Automatic human model parametrization from 3d marker data for motion recognition

Hilde Kuehne
Goethe-Universität Frankfurt am Main

Martin Pruzinec

Tobias Feldmann
JENOPTIK AG

Automatic Human Model Parametrization From 3D Marker Data For Motion Recognition

Hildegard Köhler	Martin Pruzinec	Tobias Feldmann	Annika Wörner
Institute for Algorithms and Cognitive Systems	Institute for Algorithms and Cognitive Systems	Institute for Algorithms and Cognitive Systems	Institute for Algorithms and Cognitive Systems
Universität Karlsruhe	Universität Karlsruhe	Universität Karlsruhe	Universität Karlsruhe
Kaiserstr. 12	Kaiserstr. 12	Kaiserstr. 12	Kaiserstr. 12
76131 Karlsruhe, Germany	76131 Karlsruhe, Germany	76131 Karlsruhe, Germany	76131 Karlsruhe, Germany
koehler@ira.uka.de	pruzinec@ira.uka.de	feldmann@ira.uka.de	woerner@ira.uka.de

ABSTRACT

Accurate human motion models are a prerequisite for most applications dealing with the tracking, reconstruction or recognition of human motions. Often a uniform model is used, approximating the average of the evaluated subjects. However we expect most applications can be improved by using individual models for each subject with its personal body masses and features. Hence we propose an algorithm for automatic reconstruction of anatomical features of subjects from labeled 3D marker data by a parameterized generic model.

Our main contribution is a novel approach for automatically estimating skeletons of individual subjects and to transform them to a human body model by preserving its relative configuration. We show that a more accurate model can help in context of motion recognition by improving standard motion reconstruction with regard to its quantity and quality.

Keywords

Adaptive User Modeling, Model Parametrization, Motion Recognition, Human Computer Interface

1. Introduction

Kinematic human models are needed in a lot of different contexts, like motion capture, tracking, motion analysis and motion recognition. In many cases a fixed model is used and adapted by a global scaling factor. But in applications dealing with test persons data a simple scaling is not sufficient to adapt the model to the diverse anatomy of individual subjects. In other applications the model is calculated only from motion data of marker points, what normally leads to a more limited model with only those degrees of freedom which can be calculated from moving limbs, usually comprising torso, head and upper and lower extremities.

In this special context, the main focus lies on the reconstruction of human motion of different subjects from a set of 3D marker data, captured with a Vicon system by using a highly definite model with up to 108 degrees of freedom. The aim of the here presented approach is to combine the features of a static model which can have a lot of degrees of freedom and thus enabling a precise reconstruction of human motion and the features of an adaptive model representing the anatomy of the related test person.

So given static body model is transferred to a relative one by preserving its overall specification. In order to enable the adaption of the model to different body configurations, we extended the model description by 12 additional degrees of freedom, representing several segments lengths. By optimizing these segment lengths the model can be adapted to the body structure variation of the real test persons. The model refinement is completed by the adaption of the specified marker positions of the model to the real positions of the test person's body.

2. Related Work

The adaption and individualization of human models has many application areas like in medical rehabilitation, sports, entertainment industry or product design like e.g. in the automotive industry.
Industrial applications like RAMSIS [Ram07a, Meu07a] and Jack [Tra07a, Bur07a] but also free research projects like the AnyBody Project [Any07a] use adaptive human models e.g. to improve the design of security and functional components in cars. Here typically one or more standard models are used and scaled to fit the needed body configurations. They e.g. relay on configurations emerging from body screenings like the CEASAR database (Civilian American and European Surface Anthropometry) which consist 3d-scan data from about 4400 civilians

to get the opportunity for creating realistic models with a minimum of parameters [Seo03a, Azo05a].
Another approach to adaptive modeling can be found in the domain of estimating joint position from 3D marker data. A general approach to create a fully adapted human figure by combining a local technique based on relative marker trajectories and a global optimization of a skeleton model can be found in [Sil98a]. In the research of [Zha03a] a method for locating elbow and shoulder joint center from a reduced set of markers is analyzed. The method uses an optimization algorithm proposed in [Nus00a]. In [Cer06a] four selected pose estimators, a geometrical method, a SVD-based method, and the Pointer Cluster Technique (PCT) in the optimized and non optimized version, were analyzed. The study took into account all sources of errors typically affecting joint kinematics estimation like instrumental errors, soft tissues artifacts and marker dislocation.

