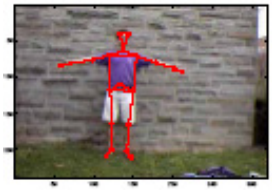
3D Model Human motion analysis

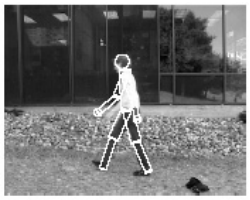
3D Model Human motion analysis has been an active research field over the past decades. It concerns the analysis of image sequences (video) of moving human subjects in order to extract useful information. Such information can be the position and heading of a moving person, the velocity of the motion, the pose of a person and the 3D properties of a human body. Moreover many current researchers have focused on higher-level information extraction as human activity recognition, event detection, automated video data annotation, avatar animation and more.

Model

3D Model Human motion analysis techniques are model-based approaches that utilize a human body 3D model. There is a large variety of human body models used. The simplest representation of a human body is the stick figure, which consists of line segments linked by joints. This representation is based on the observation that human motion is essentially the movement of the supporting bones and the model can then be enhanced with the deformable flesh surrounding the skeletal structure through higher level processing. More complex models include volumetric representations as generalized cones elliptical cylinders and spheres. The selection of the model employed usually depends on the application at hand.



Picture 1. Stick figure model. [5]



Picture 2. Model with cylinders and spheres. [3]

Framework

The general strategy for 3D Human Motion Analysis consists roughly of three steps:

- Motion capture
- Motion extraction
- Motion analysis

Motion Capture

Plenty motion capture techniques can be used. Such techniques can be divided into three major categories which are mechanical, electro-magnetic and image-based approaches. The mechanical and electromagnetic approaches are invasive and usually employ a suit with motion sensors that capture the motion of the actor. The image based approaches can be further divided into techniques that use active markers, techniques that use passive markers and techniques that do not use markers, in single and multi camera setups. Generally the vision based techniques with no markers are preferred as they are non intrusive and their applicability is far broader.

Motion extraction

The motion extraction step is concerned with recognizing the movement of the subject in the image sequence and isolating that motion. This task can be accomplished with various methods such as interest point feature detection and matching, figure-background separation, human silhouette detection, optical flow models and more. Typical problems that arise in this step are dealing with varying viewing conditions and occlusions.

Motion analysis

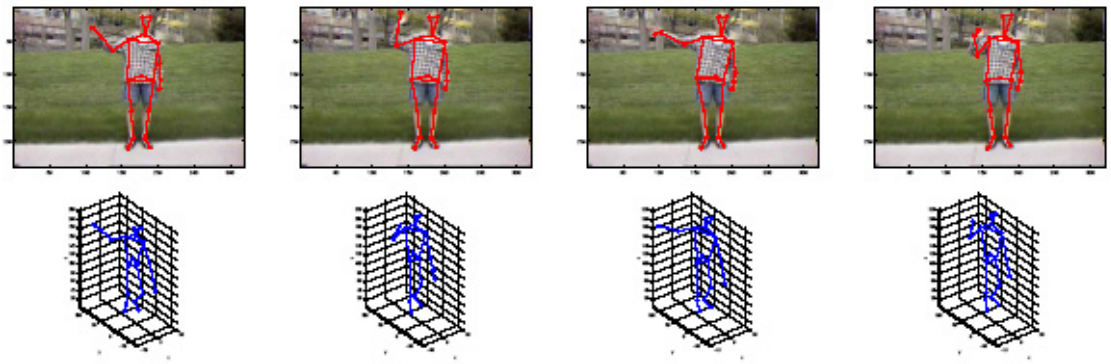
Motion analysis is then preformed on the extracted motion data where the 3D human body model is fitted to the observed motion. Typically the output of the motion analysis step is the pose estimate, the position and orientation, of the model's body parts. That information can then be used in order to compute trajectories, velocities and accelerations of the 3D model. Higher level processing of the previous data can be then carried out for human activity recognition, motion reconstruction, event detection and more.

Example of 3D Model Human motion analysis

The following example is adopted from [5] and demonstrates a probabilistic approach to 3D Model Human Motion Analysis. In this work motion analysis is treated as a Bayesian inference problem where the 3D properties of the model are inferred by the 2D representation of the subject in the image sequences. The procedure is as follows:

- 2D motion tracking in the video image extracts the 2D limb coordinates in each consecutive frame.
- A model of plausible human motions is created beforehand as a mixture of high dimensional Gaussians that, in other words, encodes the prior probability of a short motion.
- The tracking data are concatenated into short motions, referred to as 'snippets'.
- The algorithm calculates the most probable 3 dimensional snippet sequence based on the tracking data and the learned snippet priors.

The result is a Bayesian estimate of the motion in 3D space. (Due to lack of space for mathematical details refer to [5].)



Picture 3. A clip and its 3D reconstruction [5].

References

[1] J.K. Aggarwal and Q. Cai, ‘Human Motion Analysis: A Review’, Proc. IEEE Nonrigid and Articulated Motion Workshop, pp. 90±102, June 1997.

[2] Kakadiaris and D. Metaxas, ‘Model-Based Estimation of 3D Human Motion’ IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No 12, pp. 1453-1459, December 2000.

[3] H. Sidenbladh, M.J. Black and D.J. Fleet, ‘Stochastic Tracking of 3D Human figures Using 2D Image Motion’. European Conf. on Computer Vision, pp. 702-718, ECCV’2000.

[4] W.W. Lok and K.L. Chan, ‘Model-Based Human Motion Analysis in Monocular Video’, Vol 2, pp ii/697- ii/700 IEEE ICAASP, 2005.

[5] N.R. Howe, M.E. Leventon and W.T. Freeman, ‘Bayesian Reconstruction of 3D Human Motion from Single-Camera Video’, Advances in Neural Information Processing Systems 12, 1999, edited by S. A. Solla, T. K. Leen, and K-R. Muller, 2000.