N. Pavešić, H. Niemann, D. Paulus, S. Kovačič (Eds.)

3-D Scene Acquisition, Modeling and Understanding

Proceedings of the Second German-Slovenian Workshop
9. June 1994
Ljubljana, Slovenia

IEEE Slovenia Section

Contents

Image sequence analysis of 3D human arm movements

V. Filova[1], F. Solina[2], J. Lenarčič[3]

[1]IBL-Sistemi, Železna cesta 18, 61000 Ljubljana, Slovenia
[2]Faculty of EE&CS, Tržaška 25, 61000 Ljubljana, Slovenia
[3]Jožef Stefan Institute, Jamova 39, 61000 Ljubljana, Slovenia

Valentina.Filova@ijs.si

Abstract

The paper addresses the problems introduced by the single camera approach to human arm motion study. The arm is considered as a kinematic chain with a known kinematic structure. The projection is assumed to be orthographic. In order to solve the motion reconstruction problem, the structure of the arm has to be determined in each frame. This involves extracting the segments' axes from the projected contours of the arm. A robust algorithm for modeling the segments' projections was developed. The problems of extracting the linked segments' axes are discussed. Assuming that the 2D stick figure of the arm can be identified in each image in the sequence, the equations for motion reconstruction are derived in terms of matrix algebra. The time-varying sequences of the joint angles of the kinematic model are used to get a consistent description of the reconstructed motion.

1 Introduction

The analysis of human arm motion is important for applications in many areas. Understanding of the human arm motion control can help: (a) in robotics to design anthropomorphic mechanisms [10], in rehabilitation to evaluate the injury of the motion mechanism [11], (b) in behavioral neuroscience to study motor disorders caused by some disease [9], (c) in sports for gesture performance measurement [4,14], and (d) in designing new human-computer interfaces [5,17].

A traditional approach to reconstruction of human body motion in 3D is to reconstruct the movement of markers attached to the human body [1,8,22]. The 3D coordinates of the markers are computed by triangulation, using two or three cameras. There are problems with tracking the markers due to obstruction during motion and to the displacement of markers caused by the elasticity of the skin. Consequently, the need for different approaches arises. One way to track motion of the human body is to model the body surface from range data [2,12,16,19,20]. However, the approach involves sophisticated shape modeling, which is time consuming.

In this paper, we propose a single camera approach for studying the human arm movements. Our study is based on the following assumptions: (1) the human arm can be considered as a multi-joint object with a known kinematic model, (2) in each image frame, the 2D positions of the joint points of the kinematic model can be identified from projected contour of the arm, and (3) camera projection is orthographic.

The remainder of the paper is organized as follows. The next Section addresses the problems introduced by the approach. Section 3 gives the mathematical background of the process of fitting segment axes into projected contours of the arm. Some results of extracting the segments' axes are presented in Section 4. Section 5 discuss the problem of estimation the motion of the arm from linked segment axes.

2 Model based recovery from images

The goal of the research is (a) to reconstruct 3D structure and motion of the arm from projected image sequence, (b) to give an interpretation of the overall reconstructed motion, and (c) to predict the future motion from given image sequence. There are two basic approaches to 3D recovery of structure and motion of an object which model is available from projected images: the "3D Euclidean approach" based on geometrical constraints in 3D Euclidean space and the "2D non-Euclidean approach" based on analysis on the the image plane viewed as a 2D non-Euclidean space [13]. The 3D Euclidean approach begins with image features, then backprojects them into the scene, and applies object constraints expressed in terms of 3D Euclidean geometry. In contrast, the 2D non-Euclidean approach begins with object modeling, then projects the model onto the image plane, and defines the image features in terms of the 2D non-Euclidean geometry resulting from the assumed object model.

