PLOS | ONE

RESEARCH ARTICLE

Fast Fight Detection

Ismael Serrano Gracia[1]*, Oscar Deniz Suarez[1]*, Gloria Bueno Garcia[1]*, Tae-Kyun Kim[2]

1 Department of Systems Engineering and Automation, E.T.S.I. Industriales, Ciudad Real, Castilla-La Mancha, Spain, **2** Department of Electrical and Electronic Engineering, Imperial College, London, UK

* ismael.serrano@uclm.es (ISG); oscar.deniz@uclm.es (ODS); gloria.bueno@uclm.es (GBG)

OPEN ACCESS

Citation: Serrano Gracia I, Deniz Suarez O, Bueno Garcia G, Kim T-K (2015) Fast Fight Detection. PLoS ONE 10(4): e0120448. doi:10.1371/journal.pone.0120448

Academic Editor: Gonzalo G. de Polavieja, Cajal Institute, Consejo Superior de Investigaciones Científicas, SPAIN

Received: September 19, 2014

Accepted: January 22, 2015

Published: April 10, 2015

Data Availability Statement: All relevant data were obtained from third parties, and can be accessed at the following URLs: http://crcv.ucf.edu/data/UCF101.php (UCF-101 dataset); http://visilab.etsii.uclm.es/personas/oscar/FightDetection/index.html (Hockey dataset); http://visilab.etsii.uclm.es/personas/oscar/FightDetection/index.html (Movies dataset).

Funding: This work has been partially supported by Project TIN2011-24367 from Spain's Ministry of Economy and Competitiveness. Also it has been partially supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 643924. The funders had no role

Abstract

Action recognition has become a hot topic within computer vision. However, the action recognition community has focused mainly on relatively simple actions like clapping, walking, jogging, etc. The detection of specific events with direct practical use such as fights or in general aggressive behavior has been comparatively less studied. Such capability may be extremely useful in some video surveillance scenarios like prisons, psychiatric centers or even embedded in camera phones. As a consequence, there is growing interest in developing violence detection algorithms. Recent work considered the well-known Bag-of-Words framework for the specific problem of fight detection. Under this framework, spatio-temporal features are extracted from the video sequences and used for classification. Despite encouraging results in which high accuracy rates were achieved, the computational cost of extracting such features is prohibitive for practical applications. This work proposes a novel method to detect violence sequences. Features extracted from motion blobs are used to discriminate fight and non-fight sequences. Although the method is outperformed in accuracy by state of the art, it has a significantly faster computation time thus making it amenable for real-time applications.

Introduction

In the last few years, the problem of human action recognition from video has become tractable by using computer vision techniques, see surveys [1], [2], [3]. Within this topic, there is a vast literature in which experimental results are given for recognition of human actions like walking, jumping or hand waving [4]. However, action detection has been devoted less effort. Action detection is a related task in which only a specific action must be detected. Action detection may be of direct use in real-life applications, fight detection being a clear example. Whereas there is a number of well-studied datasets for action recognition, significant datasets with violent actions (fights) have not been made available until the work [5]. A violence detector has, however, immediate applicability in the surveillance domain. The primary function of large-scale surveillance systems deployed in institutions such as schools, prisons and psychiatric care facilities is for alerting authorities to potentially dangerous situations. However, human operators are overwhelmed with the number of camera feeds and manual response times are slow, resulting in a strong demand for automated alert systems. Similarly, there is increasing

19. Demarty C, Penet C, Gravier G, Soleymani M (2012) MediaEval 2012 affect task: Violent scenes detection in Hollywood movies. MediaEval 2012 Workshop. Pisa, Italy.
20. Goto S, Aoki T (2014) Violent scenes detection using mid-level violence clustering. Computer Science. CSCP. pp. 283296
21. Deniz O, Serrano I, Bueno G, Tae-Tyun K (2014) Fast *violence detection in video.* The 9th International Conference on Computer Vision Theory and Applications (VISAPP).
22. Chen M, Mummert L, Pillai P, Hauptmann A, Sukthankar R. (2010) *Exploiting multi-level parallelism for low-latency activity recognition in streaming video.* MMSys '10: Proceedings of the first annual ACM SIGMM conference on Multimedia systems. New York, NY, USA: ACM, pp. 1-12.
23. Patnaik S, Yang Y (2012) Soft Computing Techniques in Vision Science. Springer.
24. Soomro K, Zamir A, Shah M (2012) A dataset of 101 human action classes from videos in the wild. CRCV TR. Technical report
25. Wang H, Klaser A, Schmid C, Cheng-Lin L (2011) *Action Recognition by Dense Trajectories.* IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, United States, pp. 3169-3176. MSR - INRIA.
26. Wang K, Zhang Z, Wang L (2012) Violence video detection by discriminative slow feature analysis. Pattern Recognition, Springer Berlin Heidelberg, volume 321. pp. 137–144.
27. Ward R, Guha T (2012) Learning sparse representations for human action recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 34: 1576–1588. doi: 10.1109/TPAMI.2011.253 PMID: 22745001