THREE-DIMENSIONAL RECONSTRUCTION OF HUMAN MOTION BASED ON IMAGES FROM A SINGLE CAMERA

Gabriel A. D. Lopes, José A. O. Costa, Jorge A. C. Ambrósio, João Abrantes, Universidade Técnica de Lisboa, Portugal

INTRODUCTION: The use of the photograms of two or more calibrated cameras is well accepted by the biomechanics community as the minimum requirement to perform the spatial reconstruction of human motion (Allard et al. 1995). The projection of the spatial coordinates of a given anatomical point of the human body is described by two equations. Then, two or more cameras are necessary to obtain at least three equations required to get the original coordinates of the point (Abdel-Aziz and Karara, 1971). Regardless of the number of cameras used in the reconstruction, the kinematic characteristics of the biomechanical model are not generally taken into account during the reconstruction. The kinematic constraints of a biomechanical model cannot be guaranteed and the data must be processed before being used.

A biomechanical model of 16 segments, based on that by Silva, et al. (1997), but with body extremities, is used to support the motion reconstruction. The dependency between the Cartesian coordinates of the points describing each component of the biomechanical model is represented by a set of nonlinear kinematic equations. These equations are used to substitute the linear equations representing the projections of the anatomical points with the second camera.

As the objective of the work is to show that the human motion reconstruction can be mathematically obtained from the images of a single camera, in what follows it is assumed that the dimensions of the biomechanical segments and the projected coordinates of the anatomical points are known and that the camera is calibrated. The reference motion used to present the methodology corresponds to an athlete subjected to an heavy side impact at the upper torso level.

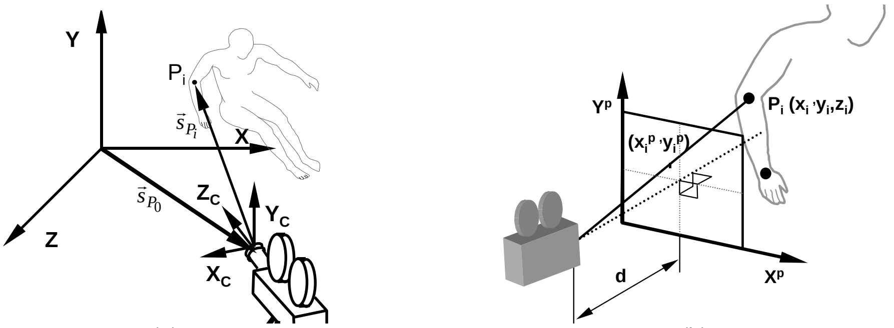
Direct Linear Transformations and Biomechanical Models

The projection of a three dimensional object onto a two dimensional space, described in Figure 1, requires the application of successive transformations, known in computer graphics as virtual camera (Hearn and Baker, 1994). The relation between the projected (x_i^P,y_i^P) and the original spatial coordinates (x_i,y_i,z_i) of P is

$$\begin{aligned} x_i^p &= -S_x - S_s d\, \frac{x_i\left(e_0^2+e_1^2-\frac{1}{2}\right)+y_i\left(e_1e_2-e_0e_3\right)+z_i\left(e_0e_2+e_1e_3\right)-\frac{x_0}{2}}{x_i\left(e_1e_3-e_0e_2\right)+y_i\left(e_0e_1+e_2e_3\right)+z_i\left(e_0^2+e_3^2-\frac{1}{2}\right)-\frac{z_0+d}{2}} \\ y_i^p &= -S_y - S_s d\, \frac{x_i\left(e_1e_2+e_0e_3\right)+y_i\left(e_0^2+e_2^2-\frac{1}{2}\right)+z_i\left(e_2e_3-e_0e_1\right)-\frac{y_0}{2}}{x_i\left(e_1e_3-e_0e_2\right)+y_i\left(e_0e_1+e_2e_3\right)+z_i\left(e_0^2+e_3^2-\frac{1}{2}\right)-\frac{z_0+d}{2}} \end{aligned} \tag{1}$$

which are similar to the DLT equations written as (Abdel-Aziz and Karara, 1971)

$$\begin{cases} a_i x_i + b_i y_i + c_i z_i + d_i = 0 \\ e_i x_i + f_i y_i + g_i z_i + h_i = 0 \end{cases} \tag{2}$$



