Robust Object Tracking Based on Adaptive Templates Matching via the Fusion of Multiple Features

Zhiyong Li, Song Gao, Ke Nai

PII: S1047-3203(17)30011-1
DOI: http://dx.doi.org/10.1016/j.jvcir.2017.01.012
Reference: YJVCI 1932

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

Received Date: 18 June 2016
Revised Date: 7 November 2016
Accepted Date: 11 January 2017

Please cite this article as: Z. Li, S. Gao, K. Nai, Robust Object Tracking Based on Adaptive Templates Matching via the Fusion of Multiple Features, *J. Vis. Commun. Image R.* (2017), doi: http://dx.doi.org/10.1016/j.jvcir.2017.01.012

Robust Object Tracking Based on Adaptive Templates Matching via the Fusion of Multiple Features

Zhiyong Li[1,2]*, Song Gao[1,2], Ke Nai[1,2]
[1] College of Computer Science and Electronic Engineering, Hunan University,
[2] Key Laboratory for Embedded and Network Computing of Hunan Province,
Changsha, China
*zhiyong.li@hnu.edu.cn

Abstract: Moving object tracking under complex scenes remains to be a challenging problem because the appearance of a target object can be drastically changed due to several factors, such as occlusions, illumination, pose, scale change and deformation. This study proposes an adaptive multi–feature fusion strategy, in which the target appearance is modeled based on timed motion history image with HSV color histogram features and edge orientation histogram features. The variances based on the similarities between the candidate patches and the target templates are used for adaptively adjusting the weight of each feature. Double templates matching, including online and offline template matching, is adopted to locate the target object in the next frame. Experimental evaluations on challenging sequences demonstrate the accuracy and robustness of the proposed algorithm in comparison with several state-of-the-art algorithms.

Index terms: Visual tracking, Feature fusion, Double templates matching, Timed motion history image

1 Introduction

Object tracking [1, 2] is one of the most active research topics in computer vision because of its importance in applications, such as automated video surveillance [3], video indexing, traffic monitoring [4], and human-computer interaction [5]. Although object tracking has been studied in the past decades and numerous algorithms have been proposed [6-23], it remains a very challenging problem to design a robust and efficient tracking system since the appearance of a target object can be drastically changed due to several factors, such as pose variations, illumination changes, partial or full occlusion, abrupt motion, and background clutter. Therefore, how to design a robust appearance model that can adaptively handle these factors mentioned above is a key task in most recently proposed algorithms [10, 13, 14, 17, 18, 23, 45, 46].

To overcome these difficulties, many object tracking algorithms have been proposed in recent years. Compressive tracking (CT) [23] proposed an appearance model based on non-adaptive random projections that preserved the structure of the original image space. A very sparse measurement matrix was adopted to efficiently compress the features from the foreground and background targets. Furthermore, multiple instance learning (MIL) [10] had also been applied to object tracking, where all ambiguous positive and negative samples were put into bags to learn a