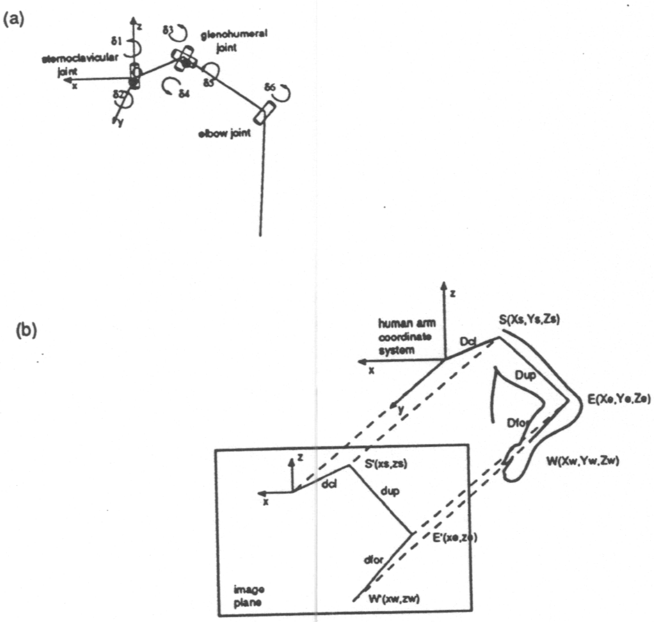

Figure 1: 3D Structure recovery using the kinematic model: (a) kinematic model of the left arm, (b) backprojection of the segments' axes into the scene.

The research presented in the paper is carried out following the first approach. The human arm is considered as a kinematic chain of three rigid objects connected by joints (see Figure 1a) . The kinematic model [1] specifies the lengths of the arm segments, the ranges of the joint angle values and the correlation of joints due to functional connection [10]. The 3D structure of the human arm is defined as 3D position of the joint points of the kinematic model relative to each other. Consequently, the 3D structure recovery from a projected contour of the arm can be done without reconstructing the shape of the arm segments. It suffices to fit an axis into separated arm segments' projections, and then backproject the 2D linked axes into the scene by introducing unknown parameters i.e. the depth of the joint points (see Figure 1b) . Assuming that this can be done in each image in the sequence, the geometrical constraints specified with the kinematic model can be

[1] According to the model developed by Lenarčič [10] and Umek[21] the arm (without the hand) has six revolute degrees of freedom (DOF), two in the sternoclavicular joint, three in the glenohumeral joint and one in the elbow joint. The reference coordinate frame is placed in the sternoclavicular joint.

used to derive the 3D recovery equations in terms of vector calculus and matrix algebra. The 3D motion can be reconstructed as transition between two postures of the arm in space. This requires backprojection of each image in the sequence. The reconstructed 3D motion can be described with the trajectories of the joint points of the kinematic chain or with the time-varying sequences of the joint angles.

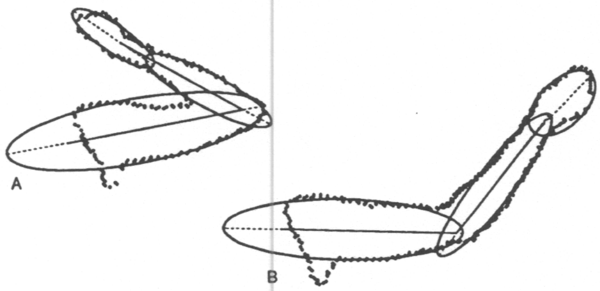

Figure 2: Extracting the segments' axes.

3 Extracting image features

The 3D structure recovery involves fitting 2D linked axes to the projected contours of the arm (see Figure 2). In order to do this the image of the arm should be first segmented according to the model of the arm and then an axis should be fitted into each segment separately. Intersections of the fitted segment axes give the image coordinates of the joint point projections. Taking into consideration the recent results that the segmentation and modeling are better solved in parallel [18], we first map portions of the image data to the model manually using window of interest technique (see Figure 3). Next , we model the contour points with second order curve using nonlinear maximum likelihood estimation (robust M-estimation). At the end, we resegment the data according to the results of the modeling process. The second order curve, which is used as model function, has many good properties: (a) is symmetrical, (b) is described with a small number of parameters, (c) the curve axis give the orientation of the arm segment, and (d) by adding new parameters other forms can be obtained which better fit the shape of the arm segments [12,18,19] .