(a) (b)

Figure 1 a) Camera location and orientation; b) Point projection

The DLT parameters a_i, b_i, …, h_i are related with the visualization parameters, which are the camera position (x_0,y_0,z_0), camera orientation (e_0,e_1,e_2,e_3), focal distance (d), scale variation (S_s) and film origin shift (S_x, S_y).
A 16 segment biomechanical model, shown in Figure 2, is used to support the spatial motion reconstruction. To define its full motion, i.e., position and orientation of its rigid segments, it is necessary to reconstruct the spatial positions of 22 anatomical points defined at the articulations (Celigueta, 1996). The distance between two points, for a biomechanical segment, must remain constant throughout the motion

$$\mathbf{r}_{ij}^T \mathbf{r}_{ij} - d_{ij}^2 = 0 \tag{3}$$

where the vector connecting P_i to P_j is denoted by $\boldsymbol{r}_{ij}$ and its length by d_{ij}. For the complete biomechanical model 22 of these independent kinematic constraints are defined. This system of nonlinear algebraic equations, together with the system of 44 linear equations representing the single camera projections, forms the set of equations required to obtain the 66 spatial coordinates of all anatomical points.

Motion Reconstruction

For a single frame, there are up to $2\text{x}10^6$ possible solutions for the nonlinear system, and the selection of the most probable position of the biomechanical model is

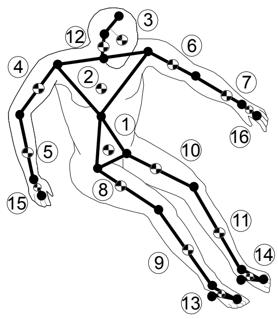


Figure 2. Biomechanical model used in the reconstruction

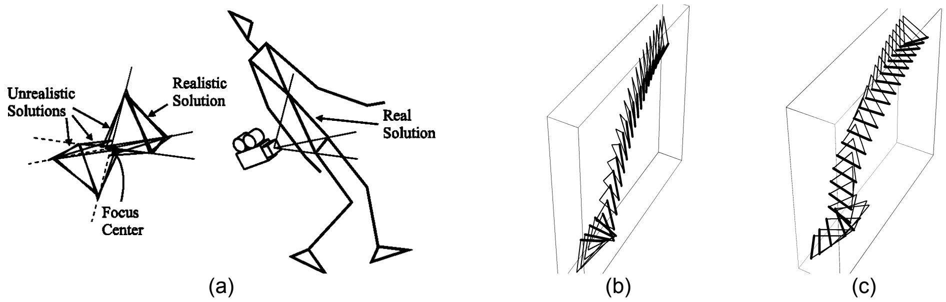


Figure 3 a) For a single frame; b) Correct branch; c) Degenerated branch

basically impossible. The lower-torso is a segment represented by a triangle that has a sub-set of 9 equations involving coordinates of 3 anatomical points. For a single frame this system of equations has 8 solutions, shown graphically in Figure 3a. The solutions are symmetric with respect to the focal point, and due to their unfeasible physical location, 4 are automatically discarded. For each following frame, up to 4 more solutions are obtained and up to 4^n branches of motion are possible for the lower torso triangle. In order to keep the selection process under control most of the branches are eliminated from the calculations at an early phase of the reconstruction process. The most feasible branches of motion must correspond to the minimum of a cost function associated for instance with motion smoothness.

Let a branch of solutions corresponding to the lowest cost function for the first few frames be chosen. The reconstruction of the lower torso position for the next frame can now be supported using linear extrapolation to approximate its position and correcting it with the analytical solution of the nonlinear equations. If two branches of motion have similar low cost functions and no choice between them is reasonable for a small number of frames, their motion reconstruction continues in parallel until a branch has a lower cost function or the process cannot proceed.

