Automatica 90 (2018) 271–285

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Image feedback based optimal control and the value of information in a differential game[☆]

Vladimir Macias [a], Israel Becerra [b], Rafael Murrieta-Cid [a,*], Hector M. Becerra [a], Seth Hutchinson [b]

[a] Centro de Investigación en Matemáticas, CIMAT, Guanajuato México, Mexico
[b] University of Illinois at Urbana-Champaign, Urbana, IL, USA

ARTICLE INFO

Article history:
Received 24 November 2016
Received in revised form 28 June 2017
Accepted 30 November 2017
Available online 15 February 2018

Keywords:
Pursuit/Evasion
Optimal control
Projective geometry
Feedback
Motion strategies

ABSTRACT

In this paper, we address pursuit-evasion problems in which the pursuer is a Differential Drive Robot (DDR) that attempts to capture an omnidirectional evader. From the Nash property it follows that if the evader deviates from its maximum potential speed then the capture time shall not increase for a pursuer that does not deviate from its Nash equilibrium motion strategy. However, it is not immediately clear how the pursuer could exploit that evader's deviation from its maximum potential speed, which might correspond to situations where the evader's capabilities may degrade with time, for example, battery depletion in an autonomous vehicle, or fatigue in an animal evader. This can be considered as a scenario of an evader in which the set of admissible controls varies with time. In the present paper we consider such scenario. In our first result, we propose an alternative strategy for the pursuer, which, for certain scenarios, further reduces the capture time compared to the strategy based on the maximum potential evader's speed. In our second result, we show that, under non-anticipative strategies, a pursuer strategy that uses the instantaneous evader speed alone, *does not always guarantee* to improve the payoff for the pursuer, nor the capture of the evader. Hence, we conclude that the evader's location is the relevant information for the pursuer to know. Later, we present vision-based control laws that implement the optimal pursuer strategy. The optimal pursuer strategy is characterized by a partition of the reduced space (a representation of the game in the pursuer's body-attached coordinate system) in which each region maps to an optimal pursuer action. We consider the case for which the pursuer is equipped with an omnidirectional catadioptric camera. Finally, in our third result we show that the location of the evader on the image can be directly used by the pursuer to define its motion strategy, in spite of the distortion of the state space suffered on the image. That is, the pursuer is able to apply its motion strategy using the image without explicitly reconstructing the evader's position. This approach is computationally efficient, and robust to occlusions and noise in the image.

1. Introduction

In this paper, we consider the pursuit-evasion problem of capturing an omnidirectional evader using a Differential Drive Robot (DDR) in an obstacle-free environment. More specifically, given an initial condition in which the evader is at distance L from the pursuer, the pursuer's goal is to reduce this distance to $L < l$ (the capture condition) as quickly as possible, while the evader's goal is to delay capture for as long as possible. The classical solution to this problem yields so-called Nash equilibrium strategies. While it is well known that neither player can improve its guaranteed payoff by unilaterally deviating from its Nash strategy, it is not always immediately clear how one player can exploit deviation from the Nash strategy by the other player. In this paper, we address this issue. We then present a vision-based control law that implements the optimal pursuer strategy (or policy).

The results in this paper are related to previous work presented in Jacobo, Ruiz, Murrieta-Cid, Becerra, and Marroquin (2015) and Ruiz, Murrieta-Cid, and Marroquin (2013). In Ruiz et al. (2013), the optimal strategies for each player are expressed in terms of a

---

[☆] This work was partially funded by CONACYT projects 220796 and 264896, and post-doctoral fellowship 277028. The authors would also like to acknowledge the financial support of Intel Corporation for the development of this work. The material in this paper was partially presented at the IEEE International Conference on Robotics and Automation, May 26–May 30, 2015, Washington, USA. This paper was recommended for publication in revised form by Associate Editor C.C. Cheah under the direction of Editor Thomas Parisini.

* Corresponding author.
E-mail addresses: maciasandoval@cimat.mx (V. Macias), israelb4@illinois.edu (I. Becerra), murrieta@cimat.mx (R. Murrieta-Cid), hector.becerra@cimat.mx (H.M. Becerra), seth@illinois.edu (S. Hutchinson).

https://doi.org/10.1016/j.automatica.2017.12.045
0005-1098/© 2017 Elsevier Ltd. All rights reserved.

partition of the playing space into disjoint regions, and open-loop, time optimal strategies of the players are defined for each region. A feedback control strategy to implement the optimal pursuer strategy was presented in Jacobo et al. (2015). For this approach, computer vision-based state estimation was implemented using the 1D trifocal tensor (1D because only bearing information is used to compute it). The approach of Jacobo et al. (2015) was motivated by the possibility that an evader could avoid capture in cases where the pursuer executed its Nash strategy in open loop; nevertheless, Jacobo et al. (2015) did not address the issue of how the pursuer could exploit deviations by the evader from its maximum potential speed, which is addressed in this paper. Further, the approach of Jacobo et al. (2015) relied on position-based visual servo methods, which are known to be sensitive to state estimation errors or calibration inaccuracies (Chaumette & Hutchinson, 2006).

