Acta Astronautica 122 (2016) 209–218

Contents lists available at ScienceDirect

Acta Astronautica

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

Autonomous robotic capture of non-cooperative target by adaptive extended Kalman filter based visual servo

Gangqi Dong, Zheng H. Zhu*

Department of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada

ARTICLE INFO

Article history:
Received 1 August 2015
Received in revised form
11 January 2016
Accepted 5 February 2016
Available online 12 February 2016

Keywords:
Adaptive extend Kalman filter
Visual servo
Robotic manipulator
Autonomous capture
Non-cooperative target

ABSTRACT

This paper presents a real-time, vision-based algorithm for the pose and motion estimation of non-cooperative targets and its application in visual servo robotic manipulator to perform autonomous capture. A hybrid approach of adaptive extended Kalman filter and photogrammetry is developed for the real-time pose and motion estimation of non-cooperative targets. Based on the pose and motion estimates, the desired pose and trajectory of end-effector is defined and the corresponding desired joint angles of the robotic manipulator are derived by inverse kinematics. A close-loop visual servo control scheme is then developed for the robotic manipulator to track, approach and capture the target. Validating experiments are designed and performed on a custom-built six degrees of freedom robotic manipulator with an eye-in-hand configuration. The experimental results demonstrate the feasibility, effectiveness and robustness of the proposed adaptive extended Kalman filter enabled pose and motion estimation and visual servo strategy.

1. Introduction

Robotic manipulators have achieved great successes in industrial and space applications to perform tasks that are dangerous, complex, and even impossible to be conducted by human beings [1–4]. Specifically, the autonomous capture of non-cooperative targets by robotic manipulators is one of the research highlights in the robotic field [5–8]. For instance, the control for on-orbit space manipulators during grasping, docking and post-docking was studied in [7,8]. Regarding to non-cooperative target, one of the critical challenges arises from such a task is the accurate pose and motion estimates of the target [5,9] in order to define the end-effector's pose and trajectory [10,11]. Generally, the target is unknown to the robot and the vision system is extensively used for pose and motion estimation of the target due to its non-intrusive, non-damaging and non-contact nature, seen in [2,5,9,12–15]. The configuration of the vision sensing in a visual servo robotic system can be either eye-in-hand or eye-to-hand [16]. The eye-in-hand provides a detailed and accurate scene of the target while the eye-to-hand monitors the whole workspace with less accuracy [17]. Based on the error used in the feedback control loop, the visual servo can be cataloged as position-based, image-based and hybrid visual servo as per [18,19]. The critical issue to perform autonomous capture of non-cooperative targets by visual servo is the precise estimation of pose and motion of the target in a dynamic environment.

Considerable effort has been devoted to extract information from visual images in the past [20–23]. Different methodologies have been developed, which can be distinguished generally into four categories: analytic or geometric method, optimization-based method, offline method and filtering method. The geometric method, such as the photogrammetry, is widely adopted if the camera is

* Corresponding author. Tel.: +1 416 7362100x77729.
E-mail address: gzhu@yorku.ca (Z.H. Zhu).

