

A system for marker-less motion capture

B. Rosenhahn, T. Brox, U. G. Kersting, A. W. Smith, J. K. Gurney, R. Klette

In this contribution we present a silhouette based human motion capture system. The system components contain silhouette extraction based on level sets, a correspondence module, which relates image data to model data and a pose estimation module. Experiments are done in different camera setups and we estimate the model components with 21 degrees of freedom in up to two frames per second. To evaluate the stability of the proposed algorithm we perform a comparison of the motion estimation system with a marker based tracking system. The results show the applicability of the system for marker-less sports movement analysis. We finally present extensions for motion capture in complex environments, with changing lighting conditions and cluttered background. This paper is an extended version of [21] which was awarded the DAGM Main Prize on the annual symposium of the German pattern recognition society (DAGM) in Vienna, 2005.

1 Introduction

Classical motion capture (or MoCap) comprises techniques for recording the movements of real objects, usually humans [25]. A common application aims to analyze the captured data for subsequent motion analysis purposes, e.g., clinical studies, diagnostics of orthopaedic patients or to help athletes to understand and improve their performances. It has also grown increasingly important as a source of motion data for computer animation. Surveys on existing methods for MoCap can be found in [16, 12]. Well known (commercially available) marker-based tracking systems exist, e.g. provided by Motion Analysis, Vicon or Simi [15]. There are intrinsic problems in using surface markers, e.g. incorrect tracking of markers, tracking failures, the need for special lab environments and lighting conditions and the fact, that people do not feel comfortable with markers attached to the body. This often leads to unnatural motion patterns. Marker-based systems are also designed to track the motion of the markers themselves and thus it must be assumed that the recorded motion of the markers is identical to the motion of the underlying human segments. Since human segments are not truly rigid, this assumption may cause problems, especially in highly dynamic movements typically seen in sporting activities. For these reasons, marker-less tracking is an important field of research. It requires knowledge in biomechanics, computer vision and computer graphics.

Typically, researchers working in the area of computer vision prefer simplified human body models, e.g., based on stick, ellipsoidal, cylindrical or skeleton models [1, 14, 11, 13]. In computer graphics advanced object modelling and texture mapping techniques for human motions are well known [24, 8, 6, 26] but the image processing and pose estimation techniques are often simplified. Therefore we started to combine silhouette based pose estimation with more realistic human models: These are represented by free-form surface patches and local morphing along the surface patches is applied to gain a realistic human model within silhouette based MoCap. Our previous works [22, 23] are now extended to a complete human motion capture system. The system consists of an advanced image segmentation method, dynamic

occlusion handling and kinematic chains of higher complexity (21 degrees of freedom). We perform a comparison of the system with a commercial marker based tracking system [17] used to analyze sports movements¹. Exercises, such as push ups or sit ups are analyzed. We further present an approach which allows to track humans in cluttered scenes and changing lighting conditions. This is an extension of [2], where previously used rigid objects are replaced with multiple free-form surface patches connected by joint axes of various degrees of freedom.

The contribution is organized as follows: We will start with the basic setup of the motion capture system. Then we will continue with the system modules. Here we will briefly describe image segmentation based on level sets, pose estimation and the dynamic occlusion handling to deal with partial occlusion in certain frames. The next section presents the experimental results, the quantitative error analysis and motion capture results in scenes with cluttered background. The contribution ends with a brief summary.

2 The human motion tracking system

A 3D object model builds the a priori knowledge of the system, which is in this case given as two free-form surface patches with two kinematic chains. Each kinematic chain consists of seven joints (three for the shoulder, two for the elbow and two for the wrist). Furthermore we added one back segment joint to the torso surface patch. The estimation procedure is dealing with 21 unknowns, six for the pose parameters (three for rotation and three for translation), seven for each arm and one backbone joint. During correspondence estimation (along four frames) we collect around 5000 point correspondences (slightly varying dependent on the visible information) and still track in two frames per second for four camera sequences. Using the 3D model and images from a (triggered) calibrated multi-camera sequence,

¹ Motion Analysis Corporation is one of the leading providers of optical motion capture systems in entertainment, video-games, film, broadcasting, virtual reality, medicine, sports, and research.

