Deterministic 3D Human Pose Estimation Using Rigid Structure

Jack Valmadre[1] and Simon Lucey[2]

[1] University of Queensland, Australia
[2] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
jack.valmadre@gmail.com, simon.lucey@csiro.au

Abstract. This paper explores a method, first proposed by Wei and Chai [1], for estimating 3D human pose from several frames of uncalibrated 2D point correspondences containing projected body joint locations. In their work Wei and Chai boldly claimed that, through the introduction of rigid constraints to the torso and hip, camera scales, bone lengths and absolute depths could be estimated from a finite number of frames (i.e. ≥ 5). In this paper we show this claim to be false, demonstrating in principle one can never estimate these parameters in a finite number of frames. Further, we demonstrate their approach is only valid for rigid sub-structures of the body (e.g. torso). Based on this analysis we propose a novel approach using deterministic structure from motion based on assumptions of rigidity in the body's torso. Our approach provides notably more accurate estimates and is substantially faster than Wei and Chai's approach, and unlike the original, can be solved as a deterministic least-squares problem.

Keywords: Human pose estimation, Structure from motion.

1 Introduction

The task of estimating 3D non-rigid structure from a small number of 2D point correspondences is, in general, an inherently ill posed problem. Recently, Wei and Chai [1] proposed an approach for solving the non-rigid structure from motion problem specifically for bodies. Their approach took advantage of assumed, and empirically validated, rigid constraints in the human body's torso and hip. Their approach is notable in comparison to previous literature [2,3,4,5] in the area as it: (i) makes no assumptions about bone lengths or camera scale, (ii) is not limited/constrained to modeling shapes previously seen in a train set, and (iii) can handle missing body points.

In [1] Wei and Chai claim their approach requires a minimum of five frames with 2D point correspondences (under a weak perspective assumption) for a 17 bone body containing 4 rigid triangles and 7 symmetry constraints, from which they can then estimate bone lengths, camera scale, and articulated pose. Their argument is made elegantly under the assumption that if one wants to reconstruct all unknowns without any ambiguity, one needs at least the same number of constraints. The authors claim that the introduction of rigid constraints across

a finite number of frames (i.e. ≥ 5) balances the number of unknowns with constraints. In this paper we demonstrate that this, rather strong assumption, is false. This dilemma forms the central thesis of our paper. The contributions of our paper are as follows:-

– We demonstrate that Wei and Chai's approach is only valid for rigid substructures of the human body (e.g. torso) rather than the entire body's non-rigid structure. We further demonstrate that the assumption that this approach was valid for non-rigid bodies led to the false claim that camera scales, bone lengths and absolute depths can be solved in minimum of 5 frames for a 17 bone body model. (Section 4)
– A fast and deterministic solution to the problem of estimating camera scale and bone lengths for the body's rigid torso is proposed. This solution differs substantially to the approach employed in [1] which attempted to solve a quartic objective function. Unlike our fast deterministic approach based on canonical rigid structure from motion [6], the method of Wei and Chai requires a slow non-linear optimizer with a number of heuristics needing to be chosen to estimate a satisfactory answer. Our approach also exhibits superior empirical performance to [1]. Further, our approach solves for the full 3D structure of the torso up to a reflection, whereas [1] has inherent ambiguity with respect to the sign of each bone depth. (Section 5)
– A remaining problem now exists, however, with respect to estimating the bone lengths and joint angles of the residual non-rigid structures of the body (e.g, arms, legs, etc.). For this task, we propose an elegant approach based on: (i) making an assumption that all non-rigid bones (taking into account symmetry) will be parallel to the image plane in at least one frame, and (ii) using prior statistics on allowable human joint angles. (Section 6)

The most important contribution, however, of our work is to highlight where assumptions of rigidity help, during 3D human pose estimation, and where they do not. First, our work demonstrates the absolute importance of the accurate labeling of the 4 rigid points on the body's torso. Without these accurate labels it is impossible to estimate camera scale and therefore the lengths of any of the other non-rigid bones. Second, we demonstrate that, with the exception of camera scale, assumptions of rigidity in the body's torso and hip cannot aid in the estimation of bone lengths, or joint angles, in the remaining non-rigid substructures. Further assumptions are required, such as those discussed in Section 6, in order to estimate these additional lengths and angles.

1.1   Background

Factorization approaches, first proposed for recovering rigid structure by Tomasi and Kanade in [6], were extended to handle non-rigidity in the seminal paper by Bregler et al. [5]. In this notable work the authors realized that the only way to make the non-rigid structure from motion problem tractable is to impose some constraint on the object being analyzed (e.g. assuming we are looking at a body). Bregler et al. imposed this constraint as a linear shape basis. A number of

approaches that develop the use of shape basis have subsequently been proposed, including [7,8,9]. A fundamental criticism, however, of all these approaches is the specificity, or more critically the poor generalization properties, of the shape basis. For example, the shape basis of a "person walking" will not be the same as that of a "person bending down to pick something up". Akhter et al. [10] recently proposed an approach to non-rigid structure from motion that is shape agnostic, and instead places a constraint on how individual points are allowed to move through space over time. The authors frame this work in a manner similar to Bregler et al. through the introduction of a trajectory basis. A criticism of this approach, however, is that it requires each frame of a video sequence to have 2D point correspondences and cannot handle 2D points of an object sampled at random points in time.

