Abstract:
A common approach to the problem of 3D human pose estimation from video is to recursively estimate the most likely pose via particle filtering. However, standard particle filtering methods fail the task due to the high dimensionality of the 3D articulated human pose space. In this paper we present a thorough evaluation of two variants of particle filtering, namely Annealed Particle Filtering and Partitioned Sampling Particle Filtering, that have been proposed to make the problem feasible by exploiting the hierarchical structures inside the pose space. We evaluate both methods in the context of markerless model-based 3D motion capture using silhouette shapes from multiple cameras. For that we created a simulation from ground truth sequences of human motions, which enables us to focus our evaluation on the sampling capabilities of the approaches, i.e. on how efficient particles are spread towards the modes of the distribution. We show the behaviour with respect to the amount of cameras used, the amount of particles used, as well as the dimensionality of the search space. Especially the performance when using more complex human models (40 DOF and above) that are able to capture human movements with higher precision compared to previous approaches is of interest in this work. In summary, we show that both methods have complementary strengths, and propose a combined method that is able to perform the tracking task with higher robustness despite reduced computational effort.