In this paper, we consider non-anticipative strategies (Elliott & Kalton, 1972) for both players (each player has complete up-to-date information concerning the control functions employed by the other player, however, it does not know the controls that the other player will apply in the future) and we investigate the scenario of an evader in which the set of admissible controls varies with the time. Namely, we consider an evader whose maximum speed varies as time elapses; to the best of our knowledge this problem has not been studied before in the context of differential games. We assume that the maximum potential speed of the evader V_e^{max} is known for the pursuer before the game commences. We consider $V_e(t)$ as the instantaneous maximum speed at which the evader can travel as time elapses, and V_e^{max} as the upper bound for that speed for all t. In Section 5 we investigate pursuer strategies exploiting deviations by the evader from its maximum potential speed and in Section 6 we present a vision-based control law that implements the optimal pursuer strategy. More precisely, in Section 5 we show that under a non-anticipative strategies framework, using the instantaneous evader speed *does not always guarantee* to improve the payoff for the pursuer, nor the capture of the evader, hence, the only information required for the pursuer is the evader's location. Based on this result, in Section 6, we show how to obtain this information, i.e., the evader's location directly in an image, without the estimation of the evader's state on the state space. Therefore, the connection between Section 5 and Section 6 consists in first proving that the evader's location is the relevant information for the pursuer and then, in order to retrieve the evader's location, we use the projection from the state space to the image space and we prove that under this projection, a partition of the state space defining the pursuer strategy can be mapped to the image space *in spite of* the distortion of the state space suffered on the image, allowing immediate determination of the optimal pursuer action once the evader is detected in the image. This approach is computationally efficient, and robust to occlusions and noise.

A preliminary version of a portion of Section 5 of the present work appeared in Becerra, Macias, and Murrieta-Cid (2015). The main results of the present work are the following:

- From the Nash property it follows that if the evader deviates from its maximum potential speed V_e^{max}, then the capture time shall not increase for a pursuer that follows a motion strategy generated by V_e^{max} (we refer to this pursuer strategy as $\Pi_{P(V_e^{max})}$). However, Theorem 1 proposes another strategy for the pursuer, called $\Pi_{P(V_e)}$, which, for certain scenarios, further reduces the capture time compared to $\Pi_{P(V_e^{max})}$.
- Theorem 2, which shows that under non-anticipative strategies, using instantaneously the partition obtained based on the instantaneous evader speed *does not always guarantee* to improve the payoff for the pursuer, nor the capture of the evader. In Lemma 2 (which is used in Theorem 2) we *exhibit* a case in which the evader escapes if the pursuer uses the strategy based on the instantaneous evader speed, that is strategy $\Pi_{P(V_e)}$. Therefore, under non-anticipative strategies, the evader's instantaneous speed cannot be used alone to improve the payoff and the pursuer must stick to the worst case corresponding to assuming that the evader moves at V_e^{max}.
- Finally, Theorem 3 shows that the location of the evader on the image can be directly used by the pursuer to define its motion strategy. That is, the pursuer is able to apply its motion strategy using the image without explicitly reconstructing the evader's position.

The remainder of this paper is organized as follows. In Section 2, we provide a review of related work. In Section 3, we give a formal description of the problem, and in Section 4 we describe the motion strategies for both pursuer and evader. In Section 5, we present the primary theoretical results in the paper, concerning the role of information in the optimality of Nash pursuer strategies when the evader moves at suboptimal speed. Finally, in Section 6, we derive image-based control strategies that implement the optimal pursuer strategy.

2. Related work

Our work is related to optimal control methods used in robotics, for instance Balkcom and Mason (2002), Soueres and Laumond (1996) and Wang, Chen, and Soueres (2009), however those methods typically execute the motion in open loop. Our work proposes a state feedback-based motion strategy, but using information directly from an image. Our work is also related to image-based visual servo (Chaumette & Hutchinson, 2006; Lopez-Nicolas, Gans, Bhattacharya, Sagues, & Hutchinson, 2010), in the sense that the feedback is directly based on an image, however, in contrast to the classical image-based visual servoing approach, in our proposed approach, the goal for the robot is not to see a target image, but instead its objective is to bring the evader to a specific locus of points called the *usable part* (Isaacs, 1965).

The problem addressed in this paper is a pursuit-evasion game. There has been a considerable amount of research in the area of pursuit and evasion, particularly in the area of control (Başar & Olsder, 1999; Isaacs, 1965; Merz, 1971). The pursuit-evasion problem can be framed as a problem in noncooperative dynamic game theory (Başar & Olsder, 1999).

A pursuit-evasion game can be defined in several ways. One variant considers one or more pursuers, which are given the task of *finding* an evader in an environment with obstacles (Guibas, Latombe, LaValle, Lin, & Motwani, 1999; Hollinger, Singh, Djugash, & Kehagias, 2009; Tovar & LaValle, 2008; Vidal, Shakernia, Jin, Hyunchul, & Sastry, 2002). A recent survey of this kind of problem is presented in Chung, Hollinger, and Isler (2011).

Other variant consists of *maintaining visibility of a moving evader* also in an environment with obstacles (Bandyopadhyay, Ang, & Hsu, 2007; Bhattacharya & Hutchinson, 2010; Jung & Sukhatme, 2002; LaValle, González-Baños, Becker, & Latombe, 1997; Murrieta-Cid, Muppirala, Sarmiento, Bhattacharya, & Hutchinson, 2007; O'Kane, 2008). Game theory is proposed in LaValle et al. (1997) as a framework to formulate the tracking problem, and an online algorithm is presented. In Bhattacharya and Hutchinson (2010), the authors address the problem of maintaining visibility of the evader as a *game of degree* (i.e. the emphasis is over the optimization of a given criterion and not over the problem of deciding what player is the winner). The pursuer and the evader are omnidirectional (holonomic) systems. In Bhattacharya and Hutchinson (2011), the problem of maintaining visibility of