Optik 126 (2015) 2684–2689

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Three-dimensional planar object tracking with sub-pixel accuracy

A.B. Roig[a], J. Espinosa[a], J. Perez[a], B. Ferrer[b], D. Mas[a,*]

[a] Inst. Fisica Aplicada a las Ciencias y Tecnologías, University of Alicante, PO Box 99, 03080 Alicante, Spain
[b] Dep. Ing. Construccion, University of Alicante, PO Box 99, 03080 Alicante, Spain

ARTICLE INFO

Article history:
Received 8 May 2014
Accepted 26 June 2015

Keywords:
Target tracking
Shape detection
Subpixel techniques
High speed camera

ABSTRACT

Subpixel techniques are commonly used to increase the spatial resolution in tracking tasks. Object tracking with targets of known shape permits obtaining information about object position and orientation in the three-dimensional space. A proper selection of the target shape allows us to determine its position inside a plane and its angular and azimuthal orientation under certain limits. Our proposal is demonstrated both numerical and experimentally and provides an increase the accuracy of more than one order of magnitude compared to the nominal resolution of the sensor. The experiment has been performed with a high-speed camera, which simultaneously provides high spatial and temporal resolution, so it may be interesting for some applications where this kind of targets can be attached, such as vibration monitoring and structural analysis.

1. Introduction

Computer vision can be defined as the science and technology of machines that are able to see, which means that machines can extract information from images. Among many other applications, computer vision systems are used for object segmentation and tracking [1], and cover tasks from simple presence detection to non-invasive detection of small movements [2].

Precise tracking of fast moving objects usually requires high temporal and spatial resolution. Unfortunately, spatial resolution of sensors is strongly linked to the processing speed. Due to hardware limitations, CCDs and CMOS sensors are unable to simultaneously acquire, process and store large amount of data, thus impeding recording large video sequences at very high speeds. Therefore, if temporal resolution is priorized, spatial resolution has to be reduced in order to maintain the data-transfer rate.

Apart from the loss of fine image details, reduced spatial resolution also limits the accuracy in the detection tasks. Notice that motion is detected by changes in the excited pixels so, the less and larger the pixels are, the less detection accuracy is obtained. Depending on the particular application, improving the spatial resolution of the camera may be uneconomical or even impractical. However, one can take advantage of some “a priori” information about the system and the target to be tracked and, afterwards, use image processing techniques to increase the performance of the system. These methods are knowns as subpixel techniques.

Although there are alternate theoretical proposals [3], the most common methods for subpixel object tracking consist of smart interpolation of image features and it can be done in both spatial and frequency domain. In [4,5], the authors summarize the different methods for target recognition and location. From both tasks, target recognition is the most complicated since it requires of a proper analysis and segmentation of the target from the whole scene. This process can be simplified and accelerated by using specific targets that are easy to recognize. Among all possible targets, elliptical objects (including circles) are usually preferred since its shape is preserved under lateral shifts and rotations [6,7].

The first step of all algorithms for the object location in a scene consists of target isolation from the background image, which allows obtaining a distinguishable “blob” object. This step is usually performed through binary thresholding, hard-clipping or edge extraction. Once the target is isolated, several techniques can be applied for subpixel location of the blob [4,7,8]. Among them, those based on centroid calculation seem to be the most accurate and reliable methods. Unfortunately, the method is very sensitive to illumination conditions and its accuracy can be compromised in presence of noise and uneven or changing illumination [7]. Other methods based on image intensity, like those based on image correlation or template matching, may be affected of similar problems [9]. These drawbacks become more evident under the particular characteristics of image sequences obtained from fast cameras since acquisition at high frame rates needs of low exposition (typically few milliseconds), so captured images suffer from low contrast and high noise-level.

The methods based on the detection and fitting of a contour are a reliable and robust alternative for object tracking.

* Corresponding author. Tel.: +34 965 903 400; fax: +34 965 903 464.
E-mail address: david.mas@ua.es (D. Mas).

http://dx.doi.org/10.1016/j.ijleo.2015.06.070
0030-4026/

Their performance may not be as good as centroid-based techniques when dealing with static images under ideal conditions. Nevertheless, for dynamic scenes or non-controlled illumination, contour based methods are more robust since irregularities in the object shape can be detected and controlled through the mean squared error of the fitting. Additionally, they can provide information about the target geometry that may be interesting in case of shape changes.

Three dimensional object tracking adds an additional difficulty to the problem since targets may change not only their position but also their shape [10,11]. Therefore, limiting the object geometry and the movement degrees of freedom will facilitate the tracking task. Furthermore, algorithms will be faster and easier to implement. We propose to use a planar object of known geometry and with restricted movements as target, so detection and movement tracking leads to straightforward algorithms. It consists of a high-contrast elliptical line. Its geometrical parameters can be used to determine both the target position and orientation. Although the imposed restrictions limit the applicability of the method, there still is a wide range of situations where it can be used. In our case, we are interested in fields as different as structural vibrations [12] or eye tracking [2], where the object movement is clearly restricted.