The algorithm for modeling data is based on iteratively reweighted least square paradigm. The Hampel redescending function is applied for weighting data. The Hampel redescending function has the property that the image data with very large residual error (residual error is deviation of the observed point from the fitted curve) are treated as outliers. The experimental results have shown very high robustness of the algorithm to outliers [6]. Thus, the image points which belong to the forearm are rejected while modeling the upper arm. This enables resegmentation of the image data.

The Section proceeds with brief summary of the mathematical background of the robust estimation process.

3.1 Robust M estimation

A robust M-estimate for the parameter vector $\vec{p} = \{a, b, c, d, e\}$ of the model $\hat{f}(\vec{p}, x, y)$ (1), defined as

$$f(\vec{p}, x, y) = p_1\phi_1 + p_2\phi_2 + p_3\phi_3 + p_4\phi_4 + p_5\phi_5 + 1. \tag{1}$$

where $\{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5\} = \{x^2, xy, y^2, x, y\}$, minimizes the error function $\epsilon(\vec{p})$ (see Equation 2) that sums the deviations $e(x_i, y_i)$ of the observations from the fitted curve (1).

$$\epsilon(\vec{p}) = \sum_{i=1}^{N} \rho(\frac{e(x_i, y_i)}{s}). \tag{2}$$

The parameter s is a known or previously computed scale parameter and ρ is a robust loss function. This is more general than the sum of squared deviations (for the L2 regression problem, we have $\rho(x) = x^2$), or the sum of absolute deviations (L1 regression problem, where we have $\rho(x) = |x|$).

If $\psi(\vec{p}, x, y) = \frac{\partial(\rho(\vec{p}, x, y))}{\partial(\vec{p})}$, then a necessary condition for a minimum of the function $\epsilon(\vec{p})$ is that $\vec{p}$ satisfies

$$\sum_{i=1}^{N} \psi(\frac{e(x_i, y_i)}{s}) \quad \phi_m(x_i, y_i) = 0, \quad m = 1, 2, .., 5. \tag{3}$$

Equation (3) is a system of five nonlinear equations where ψ plays the role of a weighting function. It can be solved iteratively via several different methods, one is *iteratively reweighted least squares* [3]. Before the iteratively reweighted least squares scheme is applied we have to select an appropriate ψ function . We choosed the *Hampel redescending* function [6]. The Hampel M-estimator was implemented in *Mathematica* programming language.

4 Results of extracting joint point positions

The robust M estimation was used to study (a) the influence of outliers on the orientation of segments' axes, and (b) the possibilities of extracting the 2D positions of the joint points as intersections of the fitted segments' axes when the arm is performing 2D and 3D motion.

4.1 Influence of outliers on orientation of segments' axes

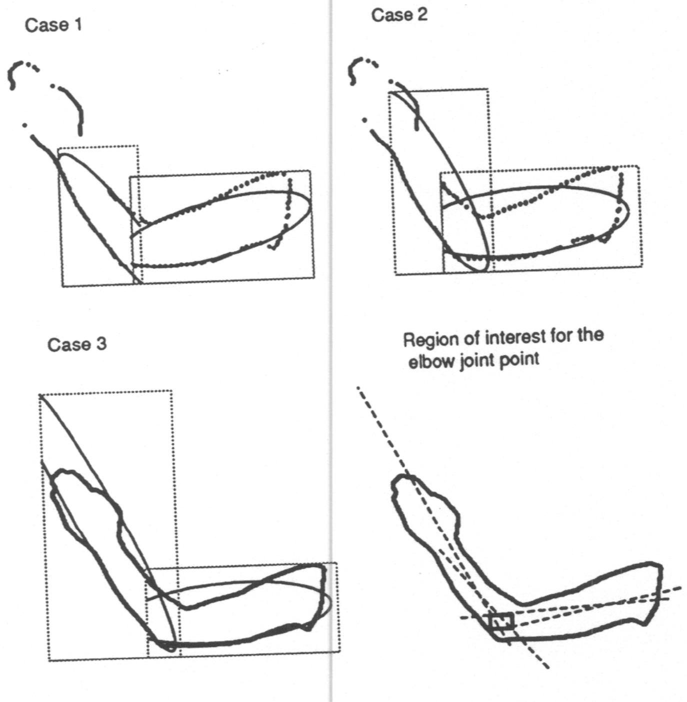

Figure 3: Influence of outliers on orientation of segments' axes.

