

# KINEMATIC SYNTHESIS OF AVATAR SKELETONS FROM VISUAL DATA

Mari Cruz Villa-Uriol, Falko Kuester, Nader Bagherzadeh

*Visualization and Interactive Systems Group,*

*University of California, Irvine, CA 92697*

mvillaur, fkuester, nader@uci.edu

Alba Perez, J. Michael McCarthy

*Robotics and Automation Laboratory,*

*University of California, Irvine, CA 92697*

maperez, jmmccart@uci.edu

**Abstract** The recovery of 3D models from visual data generally results in a geometric skeleton that is a simplification of the shape of the captured figure, called an avatar. In this paper, we introduce a method called hierarchical kinematic synthesis that identifies an articulated skeleton, called kinematic skeleton, which provides a compact representation of the movement of the tracked subject.

In this work, the kinematic skeleton of a human is computed from a finite number of key poses captured from full body articulated movements of an arbitrary subject, and provides location of joints, length and twist angle of links that form the limbs of the 3D avatar. We use an approximation to the human skeleton which consists of five serial chains constructed from revolute and spherical joints. To recover the kinematic skeleton, a hierarchical approximate finite-position synthesis methodology determines the dimensions of these chains limb by limb.

We show that this technique effectively recovers the kinematic skeleton for several synthetically generated datasets, and that the identification of the kinematic skeleton improves pose estimation for 3D data while simplifying the generation of avatar movement.

**Keywords:** Kinematic synthesis, Avatar, Skeleton.

## 1. Introduction

Diverse techniques have been studied for recovery of posture, motion analysis and motion tracking of articulated subjects from video sequences or sets of static images (Kakadiaris and Metaxas, 1998; Bregler and Malik, 1998). This has applications in the study of gesture, biomechanics, anthropometry, human gait analysis, human-computer interaction, realistic computer animation and control of virtual characters.

Body pose recovery methods from images frequently use geometrically defined models of humans as a reference for pose estimation (e.g. Cheung and Kanade, 2000; Weik and Liedtke, 2001; Kakadiaris and Metaxas, 1998).

Previous research has focused on creating a *geometric skeleton* that represents a simplification of the subject’s geometric properties as a set of solids roughly aligned to the medial axis of each limb. In this paper, we determine a *kinematic skeleton*, which is a compact representation of the movement of the subject. The geometric skeleton locates a set of rigid bodies in space, while the kinematic skeleton adds to that the location of joints that allow the movement of the rigid bodies that form the limbs.

Animation of an articulated model to achieve human-like movement is easier using a kinematic skeleton, where configuration changes are obtained by specifying angles at each joint. We find that anthropomorphic movements at the level of a complete avatar are easily obtained with simplified skeletons (for detailed modeling of human joints see for instance Lenarcic, Stanisic and Schearer, 2002).

The presented approach uses *kinematic synthesis* to find the kinematic skeleton from the movement of the figure. In particular, the input data is a series of positions of the limbs of the captured subject. Synthetically generated movements of a simplified human skeleton were used for testing the algorithm. We define the synthesis problem using dual quaternions to represent the chains and spatial displacements (Perez, 2003). The result is a compact formulation which suggests it will be feasible to be used in real time applications.

## 2. Data Acquisition

The actual data consists of synthetically generated movements resembling human motion, represented as a sequence of 4x4 homogeneous matrix transformations for each limb. In the future, video images will be used to obtain the motion information characterizing the subject. For each of the acquired image frames, a set of silhouettes of the actor and their corresponding camera calibration parameters are obtained, as shown in Figure 1. The visual hull (Laurentini et al., 1994) containing the actor is calculated by intersecting the camera frustums, and subdivided into voxels with a resolution defined by the user. This volumetric reconstruction can be described as the maximal shape that generates the same silhouette for every reference view outside of the object’s convex hull.

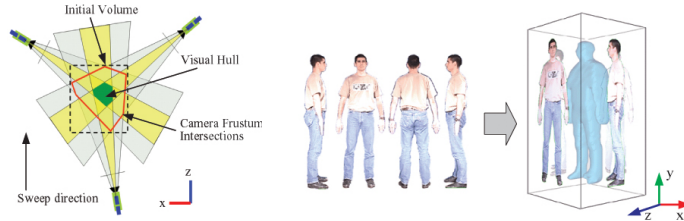


Figure 1. Visual Hull: concept (left) and example obtained from 4 images (right).

