Using Vision for Animating Virtual Humans

Ioannis Kakadiaris

University of Houston, USA

Automatic, non-intrusive vision-based capture of the human body motion opens new possibilities in applications requiring the use of geometric and kinematic data from individuals (e.g., virtual reality, teleconferencing, performance measurement). If synthesized motion is to be compelling, we must create actors for computer animations and virtual environments that appear realistic when they move.

In this talk, I will present the computer vision and computer graphics formulations and techniques that we have developed for the three-dimensional model-based motion capture and animation of unconstrained human movement from multiple cameras. Our tracking and animation system consists of a human motion analysis and a synthesis components. First, I will present novel analytical computer vision techniques to accurately recover the three-dimensional shape and pose of a subject's body parts. These techniques are based on the spatio-temporal analysis of a subject's silhouette from image sequences acquired simultaneously from multiple cameras. We employ physics-based deformable human body models to estimate the position and orientation of the subject's body parts during the analysis steps. This amounts to minimizing continuously the discrepancy between the occluding contour of the deformable model and the silhouette of the human in the acquired image sequences. To mitigate difficulties arising due to occlusion, we employ multiple cameras. For efficiency and robustness, we have devised two criteria for the active time-varying selection of the subset of the cameras that provide the most information. These criteria depend on the visibility of the subject's body parts and the observability of their predicted motion from the specific camera. For the motion synthesis component, we have developed techniques that allow the efficient and realistic animation of the subject's estimated graphical model. The advantage of our system is that the subject does not have to wear markers or special equipment. Finally, I will present motion estimation and animation results demonstrating the generality and robustness of our algorithm.