2   The Problem

In this section we will quickly review the problem of establishing 3D human poses from uncalibrated 2D point correspondences. Given a series of F 2D frames of a person in different poses, we seek to establish the 3D weak perspective structure namely, camera scale, bone lengths, and 3D pose. Under weak perspective projection, each camera has a single unknown scale parameter s_f. The length of each bone l_b is constant across all frames. The relative depth across each bone z_{fb} defines the 3D pose.

Let $\mathbf{s}$ and $\boldsymbol{\sigma}$ denote vectors of squared and inverse-squared camera parameters respectively.

$$\mathbf{s} = \begin{bmatrix} s_1{}^2 \\ \vdots \\ s_F{}^2 \end{bmatrix} \quad \boldsymbol{\sigma} = \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_F \end{bmatrix} = \begin{bmatrix} s_1{}^{-2} \\ \vdots \\ s_F{}^{-2} \end{bmatrix}$$
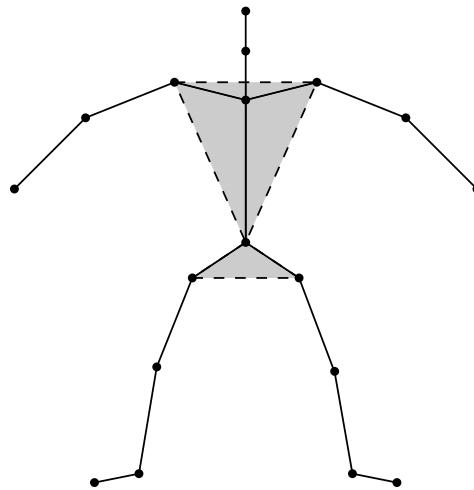
Fig. 1. Human skeleton model with $B = 17$ bones as used by [1]. Wei and Chai propose the introduction of $R = 4$ hidden bones to enforce the rigidity of points on the torso and hip (rigid sub-structures of the body are shaded).
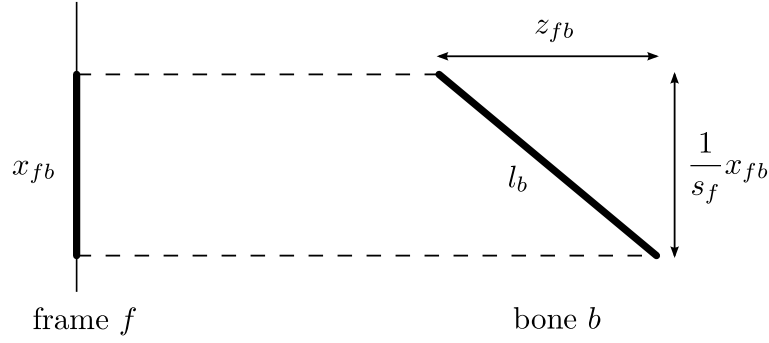
Fig. 2. Weak perspective (or scaled orthographic) projection is a good approximation for perspective projection assuming the length of each bone is small compared to the distance from the camera. In our work we assume each frame f has a single unknown scale parameter s_f, and the length l_b of each bone b is constant across all frames. x_{fb} and z_{fb} are the known projected lengths, and relative depths respectively for each frame and bone.

Let $\mathbf{l}$ be a vector of squared bone lengths. Let $\mathbf{z}_f$ and $\mathbf{x}_f$ be vectors of relative depths and known projected lengths of all bones in frame f respectively; squared.

$$\mathbf{l} = \begin{bmatrix} l_1{}^2 \\ \vdots \\ l_B{}^2 \end{bmatrix} \quad \mathbf{z}_f = \begin{bmatrix} z_{f1}{}^2 \\ \vdots \\ z_{fB}{}^2 \end{bmatrix} \quad \mathbf{x}_f = \begin{bmatrix} x_{f1}{}^2 \\ \vdots \\ x_{fB}{}^2 \end{bmatrix}$$

The problem, generically, can be expressed as minimizing the following objective function,

$$E_p(\mathbf{l}, \mathbf{z}, \boldsymbol{\sigma}) = \sum_{f=1}^{F} \|\mathbf{l} - \sigma_f \mathbf{x}_f - \mathbf{z}_f\|_2^2 \tag{1}$$

The depths for frame f are only estimated up to their absolute value $\{|z_{fi}|\}_{i=1}^{B}$. Estimating the unknown signs is discussed in Section 6, but shall be ignored for the moment. The system as described in Equation 1 has B unknown bone lengths, BF unknown bone depths and F unknown camera scales; with only BF Pythagorean constraints and 1 arbitrary scale constraint (i.e. we can assume the scale on the first frame to be unity $\sigma_1 = 1$). The system will be rank-deficient unless the number of constraints is greater than or equal to the number of unknowns. The difference between constraints and unknowns is $1 - F - B$. The condition on F therefore is,

$$F \leq 1 - B \tag{2}$$

Clearly there is no positive solution for F which satisfies the inequality. Some additional constraints/prior is required to solve the problem in Equation 1.