### 3. Synthesis of Kinematic Chains

The finite-position synthesis of a kinematic chain seeks to find locations and directions of each of its joint axes from a finite set of positions defined as rigid transformations. In approximate synthesis, the kinematic chain is dimensioned such that its final element passes approximately through this set of positions. Approximate synthesis is used to solve this problem since it is not possible to model the movement of real figures as rigid motions.

In the presented approach, we use the robot *kinematics equations* or *forward kinematics* to define the movement of a skeleton topology (Craig, 1989). We then equate the avatar kinematics equations to the input data for the limb positions from a reference configuration (Mavroidis et al., 2001; Tsai, 1999), formulated using dual quaternion algebra (Perez, 2003). The dimensions of the avatar limbs are obtained by minimizing the distance between the avatar and goal configurations.

#### 3.1 The Kinematics Equations

The kinematics equations of a robot equate the  $4 \times 4$  homogeneous matrix transformation  $[D]$  between the last element of the kinematic chain and the base frame to the sequence of local coordinate transformations along the  $m$  joints of the chain,

$$[D] = [G][Z(\theta_1, d_1)][X(\alpha_{12}, a_{12})][Z(\theta_2, d_2)] \dots [X(\alpha_{m-1, m}, a_{m-1, m})][Z(\theta_m, d_m)][H]. \quad (1)$$

The parameters  $(\theta, d)$  and  $(\alpha, a)$  are known as the *Denavit-Hartenberg parameters*. The transformation  $[G]$  defines the position of the first joint of the chain relative to a fixed frame, and  $[H]$  locates the last limb relative to the last joint.

We use a dual quaternion representation of the displacements instead of the homogeneous matrix representation. In order to create the dual

quaternion kinematics equations, the expression in Eq.(1) is transformed into successive screw displacements by choosing a reference position  $[D_0]$ . Let  $[D_i]$  be the homogeneous matrix describing the transformation from the fixed frame to a moving frame  $M_i$ , then we can compute  $[D_{0i}] = [D_i][D_0]^{-1}$ , and Equation (1) can be transformed into

$$[D_{0i}] = [T(\Delta\theta_1, \mathbf{S}_1)][T(\Delta\theta_2, \mathbf{S}_2)] \dots [T(\Delta\theta_m, \mathbf{S}_m)], \quad (2)$$

where  $\Delta\theta_j = \theta_j^i - \theta_j^0$  or  $\Delta\theta_j = d_j^i - d_j^0$  depending on whether the joint is revolute or prismatic, respectively. The displacements  $[T(\Delta\theta_i, \mathbf{S}_i)]$  define the rotations about and translations along the joint axes  $\mathbf{S}_i$  measured in the fixed frame relative to the reference configuration  $[D_0]$ .

### 3.2 Dual Quaternion Kinematics Equations

The kinematics equations of the serial chain are expressed as elements of the Clifford algebra of dual quaternions. A spatial displacement consisting of a rotation of  $\theta$  and translation of  $d$  around and along a screw axis  $\mathbf{S}$  is written as the dual quaternion  $\hat{S}(\hat{\theta}) = \sin(\frac{\hat{\theta}}{2})\mathbf{S} + \cos(\frac{\hat{\theta}}{2})$ , where  $\hat{\theta} = \theta + \epsilon d$ , and  $\mathbf{S} = \mathbf{S} + \epsilon \mathbf{p} \times \mathbf{S}$  is the expression of the screw axis as a dual vector using Plücker coordinates. The dual unit  $\epsilon$  is defined so that  $\epsilon^2 = 0$  (Bottema and Roth, 1979).

Notice that the dual quaternion expression encodes the same information as the  $4 \times 4$  screw displacement matrix  $[T(\Delta\theta, \mathbf{S})]$ . Thus, we can define the dual quaternion kinematics equations for the serial chain by simply replacing the  $4 \times 4$  matrices in (2) by their dual quaternion equivalents to obtain

$$\hat{D}^i = \hat{S}_1(\Delta\hat{\theta}_1^i) \hat{S}_2(\Delta\hat{\theta}_2^i) \dots \hat{S}_m(\Delta\hat{\theta}_m^i). \quad (3)$$

The Plücker coordinates of the screw axes  $\mathbf{S}_i$ ,  $i = 1, \dots, m$ , are defined in the base frame  $F$ .

