Accepted Manuscript

Next-Active-Object Prediction from Egocentric Videos

Antonino Furnari, Sebastiano Battiato, Kristen Grauman, Giovanni Maria Farinella

PII: S1047-3203(17)30196-7
DOI: https://doi.org/10.1016/j.jvcir.2017.10.004
Reference: YJVCI 2071

To appear in: *J. Vis. Commun. Image R.*

Revised Date: 18 August 2017
Accepted Date: 11 October 2017

Please cite this article as: A. Furnari, S. Battiato, K. Grauman, G.M. Farinella, Next-Active-Object Prediction from Egocentric Videos, *J. Vis. Commun. Image R.* (2017), doi: https://doi.org/10.1016/j.jvcir.2017.10.004

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.

ACCEPTED MANUSCRIPT

Next-Active-Object Prediction from Egocentric Videos

Antonino Furnari*, Sebastiano Battiato

University of Catania - Department of Mathematics and Computer Science

Kristen Grauman

The University of Texas at Austin - Computer Science Department

Giovanni Maria Farinella

University of Catania - Department of Mathematics and Computer Science

Abstract

Although First Person Vision systems can sense the environment from the user's perspective, they are generally unable to predict his intentions and goals. Since human activities can be decomposed in terms of atomic actions and interactions with objects, intelligent wearable systems would benefit from the ability to anticipate user-object interactions. Even if this task is not trivial, the First Person Vision paradigm can provide important cues to address this challenge. We propose to *exploit the dynamics of the scene to recognize next-active-objects before an object interaction begins*. We train a classifier to discriminate trajectories leading to an object activation from all others and forecast next-active-objects by analyzing fixed-length trajectory segments within a temporal sliding window. The proposed method compares favorably with respect to several baselines on the Activity of Daily Living (ADL) egocentric dataset comprising 10 hours of videos acquired by 20 subjects while performing unconstrained interactions with several objects.

Keywords: egocentric vision, forecasting, object interaction,

*Corresponding author

Email addresses: furnari@dmi.unict.it (Antonino Furnari), battiato@dmi.unict.it (Sebastiano Battiato), grauman@cs.utexas.edu (Kristen Grauman), gfarinella@dmi.unict.it (Giovanni Maria Farinella)

Download English Version:

https://daneshyari.com/en/article/6938428

Download Persian Version:

https://daneshyari.com/article/6938428

Daneshyari.com