matrices in one of the coordinate axis is easily extended to the other two coordinate axis.

For these synthetic tests we used $R = I$, $\Omega = 0$, $\Lambda = 0$, and S a 9×156 matrix representing the quadratic formulation of the synthetic cubic object. Examples of these experiments can be seen in Figure 4(b) and Figure 4(c).

From these experiments, we can conclude that the diagonal entries of the linear deformation matrix, Γ_{11} , Γ_{22} and Γ_{33} , are responsible for the linear expansion/compression of the object, in the direction of the x -axis, y -axis and z -axis respectively (see Figure 4(b)). The off-diagonal entries of Γ (Γ_{12} , Γ_{13} and Γ_{23}) are responsible for pure *shear* deformations (see Figure 4(c)). For instance, let us consider a vector \mathbf{a} initially parallel to the x -axis, and another vector \mathbf{b} initially parallel to the y -axis, and thus orthogonal to \mathbf{a} . Γ_{12} is responsible for increasing or decreasing the angle between \mathbf{a} and \mathbf{b} , without changing the orientation of \mathbf{a} . Thus the vectors will no

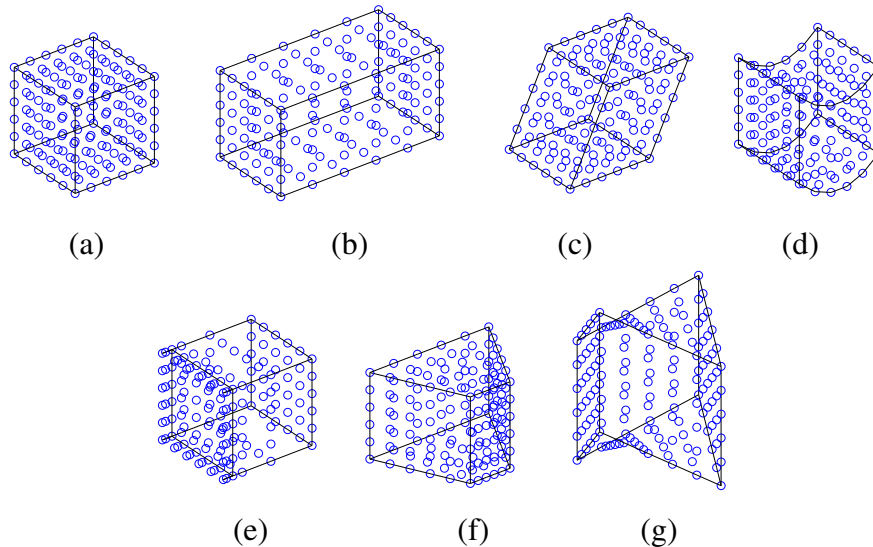


Figure 4: (a) Synthetic cubic object. (b) Example of an extension motion on the cubic object caused by $\Gamma_{11} = 1.5$. (c) Example of the shear deformation on the cubic object caused by $\Gamma_{13} = 0.5$. (d) Example of a interpenetration in the extension motion on the cubic object caused by $\Omega_{11} = 1$. (e) Example of the deformation caused by a non-diagonal entry, with $\Omega_{31} = 0.5$. (f) Example of the lateral contraction/extension deformation caused by $\Lambda_{11} = 0.5$ (g) Example of the twisting deformation caused by $\Lambda_{12} = 0.5$.

longer be orthogonal, and the object will present a shear deformation. Analogous results can be drawn for the other entries of the matrix.

3.1.2 Quadratic deformation and quadratic shape matrices

Following the same procedure as in Section 3.1.1, we examined the properties of the quadratic deformation and quadratic shape matrices by applying different transformations to the synthetic cubic object. For this case we used $R = I_{3 \times 3}$, $\Gamma = I_{3 \times 3}$, $\Delta = 0_{3 \times 3}$, and S as the 9×156 shape matrix of the synthetic cube.

Experiments show that there are two different bending/bulging possible deformations per coordinate axis. These deformations correspond to the off-diagonal entries of Ω (Ω_{12} , Ω_{13} , Ω_{21} , Ω_{23} , Ω_{31} and Ω_{32}) and they can be associated, for instance, to muscular contractions, which are one of the main sources of soft-tissue artifacts (see Figure 4(e)). Modelling these deformations is thus very important when dealing with this problem.