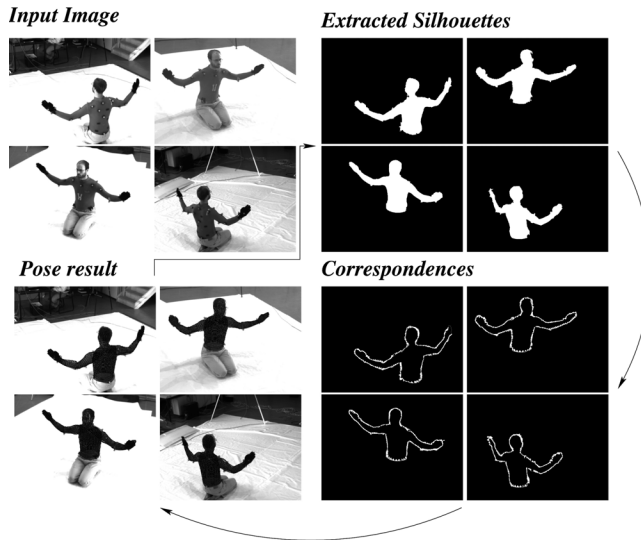


Figure 1: The capture system consists of iterating the following steps: Segmentation, correspondence estimation, pose estimation.

the motion tracking system consists of three main components, namely silhouette extraction, matching and pose estimation. All components are iterated to stabilize segmentation on the one hand and pose estimation on the other hand.

2.1 Image segmentation



Figure 2: Silhouette extraction based on level set functions. Left: Initial segmentation. Right: Segmentation result.

We refer to image segmentation as the estimation of the closed object contour in the image. This task can become very difficult, since noise, shading, occlusions, inhomogeneous objects, or background clutter may distort the segmentation. We apply the level set technique [19, 7, 9, 3], in which a level set function $\Phi \in \Omega \mapsto \mathbb{R}$ splits the image domain Ω into two regions Ω_1 and Ω_2 , the object and the background. It holds $\Phi(x) > 0$ if $x \in \Omega_1$ and $\Phi(x) < 0$ if $x \in \Omega_2$. The zero-level line thus marks the boundary between the two regions. The segmentation is sought to maximize the total a-posteriori probability given the probability densities p_1 and p_2 of Ω_1 and Ω_2 , i.e., pixels are assigned to the most probable region according to the Bayes rule. Additionally, the boundary between both regions should be as small as possible. This can be expressed by the following energy functional that is sought to be minimized:

$$E(\Phi, p_1, p_2) = - \int_{\Omega} (H(\Phi) \log p_1 + (1 - H(\Phi)) \log p_2 + \nu |\nabla H(\Phi)|) dx \quad (1)$$

where $\nu > 0$ is a weighting parameter and $H(s)$ is a regularized version of the Heaviside function, e.g., the error function. Minimization with respect to the region boundary represented by Φ can be performed according to the gradient descent equation

$$\partial_t \Phi = H'(\Phi) \left(\log \frac{p_1}{p_2} + \nu \operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) \right) \quad (2)$$

where $H'(s)$ is the derivative of $H(s)$ with respect to its argument. The probability densities p_i are estimated according to the EM principle: having the level set function initialized with some contour, the probability densities within the two regions are estimated by the gray value histograms smoothed by a Gaussian kernel K_{σ} with standard deviation σ . The density estimates are used in the gradient descent and the process is iterated until it converges.

This rather simple and fast approach is sufficient for our laboratory set-up. Figure 2 shows an example image and the contour evolution over time. Obviously, the body silhouette is well extracted, besides some deviations in the head region caused by the dark hair. Such inaccuracies, however, can be handled by the pose estimation procedure. Moreover, we have a tracking assumption. Therefore, we can initialize the silhouette with the pose of the last frame which greatly reduces the number of iterations needed. This implementation is fast; the segmentation algorithm needs 50 ms per frame, i.e., 200 ms in a four-camera setup.

In more complex scenes, the segmentation is susceptible to distracting artifacts like shadows, texture, or background clutter. A grossly bad contour, however, makes the pose estimation procedure fail. In [2] we therefore presented an approach for iterative segmentation and pose estimation, where the shape knowledge supports the segmentation procedure. The energy in (1) is extended by an additional term

$$E_{\text{Shape}}(\Phi, \xi) = \lambda \int_{\Omega} (\Phi - \Phi_0(\xi))^2 dx \quad (3)$$

that penalises deviations of the contour from the projected model surface with the pose parameters ξ given as level set function Φ_0 . This term is minimized by drawing the contour towards the projected surface on one hand, but also by adapting the pose parameters such that the projected surface better fits to the extracted contour. The energy functional therefore represents a variational model for the joint estimation of the contour and the pose parameters.