### 3.3 Design Equations

For the solution of our problem, we are given a series of positions (location and orientation) of the subject at different time frames. Let the  $n$  positions be defined as the  $4 \times 4$  transforms  $[P_i]$ ,  $i = 1, \dots, n$ . Construct the  $n - 1$  relative transformation matrices  $[P_{1i}] = [P_i][P_1]^{-1}$  and their associated dual quaternions  $\hat{P}_{1i}$ ,  $i = 2, \dots, n$ . Equating the task dual quaternions  $\hat{P}_{1i}$  to the kinematics equations (3), yields the design equations

$$\mathcal{Q}_i : \quad \hat{S}_1(\Delta\hat{\theta}_1^i) \hat{S}_2(\Delta\hat{\theta}_2^i) \dots \hat{S}_m(\Delta\hat{\theta}_m^i) - \hat{P}_{1i} = 0, \quad i = 2, \dots, n. \quad (4)$$

The solution of the total set of design equations provides with the location and orientation of the joint axes in the reference configuration. In addition, the solution provides also with the inverse kinematics, that is, the joint angles needed to reach a given position of the limbs.

## 4. Kinematic Skeleton Synthesis

The synthesis of a kinematic skeleton begins with the definition of the skeleton topology and the input data for each limb. From this information, the design equations for each of the five chains are stated. A hierarchical solving process is used to synthesize an approximated solution, including the successive averaging of solutions to obtain an acceptable final skeleton.

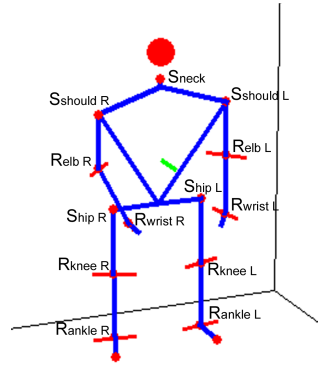


Figure 2. The topology of the kinematic skeleton.

### 4.1 The topology of the kinematic skeleton

Our approach to the kinematic human skeleton consists of five serial chains starting from the center of the chest: a chain for the neck and head, a chain for each of the arms and a chain for each of the legs. The five chains that form the skeleton, presented in Figure 2, are defined as follows. *Head* goes from the chest to the center of the head, and consists of a spherical joint (S) that performs the movements of the neck. *Left arm* goes from the chest to the left hand. It is composed of a spherical joint in the shoulder, a revolute joint in the elbow, and a revolute joint in the wrist. We denote the topology of this chain as S-R-R. *Right arm* has the same S-R-R structure as the left arm. *Left leg* consists of a spherical joint in the hip, a revolute joint in the knee, and a revolute joint in the ankle. It is hence also a S-R-R chain, and the same holds

for *Right leg*. In Figure 2, spherical joints are represented as spheres and revolute joints as lines located at the joints.

## 4.2 Design equations for the kinematic skeleton

The kinematics equations in dual quaternion form for each of the limbs are

$$\begin{aligned}\hat{Q}_{head} &= \hat{S}_{neck}(\theta_{h1}, \theta_{h2}, \theta_{h3}), \\ \hat{Q}_{arm} &= \hat{S}_{should}(\theta_{a1}, \theta_{a2}, \theta_{a3}) \hat{R}_{elb}(\theta_{a4}) \hat{R}_{wrist}(\theta_{a5}), \\ \hat{Q}_{leg} &= \hat{S}_{hip}(\theta_{l1}, \theta_{l2}, \theta_{l3}) \hat{R}_{knee}(\theta_{l4}) \hat{R}_{ankle}(\theta_{l5}).\end{aligned}\quad (5)$$

The dual quaternion expression of rotations of  $\theta$  around an axis of Plücker coordinates  $\mathbf{S} = (s_x, s_y, s_z) + \epsilon(s_x^0, s_y^0, s_z^0)$ , is

$$\hat{S}(\theta) = \begin{Bmatrix} \sin \frac{\theta}{2} s_x \\ \sin \frac{\theta}{2} s_y \\ \sin \frac{\theta}{2} s_z \\ \cos \frac{\theta}{2} \end{Bmatrix} + \epsilon \begin{Bmatrix} \sin \frac{\theta}{2} s_x^0 \\ \sin \frac{\theta}{2} s_y^0 \\ \sin \frac{\theta}{2} s_z^0 \\ 0 \end{Bmatrix}. \quad (6)$$