The experiment was carried out as following. Two sets of data were cut from the projected contour by defining the window of interest manually. The data was assigned to the forearm and the upper arm respectively. The model was fitted to the data using the robust M-estimation algorithm and the segments' axes were extracted. The intersection of the fitted axes gives the possible position of the projected elbow joint points from the kinematic model. Then the windows of interest were widen and new sets of data were obtained. Figure 3 presents the results of fitting the model in three different sets of data. The number of iteration in all the cases was the same (twelve iterations). The initial estimate was obtained with the least square method. The obtained results for the orientation of the segments' axes differ. This is caused by the different quantities of noise in the data and same number of iterations. The results shown that the estimation process in most of the cases entered a minimum and greater number of iteration would not give better results. Due to the bad initial estimate and unsuitability of the model (the projected contours are not perfect ellipses and the shape of the arm change during motion) some points are incorrectly rejected as outliers. As a result we obtain a region where the projection of the elbow joint point probably lies. This tolerance region should be taken into consideration while calculating the 3D posture of the arm segments via 3D coordinates of the joint points.

4.2 Determination of joint point positions

Figure 4 presents a sequence of orthographic projections of the arm performing 3D motion. The projected contours were processed as follows. The image data was mapped to the 3D model manually defining window of interest for the forearm and upper arm in each frame separately. The segments' axes were extracted by modeling the data with second order curves using robust M-estimation [6]. Intersections of the fitted axes give the possible position of the projected joint points from the kinematic model. The obtained results lead to the following conclusions. It is difficult to determine the projection of the shoulder joint point in a separate image. In Figure 4a this could be done by multiple frame analysis. In Figure 4c is necessary to use the dependence among clavicula and upper arm joint angles specified in the kinematic model [21]. In the cases where the elbow joint point projection cannot be extracted as intersection of the forearm axis and the upper arm axis this should be done using multiple frame analysis.

(a)

(b)

(c)

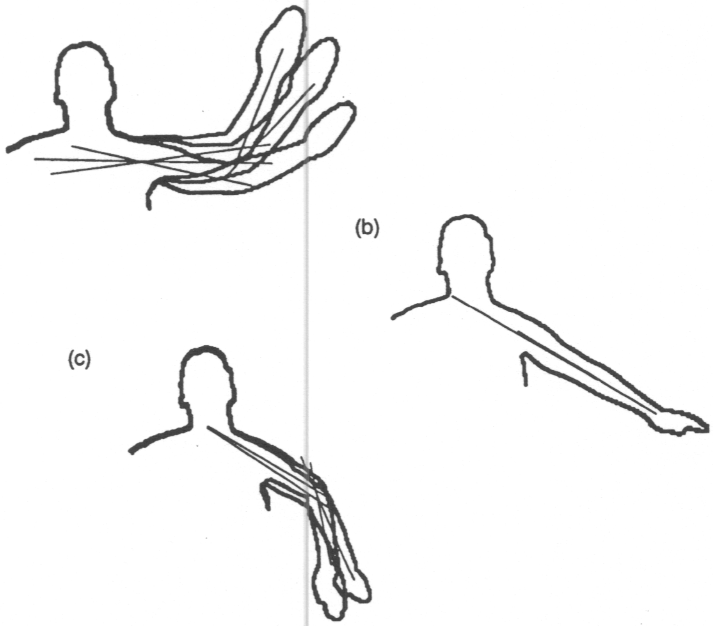

Figure 4: An orthographic projection of a 3D human arm motion.