Local optimization is achieved by alternating the minimization with respect to the contour for fixed pose parameters

$$\partial_t \Phi = H'(\Phi) \left(\log \frac{p_1}{p_2} + \nu \operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) + 2\lambda (\Phi_0 - \Phi) \right) \quad (4)$$

and the minimization with respect to the pose parameters for a fixed contour with the method described in section 2.3.

The model in [2] further extends the image gray value by color and texture features from [4] in order to allow for textured objects. These additional features can be integrated into the segmentation model via the probability density models [3]. Whereas in [2] the model has been restricted

to rigid objects, section 3.1 presents first results for motion capture of human upper torsos with cluttered background. Due to its iterative nature, the extended algorithm is considerably slower than the basic one, and real-time performance is currently not available.

2.2 Correspondence estimation

After image segmentation, correspondences between the object model and the extracted silhouettes have to be established. To this end, we follow a modified version of an ICP algorithm [22] and use a voting method to decide whether a point belongs to the torso or to one of the arms. These correspondences are used by the pose estimation module described in the subsequent section to determine new pose parameters. The correspondences are iteratively updated according to the new pose until the overall pose converges.

2.3 Pose estimation

For pose estimation we assume a set of point correspondences (X_i, x_i) , with 4D (homogeneous) model points X_i and 3D (homogeneous) image points x_i . Each image point defines a 3D Plücker line $L_i = (n_i, m_i)$ (projective ray), with a (unit) direction n_i and moment m_i [18].

Every 3D rigid motion can be represented in an exponential form

$$M = \exp(\theta \hat{\xi}) = \exp \begin{pmatrix} \hat{\omega} & v \\ 0_{3 \times 1} & 0 \end{pmatrix} \quad (5)$$

where $\theta \hat{\xi}$ is the matrix representation of a twist $\xi \in se(3) = \{(v, \hat{\omega}) | v \in \mathbb{R}^3, \hat{\omega} \in so(3)\}$, with $so(3) = \{A \in \mathbb{R}^{3 \times 3} | A = -A^T\}$. The Lie algebra $so(3)$ is the tangential space of the 3D rotations. Its elements are (scaled) rotation axes, which can either be represented as 3D vector

$$\theta \omega = \theta \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}, \quad \text{with } \|\omega\|_2 = 1 \quad (6)$$

or as screw symmetric matrix

$$\theta \hat{\omega} = \theta \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \quad (7)$$

In fact, M is an element of the one-parametric Lie group $SE(3)$, known as the group of direct affine isometries. A main result of Lie theory is that to each Lie group there exists a Lie algebra which can be found in its tangential space, by derivation and evaluation at its origin; see [18] for more details. The corresponding Lie algebra to $SE(3)$ is denoted as $se(3)$. A twist contains six parameters and can be scaled to $\theta \xi$ with a unit vector ω . The parameter $\theta \in \mathbb{R}$ corresponds to the motion velocity (i.e., the rotation velocity and pitch). For varying θ , the motion can be identified as screw motion around an axis in space. To reconstruct a group action $M \in SE(3)$ from a given twist, the exponential function $\exp(\theta \hat{\xi}) = M \in SE(3)$ must be computed. This can be done efficiently by using the Rodriguez formula [18],

$$\exp(\hat{\xi} \theta) = \begin{pmatrix} \exp(\theta \hat{\omega}) & (I - \exp(\hat{\omega} \theta))(\omega \times v) + \omega \omega^T v \theta \\ 0_{1 \times 3} & 1 \end{pmatrix}, \quad \text{for } \omega \neq 0 \quad (8)$$

with $\exp(\theta \hat{\omega})$ computed by calculating

$$\exp(\theta \hat{\omega}) = I + \hat{\omega} \sin(\theta) + \hat{\omega}^2 (1 - \cos(\theta)). \quad (9)$$

Note that only sine and cosine functions of real numbers need to be computed.

For pose estimation we combine the reconstructed Plücker lines with the screw representation for rigid motions and apply a gradient descent method: Incidence of the transformed 3D point X_i with the 3D ray $L_i = (n_i, m_i)$ can be expressed as

$$(\exp(\theta \hat{\xi}) X_i)_{3 \times 1} \times n_i - m_i = 0. \quad (10)$$

Indeed, X_i is a homogeneous 4D vector and after multiplication with the 4×4 matrix $\exp(\theta \hat{\xi})$ we neglect the homogeneous component (which is 1) to evaluate the cross product with n_i . We now linearize the equation by using $\exp(\theta \hat{\xi}) = \sum_{k=0}^{\infty} \frac{(\theta \hat{\xi})^k}{k!} \approx I + \theta \hat{\xi}$, with I as identity matrix. This results in

$$((I + \theta \hat{\xi}) X_i)_{3 \times 1} \times n_i - m_i = 0 \quad (11)$$

and can be re-ordered into an equation of the form $A \xi = b$. Collecting a set of such equations (each is of rank two) leads to an overdetermined system of equations, which can be solved using, for example, the Householder algorithm. The Rodriguez formula can be applied to reconstruct the group action M from the estimated twist ξ . Then the 3D points can be transformed and the process is iterated until the gradient descent approach converges.