Numerical simulations presented in Section 3 prove that the methods and targets used permit increased resolution since the accuracy obtained can be better than 0.03 pixel for displacement tracking, 0.05° for in-plane rotations and 0.1 pixel for axes length change. The experimental results in Section 4 show that noise, reduced contrast and other factors decrease the accuracy to 0.2 pixel, 0.1° and 0.3 pixel, respectively. Tracking accuracy is in consonance with other authors' results [7]. However, references for accuracy for tracking rotation and shape changes could not be found in the literature. Additionally, the method does not require of additional cameras or previous calibration of the imaging system [13], which makes it fast and simple to implement.

2. Method

The proposed method is simple and it briefly consists of fitting the contour of a registered blob to a known mathematical expression. Therefore, it is necessary to select a target that provides a convenient contour line. Among all possible targets, the most appropriate are those preserving the topology at all possible movements [6]. The simplest one fulfilling this characteristic is an ellipse with general equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (1)$$

By least squares fitting the blob contour to Eq. (1), we obtain several geometrical parameters of the ellipse, such as center (x_c, y_c), long and short axes (a, b), and orientation, θ [14]:

$$(x_c, y_c) = \left(\frac{-D}{2A}, \frac{-E}{2C}\right) \quad (2)$$

$$a = 2\max\left[\sqrt{\frac{D^2}{4A^2} + \frac{E^2}{4AC} - \frac{F}{A}},\ \sqrt{\frac{D^2}{4AC} + \frac{E^2}{4C^2} - \frac{F}{C}}\right]$$
$$b = 2\min\left[\sqrt{\frac{D^2}{4A^2} + \frac{E^2}{4AC} - \frac{F}{A}},\ \sqrt{\frac{D^2}{4AC} + \frac{E^2}{4C^2} - \frac{F}{C}}\right] \quad (3)$$

$$\theta = \frac{1}{2}\arctan\left(\frac{B}{C - A}\right) \quad (4)$$

The center of the ellipse locates the object inside the image plane. Orientation of the ellipse informs about object rotations in the image plane, whereas the axis lengths may both be associated to real changes in the ellipse shape (deformations) or provide information about rotations around an axis contained in the image plane. This kind of movements will be seen as a change in the perspective and, consequently, in the apparent eccentricity.

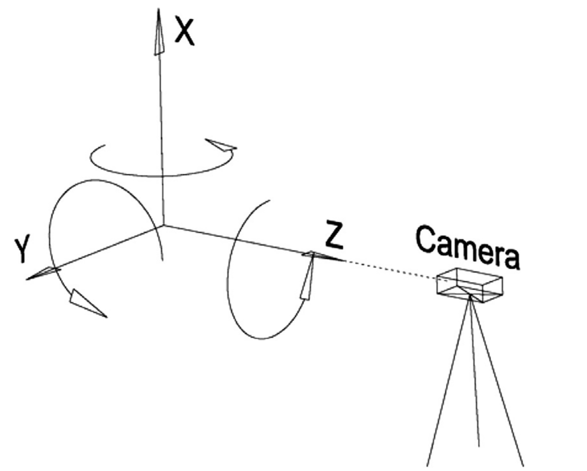

Fig. 1. Possible movements of an object in the space (three shifts and three rotations). Our method allows reporting five of the six possibilities, just loosing displacement along the camera axis (z-axis) due to the constancy of axial magnification.

The movement of a solid object in the three-dimensional space can be decomposed in six basic movements: three shifts and three rotations around the coordinate axes. The geometrical parameters in Eqs. (2–4) report on five of these six basic object movements: two shifts and three rotations around the coordinate axes (see Fig. 1). The missing shift corresponds to movements in the direction perpendicular to the image plane, which are not detectable unless they provoke a noticeable change in the object size.

The easiest case for object tracking happens when the target is attached to the object under study and a video sequence is registered. Therefore, contrast is good and the target is relatively free of noise. Nevertheless, round shapes are very common in natural scenes and thus, the method can be applied on specific details whose contour can be assumed to be elliptical. Notice that, in all cases, accurate focusing of the target is not indispensable since the unique requirement is a contour detection. The tracking is performed by fitting to (1) the contour from each frame of the sequence. All these parameters can be calculated with subpixel accuracy, which increases the nominal resolution of the sensor.

3. Numerical simulations

The accuracy of the method has been tested through several numerical simulations. The aim is to determine the sensitivity of the contour fitting to detect subpixel changes in the ellipse parameters in ideal conditions. Thus, we have calculated the response to center displacement, shape variation (axis length) and change in the orientation of the ellipse for different target sizes. Several elliptical contours have been implemented on a discrete mesh through software and then least square fitted to (1) in order to obtain the parameters from Eqs. (2)–(4).

The first simulation concerns displacement tracking. Circular targets of diameters ranging from 4 to 400 pixel have been displaced one full pixel in 100 steps. We compare theoretical center positions with those computed through Eq. (2) and obtain the error for each incremental sequence. In Fig. 2, we represent the maximum error, the mean error and its standard deviation of the whole movement sequence for each target diameter.

Small sizes present the strongest error oscillations. This effect is due to the discrete nature of the target. In those cases, the target is strongly pixelated and it may change its shape with the displacement [14]. As the target gets larger, the discrete shape