3   Wei and Chai's Method

In recent work Wei and Chai [1] proposed a novel solution to this under-constrained problem through the introduction of rigid constraints on the torso (3 bones) and

the hip (2 bones). The authors demonstrated empirically [1] that the hip and torso can be considered separate rigid structures. In their approach they enforced the rigid constraint through a hidden bone, which closes a pair of bones to form a rigid triangle, resulting in the following rigid constraint function which is quartic (in terms of the squared variables).

$$E_r(\mathbf{e}, \mathbf{z}, \boldsymbol{\sigma}) = \sum_{f=1}^{F} \sum_{i=1}^{R} \left[\left(e_i^2 - {x_{fi}^e}^2 \sigma_f - z_{fj(i)}^2 - z_{fk(i)}^2 \right)^2 - 4z_{fj(i)}^2 z_{fk(i)}^2 \right]^2 \quad (3)$$

where e_i denotes the length of the i-th hidden bone, x_{fi}^e denotes the corresponding projection length, $j(i)$ and $k(i)$ are indexes to the visible bones that form the rigid triangle with the i-th hidden bone. For R hidden bones, they introduce R extra unknown bone lengths and therefore RF new constraints. Wei and Chai also take advantage of the natural symmetry occurring in human bodies to improve the rank of the system. All visible bones (i, j) in a left-right pair are assumed to be of equal length,

$$l_i^2 - l_j^2 = 0 \quad (4)$$

with this property being encoded into $\mathbf{A}$ resulting in the following symmetric constraint function,

$$E_s(\mathbf{l}) = \|\mathbf{A}\mathbf{l}\|_2^2 \quad (5)$$

Wei and Chai proposed that the objective functions in Equations 1, 3 and 5 can be linearly combined such that the new objective function is,

$$\arg \min_{\mathbf{l}, \mathbf{z}, \boldsymbol{\sigma}, \mathbf{e}} = E_p(\mathbf{l}, \mathbf{z}, \boldsymbol{\sigma}) + \lambda_1 E_s(\mathbf{l}) + \lambda_2 E_r(\mathbf{e}, \mathbf{z}, \boldsymbol{\sigma}) \quad (6)$$

where λ_1 and λ_2 are parameters that are used to control the influence of the two constraint functions $E_s()$ and $E_r()$ respectively. These parameters are tuned through a cross-validation process.

3.1   Constraints versus Variables

At first glance the combination of objective and constraint functions in Equation 6 seems substantially more constrained than the original objective function described in Equation 1. The overall system has $(B + 1)F + (B + R)$ unknowns and $(B + R)F + 1$ constraints. Including symmetry adds another M constraints, effectively reducing the number of unknown bone lengths so that the condition on F (assuming $R \geq 2$) is,

$$F \geq 1 + \frac{B - M}{R - 1} \quad (7)$$

It seems for a skeleton with $B = 17$ bones, $M = 7$ symmetry constraints and $R = 4$ rigid bone pairs, it is always possible to find bone lengths, bone depths and camera parameters (up to a constant factor) given point correspondences from 5 distinct frames under weak-perspective projection. In the following section, however, we dissect this claim further showing it to be false.

4   A Toy Problem

Nothing limits the application of Wei and Chai's method to human bodies. More generally, their approach should hold for any connected structure, satisfying Equation 7, containing a subset of rigid points. For convenience we shall continue to refer to the edges connecting points in such structures as bones, even though we are no longer referring to the human body. Figure 3 shows a single free bone attached to a rigid tetrahedron. We define a bone as being *free* if one of its end points is not connected to a rigid structure, implying that the end-point can move non-rigidly with respect to the structure (e.g. bones on the arm or leg of the human body). The shape in Figure 3 according to Wei and Chai's approach can be described as $B = 4$, $R = 3$, and $M = 0$. Applying Equation 7 in this circumstance implies that the lengths and camera scales for the structure can be solved for $F \geq 3$. The absurdity of this claim can be understood if we attempt to solve this problem in an alternative manner using canonical structure from motion [11]. It is widely accepted that the weak-perspective parameters, camera orientations and 3D structure of the tetrahedron can be recovered, given a minimum of 4 non-coplanar points, from $F \geq 3$ frames using the structure from motion theorem [11]. Assuming these variables are known, however, the problem of finding the length of the single free bone remains, in general, fundamentally unsolvable. It is impossible to know the length of a free bone from multiple projections of unknown orientation. It is easy to show through a simple thought experiment, see Figure 4, that no minimum number of images necessarily guarantees that the length of a free bone can be determined. This clearly violates Wei and Chai's condition on F in Equation 7. This violation occurs due to the misuse of the hidden bones, introduced by Wei and Chai, to enforce rigidity. These rigid constraints only apply to those bones comprising the rigid structure. To understand this further, one can partition the objective function in Equation 6 into separate components describing rigid structures (e.g. the torso and hip) and the residual non-rigid structures. The only variable that is common across these separate partitions is camera scale. Were we to consider the problem under orthography, each free bone would emerge an independent, under-ranked
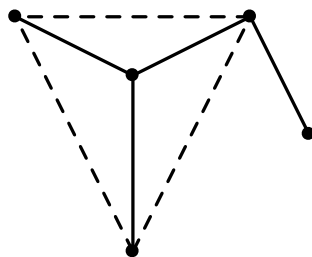
Fig. 3. This figure depicts a tetrahedron with a single free bone (i.e. the end point can move non-rigidly with respect to the rigid points defining the tetrahedron) highlighting the underlying problem in the method of Wei and Chai [1] with regard to non-rigid points. The hidden bones, denoted by a dotted line, enforce rigid structure in the tetrahedron.
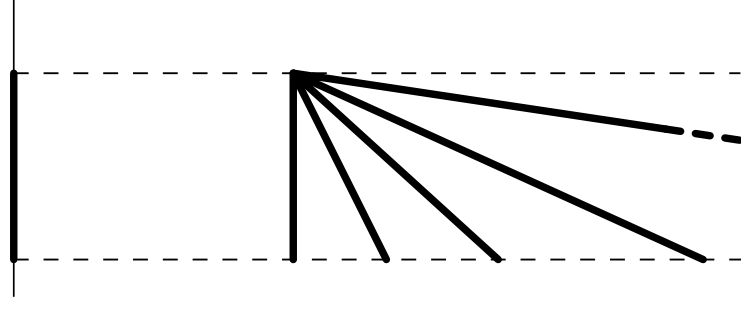
Fig. 4. This figure depicts the inherent ambiguity of a *free* bone projected onto the image plane (assuming parallel projection)