Joints are expressed as special screws with no pitch of the form $\theta_j \hat{\xi}_j$ with known $\hat{\xi}_j$ (the location of the rotation axes as part of the model representation) and unknown joint angle θ_j . The constraint equation of a j th joint has the form

$$(\exp(\theta_j \hat{\xi}_j) \dots \exp(\theta_1 \hat{\xi}_1) \exp(\theta \hat{\xi}) X_i)_{3 \times 1} \times n_i - m_i = 0 \quad (12)$$

which is linearized in the same way as the rigid body motion itself. It leads to three linear equations with the six unknown pose parameters and j unknown joint angles. Collecting a sufficient number of equations leads to an overdetermined system of equations.

Note, that since we work with reconstructed 3D lines we can gain equations from different cameras (calibrated with respect to the same global coordinate system) and merge them into one system of equations and solve them simultaneously. This is the key idea to deal with partial occlusions: A joint that is not visible in one camera must be visible in another one to get a solvable system of equations. A set of four cameras around the subject covers a large range and allows the analysis of quite complex motion patterns.

In order to deal with larger motions during pose tracking we use a sampling method that applies the pose estimation algorithm for different neighboring starting positions. From all results we then choose the one with the minimum error between the extracted silhouette and the projected surface mesh. This is important to avoid local minima during tracking.

3 Experiments

Figure 3 shows pose results of a 3-camera sequence. For visualization, the surface patches are transformed (with the estimated pose and joint angles) and projected onto the images. In this sequence the subject performs a 360 degrees rotation. It results in partial occlusions which can be handled from the algorithm.

A lack of many studies is that only a visual feedback about the pose result is given, by overlaying the pose result with the image data, e.g. [22, 14]. To enable a quantitative error



Figure 3: A 360 degrees rotation during a 3-Camera sequence. The algorithm is able to handle partial occlusions.

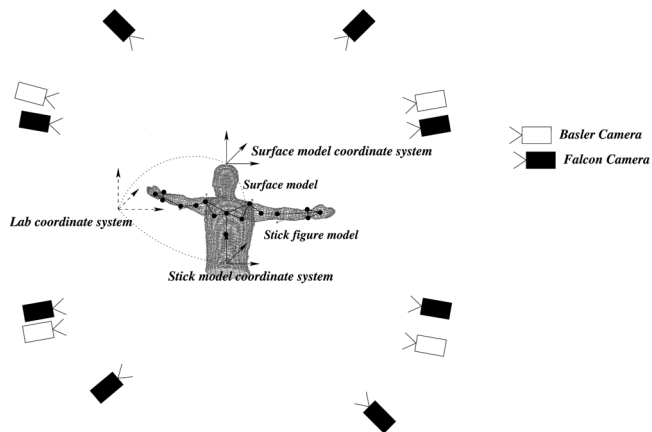


Figure 4: The coordinate systems in the lab setup.

analysis, we use a commercial marker based tracking system for comparison. Here, we use the Motion Analysis software [17], with an 8-Falcon-camera system. For data capture we use the Eva 3.2.1 software and the Motion Analysis Solver Interface 2.0 for inverse kinematics computing [17]. The solver interface extracts Euler angles which are transferred to angles representing the kinematic chain structure of the used surface model. In this system a human has to wear a body suit and retro-reflective markers are attached to it. Around each camera is a strobe light led ring and a red-filter is in front of each lens. This gives very strong image signals of the markers in each camera. These are treated as point markers which

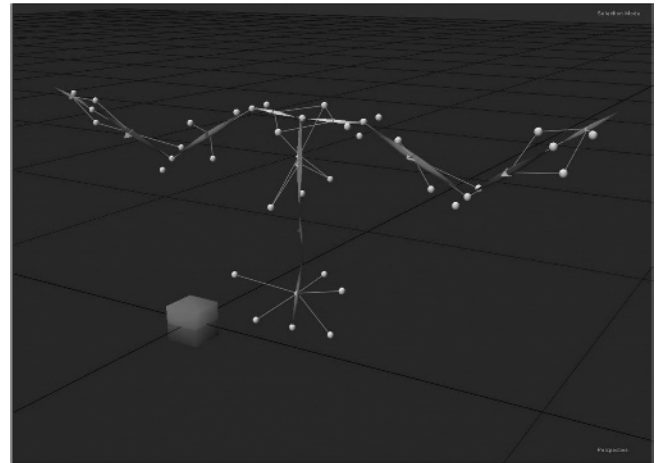


Figure 5: The coordinate systems of the markers in the lab setup. Within the Solver Interface 2.0, the markers are (manually) connected to a skeleton model.

