



Online parameter tuning for object tracking algorithms[☆]



Duc Phu Chau^{*}, Monique Thonnat, François Brémond, Etienne Corvée

STARS Team, INRIA Sophia Antipolis-Méditerranée, 2004 Route des Lucioles, 06902 Sophia Antipolis Cedex, France

ARTICLE INFO

Article history:

Received 28 October 2012
Received in revised form 20 January 2014
Accepted 13 February 2014
Available online 21 February 2014

Keywords:

Object tracking
Online parameter tuning
Controller
Self-adaptation
Machine learning

ABSTRACT

Object tracking quality usually depends on video scene conditions (e.g. illumination, density of objects, object occlusion level). In order to overcome this limitation, this article presents a new control approach to adapt the object tracking process to the scene condition variations. More precisely, this approach learns how to tune the tracker parameters to cope with the tracking context variations. The tracking context, or context, of a video sequence is defined as a set of six features: density of mobile objects, their occlusion level, their contrast with regard to the surrounding background, their contrast variance, their 2D area and their 2D area variance. In an offline phase, training video sequences are classified by clustering their contextual features. Each context cluster is then associated to satisfactory tracking parameters. In the online control phase, once a context change is detected, the tracking parameters are tuned using the learned values. The approach has been experimented with three different tracking algorithms and on long, complex video datasets. This article brings two significant contributions: (1) a classification method of video sequences to learn offline tracking parameters and (2) a new method to tune online tracking parameters using tracking context.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Mobile object tracking plays an important role in an increasing number of computer vision applications (e.g. home care, sport scene analysis and visual surveillance). The object trajectories are useful for activity recognition, learning of interest zones or paths in a scene and detection of events of interest. Unfortunately the tracking quality depends on many factors: the quality of vision tasks performed at lower levels such as object detection, object classification, and by some video features such as complexity of object movements, scene illumination intensity, low contrast, high density and occlusion frequency of mobile objects. In particular, for a long video sequence (i.e. several hours) in which the variations of these properties happen frequently, the tracking quality is still an issue. The problems we focus on are the following: How can an automatic system robustly track mobile objects in different conditions and situations such as the ones cited above. And in those complex cases, how can the user regulate the tracking parameters to get an optimal tracking quality?

In order to answer these two questions, we propose in this article a new method for controlling tracking algorithms. The objective of the proposed method is to define an automatic control algorithm which is able to adapt online the tracking task to the scene variations in a video sequence

by tuning the tracking parameters over time. We aim to build a control algorithm which is: *generic*, *flexible* and *intelligent*. The term “generic” means that our method can handle different tracking algorithm categories. In this work, our objective is to control tracking algorithms which rely on object appearance or points of interest. These algorithms are selected because their approaches are largely studied in the state of the art. The term “flexible” implies that the structure of the proposed control algorithm can be adapted for handling other tracking algorithm category (e.g. object silhouette-based tracking). The term “intelligent” means that this approach requires less human interaction than the control methods in the state of the art.

1.1. Hypotheses

The control method presented in this manuscript relies on the two following hypotheses:

1. The considered tracking algorithms have at least one tunable parameter which influences significantly the tracking quality.
2. There exist a number of contexts which have an impact on the tracking quality. Let g be a function mapping a video v_i to its context. For a tracking algorithm \mathcal{T} , we suppose that there exists a function $f_{\mathcal{T}}$ mapping a video context to satisfactory tracking parameter values (i.e. parameter values for which the tracking quality is greater than a predefined threshold s):

$$\forall v_i, \exists f : |Q(\mathcal{O}_{\mathcal{T}}(f_{\mathcal{T}} \circ g(v_i)), G_{v_i})| > s \quad (1)$$

[☆] This paper has been recommended for acceptance by Richard Bowden, PhD.

^{*} Corresponding author. Tel.: +33 49715 5310.

E-mail addresses: Duc-Phu.Chau@inria.fr (D.P. Chau), Monique.Thonnat@inria.fr (M. Thonnat), Francois.Bremond@inria.fr (F. Brémond), Etienne.Corvee@inria.fr (E. Corvée).

where G_{v_i} represents the tracking ground-truth data of video v_i ; $\mathcal{O}_{\mathcal{T}}(\cdot)$ represents the output of tracker \mathcal{T} corresponding to given parameter values; $Q(\mathcal{O}_{\mathcal{T}}(\cdot), G_{v_i})$ represents the quality of tracker \mathcal{T} output compared to the tracking ground-truth data of video v_i .

Let ϵ_1 and ϵ_2 be predefined thresholds. The function f is assumed to satisfy the following property if the temporal lengths of v_1 and v_2 are short enough (lower than 50 frames):

$$\forall v_1, v_2 : \text{if } |g(v_1) - g(v_2)| < \epsilon_1 \quad (2)$$

$$\Rightarrow |Q(\mathcal{O}_{\mathcal{T}}(f_{\mathcal{T}} \circ g(v_1)), G_{v_1}) - Q(\mathcal{O}_{\mathcal{T}}(f_{\mathcal{T}} \circ g(v_2)), G_{v_2})| < \epsilon_2. \quad (3)$$

This hypothesis means that if the contexts of two videos v_1 and v_2 are close enough, the tracking performances for v_1 and v_2 corresponding to their satisfactory tracking parameter values are also close enough.

