Synthetic Human Movement: to Mimic and Deviate

1D−6

Douglas Dow Sudhanshu K. Semwal
Matsushita Electric, Central Research Lab

Introduction

Modeling human movement is a major challenge in computer graphics. Unlike simple objects such as a ball, the human body has many degrees of freedom and nonlinear interdependencies. Its motion is difficult to model. The two main methods used to determine human motion are key-framing and algorithmic.

In key-framing the body's positions and parameters are specified for only important (key) frames. An interpolation method determines the other inbetween frames. This method is tedious, relying fully on the skill of an animator. In the algorithmic approach, the motion is algorithmically determined. Such algorithms are not only difficult to come up with, they typically require much processing time.

As an example of how hard these motion algorithms are, consider walking. Years of work by many research teams have gone into developing walking algorithms. Walking is complex, yet compared with the countless number of nonperiodic, nonlinear motions that humans do, it is a small task. When will the algorithms be developed for the motions of dance, boxing, hockey, a tennis serve, a volleyball spike, or a child playing with a dog?

Mimic Life

A fundamentally different way to determine movement is to have the synthetic actor mimic a real human who is moving. A human motion is detected and stored as 3D motion data. It is then used to move the synthetic actor. This method builds upon research in model based coding and virtual reality. For instance, the data glove (used in virtual reality) measures 3D information, which can be used to draw a synthetic hand.

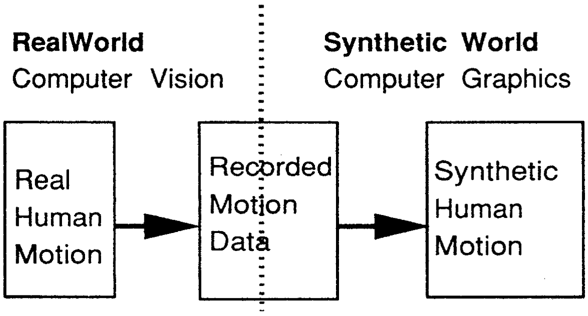

Fig. 1: Mimicing motion of real human.

Fig. 1 shows the method we are using to mimic real motion. Computer vision algorithms are used to extract 3D information of human motion from the real world. That information is used by computer graphics algorithms in the synthetic world to move and shape the synthetic actor.

Part I Computer Vision:
Analyzing the Real World

The 3D joint position information needs to be extracted from real human motion.

A computer vision technique to do this is called rotoscoping [1,2]. It uses a minimum of two cameras at orthogonal angles to obtain the 3D joint position information. The recording of the human's joint positions for every frame is a difficult computer vision task. It can be simplified by using color coded markers at the joint positions of the human. The marker's color value would identify the particular joint that it relates to.

Moreover, computer vision is not the only way to obtain 3D joint position information [1]. Other methods include using goniometers, time-multiplexed light-emitting diodes, or a body suit (similar to the VPL data glove).

Since we have not yet developed a practical computer vision technique to detect the 3D joint information, we simulated the motion extraction by manually pointing a mouse cursor at each joint for each frame to record its (x,y) position.

Between the joint positions, body segment surfaces need to be created. If their shape is not symmetrical, orientation information is needed. Fig. 2 demonstrates how this is done. Each joint has an orientation vector. The elbow joint is specified by position **p1** and orientation vector **o1**. **o1** lies in the plane of the lines **l1** and **l2**, cutting that angle in half. **o1** indicates the "top" of the forearm. The orientation vector is a point of reference for the algorithm that positions and rotates the forearm so that it is in correct alignment with the rest of the body.

人体動作合成：模倣と修飾
ダグラス ダウ、シュダンシュ セマール
松下電器産業株式会社 中央研究所

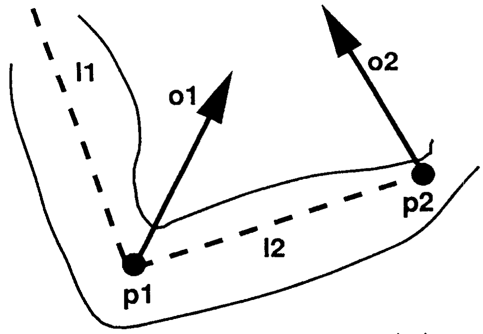

Fig. 2: Joint Data, 3D position and orientation vector

Part II Computer Graphics: Creating the Synthetic World

Information obtained from the real world is passed into the synthetic world. Each joint contains its 3D position and the orientation vector. This is used in the synthetic world to move the synthetic human.

