Spatio-temporal Tendency Reasoning for Human Body Pose and Shape Estimation from Videos

Boyang Zhang, Suping Wu*, Hu Cao, Kehua Ma, Pan Li, Lei Lin

* represents corresponding author

School of Information Engineering, Ningxia University, Yinchuan, China

Abstract

➢ The existing human pose and shape estimation from videos methods are difficult to reconstruct a reasonable human body in an unconstrained environment (extreme illumination, motion blur). Although some approaches attempt to improve performance by adding external data resources, these methods do not take full advantage of the potential information in the underlying data.

➢ While these methods improve the temporal consistency of human pose estimation in video, they lack spatial understanding and reasoning capabilities, leading to biased predictions.

Objective

Our approach aims to alleviate the problem of human reconstruction in unconstrained scenes by reasoning about the spatio-temporal tendency of moving human body.

Main Contribution

➢ We propose a spatio-temporal tendency reasoning (STR) for human body pose and shape estimation from videos, which can alleviate the problem of human reconstruction in unconstrained scenes.

➢ We design a temporal tendency reasoning module and a spatial tendency enhancing module. The tendency reasoning enhancing module consists of a temporal tendency reasoning and a spatial tendency enhancing module. The integration strategies consist of a self-integration strategy and a cross-integration strategy.

STR architecture

STR consists of two modules, tendency reasoning enhancing module and an integration strategy module. The tendency reasoning enhancing module consists of a temporal tendency reasoning module and a spatial tendency enhancing module. The integration strategies consist of self-integration strategy and cross-integration strategy.

Integration Strategies

Table 1. Comparisons of our approach with state-of-the-art methods on 3DPW(in-the-wild), MPI-INF-3DHP(outdoor), Human3.6M(indoor) testing set. We denote whether 3DPW is involved in the training process as w/ 3DPW, w/o 3DPW respectively.

Table 2. Effects of the network designs on the performance on the MPI-INF-3DHP dataset.

Experimental Results

Figure 1. Qualitative Comparison under extreme illumination.

Figure 2. The acceleration errors comparison and Qualitative visualization of STR in the unconstrained scene.