Each segment adjacent to the lower torso, is represented by two anatomical points, one of which is already known. The position of the second point must be not only on the surface of a sphere centered in the known point, but must also lie on a straight line that passes by the focus of the camera and by the projected position of the point. The resulting equation has, at most, two solutions for the spatial position of P_j. If the equation lacks any real solution, then the branch in which the position of point P_k is included is not feasible. For the branches of solutions that are maintained in a first instance, the reconstruction of the adjacent segments is attempted. The whole branch is discarded if a solution cannot be obtained. In some

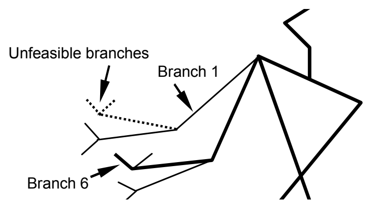
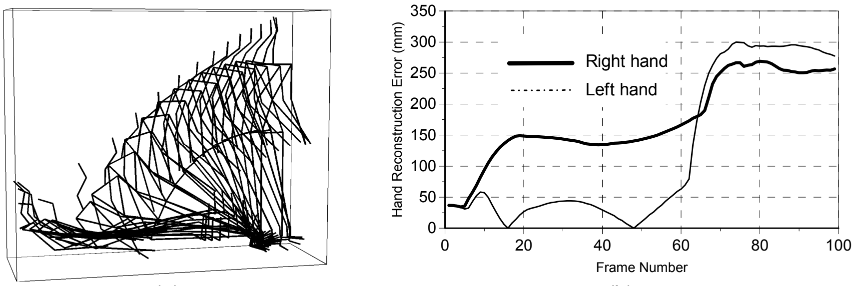


Figure 4 Multiple solutions for the position of the body extremities



(a) (b)

Figure 5 a) Reconstructed motion; b) error between reconstructed and real motion

situations, as illustrated in Figure 4, it may happen that more than one branch of feasible solutions has to be maintained until the end of the reconstruction.
The spatial motion reconstruction of the biomechanical model with the sequence of frames obtained from a single calibrated camera leads to the result presented in Figure 5a. The error between the reconstructed and the original motion for a given anatomical point is evaluated as the distance between its reconstructed and original positions. For biomechanical segments representing the torso, the reconstruction errors are meaningless. Large errors occur for terminal segments defined by two points only, in particular for the hands.

CONCLUSIONS: A methodology for the spatial reconstruction of human motion using a single stationary camera has been described and demonstrated. This is achieved by associating the kinematic constraint equations of a biomechanical model to the linear equations representing the projections of the anatomical points. Due to the multiple solutions of this system of equations, a selection strategy based on physical characteristics of the motion was devised and the complete reconstruction process has been fully automated. These results clearly show that it is mathematically possible to reconstruct motion based on the images from a single camera. Before the methodology can be applied in practice, procedures based on images of a single camera must be devised to: calibrate the camera, deal with digitalization errors, identify the length of the anatomical segments.

Acknowledgements
Funding for this research was provided by Fundação para a Ciência e Tecnologia under the project PBICT/P/CEG/2348/95. The contribution of Mr. Miguel Silva and Mr. António Veloso to this work with the data used is gratefully acknowledged.

REFERENCES:

Abdel-Aziz, Y. I., Karara, H. M. (1971). Direct Linear Transformation from Coparator Coordinates into Object Space Coordinates. In *Proc. ASP/UI Symposium on Close Range Photogrammetry* (pp. 1-18). Falls Church, VA: American Society of Photogrammetry.
Allard, P., Stokes, I. A. F., Blanchi, J.-P. (1995). Three-Dimensional Analysis of Human Movement. Champaign, Ill.: Human Kinetics.
Celigueta, J. (1996). Multibody Simulation of the Human Body Motion in Sports. In J. Abrantes (Ed.), *Proceedings of the XIV International Symposium on Biomechanics in Sports* (pp. 81-94). Lisboa: Edições FMH.
Hearn, D., Baker, P. (1994). Computer Graphics. 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
Silva, M. P. T., Ambrósio, J. A. C., Pereira, M. S. (1997). Biomechanical Model with Joint Resistance for Impact Simulation. *Multibody System Dynamics* **1**, 65-84.