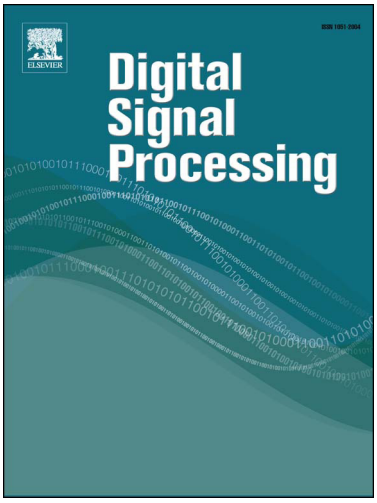
Accepted Manuscript



Visual tracking using global sparse coding and local convolutional features

Xianyou Zeng, Long Xu, Lin Ma, Ruizhen Zhao, Yigang Cen

PII: S1051-2004(17)30232-4
DOI: https://doi.org/10.1016/j.dsp.2017.10.007
Reference: YDSPR 2208

To appear in: *Digital Signal Processing*

Please cite this article in press as: X. Zeng et al., Visual tracking using global sparse coding and local convolutional features, *Digit. Signal Process.* (2017), https://doi.org/10.1016/j.dsp.2017.10.007

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

Visual tracking using global sparse coding and local convolutional features

Xianyou Zeng[a,b], **Long Xu**[c], **Lin Ma**[d], **Ruizhen Zhao**[a,b], **Yigang Cen**[a,b]
[a]Institute of Information Science, Beijing Jiaotong University, Beijing, 100044 China
[b]Key Laboratory of Advanced Information Science and Network Technology of Beijing, Beijing, China
[c]Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, 100012 China
[d]Tencent AI Lab, Shenzhen, 518052 China

Abstract. Visual tracking is a challenging task in many computer vision applications due to factors such as occlusion, scale variations, background clutter, and so on. In this paper, we present a robust tracking algorithm by representing the target at two levels: global and local levels. Accordingly, the tracking algorithm is composed of two parts: global and local parts. The global part is a discriminative model which separates the foreground object from the background based on holistic features. In the local part, we explore the target's local representation by a set of filters convolving the target region at each position. Then, the global part and local part are integrated into a collaborative model to construct the final tracker. Experiments on the tracking benchmark dataset with 50 challenging videos demonstrate the robustness and effectiveness of the proposed algorithm, outperforming several state-of-the-art models.

Keywords: Visual tracking, Sparse representation, Local convolutional feature, Collaborative model.

*Address all correspondence to: Long Xu, lxu@nao.cas.cn

1 Introduction

Visual tracking is one of the most important research topics in multimedia processing and has been widely used in human behavior analysis, video surveillance, security, military, transportation, aerospace, and so on. Although significant progress has been made in the past years,[1–7] tracking still remains a challenging task, since only ground-truth in the first frame can be used and the target may undergo many challenges, such as illumination change, partial occlusion, pose variation, and shape deformation, etc. All these challenges may result in failed tracking.

To address the above challenges for robust tracking, various representation schemes are introduced into tracking task, such as pixel-based tracker,[38] feature-based trackers (e.g. Haar-like features,[2,4,5] HOG descriptors[7,8]), sparse-based trackers,[10,31–37] subspace representation based trackers[29,30] and multilevel quantization tracker.[39]

Download English Version:

https://daneshyari.com/en/article/6951903

Download Persian Version:

https://daneshyari.com/article/6951903

Daneshyari.com