However, not every deformation represented by this matrix has a physical counterpart in this particular case. The three diagonal entries of Ω (Ω_{11} , Ω_{22} and Ω_{33}) represent a quadratic extension/contraction along the x -axis, y -axis and z -axis respectively. Still, not only has extension and contraction been (linearly) modelled by Γ , but also this transformation has the problem of plane interpenetration *i.e.* due to its quadratic nature, the relative order among the planes might change during the deformation. This is clearly visible in Figure 4(d), where one of the outermost planes in the initial configuration, where the edge of the cubic object is represented, is now one of the inner planes. Such deformations do not happen when modelling the human body and so, we will not allow these deformations on the model. Finally, we define Ω as:

$$\Omega = \begin{bmatrix} 0 & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & 0 & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & 0 \end{bmatrix}. \quad (21)$$

3.1.3 Cross-terms deformation and cross-terms shape matrices

To evaluate the effects of the cross-terms deformation matrix, following the same procedure used in the previous sections. In this case, we will use $R = I$, $\Gamma = I$, $\Omega = 0$ and S , as before, is a 9×156 matrix representing the quadratic formulation of the synthetic cubic object.

With this transformation matrix two different kinds of deformations are observed. One of the deformations is characterized by a lateral contraction in one side of the object, while on the opposite side a lateral extension is observed. This deformation has two modes per axis, and thus we have six matrix entries which gives such deformations: Λ_{11} , Λ_{22} , Λ_{33} , Λ_{13} , Λ_{21} and Λ_{32} (see Figure 4(f)).

The other kind of deformation observed can be described as *twisting* the object around each of the coordinate axis (see Figure 4(g)). The matrix entries responsible for this kind of motion are the remaining three entries: Λ_{12} , Λ_{23} and Λ_{31} .

On the other hand, the twisting deformation mode is not expected to happen when modelling human body parts, and so it will not be allowed to vary. Summarising, we define the cross-value deformations matrix as:

$$\Lambda = \begin{bmatrix} \Lambda_{11} & 0 & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & 0 \\ 0 & \Lambda_{32} & \Lambda_{33} \end{bmatrix}. \quad (22)$$

3.1.4 Model bounds

The deformations allowed by this model have physical meaning only up to a certain degree. For instance, we do not expect a body segment to expand indefinitely, or to contract until all the points collapse on a plane. Similar issues are present in all the deformation modes allowed by the model. Additionally, when applying the model without bounds for these values, even though the motion reconstruction is accurate, the rigid component of the shape factor will be different from the shape present on the motion (see Figure 5). This happens because if any deformation is allowed, the model will generate unrealistic deformations, to which will correspond unrealistic shape factors. Therefore, we must define an upper and lower bound to the entries of the transformation matrices as to prevent meaningless deformations. On the other hand, we do not want the model to become overconstrained, as this would prevent correct modelling of the soft-tissues. Thus, based solely on empirical evaluation of the effects of the bounds on the transformation matrices, we defined the upper and lower bounds as a ± 0.5 from the value of the parameters on the rigid body case. Finally, the upper and lower bounds can be formed as:

$$UB_{\Gamma} = \begin{cases} 1.5, & \text{for diagonal entries.} \\ 0.5, & \text{for off-diagonal entries.} \end{cases};$$

$$LB_{\Gamma} = \begin{cases} 0.5, & \text{for diagonal entries.} \\ -0.5, & \text{for off-diagonal entries.} \end{cases};$$

$$UB_{\Omega} = 0.5;$$

$$LB_{\Omega} = -0.5;$$

$$UB_{\Lambda} = 0.5;$$

$$LB_{\Lambda} = -0.5;$$

where LB stands for lower bound, and UB stands for upper bound.