The hypothesis 2 is given for two objectives. First, we can compute the satisfactory tracking parameter values for a video context cluster using satisfactory parameters of contexts (see Section 3.4.2). Second, the satisfactory tracking parameters for context clusters can be used for tuning online the tracking parameters (see Section 4.2).

1.2. Article organization

This article is organized as follows. Section 2 presents a state of the art on control methods. Section 3, entitled “offline learning phase”, details a scheme to learn satisfactory tracking parameters for each video context cluster. Section 4 describes the online parameter tuning process. Section 5 is dedicated to the experimentation and validation of the proposed method. Section 6 presents concluding remarks as well as future work.

2. State of the art

Many approaches have been proposed to track mobile objects in a scene [37]. Depending on taxonomy criteria, the trackers can be classified into different categories. Fig. 1 presents a taxonomy example (the red ellipses mark the tracker categories controlled by the proposed method). However the quality of tracking algorithms always depends on scene properties such as: mobile object density, contrast intensity, scene depth and object size. The selection of a tracking algorithm for an unknown scene becomes a hard task. Even when the

tracker has already been determined, it is difficult to tune online its parameters to get a high performance.

The idea about an automatic control to adapt the performance of a system to the problem of scene variations has already been studied. However some methods limit their studies to static image and not to video processing. For example the authors in Ref. [34] present a framework which is able to integrate expert knowledge and uses it to control the image processing programs. The framework is experimented on three different applications: road obstacle detection, medical imaging and astronomy. By considering both context and evaluation criteria, the system can find the best algorithm among a predefined algorithm set and tune its parameters to obtain the best possible performance. However, the construction of a knowledge base for this system requires a lot of time and data.

The authors in Ref. [17] present a controlled video understanding system based on a knowledge base. The system is composed of three main components in which the control component performs several steps for managing all the online processes of the system (e.g. program execution and automatic parameter tuning). Different rules are defined in this component based on user goal, contextual information and evaluation results. However their approach does not address directly the tracking task.

Some methods have addressed the tracking parameter tuning, however their approaches require too strong hypotheses and expert knowledge. For example, the author in Ref. [30] proposes an approach to tune automatically tracking algorithm parameters. In this approach, the tracker quality is represented as a function of tuned parameters. The author supposes that this function has no local optimal solutions. Using this hypothesis, for each parameter and a training video, the author determines its optimal value thanks to expert knowledge. Then the parameter tendency (i.e. increase or decrease) for converging to the optimal value is learned in function of the tracker input and output. This learned parameter tendency is used in the online phase to tune automatically the corresponding parameter to improve the tracking performance. In Ref. [9], the authors compare the tracker results with corresponding ground-truth data to determine the importance of each parameter for each context and to exploit the influence of each parameter variation on tracker performance. The authors suppose that parameter variations are independent. This is a strict hypothesis because the parameters are usually dependent on each other. In Ref. [10], the authors propose a tracking algorithm whose parameters can be learned offline for each tracking context. However the authors suppose that the context

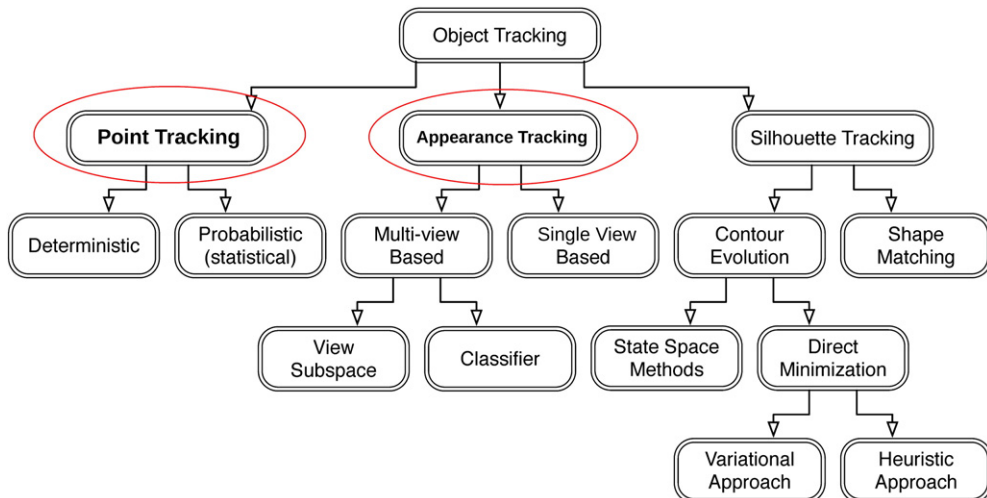


Fig. 1. Taxonomy of tracking methods (adapted from Ref. [37]). The red ellipses mark the tracker categories controlled by our method.

Download English Version:

<https://daneshyari.com/en/article/526983>

Download Persian Version:

<https://daneshyari.com/article/526983>

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