Effective and efficient human action recognition using dynamic frame skipping and trajectory rejection

Jeong-Jik Seo, Hyung-Il Kim, Wesley De Neve, Yong Man Ro

PII: S0262-8856(16)30100-7
DOI: doi: 10.1016/j.imavis.2016.06.002
Reference: IMAVIS 3517

To appear in: *Image and Vision Computing*

Received date: 1 October 2015
Revised date: 17 April 2016
Accepted date: 10 June 2016

Please cite this article as: Jeong-Jik Seo, Hyung-Il Kim, Wesley De Neve, Yong Man Ro, Effective and efficient human action recognition using dynamic frame skipping and trajectory rejection, *Image and Vision Computing* (2016), doi: 10.1016/j.imavis.2016.06.002

Draft for the preparation of submission to Image and Vision Computing

Effective and Efficient Human Action Recognition using Dynamic Frame Skipping and Trajectory Rejection

Jeong-Jik Seo[a], Hyung-Il Kim[a], Wesley De Neve[a,b], and Yong Man Ro[a,*]
[a]Image and Video Systems Lab, School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea
[b]Multimedia Lab, Ghent University-iMinds, Belgium
*Corresponding author: Yong Man Ro
Email: ymro@kaist.ac.kr, Tel: +82-42-350-3494, Fax: +82-42-350-5494

Abstract

Human action recognition (HAR) is a core technology for human-computer interaction and video understanding, attracting significant research and development attention in the field of computer vision. However, in uncontrolled environments, achieving effective HAR is still challenging, due to the widely varying nature of video content. In previous research efforts, trajectory-based video representations have been widely used for HAR. Although these approaches show state-of-the-art HAR performance for various datasets, issues like a high computational complexity and the presence of redundant trajectories still need to be addressed in order to solve the problem of real-world HAR. In this paper, we propose a novel method for HAR, integrating a technique for rejecting redundant trajectories that are mainly originating from camera movement, without degrading the effectiveness of HAR. Furthermore, in order to facilitate efficient optical flow estimation prior to trajectory extraction, we integrate a technique for dynamic frame skipping. As a result, we only make use of a small subset of the frames present in a video clip for optical flow estimation. Comparative experiments with five publicly available human action datasets show that the proposed method outperforms state-of-the-art HAR approaches in terms of effectiveness, while simultaneously mitigating the computational complexity.

Index Terms

Frame skipping, human action recognition (HAR), motion descriptor, motion trajectory, optical flow