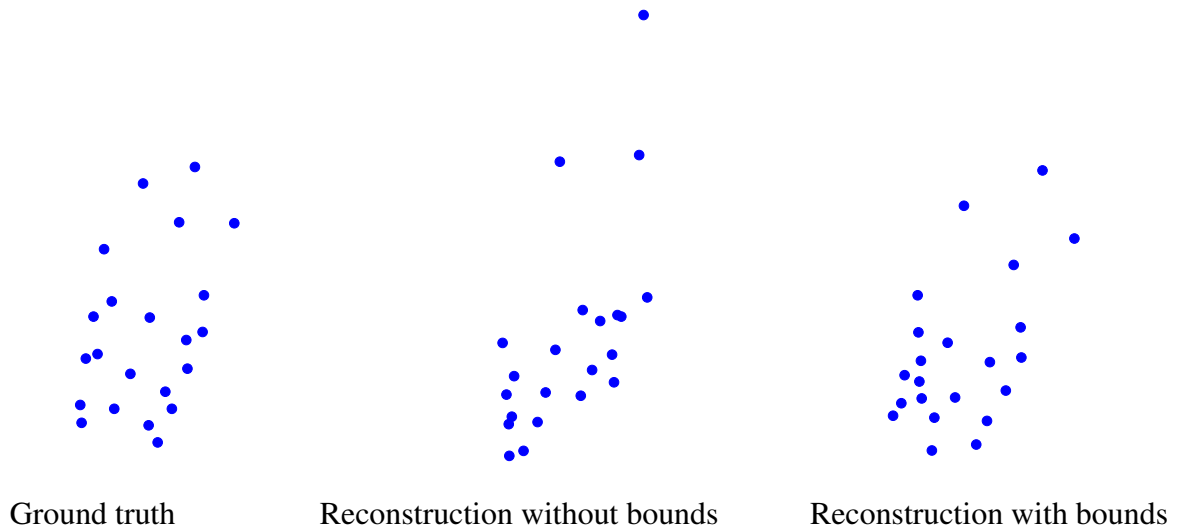


Figure 5: Left: an example of the shape observed on the measurement matrix. Middle: the rigid component of the shape matrix when using upper and lower bounds on the model. Right: the rigid component of the shape matrix if no bounds on the deformation are used. It is clear from these images that if no bounds are applied, the rigid shape recovered will be different from the observed shape during the motion.

As a further observation, biological soft-tissues are generally viewed as visco-hyperelastic, incompressible materials [32]. Thus, imposing a volume conservation constraint on our model is of the most importance, as it has a strong physical base. This can be done by requiring the determinant of the Linear deformation matrix to be unitary [31, 32]. This constraint can be written analytically by forcing the following relation:

$$\Gamma_{11}\Gamma_{22}\Gamma_{33} = 1.$$

3.2 Non-linear optimization with a quadratic model

Implementing an iterative method, similar to the one presented in Section 2.3 to solve the factorization problem with the Quadratic model proved not to be a straight forward task. The reason is mainly because the overall quadratic model is highly non-linear both in the motion and shape components. Thus, adopting a linear alternation scheme as in Section 2.3 would lead to a minimisation in which the cost functions are still non-linear at each step. For instance, in the estimation of the motion parameters we have a rotation matrix which multiplies the quadratic deformation parameters. Similarly, in the shape components, we have the squared and cross-product version of the 3D coordinates. For these reasons, we decided to directly adopt a non-linear optimization approach, initialized by the rigid-body motion estimated by the weighted factorization approach described in Section 2.3. We defined a non-rigid cost function which will then be optimized using a Levenberg-Marquardt iterative optimization scheme, generically called bundle-adjustment (BA) [33].

Our goal is to find the best representation of the data matrix W by using the quadratic model described in Section 3.1. When reconstructing the data matrix W based on a model, there will always be some residual error associated to it caused, for instance, by the noise in the data, computational errors or model inaccuracies. The 3-vector of the 3D coordinates estimated by our model for point j at frame i can thus be represented as:

$$\tilde{\mathbf{w}}_{ij}^{(rec)} = M_i \mathbf{s}_j = R_i [\Gamma_i \Omega_i \Lambda_i] \mathbf{s}_j; \quad (23)$$

while the residual error of reconstruction can be defined as:

$$\mathbf{e}_{ij} = \tilde{\mathbf{w}}_{ij} - \tilde{\mathbf{w}}_{ij}^{(rec)}, \quad (24)$$

where $\tilde{\mathbf{w}}_{ij}$ is the data acquired by the MOCAP system.