system (or each left-right pair remembering symmetry). It becomes clear that the claim of Wei and Chai is just a mirage, creating the illusion of a solvable system by combining under- (i.e. non-rigid) and over-ranked (i.e. rigid) systems.

5   Estimating Camera Scale from Rigid Sub-structure

From Section 4 it is clear that Wei and Chai's approach cannot be applied to rigid structures with additional free bones. Their approach, however, still holds for the estimation of camera scale, bone lengths and absolute depths for rigid structures. This, unfortunately, involves finding the solution to a quartic objective function, as described in Equation 6, which is susceptible to local minima and requires a slow non-linear, non-deterministic solver. As discussed in Section 4 an alternative method for estimating the camera scale, bone lengths and actual depths of the rigid structure is through the use of canonical structure from motion [11,6]. Canonical structure from motion has a number of advantages over the objective function described in Equation 6 specifically: (i) structure from motion has a deterministic, linear least-squares solution [12], and (ii) the direction of the depth across each bone can be found, not just the magnitude.

Structure from motion, under a weak perspective assumption, involves the resolution of a matrix of 2D projection coordinates $\mathbf{W}$ ($2F \times N$) into rotation $\mathbf{R}$ ($2F \times 3$) and structure $\mathbf{S}$ ($3 \times N$) matrices where $N+1$ is the number of points in the rigid structure [6]. We assemble $\mathbf{W}$ from relative vectors between projected points instead of expressing points relative to their centroid, explaining why $\mathbf{W}$ is $2F \times N$ instead of $2F \times (N+1)$ where,

$$\mathbf{W} = \mathbf{RS} = \begin{bmatrix} s_1\mathbf{i}_1^T \\ \vdots \\ s_F\mathbf{i}_F^T \\ s_1\mathbf{j}_1^T \\ \vdots \\ s_F\mathbf{j}_F^T \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \cdots \mathbf{p}_N \end{bmatrix} \tag{8}$$

Typically, singular value decomposition (SVD) is used to find a rank-3 approximation $\hat{\mathbf{W}} = \hat{\mathbf{R}}\hat{\mathbf{S}}$. The rotation and structure matrices are then determined up to a rotation and constant factor by enforcing metric constraints on $\mathbf{R}$.

$$\mathbf{W} \approx \hat{\mathbf{W}} = \hat{\mathbf{R}}\hat{\mathbf{S}} = (\hat{\mathbf{R}}\mathbf{G})(\mathbf{G}^{-1}\hat{\mathbf{S}}) = \mathbf{R}\mathbf{S} \qquad (9)$$

For the special case of $N = 3$ (such as with the human torso), the SVD is not required. We can simply substitute $\hat{\mathbf{R}} = \mathbf{W}$, $\hat{\mathbf{S}} = \mathbf{I}$. In our approach the typical weak perspective metric constraints [13] are modified to obtain a single estimate of scale for each frame. We then adopt the method presented by Morita and Kanade [12] for posing the metric constraints as a linear least squares optimisation to solve for $\mathbf{G}$, subject to $\mathbf{s} > 0$.