http://dx.doi.org/10.1016/j.actaastro.2016.02.003
0094-5765/

calibrated and the geometric features of target are known in advance. It extracts pose information about six degrees of freedom (DOF) of the target from two-dimensional (2D) images. However, the photogrammetry relies heavily on the accuracy of imaging processing and is prone to the errors of target feature measurements and camera calibration. Moreover, the geometric method estimates is solely based on the current measurements to make predictions for the future, which is a memoryless or Markov process. Thus, its results may be noisy if the image information is not smooth. The optimization-based method requires a sufficient large set of target images in different poses, which is not available when the target is non-cooperative. The offline method is performed afterwards, which is not applicable for real-time tracking, approaching and capturing the target. The filtering method is widely used in pose and motion estimation. Among all filters, the Kalman filter (KF) is widely used in analyzing and solving pose estimation problems based on a set of noisy observations over time. Since KF being proposed in 1960 [24], it has been adopted in almost every engineering field, especially in the robotic vision, and has been extended from the original linear system to nonlinear system with different enhancing techniques. Ref. [25] provided a detailed survey of the application of KF and its various extensions in the field of robotic vision. The survey shows evidently the extended Kalman filter (EKF) is the most widely adopted nonlinear state estimation algorithm in robot vision applications. A good performance of EKF mainly depends on the good estimation of covariance matrices of system and measurement models of the filter. The challenge arises in estimating the covariance matrix of system model in dealing with the non-cooperative target where its motion is unknown in advance and unpredictable. Furthermore, all KF algorithms require the input of initial conditions and measurements over the time. Although the initial conditions do not change the convergence property of the KF, they do affect the performance of the filter, such as the convergence speed, and its results, especially when dealing with the non-cooperative target where the initial conditions are unknown. Therefore, the focus of this study will be on the impact of pose estimation of non-cooperative targets by KF in visual servo where the relative pose and motion of a non-cooperative target with respect to the camera frame is used for real-time control of the robotic manipulator, seen with limited success in [12,14]. To achieve the objective, a new approach that combines both the adaptive extended Kalman filter and photogrammetry to obtain more robust and accurate pose and motion estimates of the non-cooperative target from noisy image data in real-time is presented. The approach has been validated experimentally using a custom-built robotic manipulator with an eye-in-hand camera mounted closely to the end-effector. The experimental results demonstrate the effectiveness and robustness of the proposed algorithm as well as the control strategy by successfully tracking, approaching and capturing a non-cooperative target autonomously in both static and dynamic scenarios.

2. Camera model

Without loss of generality, the global frame is fixed in the inertia space, the camera frame is located at the center of the image plane and the target frame is attached to the target at the rotating center. The pose of a target can be described by the target frame with respect to the camera frame, such that, $\{x_{To}, y_{To}, z_{To}, \theta_x, \theta_y, \theta_z\}^T$. Here the $\{x_{To}, y_{To}, z_{To}\}^T$ is the origin of the body-fixed target frame and $\{\theta_x, \theta_y, \theta_z\}^T$ are the Euler angles of target frame with respect to the camera frame. The singularity in the frame transformation caused by the Euler angles is assumed avoided by imposing physical limits on the joint angles of manipulator in the operation space. Then, we can easily obtain the rotational matrix from target frame to the camera frame, denoted by $\mathbf{R}_{TC}$. Assume the coordinates of a feature point on the target is known in the target frame and defined as $\{x_T, y_T, z_T\}^T$. Further define $\{x_C, y_C, z_C\}^T$ as the coordinates of the corresponding feature point in the camera frame. An augmented homogeneous relationship between the target frame and the camera frame can be established as

$$\begin{Bmatrix} x_C \\ y_C \\ z_C \\ 1 \end{Bmatrix} = \begin{bmatrix} & \mathbf{R}_{TC} & & \begin{matrix} x_{To} \\ y_{To} \\ z_{To} \end{matrix} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} x_T \\ y_T \\ z_T \\ 1 \end{Bmatrix} \tag{1}$$

Consider a pin-hole camera model, the feature point on the target is projected onto the image plane, such that

$$\begin{Bmatrix} x_m \\ z_m \end{Bmatrix} = \frac{-f}{y_C - f} \begin{Bmatrix} x_C \\ z_C \end{Bmatrix} \tag{2}$$

where f stands for the focal length of the camera, and $\{x_m, z_m\}^T$ denotes the projected image coordinates. Here, it is assumed that the y_C–axis of the camera frame is pointing from the camera towards the target.

Rewrite Eq. (2) in the following form, such that,

$$\begin{cases} x_m y_C - x_m f + x_C f = 0 \\ z_m y_C - z_m f + z_C f = 0 \end{cases} \tag{3}$$

Eq. (3) indicates that there are two independent equations for one feature point, which contains six unknowns (pose information). Theoretically, one needs only three feature points to solve for the six unknowns. To eliminate the ambiguity in pose estimation with three feature points and increase the robustness of algorithm, minimum four feature points are widely adopted in literature [26]. Consequently, there will be up to eight equations with six unknowns, which are solved by an iterative least square approach with an initial guess.

The photogrammetry is a Markov process, which means that it is based only on the current measurement and prone to the measurement noise. Moreover, the computational cost of photogrammetry may increase or converge to local minima that are not true solutions, if the initial guess is far away from the real pose. As a result, the system sampling time interval may be adjusted/reduced, which is not desirable when dealing in real-time pose estimation. Another short come of the photogrammetry is