Neurocomputing 204 (2016) 222–230

Contents lists available at ScienceDirect

Neurocomputing

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

An accurate and practical calibration method for roadside camera using two vanishing points

Xinhua You [a], Yuan Zheng [b,*]

[a] Zhixing College of HuBei University, China
[b] School of Computer Science, Harbin Institute of Technology, Shenzhen Graduate School, China

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form
7 September 2015
Accepted 24 September 2015
Available online 25 April 2016

Keywords:
Roadside camera
Camera calibration
Vanishing point
Least squares optimization
Dynamic calibration

ABSTRACT

Calibrating roadside camera is essential and indispensable for intelligent traffic surveillance systems. Due to the characteristics of the traffic scenarios, the commonly used camera calibration methods based on calibration patterns are no longer suitable, since there are generally no calibration patterns (e.g., checkerboard) in traffic scenarios. In this paper, we propose an accurate and practical calibration method for roadside camera, where the vanishing point in the direction of traffic lane and the vertical vanishing point are employed that can be easily obtained from most traffic scenarios. By making full use of video information, the multiple observations of two vanishing points are available. Using these observations, we present a dynamic calibration method to correct camera parameters, where the least squares estimation is employed instead of closed-form computation. The experimental results on synthetic data show the accuracy and robustness to the noise of vanishing points; the experimental results on real traffic images demonstrate the effectiveness and practicability of the proposed calibration method.

1. Introduction

With the development of society and economy, the range of traffic monitoring is constantly expanding, and the number of roadside cameras used for traffic monitoring is explosively growing. Therefore, traditional manual monitoring no longer meets the needs and the intelligent traffic surveillance has already become an inevitable trend. By exploiting computers to complete the various monitoring tasks, the intelligent traffic surveillance not only saves the human resources but also improves the reliability of traffic monitoring. Camera calibration is to provide a mapping relationship between 3D scene and 2D image and thus is the foundation for accomplishing the various computer vision tasks. Over the past decades, camera calibration is broadly used for three-dimensional reconstruction [1,2], robot navigation [3], scene measurement [4] and visual surveillance, and thus has attracted more and more attention. For traffic surveillance application, the roadside camera calibration is foundational and essential. Once a roadside camera is calibrated, it can be used for traffic scenario analysis and target object analysis, such as object classification [5–7], object localization [8–11], object tracking [12–17] and vehicle speed estimation [18].

Up to now, a large amount of camera calibration methods has emerged, which can be broadly classified into traditional calibration methods based on calibration patterns, active vision calibration methods and self-calibration methods. However, the effective calibration methods for roadside camera are not so many. Considering that the roadside camera used for traffic monitoring is generally stationary and monocular, the active vision calibration methods and the multiview-based self-calibration methods are not suitable. Moreover, there are no calibration patterns (e.g., checkerboard) in traffic scenarios in general. Accordingly, traditional calibration-pattern-based methods are also not suitable for roadside camera calibration. According to the projective geometry theory, the vanishing point is the projection of a point at infinity and is closely related to camera parameters. In view of the characteristics of the traffic scenarios, the existing calibration methods for roadside camera exploit the information of vanishing points [19–22] that can be obtained by extracting edge features [23,24] from the traffic images.

1.1. Related work

From the projective geometry theory, it is well known that the vanishing points are closely related to camera parameters and are widely used in camera calibration methods. According to the number of the vanishing points to be used, the existing roadside camera calibration methods can be divided into three types: calibration methods based on one vanishing point, calibration

* Corresponding author.
E-mail address: zhengyuan@hitsz.edu.cn (Y. Zheng).

http://dx.doi.org/10.1016/j.neucom.2015.09.132
0925-2312/

methods based on two vanishing points and calibration methods based on three vanishing points.

The one-vanishing-point-based methods use the vanishing point in the direction of traffic lane to calibrate camera. One vanishing point is capable of estimating two camera parameters. In [25], the authors measured camera height and tilt angle in advance, and then calculated camera's focal length and the other angle using vanishing point. Song and Tai [26] employed the vanishing point and the lane width to compute camera's focal length and two rotation angles. In [27], the authors combined the vanishing point with the width and length of vehicle box model to calibrate camera. For this kind of calibration methods, Kanhere and Birchfield [28] reviewed the existing methods and proposed several new methods. All these methods employ some priori knowledge, such as camera height, lane width, the length of lane marking or the perpendicular distance from the camera to the road's edge. The latest progress has been presented in [29], where the authors obtained the vanishing point and the road's priori knowledge by fitting a road model to image data and then the length-based method in [28] is used for roadside camera calibration.

