Applied Soft Computing 46 (2016) 953–966

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Early detection of human actions—A hybrid approach

Ekta Vats, Chee Seng Chan *

Centre of Image and Signal Processing, Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 15 June 2015
Received in revised form 6 October 2015
Accepted 2 November 2015
Available online 21 November 2015

Keywords:
Human action recognition
Human activity recognition
BK subproduct
Human motion analysis

ABSTRACT

Early detection of human actions is essential in a wide spectrum of applications ranging from video surveillance to health-care. While human action recognition has been extensively studied, little attention is paid to the problem of detecting ongoing human action early, i.e. detecting an action as soon as it begins, but before it finishes. This study aims at training a detector to be capable of recognizing a human action when only partial action sample is seen. To do so, a hybrid technique is proposed in this work which combines the benefits of computer vision as well as fuzzy set theory based on the fuzzy Bandler and Kohout's sub-triangle product (BK subproduct). The novelty lies in the construction of a frame-by-frame membership function for each kind of possible movement. Detection is triggered when a pre-defined threshold is reached in a suitable way. Experimental results on a publicly available dataset demonstrate the benefits and effectiveness of the proposed method.

1. Introduction

Human action recognition has been widely studied over the years with real-time applications in video surveillance [1–3], health-care monitoring [4–6], sport analysis [7,8], etc. However, detecting ongoing human action early, i.e. as soon as it begins but before it finishes has not received much attention in the recent past. Most of the methods dealt with detection of the action after its completion. On the contrary, for early detection it is essential to detect partial action [9–13]. Early detection of human action is essential in several situations such as monitoring criminal activities, patients' fall detection, etc. Consider the example of an elderly care system in a hospital, it is crucial to accurately and rapidly detect the falling activity of the elderly patients as soon as possible, so that necessary medical care can be provided in a timely manner.

Early detection of human action is a daunting task given the vast amount of uncertainty involved therein. The conventional computer vision solutions often fall short of providing efficient solution as they are not robust enough to handle issues such as uncertainty, imprecision and vagueness. Hybrid techniques are believed to address these issues to a considerable extent by exploiting the strengths of one technique to alleviate the limitations of another [14,15]. Therefore, in this paper a hybrid technique for early detection of human action is proposed as the synergistic integration of computer vision solutions and fuzzy set theory. It is believed that computer vision methods and fuzzy approaches do not behave in a conflicting manner, but rather complimenting one another [16]. The fusion of these techniques towards performing human action recognition as early as possible can be achieved through proper hybridization. To this end, the relationship between a human and the action being performed is studied using the Bandler and Kohout's (BK) sub-triangle product (subproduct) [17], efficiently integrated with computer vision techniques including feature extraction and motion tracking to perform human action recognition effectively.

To the best of the authors' knowledge, this paper is the first attempt towards providing a solution to early human action detection using a hybrid technique, combining the benefits from computer vision and fuzzy set theory. Fuzzy BK subproduct is chosen in this work due to its flexibility and efficacy to be employed in real-world applications [18–21], and also its capability to imitate the natural human behavior, i.e. modus-ponen way [22]. Modus-ponen refers to our interpretation of available information while solving real-life problems, for example if A implies B, and A is asserted to be true, therefore B must be true. Another issue addressed by the proposed method is to handle the cumulative tracking errors and precision problem using a set of overlapped fuzzy numbers known as fuzzy quantity space, where individual distance among them is defined by a predefined metric [23,24]. We intend to provide a solution for early human action detection closest to natural human perception. The novelty lies in the hybrid based learning formulation to train the early detector such that once the detector has been trained, it can be flexibly used in

* Corresponding author. Tel.: +60 3 7967 6433.
E-mail addresses: ektavats_2608@siswa.um.edu.my (E. Vats), cs.chan@um.edu.my (C.S. Chan).

http://dx.doi.org/10.1016/j.asoc.2015.11.007
1568-4946/

several ways depending upon the application. Experiments on a standard human action dataset illustrate the capability of the proposed hybrid technique to make reliable early detection of human action. The partial human action is modeled, where the fuzzy membership function provides the basis to detect an action before it is completed when a certain threshold is attained in a suitable way. To summarize, our main contribution is one of the first attempts to employ hybrid technique, a fuse between computer vision and fuzzy sets approaches for early detection of human action. Most of the classical solutions [16] had been focusing on human action recognition.

A preliminary version of this work was presented earlier [25]. The present work adds to the initial version in significant ways. Firstly, we improve the early human action detection framework by introducing fuzzy quantity space in the tracking stage to handle the cumulative tracking errors. Secondly, considerable new analyses and intuitive explanations are added to the original results. We also extend the original experiments from using a partial Weizmann dataset to a full Weizmann dataset.

This paper is structured in the following way. Section 2 provides the background on human action recognition from the fields of computer vision and fuzzy set theory. The BK subproduct approach is revisited in Section 2.2.1. Section 2.3 reviews early event detection. The proposed hybrid technique for early detection of human action is described in Section 3. Section 4 provides an analysis of the experimental results and assesses the effectiveness and benefits of the proposed method. Finally, Section 5 concludes the paper.

