

Accepted Manuscript

Learning to detect video events from zero or very few video examples

Christos Tzelepis, Damianos Galanopoulos, Vasileios Mezaris, Ioannis Patras

PII: S0262-8856(15)00111-0
DOI: doi: [10.1016/j.imavis.2015.09.005](https://doi.org/10.1016/j.imavis.2015.09.005)
Reference: IMAVIS 3434

To appear in: *Image and Vision Computing*

Received date: 29 March 2015
Revised date: 9 September 2015
Accepted date: 28 September 2015



Please cite this article as: Christos Tzelepis, Damianos Galanopoulos, Vasileios Mezaris, Ioannis Patras, Learning to detect video events from zero or very few video examples, *Image and Vision Computing* (2015), doi: [10.1016/j.imavis.2015.09.005](https://doi.org/10.1016/j.imavis.2015.09.005)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Learning to detect video events from zero or very few video examples

Christos Tzelepis^{a,b}, Damianos Galanopoulos^a, Vasileios Mezaris^a, Ioannis Patras^b

^a*Information Technologies Institute (ITI), CERTH, Thessaloniki, 57001, Thessaloniki, Greece*

^b*Queen Mary, University of London, Mile End Campus, UK, E14NS*

Abstract

In this work we deal with the problem of high-level event detection in video. Specifically, we study the challenging problems of i) learning to detect video events from solely a textual description of the event, without using any positive video examples, and ii) additionally exploiting very few positive training samples together with a small number of “related” videos. For learning only from an event’s textual description, we first identify a general learning framework and then study the impact of different design choices for various stages of this framework. For additionally learning from example videos, when true positive training samples are scarce, we employ an extension of the Support Vector Machine that allows us to exploit “related” event videos by automatically introducing different weights for subsets of the videos in the overall training set. Experimental evaluations performed on the large-scale TRECVID MED 2014 video dataset provide insight on the effectiveness of the proposed methods.

Keywords: video event detection, textual event description, zero positive examples, few positive examples, related videos

Email addresses: tzelepis@iti.gr (Christos Tzelepis), dgalanop@iti.gr (Damianos Galanopoulos), bmezaris@iti.gr (Vasileios Mezaris), i.patras@qmul.ac.uk (Ioannis Patras)

Download English Version:

<https://daneshyari.com/en/article/4969075>

Download Persian Version:

<https://daneshyari.com/article/4969075>

[Daneshyari.com](https://daneshyari.com)