論文の内容の要旨

論文題目　Cooperative estimation of human motion and surface using multiview videos
（多視点ビデオからの人物の動きと形状の協調的推定）

氏　名　羅　衛蘭

This thesis proposes a human motion tracking method that not only captures the motion of the skeleton model but also deforms the surface model using images acquired by multiple synchronized cameras. In the motion tracking process, we extracts articulated postures with 42 degrees of freedom through a sequence of visual hulls. We seek a globally optimal solution for the likelihood using local memorization of the fitness of each body segment. Our method avoids problems of local minima efficiently by using a mean combination and an articulated combination of particles selected according to the weights of the different body segments. We deform the template surface model by using the motion tracking data with linear blend skinning. The details of the surface are recovered by fitting the deformed surface to 2D silhouettes. The extracted posture and estimated surface are refined by registering the corresponding body segments. In our experiments, the mean error between the samples of the deformed reference model and the target is about 2 cm and the mean matching difference between the images projected by the estimated surfaces and the original images is about 6%.

Segmentation for the human body is important while motion estimation for each body joint. We first segment the template mesh surface utilizing the geodesic distance and the underlying skeleton and then label the corresponding volumetric model. The skinning weights for the vertices can be calculated while the connection between the vertex and the bone is known. The visual hull can be labeled by searching for the nearest vertex in the corresponding mesh surface. The segmented volumetric model is then utilized in the motion tracking process for the next frame.

A sampling method is proposed for selecting maximally dispersed voxels to ensure that there are sufficient voxels for each body segment. We obtain the bounding box of a volumetric model and scale it to be a cube, divide it to obtain about 2% of the voxels.

We propose a simple approach to eliminating undesired human motions via self-intersection. We provide a weighting function to guide the motion process to remove bad particles that contain body segments in collision with each other.

Although global optimization approaches can avoid local-misaligned problem, it is usually time-consuming and still hard to obtain the optimal solution for tracking especially in high dimensional search space. The modified particle filter method we proposed quickens the convergence by memorizing the mean squared distance between the samples in the reference and in the current frame for each body segment, and combine them to provide more “appropriate” particles for the annealing process.

A sampling method for voxel selection is provided to reduce the computational cost, while ensuring that the selected voxels are spread across all the human segments. In addition, a self-collision detection method is utilized to search for real particles in the tracking process.

Non-rigid deformation for recovering surface detail can be realized by linear blending skinning (LBS) method while obtaining the corresponding vertices by silhouette constraints. In addition, we take samples of the segmented volumetric model to match the target, using iterative closest point (ICP) registration to smooth the errors caused by the mesh deformation. The extracted transformation for each body limb provides a good initial estimation for the ICP algorithm. The template surface is deformed to generate the reference surface for the next frame using LBS.