The two-vanishing-points-based methods employ two orthogonal vanishing points to calibrate camera. Two vanishing points are capable of estimating four camera parameters, namely camera's focal length and three rotation angles. In [18,28,30], two orthogonal vanishing points in the road plane are used, where one is in the direction of traffic lane and the other is in the orthogonal direction to traffic lane. In [31–33], the vertical vanishing point and one vanishing point in the road plane are used for camera calibration. To obtain these two vanishing points, Lv et al. [31] detected and tracked the walking humans; Lee and Nevatia [32] provided an interactive tool for the user to draw lines; Hodlmoser et al. [33] utilized both pedestrians and zebra-crossing, where the pedestrians were used to obtain vertical vanishing point and the zebra-crossing was used to determine the other vanishing point. Except for these two vanishing points, the latest work [34] exploits the information of horizon vanishing line.

The three-vanishing-points-based methods employ three mutually orthogonal vanishing points to calibrate camera. These three vanishing points consist of two orthogonal vanishing points in the road plane and the vertical vanishing point. Three vanishing points are capable of estimating six camera parameters, namely camera's focal length, principal point and rotation angles. In [35], the vertical vanishing point is obtained by walking men and the other two vanishing points are obtained by the appearance of moving vehicles.

1.2. Motivation

Through analyzing the existing calibration methods mentioned above, we observe that: (1) the priori knowledge used in the first type of methods may be unknown in practice, and moreover returning to the traffic scenario to obtain the priori knowledge is not feasible in most cases; (2) the third type of methods is capable of estimating more camera parameters including the principal point, however the requirement of obtaining three orthogonal vanishing points is too strict for most traffic scenes; (3) most of these methods give the closed-form solution of camera parameters, and thus they often poorly perform when the vanishing points obtained from traffic images contain noise. In view of these, our motivation is to present a more practical and accurate roadside camera calibration method, which not only is suitable for more traffic scenarios but also is more accurate and robust to the noise of vanishing points.

In this paper, two vanishing points, namely the vanishing point in the direction of traffic lane and the vertical vanishing point, are used in view of the characteristics of the traffic scenarios. These two vanishing points can be obtained by static scene structure [21,26,32], moving vehicle [4,27,35] or walking human [31,33]. The focus of this paper is on the roadside camera calibration method when the vanishing points are available in advance. In practice, the multiple observations of the vanishing points can be obtained from the traffic scenario and they provide the useful information for camera calibration. To improve the accuracy of calibration results, we present a dynamic calibration method that employs these observations to update camera parameters. Considering that the observation errors of the vanishing points always exist, we model both camera parameters and the observation errors of the vanishing points, and then convert the camera calibration problem into a least squares optimization problem using MAP (Maximum A Posteriori) estimation.

1.3. Paper structure

The remainder of the paper is organized as follows: Section 2 gives the preliminary knowledge about camera model. Section 3 proposes a dynamic calibration method based on least squares optimization to estimate camera parameters. Experiments on synthetic data and real traffic images are performed in Section 4. Some concluding remarks are presented in Section 5.

2. Camera model

A camera model describes the projection relationship between 3D space domain and 2D image domain, which consists of both intrinsic parameters and extrinsic parameters. The intrinsic parameters describe the inherent characteristics of the camera that are composed of focal length, principal point and lens skew. The extrinsic parameters describe the rigid motion of the camera that are comprised of rotation matrix R and translation vector T.

First, we introduce three coordinate systems, namely image coordinate system U–V, camera coordinate system X_C–Y_C–Z_C and world coordinate system X–Y–Z, as shown in Fig. 1. In the image coordinate system, the origin is allocated in the upper-left corner of the image, the U-axis points toward the right of the image, and the V-axis points downward. In the camera coordinate system, the origin is the camera optical center, the X_C– and Y_C–axes are parallel to the U- and V-axes, respectively, and the Z_C-axis is in the direction of optical axis. In the world coordinate system, the origin is located on the road plane, the X-axis is defined as the direction of the traffic lane, the Z-axis points upward vertically from the road

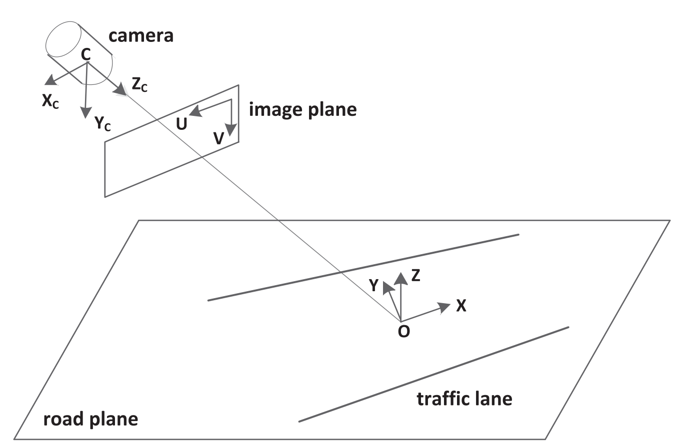

Fig. 1. Geometry relationship of camera and roadway, where the X-, Y- and Z-axes denote the world coordinate system, the X_C–, Y_C– and Z_C–axes denote the camera coordinate system and the U- and V-axes denote the image coordinate system.