Image and Vision Computing 70 (2018) 46–54

Contents lists available at ScienceDirect

Image and Vision Computing

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

Marker-based non-overlapping camera calibration methods with additional support camera views☆

Fangda Zhao*, Toru Tamaki*, Takio Kurita, Bisser Raytchev, Kazufumi Kaneda

Graduate School of Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima City, Hiroshima 739-8527, Japan

ARTICLE INFO

Article history:
Received 24 February 2017
Received in revised form 19 October 2017
Accepted 25 December 2017
Available online xxxx

Keywords:
Non-overlapping camera calibration
AR marker
Pose estimation

ABSTRACT

Simple methods to calibrate non-overlapping cameras using markers on the cameras are proposed. By adding an augmented reality (AR) marker to a camera, we can find the transformation between the fixed AR marker and the camera. With such information, the relative pose between cameras can be found as long as the markers are visible to additional support cameras. The proposed method consists of two steps: (1) use of an extra support camera and a chessboard to find the transformation between the AR marker and the camera and (2) use of the transformation between markers to calibrate non-overlapping cameras. Compared to an existing method, the proposed method works stably and uses fewer images.

1. Introduction

Camera calibration has been investigated for a long time and remains a popular topic [1,2]. Most vision algorithms, for example [3–5], require accurate intrinsic camera parameters and, if multiple cameras are used, extrinsic camera parameters [4].

Stereo camera calibration [6] is the most common case that requires extrinsic parameter estimation. Stereo cameras necessarily share the field of view (FOV); therefore, objects and markers in the shared FOV can be used to estimate the extrinsic camera parameters. Many 3D reconstruction algorithms have benefited from this type of calibration process [5].

Another type of extrinsic calibration, i.e., non-overlapping camera calibration [7], does not share the FOV, thus making stereo calibration methods inapplicable. Due to different cost and accuracy requirements, different non-overlapping camera-calibration methods have been proposed for different applications, such as surveillance [8] and autonomous vehicle navigation [9]. However, none of these methods is universal, and each has advantages and disadvantages, as we describe in Section 1.1.

In this paper, we propose portable and stable calibration methods for non-overlapping cameras using markers.[1] The basic concept is rather straightforward, i.e., we estimate the transformation between multiple cameras by estimating the transformation between markers. Specifically, we place a marker on each camera to be calibrated (i.e., *target* cameras). Then, we capture images of the target cameras using other cameras (i.e., *support* cameras) such that all target camera markers are captured simultaneously in the support cameras' FOVs. An overview of the proposed method's configuration is shown in Fig. 1. Here, the task is to estimate the transformation between the target cameras (denoted by 1 and 2 in Fig. 1) using the support cameras. The proposed method consists of two parts. First, we use the support cameras and calibration (chessboard) patterns to estimate transformations between each target camera and corresponding marker pairs. Then, we estimate the transformations between the target cameras.

Our primary contributions are as follows.

- We present methods to calibrate the extrinsic parameters of non-overlapping cameras using external support cameras with markers.
- The proposed methods are evaluated using synthetic and real data to demonstrate their robustness against various camera configurations.

☆ This paper has been recommended for acceptance by Hongdong Li.
* Corresponding authors.

E-mail addresses: zhao@eml.hiroshima-u.ac.jp (F. Zhao), tamaki@hiroshima-u.ac.jp (T. Tamaki).

[1] A conference version of this paper was presented [10]. This paper extends that version by adding two different methods to the preparation step and providing additional experimental results.