5.1   Experiments

Our experiments were carried out on the CMU Motion Capture Database.[1] Experiments were performed using 5 randomly selected frames from each of 128 motion sequences. Each sequence featured one of 20 different actors. For each sequence, the camera parameters and rigid bone lengths on the torso were calculated using our structure from motion approach and variants of Wei and Chai's non-linear optimisation. Histograms of normalized scale and length results are shown in Figures 5 and 6 respectively. Scale and length estimates have all been normalized by their ground-truth values. The more accurate the results, the more they tend towards unity.
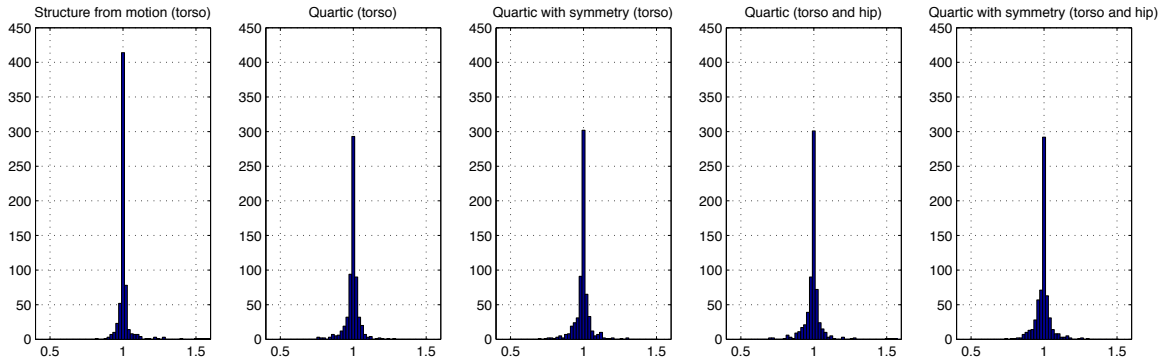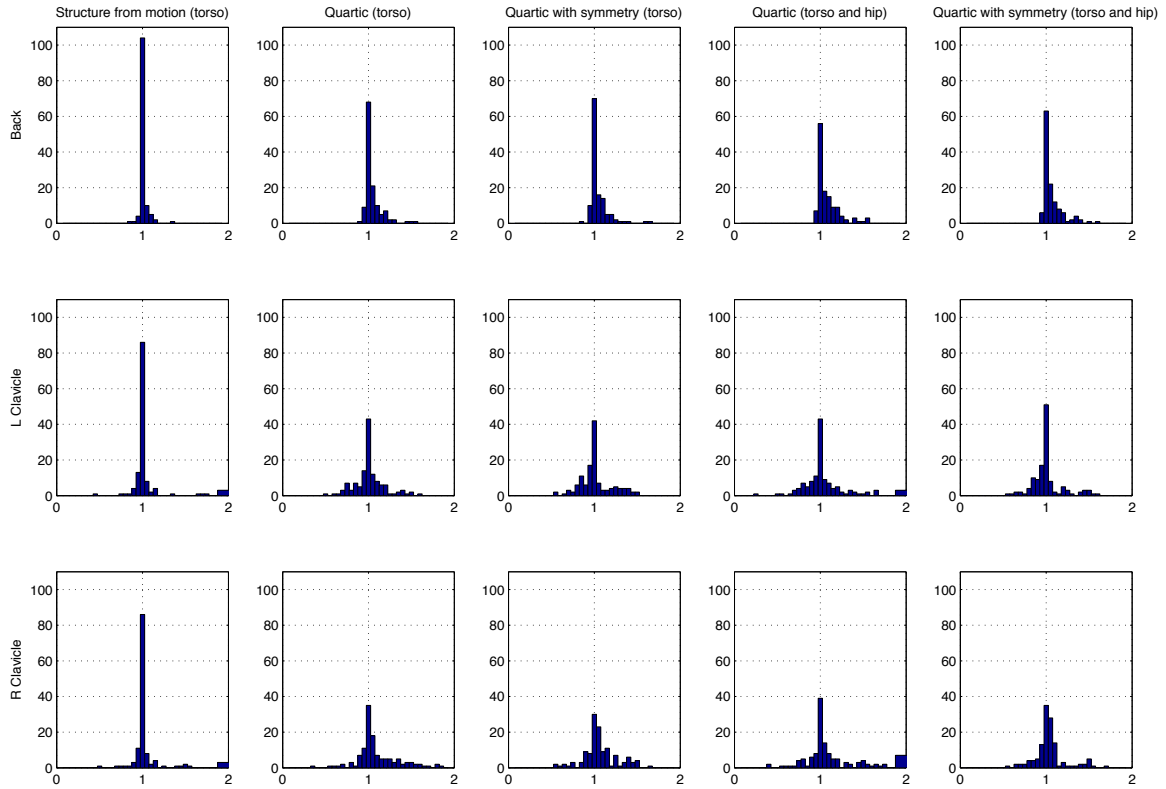

Fig. 5. Histograms of camera scale estimates using our deterministic structure from motion approach compared to Wei and Chai's computationally expensive non-linear optimisation, with and without symmetry and hip bones. All estimates have been normalized by their ground-truth values so that the more accurate the approach the more the scale estimate will tend towards unity.

