Action recognition using kinematics posture feature on 3D skeleton joint locations

Department of Intelligent Media, SANKEN (The Institute of Scientific and Industrial Research)
Professor Yasushi Yagi Researchmap https://researchmap.jp/read0013801?lang=en

Institute for Advanced Co-Creation Studies/Department of Intelligent Media, SANKEN (The Institute of Scientific and Industrial Research)
Professor Yasushi Makihara Researchmap https://researchmap.jp/yasushi_makihara?lang=en

Abstract

Action recognition is a very widely explored research area in computer vision and related fields, which has a wide variety of applications such as healthcare and surveillance. We propose to extract Kinematics Posture Feature (KPF) from 3D joint positions based on skeleton data for action recognition. We applied our method on five benchmark datasets for action recognition and obtained comparable or better results than the existing methods.

Background & Results

Human Action Recognition (HAR) is one of the most prominent and challenging research areas of computer vision and ubiquitous computing in recent years. Because of the articulated nature of human motion, there lies a major issue to model the human actions that are ambiguous, dynamic, and interactive with other objects. Modeling specific temporal structures of human actions is another challenging task. RGB videos and images also pose security and privacy issues. Recently, cost-effective depth cameras have fostered the progress of promising approaches to develop reliable and cost-effective solutions by providing 3D depth and skeleton data of the scene that largely eases the task of action recognition.

We propose Linear Joint Position Feature (LJPF) and Angular Joint Position Feature (AJPF, Fig. 1) based on 3D joint positions and angles between bone segments. We then combine these two kinematics features for each video frame for each action to create the KPF feature sets. These feature sets encode the variation of motion in the temporal domain as if each body joint represents kinematics position and orientation sensors. From the segments extracted by sliding windows, we compute the Position-based Statistical Feature (PSF) which consist of temporal domain statistical features (e.g., mean, standard deviation, variance, etc.). For performing classification, we explore Support Vector Machine (Linear), RNN, CNNRNN, and ConvRNN model (Figs. 2 and 3).

For evaluation, we explore five benchmark datasets namely UT-Kinect-Action3D, Kinect Activity Recognition Dataset (KARD), MSR 3D Action Pairs, Florence 3D, and Office Activity Dataset (OAD). To prevent overfitting, we consider the leave-one-subject-out framework as the experimental setup and perform 10-fold cross-validation. Our approach outperforms several existing methods in these benchmark datasets and achieves very promising classification performance. Besides, the proposed PSF feature sets demonstrate prominent performance in both statistical machine learning- and deep learning-based models.

Significance of the research and Future perspective

HAR plays a fundamental role in numerous relevant and heterogeneous application fields such as assistive living, health care, video-surveillance, augmented reality, patient monitoring, intelligent surgery, and so on. We can also take the help of Active and Assisted Living (AAL) tools to reduce the social-cost and mitigate the challenges of the aging population in the modern and developed society. One of future research avenues is to explore two-persons interactions and multi-view cases to extend the application ranges.

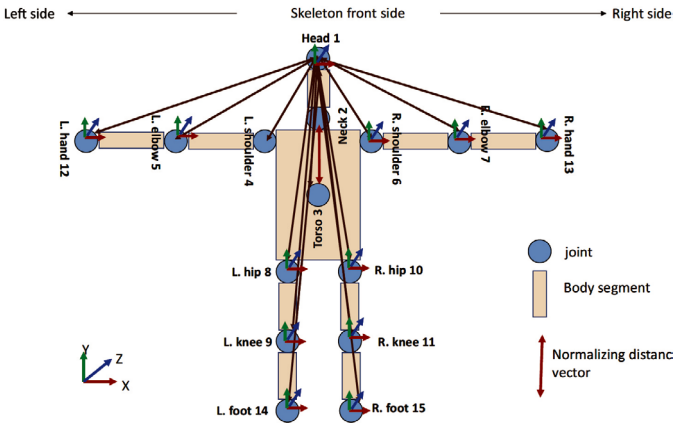

FIg.1 Distance vector representation of each joint with respect to the head joint

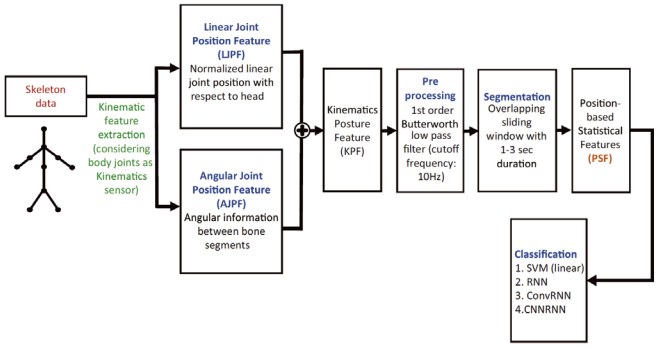

Fig.2 Pipeline of the proposed method

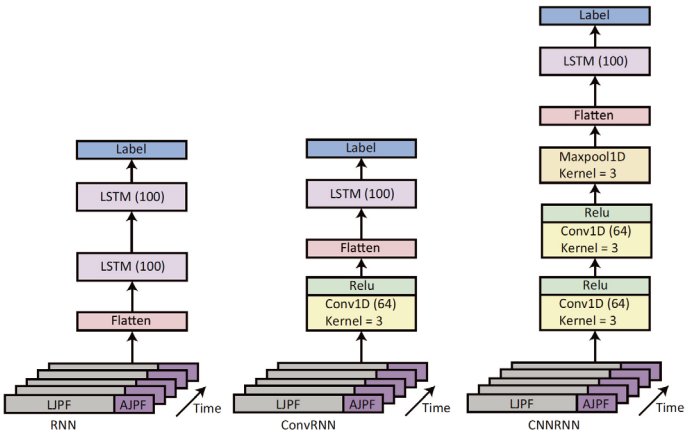

Fig.3 Architectures of deep learning models

Patent
Treatise Md Atiq Rahman Ahad; Masud Ahmed; Anindya Das Antar et al. Action recognition using kinematics posture feature on 3D skeleton joint locations. Pattern Recognition Letters, 2021, 145, 216-224. doi: 10.1016/j.patrec.2021.02.013
URL http://www.am.sanken.osaka-u.ac.jp/
Keyword action recognition, computer vision, pattern recognition, image media processing