For the spherical joint, the dual quaternion expression for the displacements is given by

$$\hat{S}(\theta_1, \theta_2, \theta_3) = \begin{Bmatrix} \alpha_1 \mathbf{g}_1 + \alpha_2 \mathbf{g}_2 + \alpha_3 \mathbf{g}_3 \\ \alpha_4 \end{Bmatrix} + \epsilon \begin{Bmatrix} \alpha_1 \mathbf{g}_1^0 + \alpha_2 \mathbf{g}_2^0 + \alpha_3 \mathbf{g}_3^0 \\ 0 \end{Bmatrix}, \quad (7)$$

where  $\mathbf{g}_1 + \epsilon \mathbf{g}_1^0$ ,  $\mathbf{g}_2 + \epsilon \mathbf{g}_2^0$  and  $\mathbf{g}_3 + \epsilon \mathbf{g}_3^0$  are the Plücker coordinates of three perpendicular axes intersecting at a point, and the  $\alpha_i$  appear as combinations of the joint variables,

$$\begin{aligned}\alpha_1 &= \sin \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \cos \frac{\theta_3}{2} + \cos \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \sin \frac{\theta_3}{2}, \\ \alpha_2 &= \cos \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \cos \frac{\theta_3}{2} - \sin \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \sin \frac{\theta_3}{2}, \\ \alpha_3 &= \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \cos \frac{\theta_3}{2} + \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \sin \frac{\theta_3}{2}, \\ \alpha_4 &= \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \cos \frac{\theta_3}{2} - \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \sin \frac{\theta_3}{2}.\end{aligned}\quad (8)$$

In order to create the design equations, we equate the expressions in Eq.(5) to the input data obtained from the images,  $\hat{P}^i$ , for as many time frames  $i$  as are needed to solve for the variables,

$$\begin{aligned}\hat{Q}_{head} &= \hat{P}_{head}^j, \quad j = 2, \dots, n, \\ \hat{Q}_{arm} &= \hat{P}_{arm}^i, \quad \hat{Q}_{leg} = \hat{P}_{leg}^i, \quad i = 2, \dots, m,\end{aligned}\quad (9)$$

and solve individually for each of the chains that form the kinematic skeleton.

### 4.3 Solution process

In order to synthesize the kinematic skeleton efficiently, we use a hierarchical kinematic synthesis approach. It consists of solving the design equations body segment by body segment, starting from the closest to the base or fixed frame and working sequentially forward along the chain. In the process, we define the location of each segment at each instant of time. The resulting equations can be easily solved.

For instance, for the kinematic chain corresponding to one of the arms in Eq.(5), we solve first for the shoulder using data from three frames (two frames and the reference or initial frame), by stating the set of equations

$$\hat{S}_{should}(\theta_{a1}^i, \theta_{a2}^i, \theta_{a3}^i) = \hat{P}_{upper}^i, \quad i = 2, 3, \quad (10)$$

that are solved numerically to obtain the coordinates of the point defining the spherical joint,  $\mathbf{c}$ , and the joint angles  $\theta_{a1}^i, \theta_{a2}^i, \theta_{a3}^i$  that most closely reach positions 2 and 3. These data allow us to create the dual quaternion that approximates the movement of the upper arm, namely  $\tilde{S}_{should}(\theta_{a1}, \theta_{a2}, \theta_{a3})$ .

To solve for the elbow, the input data of the lower arm must be equated to the series of transformations from the base to the lower arm,

$$\hat{S}_{should}(\theta_{a1}^i, \theta_{a2}^i, \theta_{a3}^i) \hat{R}_{elb}(\theta_{a4}^i) = \hat{P}_{lower}^i, \quad i = 2, \dots, n. \quad (11)$$

To isolate the dual quaternion corresponding to the elbow, Eq. (11) is left-multiplied by the inverse of the solution for the shoulder using the dual quaternion conjugate,

$$\hat{R}_{elb}(\theta_{a4}^i) = \tilde{S}_{should}^*(\theta_{a1}^i, \theta_{a2}^i, \theta_{a3}^i) \hat{P}_{lower}^i, \quad i = 2, \dots, n, \quad (12)$$

where the right part of the equality consists of known data, and  $n = 2$  for the case of a revolute joint. We obtain the Plücker coordinates of the joint axis and the joint rotation for the elbow, which can be used to create the dual quaternion of the joint  $\tilde{R}_{elb}(\theta_{a4})$ .

