

Few-Shot Human Motion Prediction via Meta-Learning

Liang-Yan Gui, Yu-Xiong Wang, Deva Ramanan, and José M. F. Moura

Carnegie Mellon University
{lgui, yuxiongw, deva, moura}@andrew.cmu.edu

Abstract. Human motion prediction, forecasting human motion in a few milliseconds conditioning on a historical 3D skeleton sequence, is a long-standing problem in computer vision and robotic vision. Existing forecasting algorithms rely on extensive annotated motion capture data and are brittle to novel actions. This paper addresses the problem of few-shot human motion prediction, in the spirit of the recent progress on few-shot learning and meta-learning. More precisely, our approach is based on the insight that having a good generalization from few examples relies on both a generic initial model and an effective strategy for adapting this model to novel tasks. To accomplish this, we propose proactive and adaptive meta-learning (PAML) that introduces a novel combination of model-agnostic meta-learning and model regression networks and unifies them into an integrated, end-to-end framework. By doing so, our meta-learner produces a generic initial model through aggregating contextual information from a variety of prediction tasks, while effectively adapting this model for use as a task-specific one by leveraging learning-to-learn knowledge about how to transform few-shot model parameters to many-shot model parameters. The resulting PAML predictor model significantly improves the prediction performance on the heavily benchmarked H3.6M dataset in the small-sample size regime.

Keywords: Human motion prediction · Few-shot learning · Meta-learning

1 Introduction

One of the hallmarks of human intelligence is the ability to predict the future based on past observations. Through perceiving and forecasting how the environment evolves and how a fellow human acts, a human learns to interact with the world [60]. Remarkably, humans acquire such a prediction ability from just a few experiences, which is yet generalizable across different scenarios [50]. Similarly, to allow natural and effective interaction with humans, artificial agents (e.g., robots) should be able to do the same, i.e., forecasting how a human moves or acts in the near future conditioning on a series of historical movements [29]. As a more concrete example illustrated in Figure 1, when deployed in natural environments, robots are supposed to predict unfamiliar actions after seeing only a few examples [27, 20]. While human motion prediction has attracted increasing