Experiments were restricted to estimating camera scale, as well as lengths, from just the body's torso not the hip. This is due to the structure from motion theorem [11] which states that you need a minimum of 4 non-coplanar rigid points. Inspecting Figure 1 one can see the torso satisfies this criteria (containing 4 non-coplanar points[2]), but the hip does not (containing only 3 points).

---

[1] http://mocap.cs.cmu.edu/
[2] Care was taken to consistently mark the torso joint at the front of the subject's chest to ensure non-coplanar points.

Fig. 6. Histograms of the torso bone lengths; namely the back, left clavicle and right clavicle bones. Length estimates depicted include our deterministic structure from motion approach compared to Wei and Chai's computationally expensive non-linear optimisation, with and without symmetry and hip bones. All estimates have been normalized by their ground-truth values so that the more accurate the approach the more the length estimate will tend towards unity.

Similarly, for Wei and Chai their approach cannot be applied solely to the hip due to there only being $R = 1$ hidden bones enforcing rigidity. Inspecting Equation 7 one can see that there is no F that satisfies the condition for $R = 1$.

Results in Figures 5 and 6 demonstrate that our structure from motion approach consistently outperforms Wei and Chai's for both scale and length estimation (in terms of the estimates tending towards unity). Results did not improve even when the symmetry constraint function (enforcing the left and right clavicle in the torso to be the same length), described in Equation 5, was introduced. Additional results are depicted in Figure 5 to empirically reinforce our claim that the employment of the hip, in conjunction with the torso, does not aid in the estimation of camera scale. All our experiments were conducted in MATLAB. The approach of Wei and Chai, which required the minimization of the quartic objective function described in Equation 6 (using lsqnonlin in MATLAB), empirically took over 300 times as long to compute as our deterministic structure from motion approach.

J. Valmadre and S. Lucey

6 The Remaining Problem

It is clear from Section 5 that we are now able to reliably estimate the rigid structure of the torso along with camera scale using our structure from motion approach. A problem still remains, however, in how to estimate the lengths and joint angles of the remaining free bones.

6.1 Estimating Free Bone Lengths

Given F frames of a free bone in random orientations, as F tends toward infinity the longest observed projection provides an increasingly accurate estimation of bone length,
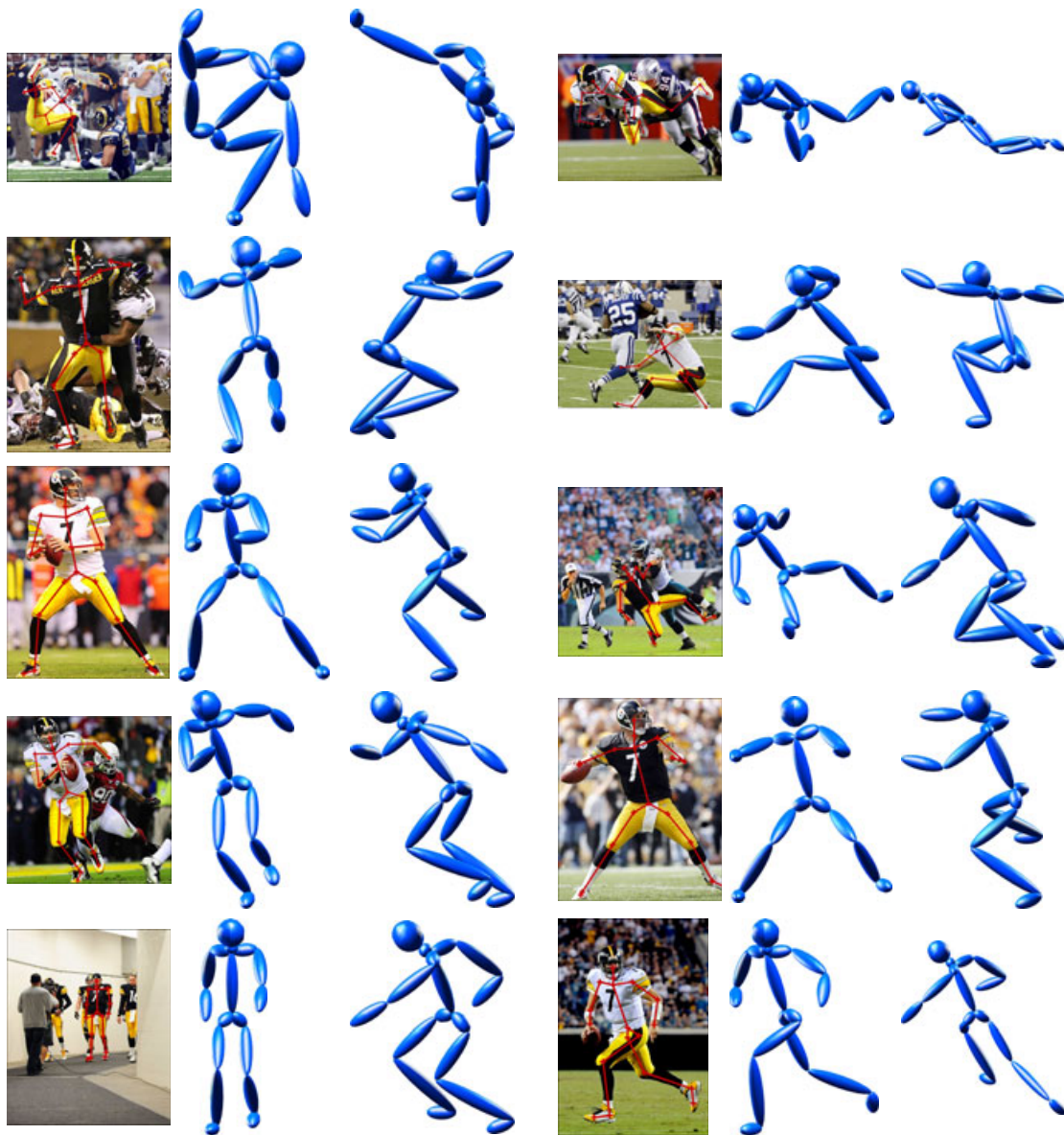
Fig. 7. 3D pose reconstructed from ten manually-labelled images of Steelers quarterback Ben Roethlisberger. Shown from the original front-view and an elevated side-view.