5 Arm motion reconstruction using the kinematic model

The human arm is assumed a kinematic chain consisting of three objects which kinematic model is known. The kinematic model facilitates the arm motion reconstruction process in many ways. The geometrical constraints specified by the model allow 3D posture recognition. The process of tracking and estimation of multiple motion in the scene is easier. The kinematic model can especially helps when multiple frame analysis is needed due to the problems of occlusion of segments or other problems connected with extraction of joint point coordinates. According to the model introduced in Section 2 the arm has six revolute degrees of freedom (DOF). For each DOF a homogeneous matrix can be derived, which describes the transformation linking the arm segments connected by joints. Using matrix algebra the 3D structure of the arm can be expressed as following:

$$r_{sh} = Rz(\delta_1).Ry(\delta_2).d_1 \tag{4}$$

$$r_{el} = r_{sh} + Rz(\delta_1).Ry(\delta_2).Ry(\delta_3).Rx(\delta_4).Rz(\delta_5).d_2 \tag{5}$$

$$r_{wr} = r_{el} + Rz(\delta_1).Ry(\delta_2).Ry(\delta_3).Rx(\delta_4).Rz(\delta_5).Rx(\delta_6).d_3 \tag{6}$$

5.1 Extracting 3D structure from 2D joint point positions

Due to projection of the arm onto the image plane, the depth information is missing. Using the geometrical constraints in 3D space given with the matrix equations 4, 5 and 6 and the 2D end-points of the arm segments (2D joint point positions), the 3D structure recovery equations can be derived as following. The position in depth of the shoulder joint point is obtained and the position of connected segments are expanded from this point. There are two possible solutions for each body segment, corresponding to a shortening due either to a forward or backward tilt of the segment. Since the arm consists of three segments, there are eight possible solutions for the structure of the arm in space. To resolve the problem, it is assumed that the initial position of the arm is known and the motion is smooth. The 3D recovery equations for the rotation parameters $\delta_1..\delta_6$ are given in Figure 5).

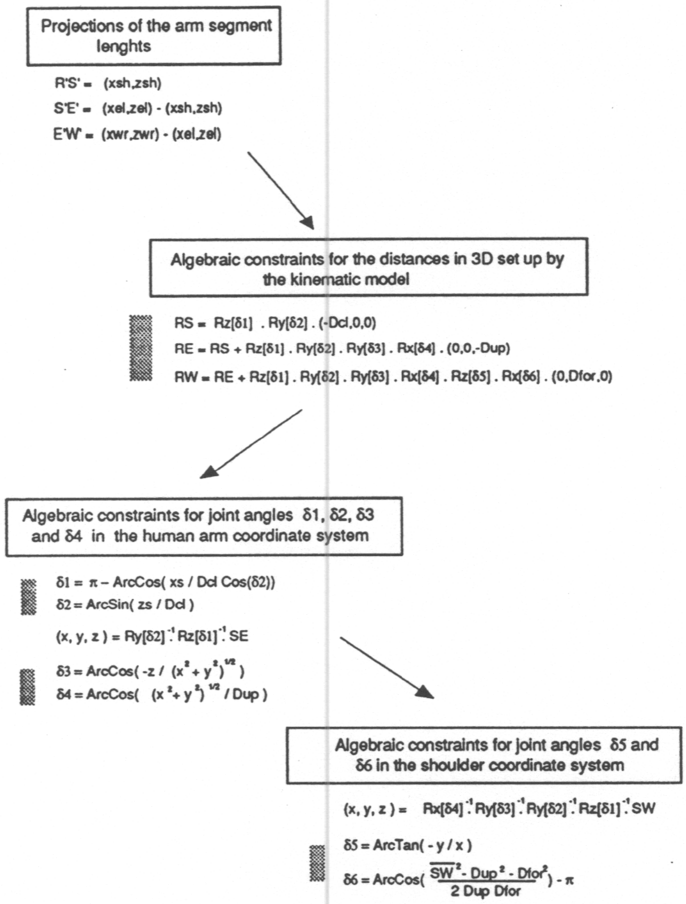

Figure 5: 3D structure recovery constrained with the kinematic model.