are reconstructed in the eight-camera system. The system is calibrated by using a wand-calibration method. Due to the filter in front of the images we had to use a second camera set-up which provides *real* image data. This camera system is calibrated by using a calibration cube. After calibration, the offsets and rotations between both world coordinate systems are calculated. Then we generate a stick-model from the point markers including joint centers and orientations. This results in a complete calibrated set-up we use for a system comparison. It is visualized in figure 4. Some reconstructed markers of the Motion Analysis system are shown in figure 5. The skeletons are connected manually in Eva 3.2.1 by connecting the reconstructed points. Using this setup we then grabbed a series of test sequences.

The images in the upper left of figure 1 show the body-suit with the attached markers. These lead to minor errors during silhouette extraction, which are omitted here. Figure 6 shows the first test sequence, where the subject is just moving the arms forwards and backwards. The diagram shows the estimated angles of the right elbow. The marker results are given as dotted lines and the silhouette results in solid lines. The overall error between both angles diagrams is 2.3 degrees, including the tracking failure between frames 200 till 250.

Figure 7 shows the second test sequence, where the subject is performing a series of push-ups. Here the elbow angles are much more characteristic and also well comparable. The overall error is 1.7 degrees. Both sequences contain partial occlusions in certain frames. But this can be handled by the algorithm.

3.1 Tracking in complex environments

While the basic segmentation and pose estimation model is sufficient to capture the human upper torso under controlled lab conditions nearly in real-time, more challenging scenes with cluttered background need the enhanced model with joint segmentation and pose estimation to yield reliable tracking results. Figure 8 shows an example stereo frame and highlights some properties of the sequence which are challenging during tracking: the subject is wearing a loose shirt which causes problems in the segmentation and does not fit

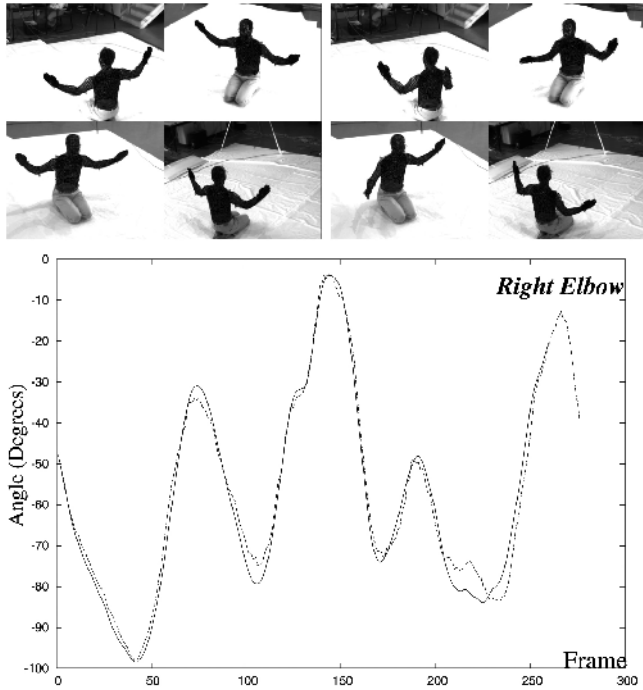


Figure 6: Tracked arms: The angle diagrams show the elbow values of the Motion analysis system (dotted) and the silhouette system (solid).