$$l_b^2 \approx \max_{f=1}^{F} \left\{ \frac{1}{s_f^2} x_{fb}^2 \right\} \tag{10}$$

where an estimate of s_f is known from our structure from motion approach. That is, the likelihood of the bone lying parallel to the camera plane in at least one image approaches certainty as F increases. In our approach, instead of requiring a large number of images to find bone lengths, we consider the "human in the loop" responsible for choosing images such that each bone is near-parallel in at least one image. We assume that a finite F well-selected frames will tend towards the true bone length. Symmetry constraints reduce the number of degrees of freedom, but even with symmetry the free bone problem remains under-ranked (by at least one unknown bone length). The key advantage of considering symmetry in free bones is that only one bone in each symmetrical pair need lie parallel to the image plane.

We note from Equation 10 that the accuracy of the solution for free bones depends critically on the accuracy of all scales. If the scale of one image is significantly under-estimated, the length of a free bone could potentially be over-estimated, which affects the 3D pose for all frames. This effect is most obvious for poses where an elbow or knee joint is close to fully extended, but has to be bent in the reconstruction to fit the projection (see the bottom-left image in Figure 7 and the middle image in Figure 9).

Fig. 8. 3D pose of cricket all-rounder Michael Clarke. Generated from five human-labelled images.

6.2   Estimating Joint Angles

Once we have an estimate of bone length we then also have the absolute bone depth,

$$z_{fb}^2 = l_b^2 - \frac{1}{s_f^2} x_{fb} \tag{11}$$

although we have ambiguities over the sign of z_{fb}. In our approach, we borrow upon Wei and Chai's approach for reducing this ambiguity by enforcing joint angle limit constraints from the biomechanics community. The joint angle limit constraints significantly reduce the ambiguity, but for some poses they are not sufficient to remove all ambiguity. When this happens, we allow the user to specify the sign of z_{fb} for bones that still have ambiguity until we have a solution.

