Journal of Information Technology Research
Volume 12 • Issue 1 • January-March 2019

Human Action Recognition Based on Inertial Sensors and Complexity Classification

Lijue Liu, School of Information Science and Engineering, Central South University, Changsha, China
Xiaoliang Lei, School of Information Science and Engineering, Central South University, Changsha, China
Baifan Chen, School of Information Science and Engineering, Central South University, Changsha, China
Lei Shu, School of Information Science and Engineering, Central South University, Changsha, China

ABSTRACT

In this article, a human action recognition technique based on complexity classification is proposed. Considering the features of human actions such as continuity, individuality, variety randomness, the demands for recognition of different types of actions are different, the problem of action recognition can be classified into simple action recognition and complex action recognition -- the classification criterions are given respectively. Meanwhile, the hardware design of data acquisition device is introduced and the angle variation is chosen to represent the user's body state changes. For simple actions, a real-time recognition algorithm based on template matching performed well on cost control, and a method based on BLSTM-RNN is used for complex motion recognition to improve the accuracy of identification.

KEYWORDS

Complexity classification, Human action recognition, Inertial sensors, LSTM-RNN, template matching

INTRODUCTION

Human action recognition has been applied in multiple fields, such as human computer interface, entertainment, visual reality (Barnachon et al., 2014, Burgner-Kahrs et al., 2015, Pietro Cipresso et al., 2014), and even rehabilitation (Sheng et al., 2016, Meng et al., 2015), remarkable results are obtained. It can be done in varieties of ways, including video (Gupta et al., 2013, Guo et al., 2014), depth cameras (Wang et al., 2012, Shao et al., 2012), and body-worn inertial sensors (Feng Cai et al., 2016) which can be used to measure and detect the inclination, shock, vibration, rotation, and multi degree of freedom. Since the body-worn inertial sensors are less expensive and more sensitive (Zhang et al., 2011, Slama et al., 2015), this paper is based on wearable inertial sensors.

Human action is continuous, that means when one wants to make some kinds of actions, various unrelated-actions must exist before and after it. Besides, the brain's control of the body is fuzzy and subjective, it's difficult to make exactly the same actions at different times. If the existence of individual difference is taken into account, the characterized difference can be even obvious. Additionally, compared with the standard model, user actions are often not standardized and incomplete. Undoubtedly, these factors will increase the difficulty of action recognition.

DOI: 10.4018/JITR.2019010102

The key of action recognition is to represent changes of the body. Since there are serval limited such as acceleration is too sensitive and the geomagnetic signal is susceptible to interference, angle variation fusion by accelerometer, gyroscope and magnetometer is chosen to represent the user's body state changes.

Besides, the difficulty of recognizing different kind of actions could not be confused, for their requirements for the performance of the processor are different, this paper discusses the problems of motion recognition into two cases -- the simple action and the complex action, and gives the classification criterion.

The remainder of this paper can be organized as follows: The next section analyzes the related literatures and comprehensively demonstrates the scope of the paper. Methodology elaborates the recognition algorithm in detail. Experiments provides the experiments to evaluate the methodology, and Results discusses the finding of this experiment. Finally, conclusions are drawn.

RELATED WORK

Z.Zhang et al (Zhang Z et al., 2009, Zhang Z et al., 2011) raised a hierarchical information fusion algorithm, which can estimate the human gestures. In the analysis, the algorithm has geometric constraints on human bone, and they introduced particle filter algorithm in the information fusion process. Although this method recognition rate is higher, its calculation is relatively complicated. Especially for the embedded systems with few resources and low computing, its implementation is difficult. Furthermore, Hao Yang (Yang et al., 2011) and Luinge (Luinge et al., 2007) gained the gesture relationship between the carrier sensing coordinate and body coordinate through human place under different statuses of predefined acceleration and angular velocity information in the human body. Besides, Huiyu Zhou (Zhou et al., 2008) adopted body skeleton model linked by the human joint and got the space position between arm elbow and wrist joints through the transformation between the coordinate system. Even though the method can locate the joint position, it cannot get the corresponding flip Angle of arm turning.

The primary work in recognition is to segment the human actions. Deng et al (Deng L et al., 2012) proposed a model based on motion segmentation method of generalized division index figure and the human body in terms of human continuous action segmentation. According to the spatial location relations, this method divided human skeleton into five parts, and then extracted molecular motion feature of everybody with the clustering algorithm. After that, they combined these features of each sport training and set up a general model. Lastly, it is to realize motion segmentation using a similarity measure method. Additionally, Ramakrishnan (Babu et al., 2004) proposed a gesture movement data segmentation method based on support vector machine, and utilized the support vector machine to train the joint angles as well as the positions of the fingers and hands data, in order to detect the boundaries between different types and split different gestures. The above methods have high requirements to the processor, while in the practical application, the embedded operating systems are more extensive, which cannot satisfy the above algorithm.

Regarding the recognition, Trung Thanh Ngo et al. (Ngo et al., 2015) presented an algorithm to recognize similar gait actions with three inertial sensors. To be specific, they incorporated the interclass relationship in the feature vector for recognition. Furthermore, they adopted 3D acceleration signal data to detect the gait period, and used acceleration and rotational velocity data to calculate the dissimilarity between the action and the gallery action, separately. However, the detection range of this method is limited to gait. It does not have a good fusion of the 6D data of inertial sensor. Based on dynamic time alignment of Gaussian mixture model clusters, Matthew Field et al (Field et al., 2015) proposed a method for matching actions in an unsupervised temporal segmentation. In facilitation of this, an extensive corpus of continuous motion sequences composed of everyday tasks was recorded as the analysis scenarios. Also, they utilized 17 sensors to detect human actions. However, the operation process is complex and the flexibility is poor. Beyond that, there are some

Related Content