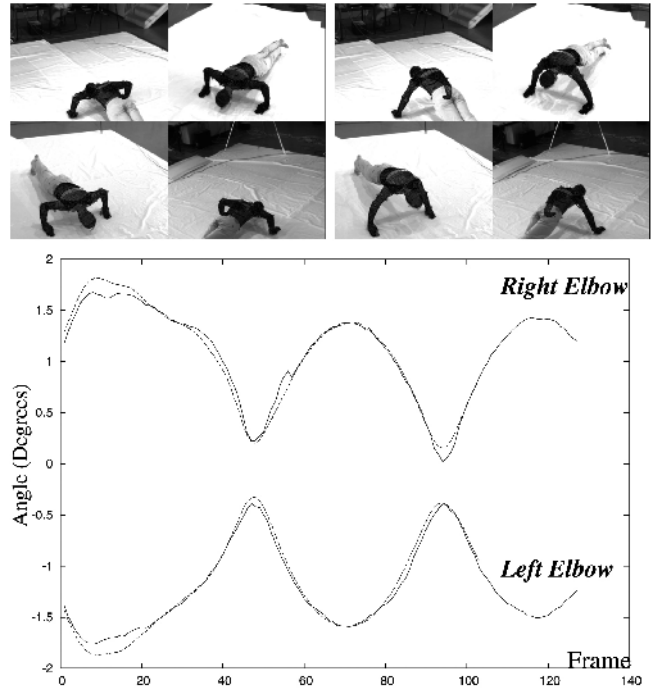


Figure 7: Tracked Push-ups: The angle diagrams show the elbow values of the Motion analysis system (dotted) and the silhouette system (solid).

accurately to the surface model which is generated from a person wearing a slim body suit. The hands are sometimes difficult to separate from the background and clutter offers many possibilities for the contour to be distracted. Finally, light from the right leads to a brightly illuminated face on one side, while it is dark on the other. Pose results of the sequence are shown in figure 9. The person is moving in the scene and shaking his arms. The upper images show the pose results overlaid with the image data and the lower images show the segmentation results. The experiment indicates the possibility to track humans even under more difficult conditions with cluttered background, which is required, for instance, in outdoor scenes.

4 Summary

The contribution has presented a human motion estimation system. It extracts silhouettes by using level set functions. The silhouettes are used for model fitting with 21 degrees of freedom in a four-camera set-up. We have further performed a comparison of the marker-less approach with a commercial marker based tracking system. In [20] eight bio-mechanical measurement systems are compared (including the Motion Analysis system) showing that the root mean square (RMS) errors are typically within three degrees. The errors we achieve with our system fit in this range quite well. The method thereby achieves almost real-time performance.

We have further presented an extended version of the system that allows for human motion tracking in more complex environments with textured objects, shadows and cluttered background. This shows that our work in [2] can be extended to kinematic chains and that we have the foundations for motion capturing in complex outdoor environ-



Figure 8: Tracking with cluttered background.

ments. Marker-less human motion tracking is highly challenging for sports, exercise, and clinical analysis. The system evaluation of our approach revealed very promising results. Future works will continue with experiments on high-speed sports movement tracking in outdoor scenes.

Acknowledgments

We gratefully acknowledge funding by the German Research Foundation (RO 2497/1-1 and RO 2497/1-2) and the Max Planck Center for Visual Computing and Communication.

References

- [1] Bregler C. and Malik J. Tracking people with twists and exponential maps. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Santa Barbara, California, pp. 8-15, 1998.
- [2] Brox T., Rosenhahn B. and Weickert J. Three-Dimensional Shape Knowledge for Joint Image Segmentation and Pose Estimation. *Pattern Recognition 2005, DAGM*, W. Kropatsch, R. Sablatnig and A. Hanbury (Eds.), Springer-Verlag, Berlin Heidelberg, LNCS 3663, pp. 109-116, Wien, 2005.
- [3] Brox T., Rousson M., Deriche R., Weickert J. Unsupervised segmentation incorporating colour, texture, and motion. In *Computer*