Fig. 3 shows the process. The human's position and orientation data is obtained from the video by using the known skeleton structure. Based on an animation script (that specifies position, speed, etc.) the raw joint positions are manipulated. Then surface models are used to build the surface upon the correctly placed skeleton for this frame.

If only the motion information is extracted from the video, the computer graphics program can have much flexibility during the normalize and manipulate step. It can make the synthetic actor mimic the recorded motion, or deviate from it in a controlled way.

In Fig. 3, the normalize and manipulate step can easily make certain deviations, or changes, from the recorded motion. These deviations include modifying the size, position, speed, length proportions of the different body segments, and the abilility to cut and paste different motions together. In this way new synthetic motions are created. For example, a real human could walk across the floor of a room. The 3D motion would be recorded. Then the synthetic actor could walk in the same way, but it could be a demon, red in color, with horns out of its head, walking through the abyss of hell. Extracting only the 3D motion, gives the computer graphics program great flexibility and control.

Conclusion

Motion of a synthetic actor is usually specified by either key-frames or an algorithmic approach. These approaches require much human effort to generate realistic motions. Rotoscoping promises to be a way to generate new motions faster. It generates motion by mimicing the motion of real humans, and allowing controlled deviations from that motion.

Our method of detecting human motions from video, normalizing and manipulating the detected motion, and building a body surface over the generated skeleton allows an animator to generate complex human motions.

References

[1] Calvert T., "Composition of Realistic Animation Sequences for Multiple Human Figures", *Making Them Move*, Morgan Kaufmann Publishers, San Mateo, CA, USA, 1991, pp 35–50.

[2] Unuma M., Takeuchi R., "A Method of Human Motion Generation for CG", (in Japanese), NICOGRAPH, 1990, pp. 136–43.

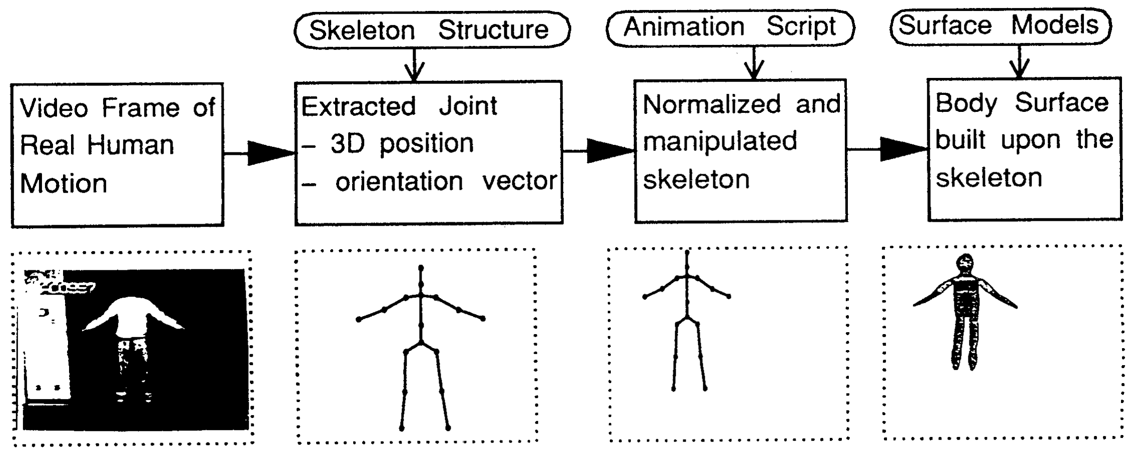

Fig. 3: Detect motion from video, manipulate it, and build body surface over it.