3. Model Specification

The model used for the presented approach is based on the human model of the SFB 588 proposed by [Sim07a], [See05a] and [Ste07a]. The model has a full definition of the human body with a maximum of 108 degrees of freedom.

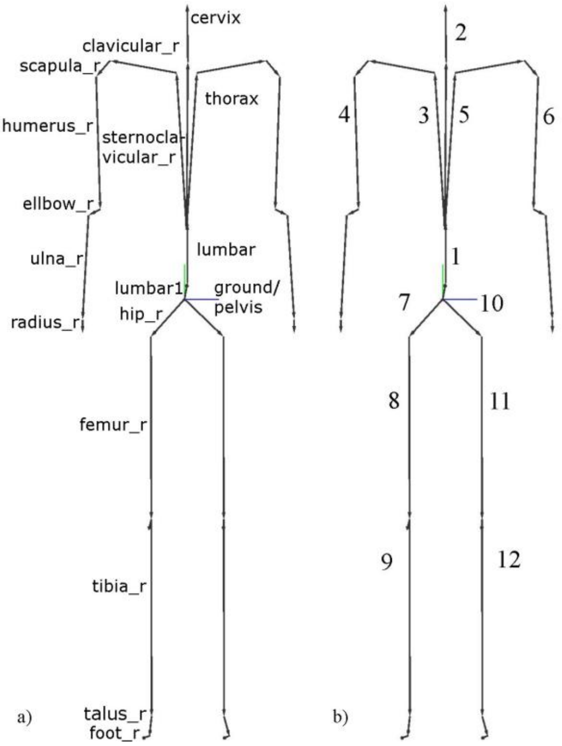

Figure 1: Model configuration a) shows the model segments as described in the overview and model configuration b) shows the extended degrees of freedom

For the presented approach it has been reduced to 35 body segments with 34 body joints, 44 degrees of freedom and 29 defined marker points. Figure 1 shows the configuration of the model segments. The left side of the model shown in Figure 1is not named, but is identically to the right side with only changing the nomenclature index of the segments.
The model definition is realized as a tree structure starting from the pelvis segment. The upper body starts from there with the lumbar structure which branches out to the thorax segment and the left and right sternoclavicular which are connected to the upper extremities. The lower extremities, starting with the left and right hip segment are connected to the pelvis too.

Segment	Translation	Rotation
ground	'torso_TX'; 'torso_TY'; 'torso_TZ'	'lower_torso_TX'; 'lower_torso_TY'; `lower_torso_TZ'
pelvis	-	'lumbar_pitch'; 'lumbar_roll';
lumbar	-	'thorax_roll'; 'thorax_yaw'; 'thorax_pitch'
lumbar1	'thorax_TY'	-
thorax	-	'neck_pitch'; 'neck_roll'
cervix	'skull_TY'	['skull_pitch'; 'skull_roll'; 'skull_yaw']
sternocla-vicular_r	'clav_r_TY'	-
clavicu-lar_r	-	-
scapula_r	-	'arm_add_r'; 'arm_flex_r'; 'arm_rot_r'
humerus_r	`ulan_r_TY'	
elbow_r		'elbow_flex_r'
ulna_r	-	'pro_sup_r'
radius_r	-	'wrist_flex_r'
wflex_r	-	'wrist_dev_r'
hip_r	'femur_r_TZ'	'hip_flex_r'; 'hip_add_r'; 'hip_rot_r'
femur_r	'tibia_r_TY'	'knee_flex_r'
tibia_r	'talus_r_TY'	'ankle_flex_r'
talus_r	-	'subt_angle_r'
foot_r	-	'toe_angle_r'

The overall degrees of freedom of model specification are listed by indicating the segment and its related degrees of freedom. Because of the symmetry of the model, only the right side is named and described here. The left side is build identically with only changing the nomenclature of the segments and degrees of freedom.