The best fit of the quadratic model to the MOCAP data is then found by minimizing the norm of the residual errors such that:

$$\arg \min_{\mathbf{R}_i, \Gamma_i, \Omega_i, \Lambda_i, \mathbf{S}_j} \sum_{i,j}^{F,P} \|\mathbf{e}_{ij}\|^2 = \arg \min_{\mathbf{R}_i, \Gamma_i, \Omega_i, \Lambda_i, \mathbf{S}_j} \sum_{i,j}^{F,P} \|\mathbf{w}_{ij} - \tilde{\mathbf{w}}_{ij}\|^2. \quad (25)$$

Note that this minimization scheme only guarantees a local minimum. Still, when a close enough initial point is given, the algorithm often converges to the global minimum. The initialization scheme used in our algorithm proved to return reasonable results on all the sequences tested.

4 EXPERIMENTS

4.1 Synthetic tests

The first step in the evaluation of this new model was to assess the 3D reconstruction errors using synthetic data. These tests used a synthetic cube with 156 feature points (see Figure ??(a)) over 500 frames. This object was then subjected to random deformations varying within the ranges defined in Section 3.1.4. Six different deformation strength ranges were chosen, where each range was defined by the maximum allowed variation of the deformation coefficients, in absolute value, when compared to rigid case. Each deformation level was tested using 50 different sequences. Performance was assessed based on the accuracy of the 3D reconstruction errors and results were compared against the Weighted Factorization approach used as initialization of this method. A plot of the average reconstruction error for each deformation level can be seen in Figure 6.

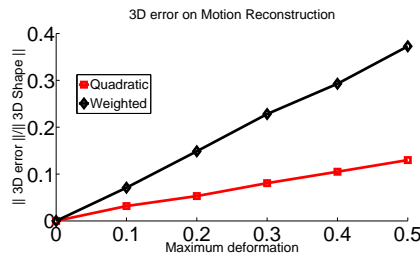


Figure 6: Results with synthetic data: 3D reconstruction errors averaged over 50 random tests for increasing levels of deformation strengths comparing the Quadratic Model and the Weighted Factorization methods.

The experiments show that the Quadratic model improves, as expected, the reconstruction provided by the Weighted Factorization approach. Even though the reconstruction error can go up to 13% when the highest level of allowed deformation is considered, we found that this case corresponds to very strong deformations which are usually not expected in soft-tissue artifacts. Realistic soft tissue deformations would usually correspond to strength levels of up to 0.3, bringing the reconstruction error to 8% in the worst case. By improving the reconstruction of the non-rigid motion of the soft-tissue artifacts, a more accurate description of the body segments is achieved, improving the quality of the models and consequently the results of the analyses performed with them.

4.2 MOCAP data

The next step on the model evaluation was to assess the reconstruction error on real human motions based on MOCAP data. For that purpose we used a real sequence of an upper arm, chosen due to the clear deformation of the biceps. In order to test the ability of the algorithm to handle noise we also added different levels of White Gaussian noise, from 0% to 3%. Due to the randomness of the noise generation, each noise level was tested in 50 different cases and the average computed. Results can be seen in Figure 7.

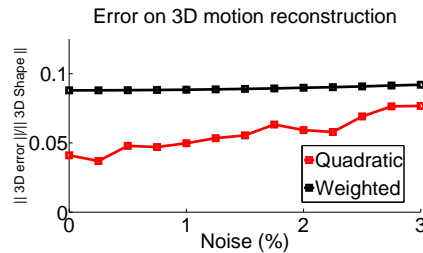


Figure 7: Plot of the average 3D reconstruction error for the MOCAP sequence of a flexing upper arm with different values of additive white Gaussian noise.

Although 3% of noise might seem a low value, a qualitative evaluation of the data revealed that the amount of noise was already relatively high, not being reasonable to expect any kind of meaningful result from tests with higher levels of noise. We also note that since we are using a real sequence, this data is already noisy due to the acquisition process of the MOCAP system. As we can see from Figure 7, when comparing the results in realistic noise levels, the performance of the Quadratic deformation model is significantly better than the Weighted Factorization approach. We also note that for the expected amount of noise in a real MOCAP test (0% added noise) the 3D reconstruction error for the Quadratic model is 4.11%, against the 8.80% of the Weighted Factorization approach.

To show the potential applications of this algorithm, we applied it on sequences of a forearm, an upper arm and a torso of a real human subject flexing his arm. The reconstructions for the three segments can be seen in Figure 8. The errors in the reconstruction of the forearm, upper arm and torso were 8.76%, 4.11% and 3.93% respectively.

