Geometric Consistency-Based Self-Supervised Neural Network: A Novel Deep Learning Framework for 3D Human Shape and Motion Reconstruction

Hua, Michelle (School: Cranbrook Kingswood School)

3D human motion reconstruction from a monocular video is one of the most attractive yet challenging research fields. It has the potential to enable 3D broadcasting, advance virtual and augmented reality, conduct sport analysis, deliver telepresence, etc. Existing machine learning methods for 3D reconstruction require a large number of hard-to-obtain training pairs, e.g., human images/videos and their corresponding 3D human models, and often suffer from performance degradation in practice due to appearance variations between the training and testing data. Therefore, I propose a novel geometric consistency-based self-supervised neural network (GC-SSN) for 3D human shape and motion reconstruction from a monocular video. In GC-SSN, the representation of a moving human is modeled with a geometric representation based on joints and silhouettes extracted from each frame of the video, thus avoiding the instability of appearance-based representations and constraints. During training, the joints and silhouettes of the reconstructed 3D human model are automatically extracted, rendered, and fed back to the reconstruction network to form a complete cycle. By enforcing the reconstructed 3D human model to align consistently with the extracted joints and silhouettes constraints from the input and output geometric representations in both the forward and backward directions, the generator, consisting of a feature encoder and a regressor, in GC-SSN can build the 3D human model with a high accuracy. The GC-SSN is self-supervised with automatically extracted joints and silhouettes without any manual annotations or ground truth 3D human shapes. It significantly improves the domain adaption and outperforms other state-of-the-art algorithms.

Awards Won:

Second Award of $2,000