We use data from the movement of the hand to solve for the last joint of the chain. The design equations for the wrist are, after pre-multiplying by the solved joints,

$$\hat{R}_{wrist}(\theta_{a5}^i) = \tilde{R}_{elb}^*(\theta_{a4}^i) \tilde{S}_{should}^*(\theta_{a1}^i, \theta_{a2}^i, \theta_{a3}^i) \hat{P}_{hand}^i, \quad i = 2, \dots, n, \quad (13)$$

where again,  $n = 2$  and we obtain the solution  $\tilde{R}_{wrist}(\theta_{a5})$ .

A similar operation is performed independently for each of the five chains that define the kinematic skeleton.

To obtain a kinematic skeleton with good accuracy, we need to repeat this process several times, due both to the lack of accuracy of the input data, and the disparity between a model made of rigid rotations and the real movement of the human body. Also the input data has to be general enough to show the movement of all joints. We perform the synthesis for a number of steps and average the solutions for each joint, until the error between solutions is small enough, according to the user needs.

#### 4.4 Minimization and measuring

Approximate synthesis is performed by minimizing the distance between the position obtained from the kinematics equations and the input data in Eq.(9). The derivative of the 8-dimensional distance vector,

$$d_{norm} = (\hat{Q}_{arm} - \hat{P}_{arm}^i) \cdot (\hat{Q}_{arm} - \hat{P}_{arm}^i) \quad (14)$$

is set equal to zero to create a system of equations

$$\left[ \frac{\partial Q_{arm}}{\partial s_k} \right]_i (\hat{Q}_{arm} - \hat{P}_{arm}^i) = 0 \quad (15)$$

that are solved for the parameters  $s_k$  of the joints.

The length of the difference vector is a metric for spatial displacements that, as any other metric that we can define, is not bi-invariant, see Martinez and Duffy, 1995. However, since the reference frame is defined in the center of the body, the displacements of all five chains are close to the standard dimensions of the human body, thus the reference frame error can be bounded. We expect to explore this in more detail with real subject data.

### 5. Results

The present results are based on synthetically generated data and are used to test the procedure. In Figure 3 we present the results obtained for a specific subject in 8 different poses, extracted from a sequence of 80 frames, which are the input data for the synthesis method. The first row shows the input synthetic data at frames  $i = 41$  to 48, the second row the corresponding recovered pose of the synthesized skeletons approximating the task configuration and the third row shows the kinematic skeleton at the reference position. We can see that a good match is obtained when using synthetic data. For this dataset, 15 frames were needed to obtain a good approximation of the whole skeleton.



