Silhouette-based human motion estimation

Uwe G. Kersting[1], Bodo Rosenhahn[2], Jason K. Gurney[1], Andrew W. Smith[1], Reinhard Klette[2]
[1]Department of Sport and Exercise Science, The University of Auckland, New Zealand
[2]Department of Computer Science and Robotics, The University of Auckland, New Zealand

INTRODUCTION

Over the recent decades various marker-based video capture systems have been developed in order to estimate segment kinematics in biomechanics applications. It may not be a disproportionate statement to claim that this method is regarded as the gold standard in laboratory based settings. Inherent problems to these systems include adequate marker placement, skin movement artefacts and marker occlusion in certain body orientations. A possible alternative to this approach are marker-less image-based motion tracking systems. Previously, simplified models were used (Fua et al., 2001) providing reasonable estimates of body movements.

In the present approach several modules such as free-form surface patches to estimate segment orientations (Rosenhahn et al., 2004) as well as global and local morphing techniques were applied (Rosenhahn & Klette, 2004). Further features include an advanced image segmentation method, dynamic occlusion handling and the inclusion of kinematic chains of higher complexity (21 degrees of freedom).

AIMS AND OBJECTIVES

The aim of this study was to compare resulting kinematics determined by the proposed marker-less motion estimation system to a commercially available marker-based tracking system. In this study upper body movements were compared from simultaneously recorded image data.

METHOD

One male subject was tested for this study. A three dimensional surface mesh of the upper body was created beforehand. In the model segments were connected by 2 and 3 degrees of freedom joints.

For video capture a digital four camera system was used (Basler A602f using SIMI motion software for video recording, 30 Hz). Cameras were set up in a rectangle arrangement. To attain optimal contrast a white sheet was spread out in the test area with the subject wearing a tight black t-shirt and black gloves. On top of the t-shirt reflective markers were placed on anatomical landmarks according to an existing upper body model (Ferdinands & Kersting, 2005).

An eight camera high resolution video system was used for the marker-based motion capture (Motion Analysis Crop. with Falcon cameras, 30 Hz). A sketch of the setup is given in Figure 1.

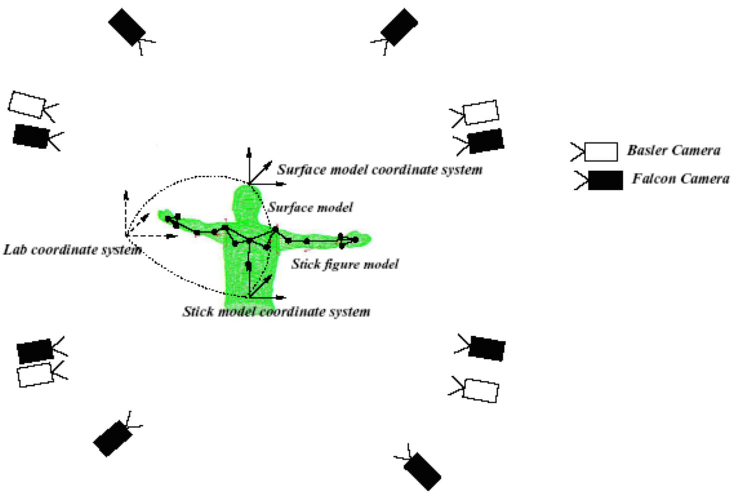

Figure 1: Schematic setup of the two camera systems.

The movement estimation sequence consists of three steps: segmentation, correspondence estimation and pose estimation. Image segmentation was realized using a level set function approach (Cremers et al., 2003). Using the silhouette from each preceding frame as a tracking assumption the algorithm needed just a few iterations per frame which makes it sufficiently fast (200 ms processing time per frame in a 4-camera setup). Pose estimation is carried out by assuming a set of correspondences between 4D model points and 3D image points. Image points are reconstructed to Plücker lines. The representation of a rigid 3D motion in its exponential form allows to express the motion as a screw motion around a given axis in space. The combination of the reconstructed Plücker lines with the screw representation for rigid motion using a gradient descent method allows to derive a system of linear equations to be solved. With reconstructing 3D lines from different cameras calibrated to the same lab co-ordinate system the system of equations can be solved for each joint position as long as this joint is visible in at least one of the camera views.

Finally, correspondences between the object model and extracted silhouettes are determined using a modified ICP (Rosenhahn et al., 2004) algorithm and a voting method for deciding which point belongs to which segment. This process is repeated on slightly transformed model configurations and iterated until the overall pose converges.