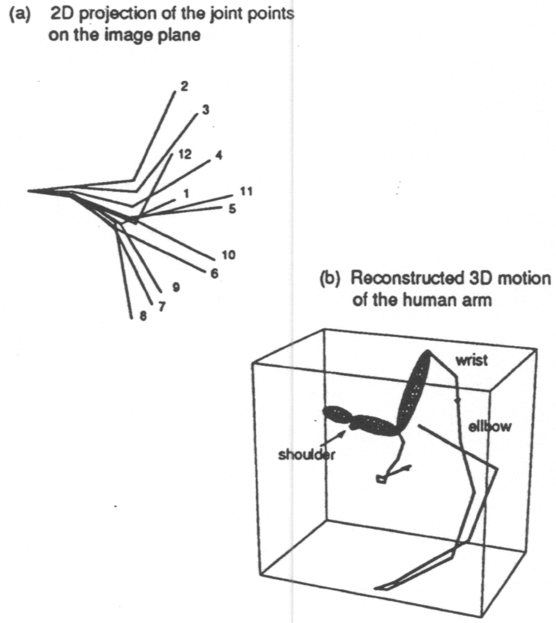

Figure 6: 3D motion of the stick figured human arm presented with: (a) its othographic projections, (b) trajectories of the joint points.

5.2 Experimental results

Once the sequence of orthographic projections of the arm motion is given and the 2D coordinates of the joint points is identified in each image (see Figure 6a), the motion can be derived as transition between two reconstructed postures of the arm in space (see Figure 6b). The reconstructed 3D motion can be also described with the values of the joint angle parameters as functions of time (see Figure 7). As the motion is assumed smooth, the next position of the arm in space and the image features can be predicted.

6 Conclusions

The paper presented an approach to 3D human arm motion and structure recovery from projected image sequence taken under orthographic projection. Because the time evolution of motion is dependent on object structure, the estimation of object motion and

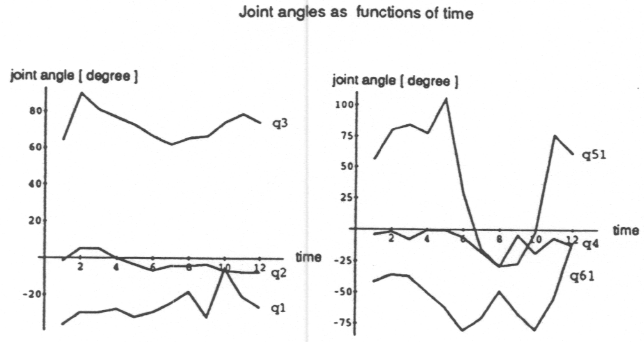

Figure 7: 3D motion described with time-varying joint angles.

structure cannot be separated. As the arm is assumed an articulated object with a known kinematic model, the 3D structure is defined either with the 3D coordinates of the joint points or with the values of the joint angles of the kinematic model. The 3D structure recovery of kinematic chain involves two processes, one is fitting an axis into separated arm segments' projections, and another is backprojection of the 2D linked axes into the scene. In order to extract the segments' axes from projected contours, we developed a robust algorithm for modeling the image data with second order curves. The paper presented some of the results obtained with the algorithm and discuss the problems connected with the extraction of the 2D joint point coordinates. Assuming that the 2D linked axes can be extracted in each image in the sequence, we developed an algorithm for 3D structure and motion recovery. The 3D structure recovery equations were derived in terms of matrix algebra. Both algorithms were implemented in *Mathematica* programming language and tested on real data.

References

[1] C. I. Attwood, G. D. Sullivan, K. D. Baker, "Model-based recognition of human posture using single synthetic images", In *Proc. of the 5th Alvey Vision Conference*, pp. 25-30, 1989

[2] N. I. Badler, B.B. Hunter, Jr. and C. Philips, "Slice display for human models construction with spheres", *Research Report MS-CIS-88-84*, Graphics Lab., Dept. of Computer and Information Science, School of Engineering and Applied Science, U-

niversity of Philadelphia

[3] P. J. Besl, J. B. Birch, L. T. Watson, "Robust window operators", *Proc. of the 2nd International Conference on Computer Vision*, Tampa, Florida, pp. 591-600, IEEE Computer Society Press, Washington.D.C, 1988