The reconstructions obtained by the Quadratic model can also be combined with the methods described in [26] to form an articulated skeleton with non-rigid segments. Given the reconstruction obtained by our algorithm on this same sequence, the joint parameters computed with the aforementioned methods for joint parameter estimation can be seen in Figure 9.

As we can see from Figure , the shoulder joint is modeled as an universal joint with the joint centre represented with the large red circle while the elbow joint is modeled as a hinge joint and its axis is represented in green. The recovered joint centre and joint axis are both consistent with the motion of the non-rigid segments provided by our Quadratic deformation model.

5 CONCLUSION

In this paper we propose a new quadratic deformation model for human body motion analysis as an extension of the rigid Structure from Motion approach. This model expands the formulation to include linear, quadratic and cross-terms deformation terms in order to model object deformations. Tests performed on synthetic data show that the model is able to handle a considerable amount of deformation. 3D reconstruction errors on synthetic data for realistic

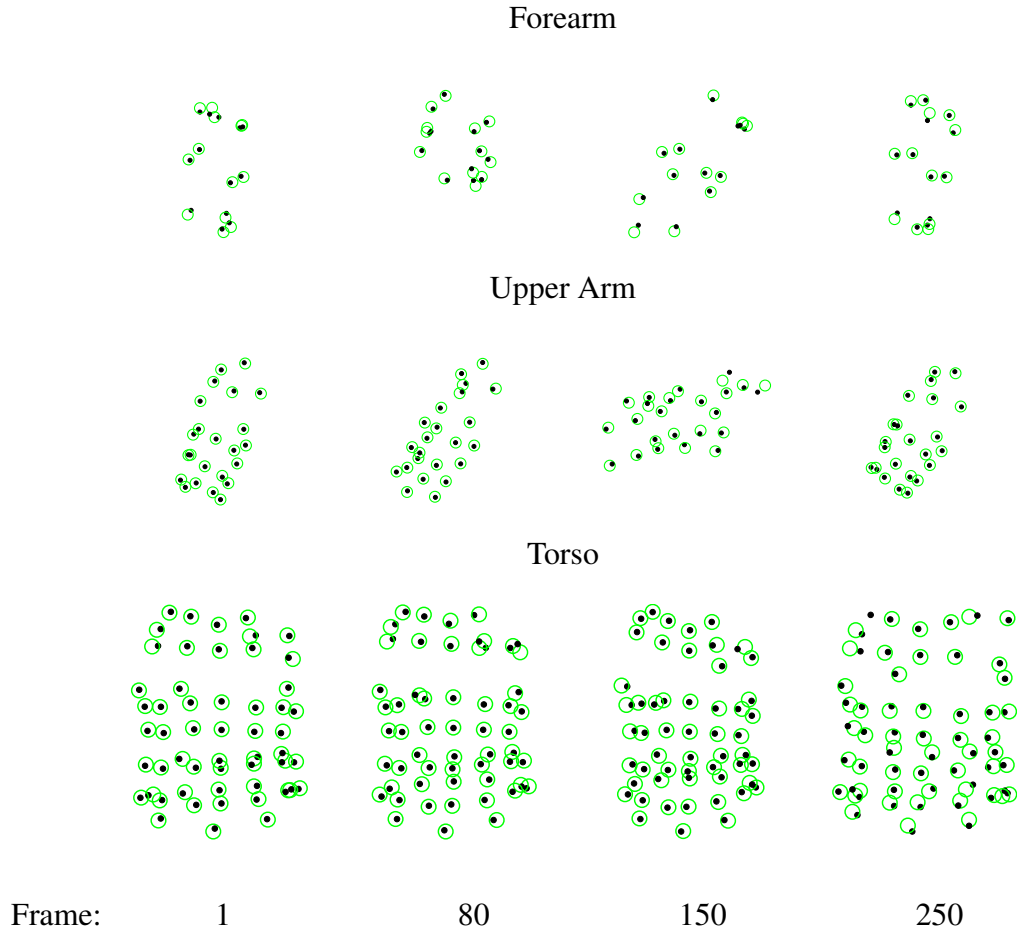


Figure 8: Four example frames of the MOCAP sequence of a subject flexing his arm. The top row shows the reconstruction of the forearm, the middle row the reconstruction of the upper arm, and the bottom row shows the reconstruction of the torso. Ground truth data is represented by the green circles, while the reconstruction is represented by the black dots.

levels of deformations were found to be between 5% and 8%. Tests performed on real MOCAP data of human body motions have shown that the new model is able correctly to model the deformations present in different segments of the human body. Additionally, it was also shown that the model can return reasonable reconstructions when dealing with noisy data.