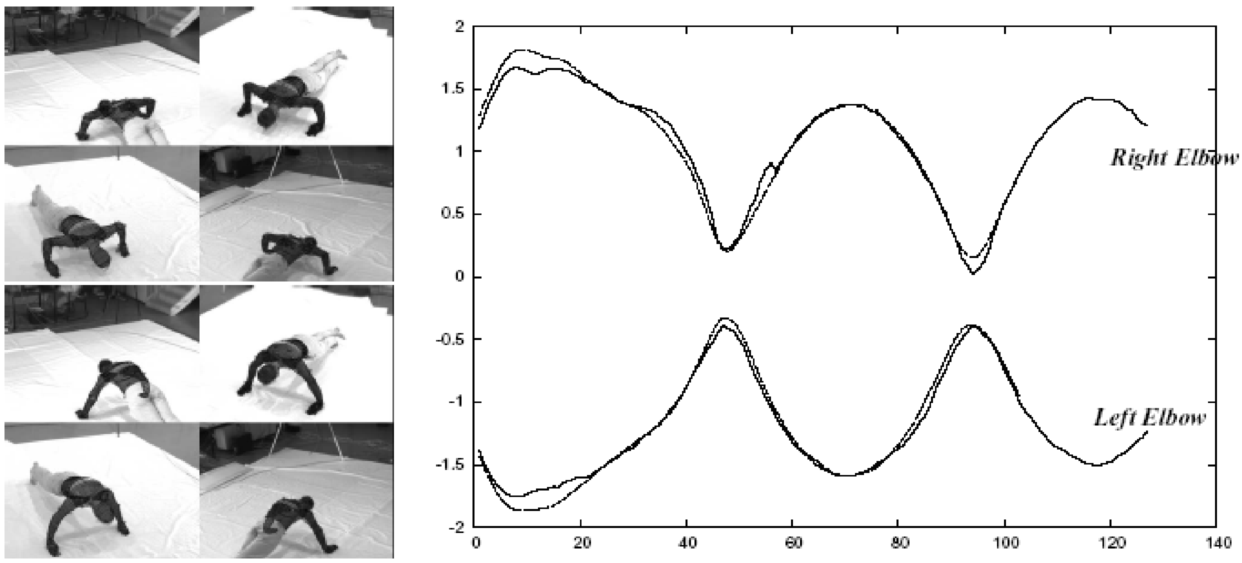

Figure 2: Example from a push-up sequence. Dotted lines: Marker-based system; solid lines: Silhouette system.

RESULTS AND DISCUSSION

Figure 5 shows the results for a push-up sequence including two sets of four images taken from the video recordings. The graph on the right gives elbow flexion-extension angles for the left and right arm. The overall error comparing the angle values is 1.7 degrees. Other complex arm movements resulted in overall errors of up to 2.3 degrees. All sequences contained partial occlusions which are obviously handled well by the algorithm (Figure 2, left).

Richards (1999) has shown that RMS errors of marker-based systems typically lie below 3 degrees. It can therefore be argued that the proposed system is comparable to marker-based approach. However, if a silhouette based movement estimation system better represents skeletal motion cannot be answered based on the current results.

CONCLUSIONS

A system to perform marker-less motion estimation has been proposed and compared to a marker-based tracking system. Deviations between approaches were of similar magnitude as comparisons between different marker systems. Currently we are conducting a similar comparison on human gait. Further research is required to test its applicability in clinical and sports biomechanics environments.

REFERENCES

Fua, P., Plänkers, R. Thalmann, D. (2001): *Computer Vision and Understanding* (81) 3: 285-302

Cremers, D., Kohlberger, T., Schnörr, C. (2003): *Pattern Recognition* (36) 9: 1929-1943

Rosenhahn, B., Klette, R., Sommer, G. (2004): in Proc. of DAGM (Rasmussen et al., eds.), Springer, 294-301.

Rosenhahn, B., Klette, R. (2004): In Proc. of 10th IWCIA (Klette & Zunic, eds.), Springer, 583-596
Richards (1999): *Human Movement Science,* 18: 589-602.
Ferdinands, R.E.D, Kersting, U.G. (2005): http://www.coachesinfo.com/category/cricket/351/

ACKNOWLEDGEMENTS

The presented work has received support by DFG grants RO 2497/1-1 and 2497/1-2.