2. Background

This section gives an overview of the background of human action recognition, with reference to computer vision methods and fuzzy set oriented approaches with a short description of the BK subproduct inference mechanism. Fig. 1 represents a general framework for human action recognition, where for an input video, firstly, the human object is detected as low-level vision task, followed by human motion tracking in the mid-level processing. Furthermore, the literature on early detection of human action is reviewed with highlight on the state-of-the-art methods along with their limitations.

2.1. Human action recognition in computer vision

There exist several surveys of human action recognition in computer vision literature [26–29], focusing on various methods employed in the analysis of human body motion. Some of the recent works on human action classification includes [30–33], and for human activity recognition includes [34–37]. However, they are only capable of detecting complete human action. In the case of early human action detection, it is essential to detect partial action, as the concern is to recognize the activity being performed as soon as possible. Another limitation of these works is their inability to handle the uncertainties that exist in a real-world environment, which is taken into account by fuzzy approaches.

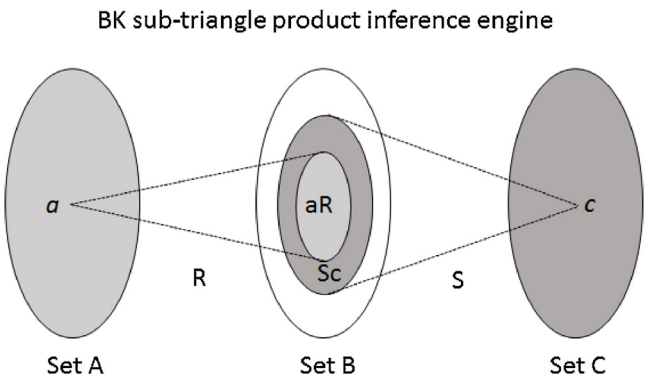

Fig. 2. Overview of BK subproduct: element a in set A is in relation with element c in set C if its image under R (aR) is a subset of image Sc.

2.2. Fuzzy human action recognition

In recent times, the fuzzy approaches such as type-1 fuzzy inference system [38,39], fuzzy HMM [40] and hybrid techniques [14,15],have proven to be beneficial in human action recognition. Fuzzy human action recognition techniques can be efficiently used to distinguish the human motion patterns, and recognize the human activities with their capability to model the uncertainties involved therein. Nonetheless, fuzzy vector quantization [41] and qualitative normalized template [24,42] provide the capability to handle the complex human activities occurring in everyday life. However, these approaches are tailor-made for human action recognition and classification tasks only, and lacking in ability to detect an action early.

BK relational products have been successfully employed in developing the inference engine for several applications such as in the medical expert system [43], information retrieval [44], autonomous underwater vehicles' path navigation [19], land evaluation [20], scene classification [21,45], etc. In this paper, a hybrid technique of fuzzy BK subproduct and the computer vision solutions is employed for human action recognition. In order to provide a better understanding of the concept, the following section revisits BK subproduct.

2.2.1. BK subproduct revisit

Bandler and Kohout [17] proposed that the relationship between two indirectly associated sets can be studied with the BK relational product that defines the relationship between the elements within the two indirectly associated sets as the overlapping of their images in a common set. Fig. 2 gives an overview of the BK subproduct for crisp relations. Let us assume that there exist three sets: set $A = \{a_i | i = 1, \cdots, I\}$, set $B = \{b_j | j = 1, \cdots, J\}$ and set $C = \{c_k | k = 1, \cdots, K\}$. If a relation R is defined between A and B such that $R \subseteq \{(a, b) | (a, b) \in A \times B\}$, and a relation S is defined between B and C such that $S \subseteq \{(b, c) | (b, c) \in B \times C\}$, then the BK subproduct can be defined as:

$$R \lhd S = \{(a, c) | (a, c) \in A \times C \quad \text{and} \quad aR \subseteq Sc\} \tag{1}$$

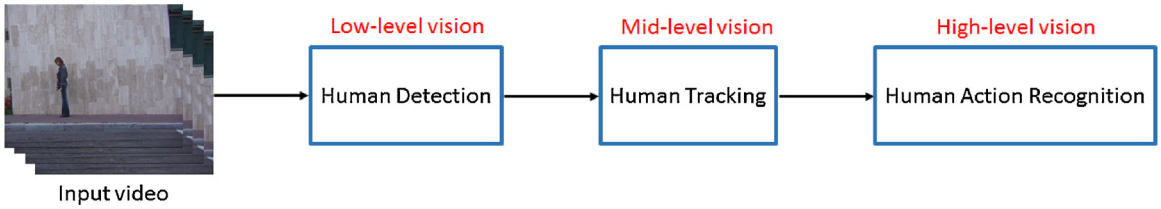

Fig. 1. A general framework for human action recognition.