Since the model is developed within the Structure from Motion factorization approach, it can be combined with other similar approaches to provide a full articulated skeleton with non-rigid motion on the segments. In this way human body models can be created which account for soft-tissue artifacts, an unwanted source of error for existing techniques.

Despite the promising results, this model still needs further development as it relies on empirical bounds for the deformations. A full validation with extensive testing on real sequences must also be carried out before the model can be used in clinical applications.

6 ACKNOWLEDGEMENTS

This work was partially funded by Fundacao para a Ciênciã e a Tecnologia (ISR/IST pluri-annual funding) through the POS_Conhecimento Program (include FEDER funds) and grant PTDC/EEA-ACR/72201/2006, and by the European Research Council under ERC Starting Grant agreement 204871-HUMANIS.

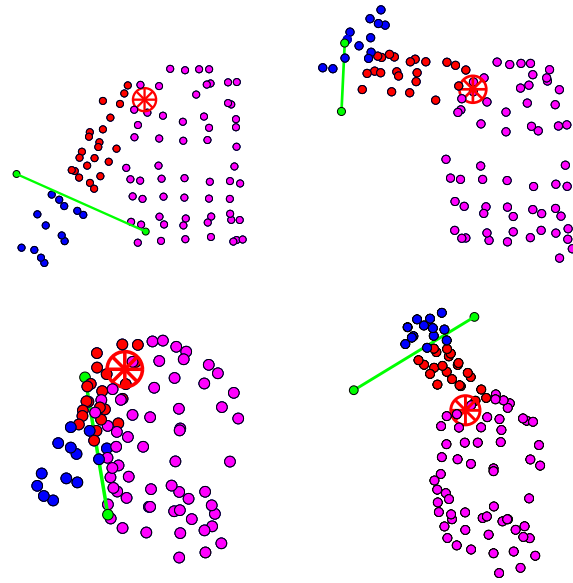


Figure 9: Multiple joint parameter estimation on a flexing arm. The elbow is modelled as a hinge joint and the shoulder as a universal joint. The rotation axis is presented in green while the joint centre of the universal joint is presented as a red circle.

We thank Dr. Sang Il Park and the Carnegie Mellon Graphics lab for the 3D MOCAP sequence used in Section 4.2.

REFERENCES

- [1] R. Baker, Gait analysis methods in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 3(1):4, 2006.
- [2] D. Knudson, *Fundamentals of Biomechanics*. Springer Science, second edition edition, 2007.
- [3] A. V. Ferreira, B. G. Rosa, J. Fayad and M. T. Silva, Análise de dinâmica directa do movimento de natação: Crawl. *Conferência Nacional de Dinâmica de Sistemas Multicorpo, Guimarães*, 269–274, 2007. In Portuguese.
- [4] T. B. Moeslund, A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, **81**, 231–268, 2001.
- [5] S. I. Park and J. K. Hodgins, Capturing and animating skin deformation in human motion. *SIGGRAPH*, 881–889, 2006.
- [6] B. Dariush. Human motion analysis for biomechanics and biomedicine. *Machine Vision Applications*, textbf14, 202–205, 2003.
- [7] J. R. Gage, P. A. Deluca and T. S. Renshaw. Gait analysis: Principle and applications with emphasis on its use in cerebral palsy. *The Journal of Bone and Joint Surgery*, 1607–1623, 1995.