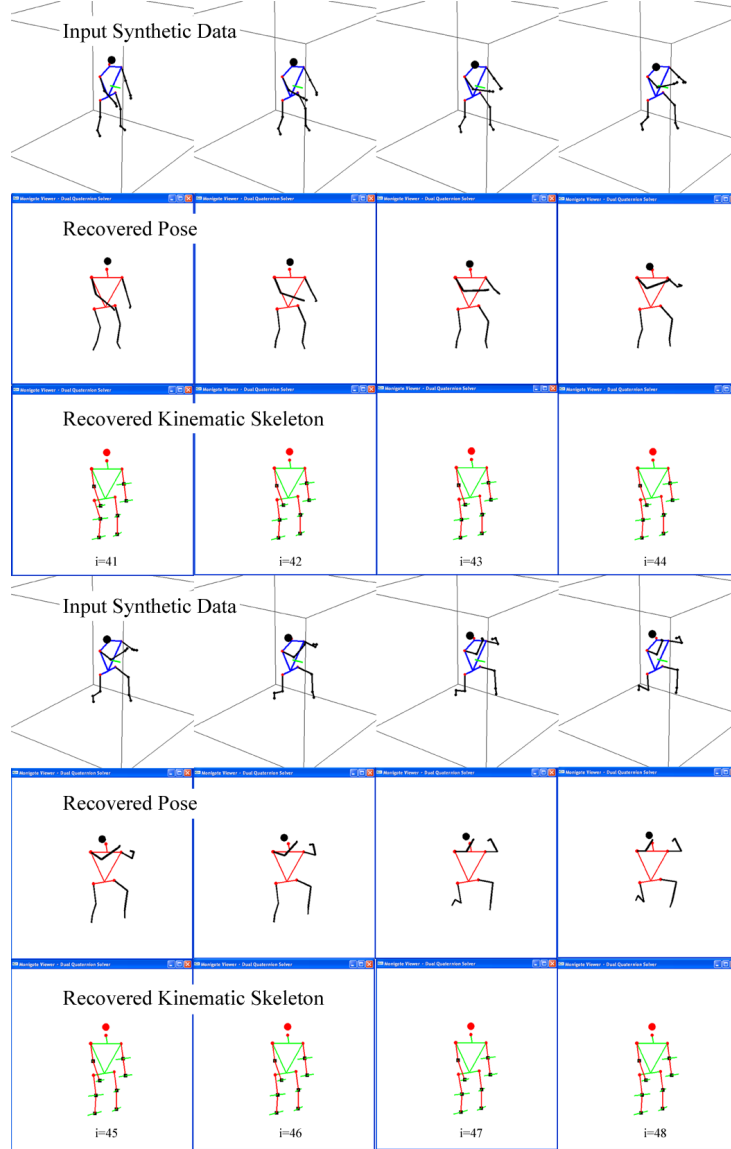


Figure 3. Synthesized kinematic skeleton at frames  $i = 41$  to 48.

## 6. Conclusions

This paper presents a new method for determining a kinematic skeleton and poses of an arbitrary human subject. Finite-position kinematic synthesis is used to fit a kinematic skeleton defined in dual quaternion

coordinates to the data that defines movement of the subject. A major advantage of this approach is that it accepts arbitrary poses of human subjects. In addition, this technique eliminates the need for manual intervention commonly found in geometric skeleton identification methods.

## 7. Acknowledgements

This research was supported in part by the National Science Foundation under contracts CCR-0083080, EIA-0203528 and National Science Foundation award DMII 0218285.

## References

- Bottema, O. and Roth, B., *Theoretical Kinematics*. NY: North Holland Press, 1979.
- Bregler, C. and Malik, J. "Tracking people with twists and exponential maps", in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8–15, 1998.
- Cheung, K. M., Kanade, T., Bouguet, J.-Y. and Holler, M., "A real time system for robust 3d voxel reconstruction of human motions", in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 714–720, 2000.
- Craig, J., *Introduction to Robotics, Mechanics and Control*, Ed. Addison Wesley Publ. Co, 1989.
- Kakadiaris, I. A. and Metaxas, D. N., "Three-dimensional human body model acquisition from multiple views", *International Journal of Computer Vision*, vol. 30, no. 3, pp. 191–218, 1998.
- Laurentini, A., "The visual hull concept for silhouette-based image understanding", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 150–162, 1994.
- Lenarčič, J., Stanišić, M. and Schearer, E., "Humanoid Humeral Pointing Kinematics", *Advances in Robot Kinematics*, eds. J. Lenarcic and F. Thomas, Kluwer Academic Publishers, 2002.
- Martinez, J. M. R. and Duffy, J., "On the metrics of rigid body displacements for infinite and finite bodies", *ASME Journal of Mechanical Design*, vol. 117, pp. 41–47, 1995.
- Mavroidis, C., Lee, E. and Alam, M., "A new polynomial solution to the geometric design problem of spatial rr robot manipulators using the denavit-hartenberg parameters", *ASME J. Mechanical Design*, vol. 123, no. 1, pp. 58–67, 2001.
- Perez, A., "Dual quaternion synthesis of constrained robotic systems", Ph.D. dissertation, Department of Mechanical and Aerospace Engineering, University of California, Irvine, September 2003.
- Tsai, L. W., *Robot Analysis: The Mechanics of Serial and Parallel Manipulators*. New York, NY: John Wiley and Sons, 1999.
- Weik, S. and Liedtke, C.-E., "Hierarchical 3d pose estimation for articulated human body models from a sequence of volume data", in *International Workshop Robot Vision*, ser. Lecture Notes in Computer Science, vol. 1998. Springer-Verlag, pp. 27–34, 2001.