To adapt the model the original specification is extended by 12 additional degrees of freedom, defining individual segment lengths of the test person, as the can be seen in figure 1. They represent the following connections:

- Lumbar and thorax - the connection between L3 and T10 over the y-axis to adjust the overall height of the upper body (1)
- Cervix and skull - the connection between C3 and the skull base over the y-axis to adjust the neck (2)
- Thorax and right clavicle - the connection between T10 and the right sternoclavicular over the y-axis to adjust the upper body height (3)
- Right humerus and ulna - the connection between right shoulder and elbow over the y-axis to adjust the length of the upper arm (4)
- Thorax and left clavicle - the connection between T10 and the left sternoclavicular over the y-axis to adjust the upper body height (5)
- Left humerus and ulna - the connection between left shoulder and elbow over the y-axis to adjust the length of the upper arm (6)
- Pelvis and right femur - the connection between the pelvis and the right hip joint over the z-axis to adjust the pelvis width (7)
- Pelvis and left femur - the connection between the pelvis and the left hip joint over the z-axis to adjust the pelvis width (8)
- Right femur and tibia - the connection between right hip joint and knee joint over the y-axis to adjust the right thigh length (9)
- Right tibia and talus - the connection between right knee joint and foot ankle over the y-axis to adjust the right lower leg length (10)
- Left femur and tibia - the connection between left hip joint and knee joint over the y-axis to adjust the right thigh length (11)
- Left tibia and talus - the connection between left knee joint and foot ankle over the y-axis to adjust the left lower leg length (12)

The model misses the adjustment of the lower arm length because in this special case there were no markers in this region. So, an adjustment would not be necessary. The joint angles as well as the segments lengths are approximated to the real marker data by an optimization algorithm in order to find the best fitting configuration for the model with respect to the given marker positions.

Second, the new absolute position from the marker to the related joint center is estimated. The deviation of the real marker position to the new reconstructed marker position is calculated over a given number of frames and the mean deviation is calculated. Than this mean deviation is added to the relative marker position defined in the model description.

4. Algorithmic approach

The implementation is based on the simulation frame work by Seemann et al. [See05a]. The model itself consists of the described set of body segments which are connected by joint elements representing the related degrees of freedom in human anatomy. The marker points are defined relative to the joint and segment positions. Usually, the motion of a test person is reconstructed by optimizing the degrees of freedom of the overall model to find the best fitting pose for the actual marker positions.

For the presented approach, the set of degrees of freedom, which usually represent the human joints, has been extended by the defined segments. The optimization includes not only possible joint angle configurations but also the defined segments lengths.. By this the model is adapted to the real body structure of the test person. This adaption can be done for every frame as well as for an initial sequence to determine the overall configuration.

The optimization is done using the Matlab implementation of the interior-reflective Newton method [Col96a]. Using this method the required number of iterations grows only very slow at increasing dimensions. While incorporating upper and lower boundaries, it allows integrating the limits of joint angles. It is guaranteed that the optimization algorithm produces only strictly feasible iterates in the bounded region. A reflective transformation is used to maintain this feasibility. To gain second order convergence a Newton system is used. By integrating an affine scaling transformation as part of the mapping function, a global convergence is achieved.

Then, the marker configuration of the model must be adapted to the marker positions of the test person. Even if the marker positions are associated to anatomical landmarks, it is not possible to define the exact point for every marker in advance because the overall marker positions can change from one capture session to another. To adjust the marker position for any session, we introduced a correction parameter for every marker, which describes the deviation of the marker from the reconstructed configuration. The deviation is calculated as the mean value of the distance of the measured and the reconstructed marker position. By adding the correction factor of every marker to the overall model, the reconstruction