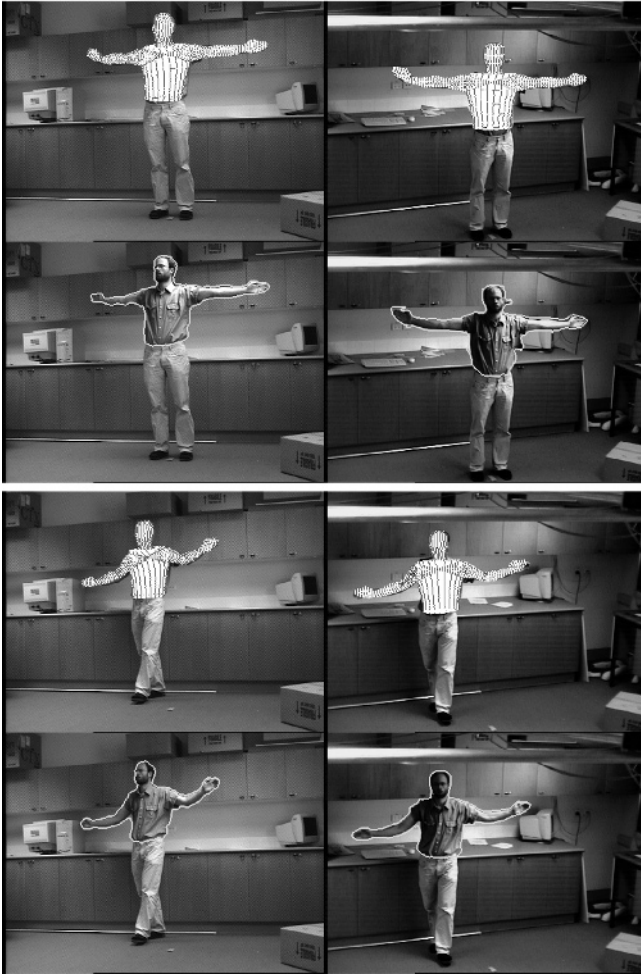


Figure 9: Pose and segmentation results.

Analysis of Images and Patterns, Springer LNCS 2756, N.Petkov, M.A.Westenberg (Eds.), pp. 353-360, Proc. CAIP 2003 Conference, Groningen, The Netherlands, 2003.

- [4] Brox T. and Weickert J. A TV flow based local scale measure for texture discrimination. In T. Pajdla and J. Matas, editors, *Computer Vision - Proc. 8th European Conference on Computer Vision*, LNCS 3022, pages 578-590. Springer, May 2004.
- [5] Campbell R.J. and Flynn P.J. A survey of free-form object representation and recognition techniques. *Computer Vision and Image Understanding (CVIU)*, Vol. 81, pp. 166-210, 2001.
- [6] Carranza, J., Theobalt, C., Magnor, M. A., and Seidel, H.-P. Free-viewpoint video of human actors. In *Proceedings of SIGGRAPH 2003*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 569-577, 2003.
- [7] Caselles V., Catté F., Coll T. and Dibos F. A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1-31, 1993.
- [8] Chadwick J.E., Haumann D.R. and Parent R.E. Layered construction for deformable animated characters *Computer Graphics*, Vol. 23, No. 3, pp. 243-252, 1989.
- [9] Chan T. and Vese L. An active contour model without edges. In M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision*, volume 1682 of *Lecture Notes in Computer Science*, pages 141-151. Springer, 1999.
- [10] Cremers D., Kohlberger T., and Schnörr Ch. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, No. 36, Vol. 9, pp. 1929-1943, 2003.
- [11] Fua P, Plänkers R., and Thalmann D. Tracking and modeling people in video sequences. *Computer Vision and Image Understanding*, Vol. 81, No. 3, pp.285-302, March 2001.
- [12] Gavriila D.M. The visual analysis of human movement: A survey *Computer Vision and Image Understanding*, Vol. 73 No. 1, pp. 82-92, 1999.
- [13] Herda, L., Urtasun, R., and Fua, P. Implicit surface joint limits to constrain video-based motion capture. In *Proc. 8th European Conference on Computer Vision*, Springer, Prague, T. Pajdla and J. Matas, Eds., vol. 3022 of *Lecture Notes in Computer Science*, 405-418, 2004.
- [14] Mikic I, Trivedi M, Hunter E, and Cosman P. Human body model acquisition and tracking using voxel data *International Journal of Computer Vision (IJCV)*, Vol. 53, Nr. 3, pp. 199-223, 2003.
- [15] MoCap-Systems Motion analysis, vicon, simi: Marker based tracking systems. www.motionanalysis.com, www.vicon.com, www.simi.com/en/, June, 2005.
- [16] Moeslund, T., and Granum, E. A survey of computer vision based human motion capture. *Computer Vision and Image Understanding* 81, 3, 231-268, 2001.
- [17] Motion Analysis Corporation www.motionanalysis.com last accessed February 2005.
- [18] Murray R.M., Li Z. and Sastry S.S. A Mathematical Introduction to Robotic Manipulation. *CRC Press*, 1994.
- [19] Osher S. and Sethian J. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, Vol.79, pp. 12-49, 1988.
- [20] Richards J. The measurement of human motion: A comparison of commercially available systems *Human Movement Science*, Vol. 18, pp. 589-602, 1999.
- [21] Rosenhahn B., Kersting U., Smith A., Gurney J., Brox T. and Klette R. A system for marker-less human motion estimation *Pattern Recognition 2005, DAGM*, W. Kropatsch, R. Sablatnig and A. Hanbury (Eds.), Springer-Verlag, Berlin Heidelberg, LNCS 3663, pp. 230-237, Wien, 2005.
- [22] Rosenhahn B., Klette R. and Sommer G. Silhouette based human motion estimation. *Pattern Recognition 2004, 26th DAGM-symposium*, Tübingen, Germany, C.E. Rasmussen, H.H. Bülthoff, M.A. Giese, B. Schölkopf (Eds), LNCS 3175, pp 294-301, 2004, Springer-Verlag Berlin Heidelberg.
- [23] Rosenhahn B. and Klette R. Geometric algebra for pose estimation and surface morphing in human motion estimation *Tenth International Workshop on Combinatorial Image Analysis (IWCIA)*, R. Klette and J. Zunic (Eds.), LNCS 3322, pp. 583-596, 2004, Springer-Verlag Berlin Heidelberg. Auckland, New Zealand,
- [24] Magnenat-Thalmann, N., Seo, H., and Cordier, F. Automatic modeling of virtual humans and body clothing. *Journal of Computer Science and Technology* No. 19 vol. 5, 575-584, 2004.
- [25] Wikipedia Motion capture. http://en.wikipedia.org/wiki/Motion_capture, September, 2005.
- [26] You, L., and Zhang, J. J. Fast generation of 3d deformable moving surfaces. *IEEE Transaction on Systems, Man, and Cybernetics, Part B: Cybernetics* Vol. 33, No. 4, 616-615, 2003.
- [27] Zang Z. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, Vol. 13, No. 2, pp. 119-152, 1999.