https://doi.org/10.1016/j.imavis.2017.12.006
0262-8856/

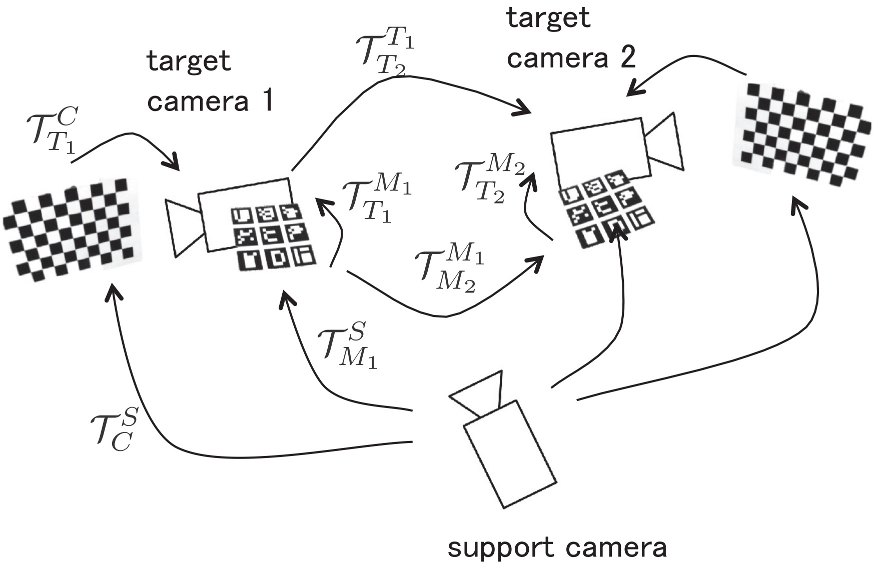

Fig. 1. Configuration overview.

The remainder of this paper is organized as follows. In the following subsection, we review related work. We describe the proposed method in Sections 2 and 3. We then provide and evaluate experimental results in Section 4. Finally, conclusions and suggestions for future work are presented in Section 5.

1.1. Related work

We categorize previous work on non-overlapping calibration into trajectory-, SLAM-, mirror-, and tracking-based methods. Trajectory- and SLAM-based methods focus on calibrating cameras attached to a mobile camera rig, whereas mirror- and tracking-based methods focus on calibrating a set of stationary cameras.

In trajectory-based methods [9,11], the transformation between fixed rig cameras should not change regardless of how the camera rig moves. Each camera on a moving rig captures image sequences, and camera motions (trajectories) are estimated for each camera. Then, the method attempts to compute extrinsic camera parameters by matching the camera trajectories. However, matching trajectories can suffer from degenerated cases, e.g., when motion occurs along a straight line. This restricts trajectory-based methods to applications with a large working space where complex camera motions can occur. In addition, accuracy is affected by the camera-motion estimation, which can be unstable in some cases. Furthermore, methods of this kind [11,12] need non-overlapping cameras to be fixed on a mobile rig. Such a requirement would be reasonable if cameras were on a vehicle for autonomous driving, but in some applications cameras are expected to be fixed on static solid walls, pillars or ground surface so that cameras mounted on them cannot be moved. Such a situation limits the usage of methods of this type.

SLAM-based methods [13,14] benefit from recent advancements in SLAM methods [3,15]. If we have a large calibration pattern that covers all non-overlapping cameras' FOVs, we can theoretically obtain the extrinsic parameters. However, generating such a large calibration pattern is impractical. By contrast, SLAM 3D reconstruction of a scene can be used as a type of calibration pattern, and transformation between cameras can be obtained by registering the 3D geometries of the same scene reconstructed from each camera. SLAM-based methods do not suffer from degenerate cases; however, scenes must be initialized in advance. Combinations of SLAM- and trajectory based methods that employ 3D information to handle degenerate cases associated with trajectory-based have been proposed [12].

The idea behind mirror-based methods [7,16,17] is interesting. The main difficulty with non-overlapping calibration is non-overlapped FOVs. However, a mirror can be used to reflect a single calibration pattern to all cameras, which makes it possible to estimate the transformation between the camera and the reflected calibration pattern. Since the camera and the calibration pattern are fixed, the pose of the mirror can also be estimated. Poses between cameras can be obtained by computing the transformations between the pattern and cameras individually. However, with mirror-based methods, it is difficult to handle mirror positions to obtain good and stable accuracy. It is known that calibration patterns should be captured in an image screen sufficiently large to increase accuracy; however, this is difficult because the results are sensitive to mirror poses, and there are various degenerated mirror poses. Therefore, mirror poses must be arranged carefully. In addition, mirror poses must satisfy a physical constraint of the mirror–camera configuration in the working space of the real environment.

Tracking-based methods [8,18] attempt to track the same object in 3D scenes captured by different cameras. These methods are typically applied to camera networks, such as surveillance systems, with multiple cameras. The motion between cameras can be predicted using a Kalman filter, and different camera poses can be estimated. Typically, tracking-based methods do not have high accuracy; however, they may be the easiest approach to calibrating relative poses of a large camera network.

We compare the existing methods in Table 1, which shows that for stationary camera calibration, existing methods suffer from either degenerate cases or do not have sufficiently high accuracy. Therefore, we propose a highly accurate method that does not suffer from degenerate cases. Experimental results indicate that the proposed method demonstrates state-of-the-art accuracy and high stability. Thus, it is expected that the proposed method is suitable for situations in which existing methods are inappropriate.

2. Formulation

The proposed method involves two steps: 1) finding the transformation between a target camera and an augmented reality (AR)