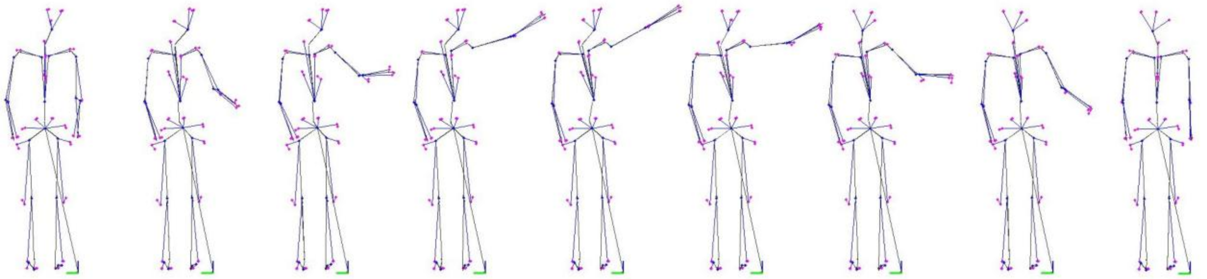

Figure 2: Example for a motion sequence of a pointing gesture. The sequence shows reconstruction results of the frames nr. 1, nr. 65, nr. 73, nr. 97, nr. 104, nr. 120, nr. 153, nr. 172 and nr. 197 (last frame)

becomes much more accurate and results in a realistic estimation of the test person´s body structure.

5. Evaluation/Results

The evaluation is done with respect to different aspects of motion reconstruction. As there is no ground truth for the joint angle configuration of a certain pose we have to define indirect criterions to measure the overall accuracy of the joint angle reconstruction of a pose.

In the manual inspection the motion of the reconstructed joint angle figure is compared with the related video sequence to determine wrong pose estimations. In order to evaluate the poses of the original static model as well of the adaptive model we animated several motion sequences of pointing gestures (Figure 2) in order to compare the reconstructed motion to the one of the video sequence. In Figure 3 the test person was smaller than the original model. As a result, the joints of the static model tend to bend in order to adapt to a specific height. The first image shows the test person standing upright. The pose of the static model shown in the second image is compensating the incorrect height by bending the knee and thorax joint. In the adapted model pose, the related joint are straight, what corresponds to the test persons pose.

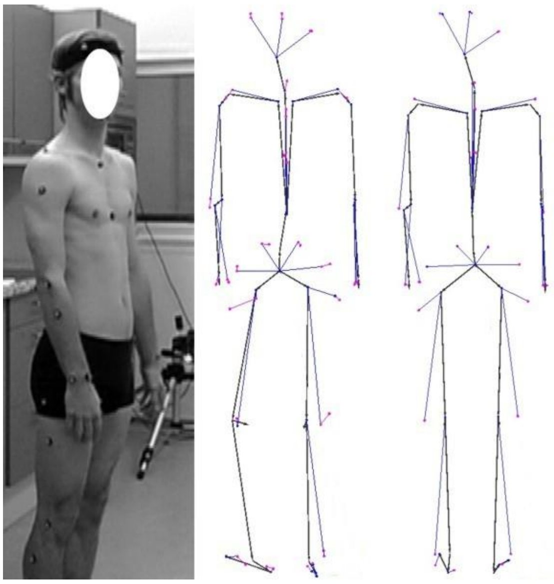

Figure 3: Test person and its pose, reconstructed with the original static model (middle) and the adaptive model (right).

This effect can also be seen in Figure 4 regarding the gradient of the knee flexion during a sequence, where the test person stands still. In the reconstruction with a fixed model, the knee joint flexion angle is much higher compared to the knee flexion reconstructed with the adaptive model.

As there is no absolute criterion for the statistical evaluation of the joint angles, there are two indirect components, which can be seen as indicator for the correctness of motion reconstruction.

First, it can be assumed that a reconstructed pose is more accurate the better the reconstructed marker points match the real marker positions. Assuming a perfect reconstruction of the pose with an ideal model, the gap between the reconstructed and the captured marker points would be zero.

The second criterion is based on the fact that human motion follows specific constrains like energy minimization and continuous speed and path adaption. This fact can hold as a criterion for the motion reconstruction because in a natural human motion the first and second derivate of the joint angle trajectories are continuous as long as no external forces were involved in the motion.

