Continuous Capture of Skin Deformation

Peter Sand∗ Leonard McMillan† Jovan Popović

Laboratory for Computer Science
Massachusetts Institute of Technology
†University of North Carolina, Chapel Hill

Abstract

We describe a method for the acquisition of deformable human geometry from silhouettes. Our technique uses a commercial tracking system to determine the motion of the skeleton, then estimates geometry for each bone using constraints provided by the silhouettes from one or more cameras. These silhouettes do not give a complete characterization of the geometry for a particular point in time, but when the subject moves, many observations of the same local geometries allow the construction of a complete model. Our reconstruction algorithm provides a simple mechanism for solving the problems of view aggregation, occlusion handling, hole filling, noise removal, and deformation modeling. The resulting model is parameterized to synthesize geometry for new poses of the skeleton. We demonstrate this capability by rendering the geometry for motion sequences that were not included in the original datasets.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation; I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Shape

Keywords: motion capture, skin modeling, human animation

1 Introduction

A digital replica of a moving human body has applications in video games, teleconferencing, automated news shows, and filmmaking. For example, the physical appearance of a celebrity actor could be recorded and later animated with acrobatic motions controlled by an animator or performed by a stunt double in a motion-capture suit. In current filmmaking, this application requires extensive manual labor to position and adjust skin around each bone and muscle. In some cases, months are spent matching a virtual character to an existing actor [Stokdyk et al. 2002].

Our goal is to build a skin model that replicates the skin deformations of a particular person. The technique described in this paper builds this model automatically from video of the subject and motion data that describes how the subject’s skeleton moves throughout the video recording. To build the model from this data, we exploit the idea that video of a moving person provides many observations of the same surface. A single set of silhouettes (even from several viewpoints) provides a highly incomplete characterization of the geometry. By having the subject move through many different poses, local configurations of the body parts are repeated, allowing the construction of a complete model.

Our main contribution is a method of gathering silhouette observations such that a simple reconstruction algorithm can create a complete deformable model, parameterized in a way that is useful for animation. We do not contribute new techniques in the areas of skin representation and skin interpolation, but in ways of quickly acquiring skin data. By using the right combination of prior tools, we substantially simplify the problem of generating a 3D model from moving silhouettes.

Our skin model, described in Section 3, represents a complex articulated figure using a collection of elongated deformable primitives. Our acquisition algorithm, described in Section 4, uses the silhouettes to provide constraints on the possible body geometry. The reconstruction algorithm, described in Section 5, uses these constraints to find a model of the skin deformations, parameterized with the motion of the skeleton. This parameterization allows animation of the skin with new motion data.

2 Related Work

The most general 3D reconstruction systems attempt to build a model of the scene at each successive time frame, allowing the acquisition of moving objects. These systems use vision methods such as binocular stereo [Nebel et al. 2001] and voxel coloring [Vedula et al. 2002]. For certain kinds of scenes, the geometry can be reasonably represented using a visual hull: the space carved about by silhouettes from a set of viewpoints [Matusik et al. 2000; Würmlin et al. 2002].

Some of these methods make frame-to-frame comparisons of the geometry [Würmlin et al. 2002; Vedula et al. 2002], but they do not accumulate observations to improve the geometry. The strength of gathering information from temporally distinct views is illustrated in recent work in real-time model acquisition, in which a rigid ob-
Figure 13: This model was generated from a female subject using 5 minutes of motion and silhouettes from three viewpoints. In the future we would like to capture a wide variety of people and interpolate their geometries to synthesize new deformable models.

Acknowledgements

Special thanks to Frédéric Durand, Seth Teller, William Freeman, Farid Janmahir, Barb Cutler, Tom Buehler, Adnan Sulejmanpašić, Daniel Vlasic, Eric Chan, and Eugene Hsu. We are grateful to Vicon Motion Systems for their permission to use the beta release of the Vicon iQ software. This research was sponsored in part by the NTT Corporation.

References


