Pattern Recognition Letters 71 (2016) 1–7

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A simplified two-view geometry based external calibration method for omnidirectional and PTZ camera pairs☆

Yalin Bastanlar*

Department of Computer Engineering, Izmir Institute of Technology, Izmir 35430, Turkey

ARTICLE INFO

Article history:
Received 6 July 2015
Available online 2 December 2015

Keywords:
Omnidirectional camera
Pan–tilt–zoom camera
Hybrid camera system
External calibration

ABSTRACT

The external calibration of a camera system is essential for most of the applications that involve an omnidirectional and a pan-tilt-zoom (PTZ) camera. The methods in the literature fall into two major categories; (1) a complete external calibration of the system which allows all degrees of freedom but highly time consuming, (2) spatial mapping between the pixel coordinates in omnidirectional camera and pan/tilt angles of the PTZ camera instead of explicitly computing the rotation and translation. Most methods in this category make restrictive assumptions about the camera setup such as optical axes of the cameras coincide. We propose an external calibration method that is effective and practical. Using the two-view geometry principles and making reasonable assumptions about the camera setup, calibration is performed with just two scene points. We extract rotation using the point correspondences in images. Locating the PTZ camera in the omnidirectional image is used to find the translation parameters and the real distance between the two scene points lets us compute the translation in correct scale. Results of the simulated and real image experiments show that our method works effectively in real world cases and its accuracy is comparable to the state-of-the-art methods.

1. Introduction

Hybrid camera systems consisting of an omnidirectional camera and a pan-tilt-zoom (PTZ) camera are widely used especially in surveillance applications. An omnidirectional camera provides 360° horizontal field of view with a low resolution whereas a PTZ camera provides high resolution images viewing a certain direction. A hybrid system combines the powerful aspects of both camera types and aims wide-angle high resolution surveillance. A typical task is to detect a moving object via omnidirectional camera and directing the PTZ camera towards the position of the moving object [1].

The external calibration of a hybrid system, i.e. estimation of camera poses with respect to each other, is fundamental for a cooperative use. Previously proposed calibration methods are either not practical enough to effectively determine the extrinsic parameters or they make restrictive assumptions which limit the applicability. This is the main motivation of our study.

We propose a practical calibration method that is based on two-view geometry principles and makes reasonable assumptions about the camera setup. Our method firstly extracts rotation using only two scene points and their pixel coordinates in the hybrid image pair. Afterwards, PTZ camera is located in the omnidirectional image and its pixel coordinates are used to find the translation parameters. Finally, the real 3D distance between the two scene points lets us compute the translation in correct scale, which is the distance between camera centers. Intrinsic calibrations of both cameras are obtained a priori.

The organization of the paper is as follows. In Section 2, we summarize the related work and explain the difference in our method. In Section 3, we explain the steps of the proposed two-point calibration method. The results of our experiments are presented in Section 4. The average accuracy obtained with both synthetic and real images are given together with a discussion comparing our results with other state-of-the-art methods. Section 5 summarizes the conclusions of our study.

2. Related work

A significant portion of the previously proposed methods perform a complete external calibration of the hybrid system without restricting the rotation and translation between the hybrid camera pair. Although these methods provide accurate results, the calibration procedures are time consuming due to extracting required number of point correspondences. Moreover, in most cases these methods are computationally expensive. For instance, a large pattern on the floor is required for the method in [2]. Following the internal calibration of the omnidirectional camera at the ceiling, the geometric relationship between omnidirectional and perspective camera is derived using point correspondences in both camera images. Then the perspective

☆ This paper has been recommended for acceptance by Egon L. van den Broek.

* Tel.: +90 232 750 7883; fax: +90 232 750 7862.
E-mail address: yalinbastanlar@iyte.edu.tr

http://dx.doi.org/10.1016/j.patrec.2015.11.013
0167-8655/

cameras are calibrated. For the method in [3], a calibration pattern is required to be captured at different spatial positions. In [4], the extrinsic parameters (rotation and translation between the cameras) of the hybrid system are extracted via 3D Euclidean reconstruction of scene points following projective reconstruction by factorization which is computationally involved and also expensive due to using non-linear minimization techniques such as Levenberg–Marquardt. In another method falling into this category [5], external calibration is performed by solving the epipolar geometry between the two cameras. The relative position of the active camera with respect to the static camera is solved together with the parameters associated with active camera's pan and tilt mechanism. In [6], authors focus on multi-view structure-from-motion and perform external calibration with essential matrix estimation using many points. The omnidirectional and perspective cameras are moving freely in 3D space with any orientation.