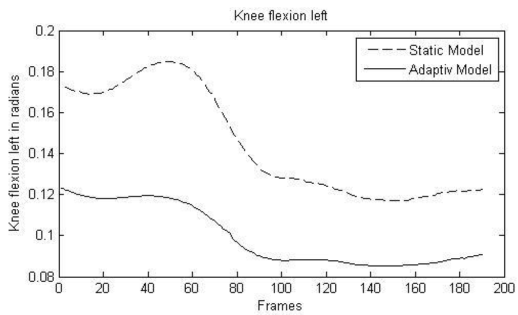

Figure 4: Joint angle trajectory of the right knee flexion reconstructed by a static and adaptive model.

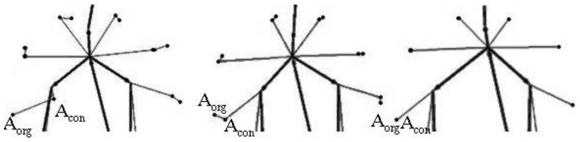

Figure 5: Pelvis configuration reconstructed with the original static model (left), an adapted model (middle) and the adaptive model (right)

Contrary to this a bad reconstruction causes discontinuities in the joint angle trajectories. If the reconstructed joint angle does no longer fit the actual pose, it needs to change significantly from one frame to another because the optimization becomes so unspecific that another pose seems to fit the model better than the older one. So we can assume that peaks in the first and second derivate of joint angle trajectories give a good indication of an inaccurate reconstruction.

To evaluate the proposed adaptive model the two defined criterions were evaluated with 380 motion sequences of pointing gestures of 19 different test persons with meanly 200 frames per sequence. First we evaluated the distance between the captured data and marker data reconstructed with the original static model, reconstructed with a static model which has been initialized with adapted model data and reconstructed with a model where the segments lengths has been adapted for every frame. As apparent in Figure 5, the gap between the original marker position A_{org} and the reconstructed marker position A_{con} can become very large by using the static model compared to the adaptive one.

It can be seen in Figure 6 that this holds for all test sequences. The overall distance of the reconstructed marker positions becomes smaller, if no rigid body segments are used.

Second, the overall discontinuities in the reconstruction of the joint angle gradients were evaluated. Here it is easy to see, that especially for the adaptive model there are less peaks (overall count 4) then for either the original static model (overall count 18) or for the adapted static model (overall count 10).

The quantitative analysis has been evaluated on a quad-core Intel PC with 4 GB ram. The mean calculation time per frame is slightly higher for the original static model (0.0622 sec/frame) than for the initially adapted model (0.0568 sec/frame) and the full adaptive model (0.0473 sec/frame) with 12 additional degrees of freedom. The mean runtime performance for all sequences is shown in Figure 7. The additional time for the original static model can be explained by the fact that more iterations are needed to find the best possible pose. The better runtime performance of the full adaptive model with 12 additional degrees of freedom can be explained by the fact that the higher dimension space allows a better overall adaption of the marker points. So the overall number of iterations declines.

6. Conclusion

We showed that it is possible to transfer a given static body model to a relative one by preserving its overall specification. We adapted the model to 19 test persons in order to gain better motion reconstructions compared to a static model. We showed that such a model can help to improve motion reconstruction in both ways, quantitatively and qualitatively. We found the interior-reflective Newton optimization strategy suitable for large scale nonlinear problems, enabling the solving of fitting problems with high accuracy in an efficient way.

An interesting effect has been shown by not just using the optimization of segments lengths for an initial frame sequence, but during the complete motion sequence, based on the idea that the human body is more a flexible than a rigid structure. So the single elements and joints can be stretched for over certain distances. The deviation of the single segments can help to simulate this stretching effect. The lower

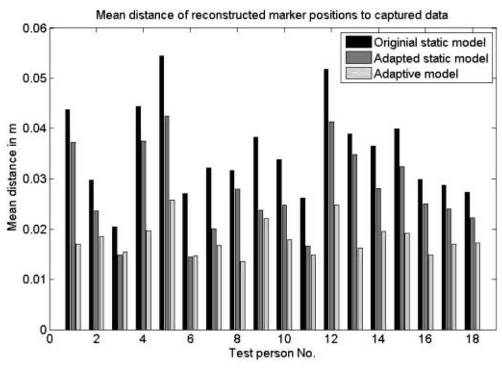

Figure 6: Mean distance of marker position reconstructed with the original static model, an adapted model and the adaptive model from captured marker positions.

