Survey of 3D Human Body Representations

Angel D. Sappa
Computer Vision Center, Spain

Niki Aifanti
Informatics & Telematics Institute, Greece

Sotiris Malassiotis
Informatics & Telematics Institute, Greece

Nikos Grammalidis
Informatics & Telematics Institute, Greece

INTRODUCTION

The problem of human body modeling was initially tackled to solve applications related to the film industry or computer games within the computer graphics (CG) community. Since then, several different tools were developed for editing and animating 3D digital body models. Although at the beginning most of those tools were devised within the computer graphics community, nowadays a lot of work proceeds from the computer vision (CV) community. In spite of this overlapped interest, there is a considerable difference between CG and CV human body model (HBM) applications. The first one pursues realistic models of both human body geometry and its associated motion. On the contrary, CV seeks more of an efficient than an accurate model for applications such as intelligent video surveillance, motion analysis, telepresence, 3D video sequence processing, and coding.

Current work is focused on vision-based human body modeling systems. This overview will present some of the techniques proposed in the bibliography, together with their advantages or disadvantages. The outline of this work is as follows. First, geometrical primitives and mathematical formalism, used for 3D model representation, are addressed. Next, a brief description of standards used for coding HBMs is given. Finally, a section with future trends and conclusion is introduced.

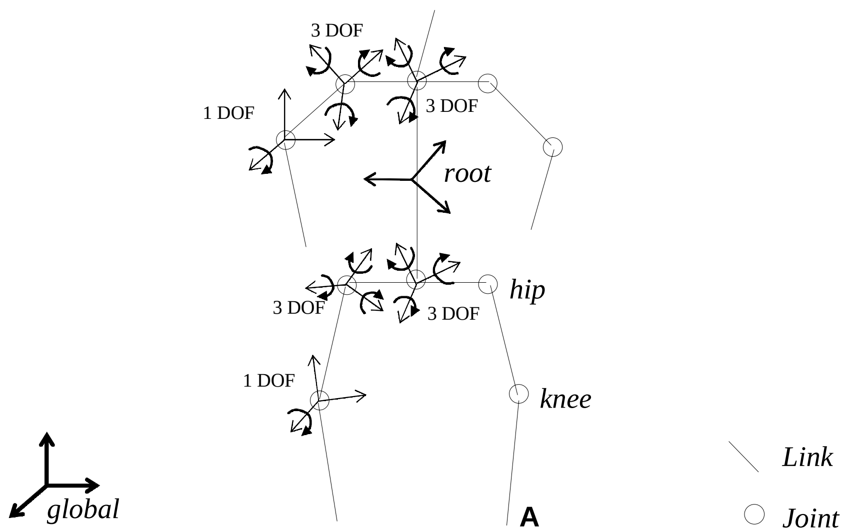
3D HUMAN BODY REPRESENTATIONS

Modeling a human body implies firstly the definition of an articulated 3D structure, in order to represent the human body biomechanical features. Secondly, it involves the choice of an appropriate mathematical model to govern the movements of that articulated structure.

Several 3D articulated representations and mathematical formalisms have been proposed in the literature to model both the structure and movements of a human body (Green & Guan, 2004). Generally, a HBM is represented as a chain of rigid bodies, called *links*, interconnected to one another by *joints*. Links can be represented by means of sticks (Yoo, Nixon & Harris, 2002; Taylor, 2000), polyhedron (Saito & Hoshino, 2001), generalized cylinders (Sidenbladh, Black & Sigal, 2002), or superquadrics (Marzani, Calais & Legrand, 2001). A joint interconnects two links by means of rotational motions about the axes. The number of independent rotation parameters will define the *degrees of freedom* (DOF) associated with a given joint. Figure 1 presents an illustration of an articulated model defined by 12 links (sticks) and 10 joints. Other HBM representations, which do not follow the aforementioned links-and-joints philosophy, have been also proposed in the literature to tackle specific applications. For example, Douros, Dekker and Buxton (1999) present a technique to represent HBMs as single entities by means of smooth surfaces or polygonal meshes. This kind of representation is only useful as a rigid description of the human body. On the contrary, Plänkers and Fua (2003) and Aubel, Boulic, and Thalmann (2000) present a framework that retains an articulated structure represented by sticks, but replaces the simple geometric primitives by soft objects. The result of this soft surface representation is a realistic model where body parts such as chest, abdomen, or biceps muscles are well modeled.

The simplest 3D articulated structure is a stick representation with no associated volume or surface (Liebowitz & Carlsson, 2001). Planar 2D representations, such as cardboard models, have also been widely used (Huang & Huang, 2002). However volumetric representations are

Figure 1. Stick representation of an articulated model defined by 22 DOF



preferred when more realistic models need to be generated. In other words, there is a trade-off between accuracy of representation and complexity. The utilized models should be quite realistic, but they should have a low number of parameters in order to be processed in real-time. Table 1 presents a summary of some of the approaches followed in the literature.

Each of the aforementioned geometrical structures is complemented by means of a motion model that governs its movements (Rohr, 1997); the objective is that the full body performs realistic movements. There is a wide variety of ways to mathematically model articulated systems from a kinematics and dynamics point of view. A mathematical model will include the parameters that describe the links as well as information about the constraints associated with each joint. A model that only includes this information is called a *kinematics model* and describes the possible static states of a system. The state vector of a kinematics model consists of the model state and the model parameters. A system in motion is modeled when the dynamics of the system are modeled as well. A *dynamics model* describes the state evolution of the system over time. In a dynamics model the state vector includes linear and angular velocities as well as position.

After selecting an appropriate model for a particular application, it is necessary to develop a concise mathematical formulation for a general solution to the kinematics and dynamics problems, which are non-linear problems. Different formalisms have been proposed in order to assign local reference frames to the links. The simplest approach is to introduce joint hierarchies formed by independent articulation of one DOF, described in terms of Euler angles. Hence, the body posture is synthesized by concatenating the transformation matrices associated with the joints, starting from the root. In order to illustrate this notation, let us express the coordinates of point **A** in the global reference frame associated with the root of the model (see Figure 1):

$$\mathbf{A}_{global} = Trans_{root\text{-}global} \; x \; Trans_{hip\text{-}root} \; x \; Trans_{knee\text{-}hip} \; x \; \mathbf{A}_{knee}$$

where $\mathbf{A}_{knee}$ represents the coordinates of points **A** relative to the local reference frame placed in the *knee-joint*; $Trans_{i\text{-}j}$ is the corresponding transformation matrix to express reference frame *i* in reference frame *j;* this matrix is defined as:

$$Trans_{i-j} = \left[\frac{R \mid T}{0 \mid 1}\right] = \begin{bmatrix} C\phi C\theta C\psi - S\phi S\psi & -C\phi C\theta S\psi - S\phi C\psi & C\phi S\theta & t_x \\ S\phi C\theta C\psi + C\phi S\psi & -S\phi C\theta S\psi + C\phi C\psi & S\phi S\theta & t_y \\ -S\theta C\psi & S\theta S\psi & C\theta & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

C and *S* represent the cosine and sine respectively, and (ϕ, θ, ω) are the Euler angles. This kind of matrix concatenation can be used to express every body part in the body global reference frame.

3D HUMAN BODY CODING STANDARDS

In order to animate or interchange HBMs, a standard representation is required. Related standards, such as

Related Content