[4] O. Kugonič, B. Širok, M. Čoh, F. Trdič. Optimization of shot-put by computer aided visualization. In *Proc. of the 7th International Conference on Mechanics in Medicine and Biology*, Ljubljana, Slovenia - Pörtschach, Austria, 1991

[5] J. Davis, M. Shah, "Recognizing hand gestures", In *Proc. of 3rd European Conference on Computer Vision*, Stockholm, Sweden, pp. 331-340, 1994

[6] V. Filova, F. Solina, J. Lenarčič, "Modeling 2D image data by robust statistics", In *Proc. of the 7th Mediterranean Electrotechnical Conference*, Antalya, Turkey, 1994

[7] V. Filova, F. Solina, J. Lenarčič, "Model-based reconstruction of 3D human arm motion from a monocular image sequence", In *Proc. of the Czech Pattern Recognition Workshop'93*, Temešvár u Pí, Czech Republic, pp. 137-143, 1993

[8] G. Ferrigno, A. Pedotti, "ELITE: A digital dedicated hardware system for movement analysis via real-time TV signal processing", *IEEE Transaction on Biomedical Engineering*, 32: 943-949, 1985

[9] T. Flash, N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model", *Research Report A.I.Memo 786*, Artificial Intelligence Lab., Center for Biological Information Processing, Whitaker Colledge, MIT, 1984

[10] J. Lenarčič, A. Umek, "Experimental evaluation of human arm kinematics", In R. Chatila and G. Hirzinger editors *Experimental Robotics II*, Springer-Verlag, London, 1993

[11] J. Lenarčič, A. Umek, V. Filova, M. Leonardi, F. Solina, "Upper extremity kinematics for studying FES induced movements", In *Proc. The Ljubljana FES Conference*, Ljubljana, 1993

[12] I. A. kakadiaris, D. Metaxas, R. Bajcsy, "Active part-decomposition, shape and motion estimation of articulated objects: A physics-based approach", In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, Washington, pp. 980-984, 1994

[13] K. Kanatani, "3D Euclidean versus 2D non-Euclidean: two approaches to 3D recovery from images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 3, pp. 329-32, 1989

[14] W. Long, Y. Yang, "Log-tracker: An attribute-based approach to tracking human body motion", *Research Report 88-19*, University of Saskatchewan, Saskatchewan, Canada, 1988.

[15] M. J. Mirza, K. L. Boyer, "Performance evaluation of a class of M-estimators for surface parameter estimation in noisy range data", *IEEE Transactions on Robotics and Automation*, Vol. 9, No. 1, pp. 75-85, 1993

[16] A. Pentland, B. Horowitz, "Recovery of nonrigid motion and structure", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13 (7): 730-742, 1991.

[17] J. M. Rehg, T. Kanade, "Visual tracking of high DOF articulated structures: An application to human hand tracking", In *Proc. of 3rd European Conference on Computer Vision*, Stockholm, Sweden, pp. 35-46, 1994

[18] F. Solina, A. Leonardis, A. Macerl, "A direct part-level segmentation of range images using volumetric models", In *Proc. of International Conference on Robotics and Automation*, San Diego, CA, 1994

[19] A. Vidmar, F. Solina, "Recovery of superquadric models from occluding contours", In Reinhard Klette and Walter Kropatsch editors *Theoretical Foundations of Computer Vision*, Mathematical Research, Vol. 69, pp. 227-240, Berlin, 1992. Akademie Verlag

[20] D. Terzopoulos, D. Metaxas, "Dynamic 3D models with local and global deformations: Deformable superquadrics", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 13 (7): 703-714, 1991

[21] A. Umek-Venturini, J. Lenarčič, "Experimental analysis of the elevation of the upper extremity in the frontal plane", In *Proc. of 3rd International Workshop on Robotics in Alpe-Adria Region*, Bled, Slovenia, 1994

[22] VICON System: User Guide, Oxford Metrics Ltd, 1985