Another major group of approaches, that use PTZ camera in a hybrid system, do not solve for extrinsic parameters explicitly. Instead, they compute a spatial mapping between omnidirectional and pan-tilt parameters of the PTZ camera. In other words, they estimate the corresponding pan and tilt angles of the PTZ camera for a given pixel coordinate in omnidirectional image. It is assumed that the pan and tilt angles of a PTZ camera are highly correlated with the corresponding pixel coordinates in the omnidirectional image. For some methods, this mapping is based on data collection and fitting (interpolation) where no geometric information about the camera setup is used ([7,8]). Other methods of spatial mapping (such as [9–13]) make assumptions to be able to use geometric constraints. One of the most common assumptions is that the optical axes of the two cameras coincide [12] (i.e. one is on top of the other). In some studies, this assumption is even further extended. For instance, in [10] authors assume that the camera origins are at the same location which is not possible in a real camera setup. They relate the distance between a pixel and the center of the omnidirectional image to the vertical angle (tilt) that is used to move the PTZ camera.

Another common restriction/assumption for a camera setup is that the relative position and orientation of the cameras are known. Examples are given in [9,11,14]. Instead of employing an external calibration method, such as ours, they use the manually measured distance between the cameras. In [11], the height of the omnidirectional camera is fixed as well to be able to measure the distance to the object of interest.

In [13], Tan et al. propose a method to calculate the relative position of the optical centers of the two cameras based on parameters extracted from two sample scene points. Then, they use this relative distance information as an input to spatial mapping. For relative position estimation, they assume that the sample points (also target object to be tracked) are on a 2D plane and optical axes of both omnidirectional and the PTZ cameras are perpendicular to this plane.

The methods in this second group are more practical than the ones in the first group (complete external calibration), however the assumptions they make can be too restrictive due to several reasons: (1) Optical axes of the two cameras may not coincide. This assumption can be satisfied only for the setups where one camera is exactly on top of each other. (2) The distance between the cameras may not be measured manually. One may not be able to put a measuring tape (or a measuring laser) between them due to an obstacle. Putting a great effort to manually measure the distance is not practical since it is to be repeated when any of the cameras is moved. (e.g. cameras in a parking lot are moved for a better or a different view [13]). This requires a practical method for re-calibration. (3) The point correspondences used for calibration may not be on a 2D plane, or this plane may not be perpendicular to the optical axes. Such an assumption restricts the method to use a suitable (or prepared) surface with planar feature points on it.

Our method does not make the assumptions made by the previously proposed practical solutions, therefore the restrictions mentioned in the previous paragraph do not exist in our approach. Only assumption we make is mounting the cameras to the surfaces that are parallel to the ground, which can be satisfied by using man-made surfaces such as ceilings (indoor or outdoor). Our method is similar to the relative position estimation method in [13] but without the restriction of point correspondences should lie on a 2D plane which is perpendicular to the optical axes of the cameras. Here, we do not propose a spatial mapping method (as done in [9–13]), although the parameters estimated with our method can be used for any spatial mapping method.

3. Our method

Our method is based on the principles of two-view camera geometry where a 3 × 3 matrix, called the fundamental matrix, encompasses the geometric relation (the translation and rotation) between the two cameras or views from two different positions of a camera. With the standard method, eight point correspondences between the two views are required to compute the fundamental matrix [15], or 7 point correspondences if the rank constraint is used.

In case of calibrated cameras, another matrix, called the essential matrix, can be computed with the point correspondences. Rotation and translation parameters can be extracted from the essential matrix easily. It is possible to compute the essential matrix with as few as five point correspondences [16], instead of eight, however the algorithm is computationally involved as it requires Gröbner basis solver to find the roots of a tenth degree polynomial.

Reasonable assumptions about the camera setup let us perform the external calibration with easier procedures and less number of point correspondences. For instance, when the optical axes of the two cameras coincide (e.g. [10,12]) there is only two degrees of freedom: the translation in the vertical direction and the rotation around the common optical axis. As mentioned previously, we do not assume a certain camera setup and we aim to develop a practical external calibration method that can be used when the cameras are moved. However, one reasonable assumption we make is mounting the cameras to the surfaces that are parallel to the ground such as the ceiling. This makes the optical axis of the omnidirectional camera perpendicular to the ground. As for the PTZ camera, there are two more rotational degrees of freedom which are the rotation between the cameras around the optical axis and the tilt angle. We set the PTZ camera to the docking reference (i.e. zero pan and tilt) however the angle between the zero pan and the coordinate system of the second camera is still to be estimated. We denote this angle with β in the rest of the paper. Tilt angle of the PTZ camera, on the other hand, is assumed to be zero relying on the docking reference. In the following, we explain how the remaining extrinsic parameters (β angle and the translation vector) of such setup are solved using only two point correspondences.

3.1. Intrinsic calibration

Our method employs an omnidirectional and a PTZ camera for which internal calibrations are obtained a priori. This is not a strong assumption since there are readily available toolboxes and it has to be done just once. Only the zoom parameter of PTZ camera affects internal calibration, which can be fixed to a certain value during the calibration task. Once the internal and external calibration tasks are over, zooming in and out does not affect extrinsic parameters.

For PTZ camera calibration, we use the method proposed in [17] and implementation is available as a MATLAB Toolbox [18]. For the omnidirectional camera, we use the sphere camera model [19] which is able to cover both catadioptric (mirrored) omnidirectional cameras and fisheye cameras. There are a few calibration methods proposed