Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 112 (2017) 554–563

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2017, 6-8 September 2017, Marseille, France

Robust human action recognition system using Laban Movement Analysis

Insaf Ajili*, Malik Mallem, Jean-Yves Didier

[a]Laboratoire d'Informatique, Biologie Integrative et Systemes Complexes (IBISC), Evry Courcouronnes, France

Abstract

Recognizing human actions from video sequences is an active research area in computer vision. This paper describes an effective approach to generate compact and informative representations for action recognition. We design a new action feature descriptor inspired from Laban Movement Analysis method. An efficient preprocessing step based on view invariant human motion is presented. Our descriptor is applied in four known machine learning methods, Random Decision Forest, Multi-Layer Perceptron and Multi-class Support Vector Machines (One-Against-One and One-Against-All). Our proposed approach has been evaluated on two challenging benchmarks of action recognition, Microsoft Research Cambridge-12 (MSRC-12) and MSR-Action3D. We follow the same experimental settings to make a direct comparison between the four classifiers and to show the robustness of our descriptor vector. Experimental results demonstrate that our approach outperforms the state-of-the-art methods.

Peer-review under responsibility of KES International

Keywords: Laban Movement Analysis; Preprocessing data; Action Recognition; Machine learning;

1. Introduction

Human action recognition in video has been an active research topic and considered as a key component for various computer vision applications, such as human-computer interaction, surveillance video analysis for security, robot control[25,15,8]. As shown in Figure 1, an action recognition system consists in labeling image sequences captured from videos with action labels. This process is still a very complex task due to the high variability of human motion. Indeed, when different subjects perform the same action, we can obtain differences between their movements or their positions regarding camera. Even for a person performing the same action multiple times, each performance can be quite different from the previous one. Therefore our system should incorporate the typical changes invariance, and extract pertinent features to faithfully represent human actions. Selecting the best features can be considered the most important step since classification can be performed by any existing machine learning after feature extraction step. Previous researches[25,21,9,8,7,19,15,23] mainly concentrate on the quantitative aspect of actions and have neglected the

* Corresponding author. Tel.: +33-169-477-515; fax: +33-169-477-599.
E-mail address: (Insaf.Ajili, Malik.Mallem, Jean-Yves.Didier)@ufrst.univ-evry.fr

1877-0509
Peer-review under responsibility of KES International
10.1016/j.procs.2017.08.168

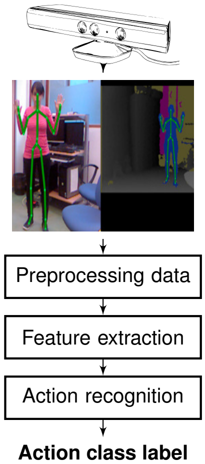

Fig. 1. Schematic view of action recognition process.

quality side of a movement which presents a very important information to recognize high level gestures. In this context, another research field appeared, the Laban Movement Analysis (LMA), developed by Rudolf Laban [11]. LMA includes the quantitative analysis but additionally addresses qualitative changes of movement in space and time. LMA is composed of the following four major movement categories: Body, Shape, Space, and Effort. To our knowledge, only few works proposed LMA approach and most of them for describing emotions or analyzing dance choreography. So they were focused on analyzing only Effort and Shape components, such as Samadani et al. [18] for describing expressive qualities of hand and arm movements, and Zacharatos et al. [24] which employed these two components to recognize players emotion for Exergames. The authors Kim et al. [10] and Suzuki et al. [22] used only the Effort quality to express emotional movements like dance motion. Others had used Laban's features defined by six features, Time, Weight, Space, Inclination, Area and Height such as Masuda [13] and Nishimura et al. [14] for robot emotion description.

Effort factors are not used in our case because our application deals with robot control (move forward, turn right, etc) triggered by human actions. Such application needs a robust action recognition system independent of many possible changes, like actions with different durations, speeds, rhythms, forces, etc). However Effort component represents the expressive quality of actions. It is often used to describe movements in terms of emotions, energy, effort, etc, which justifies its widely use in emotion recognition systems.

The remainder of this paper consists of three sections. Section II details our recognition system as well as its various stages. Section III reports our experimental results, and finally section IV presents the conclusion with some perspectives.

2. Proposed Process

In this section, the proposed human action recognition approach is detailed, three parts are presented. The first part presents the preprocessing data step, the second part is details of our proposed LMA descriptors and the last part presents the different machine learning methods used for action recognition step (Figure 1).

2.1. Preprocessing data

We use Microsoft Kinect sensor for data acquisition. Both the RGB and depth streams are recorded at 640×480 resolution. Kinect SDK tracks 3D coordinates of 20 body joints in real time (30 frames per second). For normalization stage, we define a local skeleton coordinate system (X', Y', Z') in the base $B'(i', j', k')$, its center is the midpoint of hips (P_m). As shown in Figure 2, we choose the hip axis with direction from the right hip joint P_{rsh} and the left hip joint P_{lsh} as the X'-Axis of the skeleton coordinate system. The vector starting from the midpoint of hips (P_m) and