Contact

Dr. Bodo Rosenhahn
 Max-Planck-Center Saarbruecken
 Stuhlsatzenhausenweg 85
 D-66123 Saarbruecken, Germany
 Email: rosenhahn@mpi-sb.mpg.de
<http://www.mpi-sb.mpg.de/~rosenhahn/>

Dr. Thomas Brox
 Computer Vision and Pattern Recognition Group
 Institute for Computer Science, University of Bonn
 Roemerstr. 164, 53117 Bonn, Germany
 Email: brox@cs.uni-bonn.de

Dr. Reinhard Klette
 Computer Science Department
 The University of Auckland
 Private Bag 92019, New Zealand
 Email: r.klette@auckland.ac.nz
<http://www.citr.auckland.ac.nz/~reinhard/>

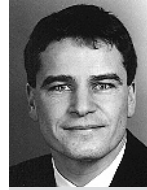
Dr. Uwe G. Kersting, Dr. Andrew W. Smith, Jason K. Gurney
 Department of Sport and Exercise Science
 The University of Auckland, Private Bag 92019
 New Zealand
<http://www2.auckland.ac.nz/tmk/ses/>



Bodo Rosenhahn gained his Ph.D. 2003 at the Cognitive Systems Group, Institute of Computer Science, Christian-Albrechts University Kiel, Germany. From 2003-2005 he was (DFG-)PostDoc at the University of Auckland. Since November 2005 he is senior researcher at the Max-Planck center in Saarbrücken, Germany. He is working on markerless motion capture, human model generation and animation, and image segmentation.



Thomas Brox received his M.Sc. degree in computer engineering from the University of Mannheim, Germany in 2002 and the Ph.D. degree in computer science from the Saarland University, Saarbrücken, Germany in 2005. Since September 2005, he has been working with the Computer Vision and Pattern Recognition Group at the University of Bonn, Germany. His research interests include image segmentation, optic flow estimation, texture analysis, shape modelling, and pose estimation.



Uwe G. Kersting is a lecturer in biomechanics at the University of Auckland (New Zealand). He has been Director of the Biomechanics Laboratory at Tamaki Campus for the previous two years. He has published in various fields ranging from sports biomechanics to tissue adaptation. He has a broad experience in video based movement analysis techniques and is currently running several projects which require three dimensional motion capture.



Andrew W. Smith obtained his PhD in Biomechanics from the University of Waterloo in Canada in 1986. He is currently Senior Lecturer in the Department of Sport and Exercise Science at the University of Auckland, New Zealand. His research interests are in three-dimensional movement analysis, particularly in human gait and balance.



Jason Gurney received the BSc (Hons) and is currently completing honours. He is currently working as researcher and technician in the Biomechanics Lab, University of Auckland.



Reinhard Klette is professor of information technology in the department of computer science at the university of Auckland (New Zealand) and director of CTR Tamaki (Centre for Image Technology and Robotics). His research interests are directed on theoretical and applied subjects in image data computing, pattern recognition, image analysis, and image understanding. He has published about 300 journal and conference papers on different topics within computer science, and books about parallel processing, image processing, shape recovery based on visual information, and digital geometry.