- [8] C. F. Runge, F. E. Zajac III, J. H. J. Allum, D. W. Risher, A. E. Bryson Jr. and F. Honegger. Estimating net joint torques from kinesiological data using optimal linear system theory. *IEEE Transactions on Biomedical Engineering*, **42**, 1158–1164, 1995.
- [9] A. Cappozzo, A. Cappello, U. D. Croce and F. Pensalfini. Surface-marker cluster design criteria for 3-D bone movement reconstruction. *IEEE Transactions on Biomedical Engineering*, **44**, 1165–1174, 1997.
- [10] M. Sati. Quantitative assessment of skin-bone movement at the knee. *The Knee*, **3**, 121138(18), 1996.
- [11] A. Leardini, A. Cappozzo, F. Catani, S. Toksvig-Larsen, A. Petitto, V. Sforza, G. Casanelli and S. Giannini. Validation of a functional method for the estimation of hip joint centre location. *Journal of Biomechanics*, **32**, 99–103, 1999.
- [12] E. J. Alexander and T. P. Andriacchi. Correcting for deformation in skin-based marker systems. *Journal of Biomechanics*, **34**, 355–361, 2001.
- [13] A. Leardini, L. Chiari, U. D. Croce and A. Cappozzo. Human movement analysis using stereophotogrammetry Part 3: Soft-tissue artifact assessment and compensation. *Gait and Posture*, **21**, 212225, 2005.
- [14] T. P. Andriacchi, E. J. Alexander, M. K. Toney, C. O. Dyrby and J. Sum. A point cluster method for in vivo motion analysis: Applied to a study of the knee kinematics. *Journal of Biomechanical Engineering*, **120**, 743–749, 1998.
- [15] C. Tomasi and T. Kanade, Shape and Motion from Image Streams under Orthography: A Factorization Approach. *IJCV*, **9**, 137–154, 1992.
- [16] C. Poelman and T. Kanade, A paraperspective factorization method for shape and motion recovery. *LNCS*, **800**, 97–110, 1994.
- [17] J. P. Costeira and T. Kanade, A multibody factorization *IJCV*, **29**, 159 – 179, 1998.
- [18] C. Bregler, A. Hertzmann and H. Biermann, Recovering Non-Rigid 3D Shape from Image Streams. *CVPR*, 690–696, 2000.
- [19] M. Brand and R. Bhotika, Flexible Flow for 3D Nonrigid Tracking and Shape Recovery. *CVPR*, 315-322, 2001.
- [20] J. Xiao, J. Chai and T. Kanade, A Closed-Form Solution to Non-Rigid Shape and Motion Recovery. *ECCV*, 573–587, 2004.
- [21] A. Del Bue, F. Smeraldi and L. Agapito, Non-rigid structure from motion using ranklet-based tracking and non-linear optimization. *IVC*, **3**, 297–310, 2007.
- [22] C. J. Taylor. Reconstruction of articulated objects from point correspondences in a single uncalibrated image. *CVPR*, **1**, 677–684, 2000.
- [23] C. Sminchisescu and B. Triggs. Covariance scaled smapling for monocular 3d body tracking. *CVPR*, **1** 447–454, 2001.

- [24] P. Tresadern and I. Reid, Articulated Structure From Motion by Factorization. *CVPR*, **2**, 1110–1115, 2005.
- [25] J. Yan and M. Pollefeys, A factorization-based approach to articulated motion recovery. *CVPR*, **2**, 815–821, 2005.
- [26] J. Fayad, A. Del Bue, P. M. Q. Aguiar, Articulated motion analysis from motion capture data. *International Symposium on Computer Methods in Biomechanics and Biomedical Engineering*, 2008.
- [27] J. Fayad, P. M. Q. Aguiar and A. Del Bue, A Weighted Factorization Approach for Articulated Motion Modelling. *To appear in Multibody Dynamics 2009*.
- [28] P. J. Rebmann and F. T. Sheehan. Precise 3d skeletal kinematics using fast phase contrast magnetic resonance imaging. *Journal of Magnetic Resonance Imaging*, **17**, 206–213, 2003.
- [29] D. S. Asakawa, G. P. Pappas, S. S. Blemker, J. E. Drace and Delp S. L. Cine phase contrast magnetic resonance imaging as a tool for quantification of skeletal muscle motion. *Seminars on Musculoskeletal Radiology*, **7**, 287–295, 2003.
- [30] A. Del Bue, F. Smeraldi and L. Agapito. Non-rigid structure from motion using ranklet-based tracking and non-linear optimization. *Image and Vision Computing*, **25** 297–310, 2007.
- [31] M. Muller, B. Heidelberger, M. Teschner and M. Gross. Meshless deformations based on shape matching. *SIGGRAPH*, **24**, 471–478, 2005.
- [32] Y. C. Fung. *Biomechanics: Mechanical Properties of Living Tissue*. Springer-Verlag, 1993.
- [33] B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon. Bundle Adjustment A Modern Synthesis. *Proceedings of the International Workshop on Vision Algorithms*, 298–372, 1999.