6.3   Experiments on Real Images

We evaluated the effectiveness of our proposed approach using freely available images of four notable athletes, Ben Roethlisberger, Michael Clarke, Yu-Na Kim and Roger Federer, in Figures 7, 8, 9 and 10 respectively.
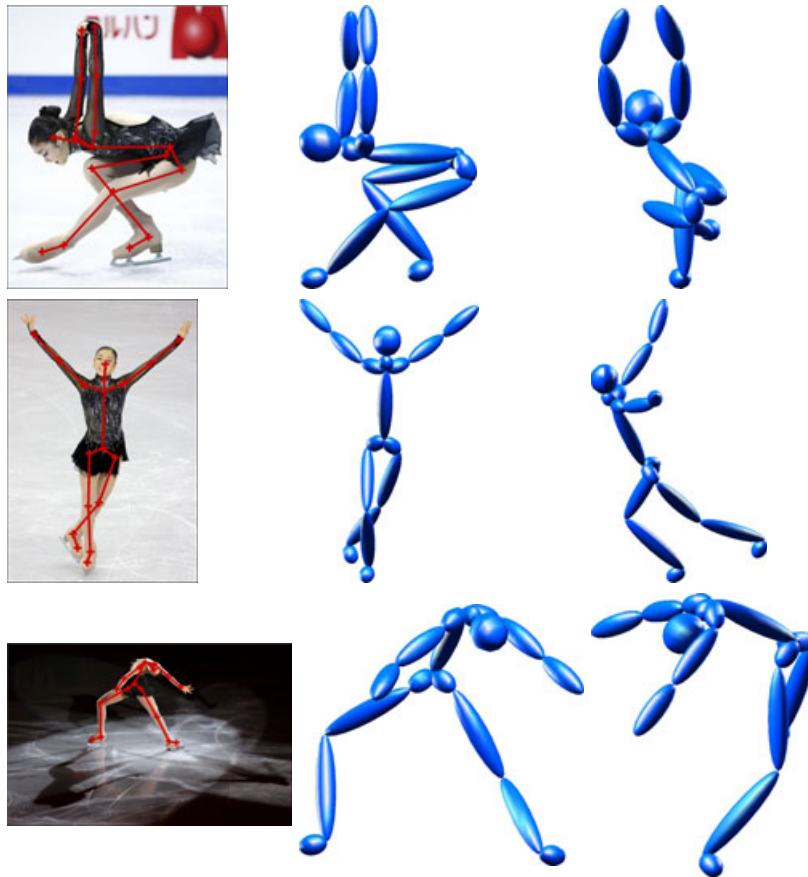
Fig. 9. Korean figure skater Yu-Na Kim, 2010 gold medallist. The 3D poses were reconstructed using only the three images shown. Wei and Chai [1] previously proposed that a minimum of five images were required.
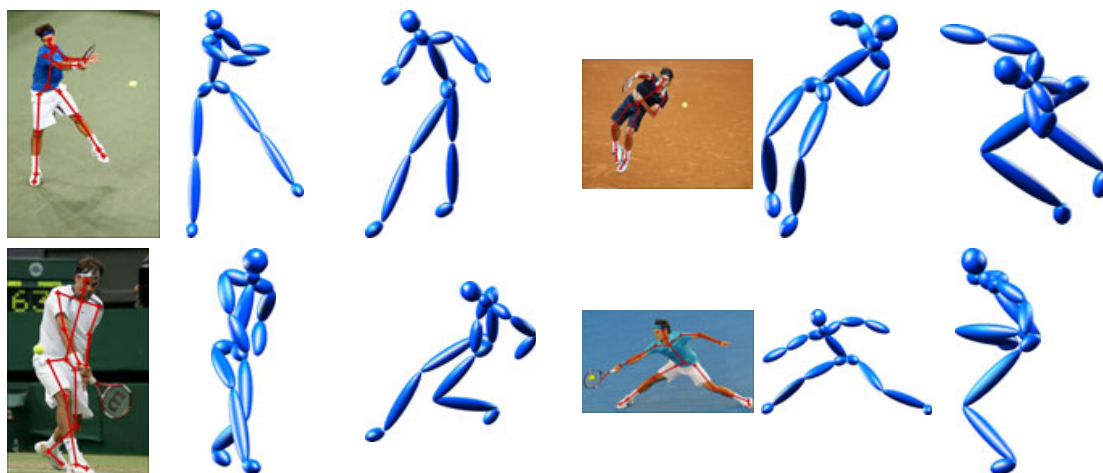
Fig. 10. The 3D pose of Swiss tennis player Roger Federer estimated from four images

7   Discussion and Conclusions

In this paper we present a novel approach for estimating 3D human pose from several frames of uncalibrated $2D$ point correspondences using a deterministic least-squares approach. Our approach, takes advanatge of rigid constraints in the human body first proposed by Wei and Chai [1]. We have additionally demonstrated that Wei and Chai's approach is only valid for rigid sub-structures of the human body (e.g. torso) rather than the entire body's non-rigid structure. We further demonstrate that the assumption that this approach was valid for non-rigid bodies led to the false claim that camera scales, bone lengths and absolute depths can be solved in minimum of 5 frames for a 17 bone body model.

References

1. Wei, X.K., Chai, J.: Modeling 3D human poses from uncalibrated monocular images. In: IEEE International Conference on Computer Vision (2009)
2. Agarwal, A., Triggs, B.: 3D Human pose from silhouettes by relevance vector regression. In: IEEE Conference on Computer Vision and Pattern Recognition (2004)
3. Taylor, C.J.: Reconstruction of articulated objects from point correspondences in a single uncalibrated image. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 677–684 (2000)
4. Barron, C., Kakadiaris, I.A.: Estimating anthropometry and pose from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition (2000)
5. Bregler, C., Hertzmann, A., Biermann, H.: Recovering non-rigid 3D shape from image streams. In: IEEE Conference on Computer Vision and Pattern Recognition (2000)
6. Tomasi, C., Kanade, T.: Shape and motion from image streams under orthography: a factorization method. International Journal of Computer Vision 9, 137–154 (1992)
7. Torresani, L., Hertzmann, A., Bregler, C.: Learning non-rigid 3D shape from 2D motion. In: NIPS (2005)

8. Xiao, J., Chai, J., Kanade, T.: A closed form solution to non-rigid shape and motion recovery. International Journal of Computer Vision 67, 233–246 (2006)
9. Torresani, L., Hertzmann, A., Bregler, C.: Non-rigid structure from motion: Estimating shape and motion with hierarchical priors. IEEE Transactions on Pattern Analysis and Machine Intelligence 30, 878–892 (2008)
10. Akhter, I., Sheikh, Y., Khan, S., Kanade, T.: Nonrigid structure from motion in trajectory space. In: NIPS (2008)
11. Ullman, S.: The interpretation of visual motion (1979)
12. Morita, T., Kanade, T.: A sequential factorization method for recovering shape and motion from image streams. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 858–867 (1996)
13. Hajder, L., Chetverikov, D., Vajk, I.: Robust structure from motion under weak perspective. 3D Data Processing, Visualization and Transmission (2004)