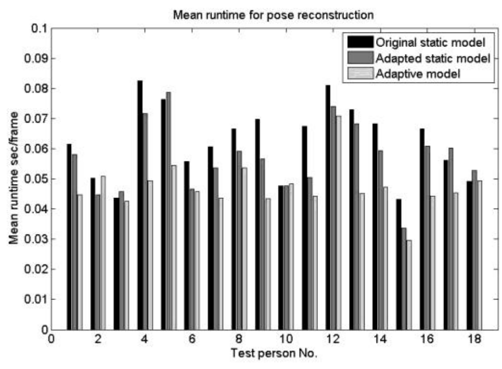

Figure 7: Mean runtime indicating the mean seconds per frame needed for the reconstruction with the original static model, an initially adapted model and the fully adaptive model.

mean distance factor in this context can be seen as an indication that a flexible model might be more accurate to simulate and reconstruct a human motion, than a static one.

Our further work will comprise the initial estimation of joint angles and related marker positions to find a better starting position and to improve the model initialization. In long-term view we will try to adapt the presented adaptive model approach to vision based motion recognition systems in order to improve vision based pose estimation.

7. ACKNOWLEDGMENT

This work was supported by the grant from the Ministry of Science, Research and the Arts of Baden-Württemberg.

8. REFERENCES

[Any07a] http://www.anybodytech.com/

[Bur07a] Burnette, C. and Schaaf, W. Issues in Using Jack Human Figure Modeling Software To Assess Human-Vehicle Interaction in a Driving Simulator. Transportation Research Record, Transportation Research Board of the National Academies, Vol. 1631, pp.1-7, 1998

[Cer06a] Ceratti, A., Croce, U. D. and Cappozzo, A. Reconstruction of skeletal movement using skin markers: comparative assessment on bone pose estimators. Journal of Neuro Engineering and Rehabilitation, 3:7, 2006

[Col96a] Coleman, T.F. and Li Y. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds, SIAM Journal on Optimization, Vol. 6, pp. 418-445, 1996

[Ram07a] RAMSIS - HUMAN SOLUTIONS, http://www.human-solutions.com/

[Meu07a] van der Meulen, Pl and Seidl, A. Ramsis – The Leading Cad Tool for Ergonomic Analysis of Vehicles. Digital Human Modeling, Lecture Notes in Computer Science, Vol. 4561, pp. 1008-1017 Springer Berlin / Heidelberg, 2007

[Nus00a] Nussbaum, M. A. and Zhang, X. Heuristics for locating upper extremity joint centres from a reduced set of surface markers. Human Movement Science, Vol. 19, pp. 797-816, 2000

[See05a] Seemann W., Stelzner G. and Simonidis C. Correction of motion capture data with respect to kinematic data consistency for inverse dynamic analysis. Proceedings of International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, ASME, Long Beach, USA, 2005

[Sil98a] Silaghi, M., Plänkers, R., Boulic, R., Fua, P. and Thalmann, D. Local and Global Skeleton Fitting Techniques for Optical Motion Capture, Lecture Notes in Computer Science, Vol. 1537, 1998

[Sim07a] Simonidis C., Stelzner G. and Seemann W. A Kinematic Study of Human Torso Motion. Proceedings of International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, ASME, Las Vegas, USA, 2007

[Ste07a] Stein T., Fischer A., Boesnach I., Gehrig D., Köhler H., Schwameder H Kinematische Analyse menschlicher Alltagsbewegungen für die Mensch-Maschine-Interaktion, J. Edelmann-Nusser, K. Witte & E.F. Moritz (Hrsg.), Sporttechnologie zwischen Theorie und Praxis: V. Aachen: Shaker, 2007

[Tra07a] Transom Technologies, Inc., http://www.transom.com/

[Zha04a] Zhang, X., Lee, S.-W. and Braido, P. Towards an integrated high-fidelity linkage representation of the human skeletal system based on surface measurement. International Journal of Industrial Ergonomics: Special Issue on Anthropometrics and Disability, Vol. 33, pp. 215-227, 2004.