J. Vis. Commun. Image R. 39 (2016) 120–131

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Simultaneous reconstruction and calibration for multi-view structured light scanning ☆

S. Garrido-Jurado, R. Muñoz-Salinas *, F.J. Madrid-Cuevas, M.J. Marín-Jiménez

Computing and Numerical Analysis Department, and Maimonides Institute for Biomedical Research (IMIBIC), Edificio Einstein, Campus de Rabanales, Córdoba University, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 9 December 2015
Revised 22 March 2016
Accepted 23 May 2016
Available online 26 May 2016

Keywords:
Structured light
Calibration
Multi-view
3D reconstruction

ABSTRACT

Structured light 3D scanning from a single viewpoint requires multiple scans and a registration process for a complete description of the scanned object. Instead, using multiple cameras and projectors simultaneously can reduce the scanning time and increase the visibility. However, a precise estimation of the extrinsic parameters of all the components is a time consuming process prone to errors. This paper proposes a method to automatically reconstruct and self-calibrate multi-view structured light systems with an arbitrary number of devices. The experimentation shows that the proposed method is precise and robust, surpassing other current state of the art approaches. The achieved calibration accuracy is similar to that obtained by a traditional chessboard pattern calibration, but being able to adapt to a wider range of situations.

1. Introduction

Reconstruction of three-dimensional objects is an important topic in many computer vision applications such as industrial inspection [1,2], cultural heritage recording [3,4] or anthropometric analysis [5,6]. Among the different techniques to obtain depth information from a scene, structured light is very popular since it allows obtaining precise measures within a reduced cost, i.e. only conventional cameras and projectors are required [7,8].

The simplest structured light systems are the single-view ones, which are comprised by a camera-projector pair. Structured light patterns are projected over the scene and identified in the camera image, producing a set of camera-projector correspondences that are triangulated to reconstruct the scene. The range of the scanned scene is limited to the intersection of the camera and projector field of views, thus, points not illuminated by the projector nor observed by the camera cannot be reconstructed.

In order to measure a complete object (e.g. a 360 degrees scan) single-view systems require to move the object (or the system) so that it is visible from many view-points and then apply a registration technique to join the scans [9,10]. On the contrary, the approach explored in this paper, Multi-View Structured Light Systems (MVSLS), consists in employing several cameras and projectors. This helps to simplify the scanning process avoiding to move the system or the object, which is of great importance in applications such as human body scanning (Fig. 1).

One of the main downsides of MVSLS is the calibration of its devices. While calibration of single-view systems is relatively simple, the process becomes more complex and tedious as the number of devices increases.

In the recent years, there has been an effort to develop multi-view calibration methods for MVSLS, however, in most cases, they require manual intervention [11], special equipment [12] or are constrained to specific conditions [13–15] such as specific device arrangements or limitations in the number of devices. Some of these proposals aim to calibrate both intrinsic and extrinsic parameters simultaneously. Intrinsic parameters are inherent to the camera and projector devices and, thus, in many applications they only need to be estimated once, unless the lens focus is modified. In practice, extrinsic parameters or poses are more likely to be modified during the day-to-day manipulation, either intentionally (e.g. device displacements to cover a specific area) or unintentionally (e.g. accidental hits or weak device supports).

This paper proposes a method to automatically reconstruct and self-calibrate the extrinsic parameters of MVSLS devices (both cameras and projectors) without being restricted to any special arrangement and without requiring any special equipment, thus avoiding a tedious calibration step each time a device is moved. The proposed method starts by sequentially projecting the light

---

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.

E-mail addresses: i52gajus@uco.es (S. Garrido-Jurado), rmsalinas@uco.es (R. Muñoz-Salinas), fjmadrid@uco.es (F.J. Madrid-Cuevas), mjmarin@uco.es (M.J. Marín-Jiménez).

http://dx.doi.org/10.1016/j.jvcir.2016.05.014
1047-3203/

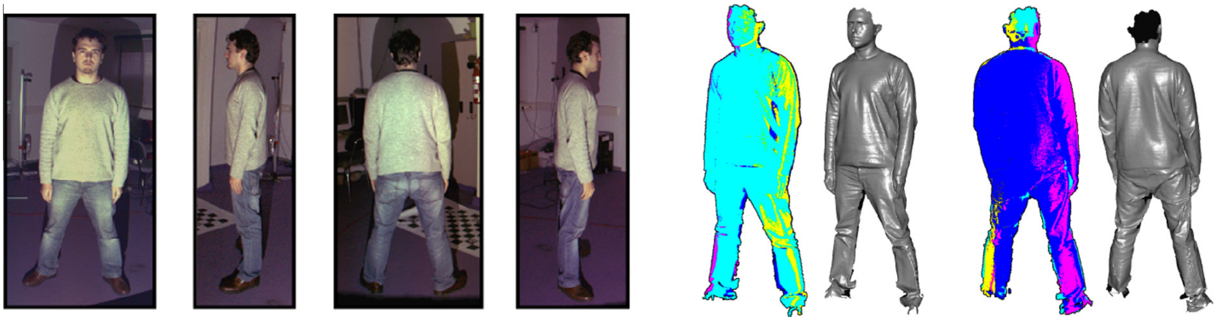

Fig. 1. Example of 360 degrees reconstruction of a person employing a MVSLS composed by 4 cameras and 4 projectors. Figure shows real camera images, colored set of point clouds of each camera-projector pair and the final complete mesh reconstruction. The MVSLS approach allows to perform the scan without moving the system or the scanned object. (Best viewed in color.) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

patterns (one projector after another) while all the cameras capture them simultaneously. Then, the intersecting field of views between the devices are automatically determined in order to find the optimal path connecting them. To that end, a Mixed Integer Linear Programming (MILP) model has been proposed which finds the optimal path based on the image correspondences between devices. MILP is an optimization technique which allows finding the optimum of a mathematical model represented by linear relationships and where some unknowns are constrained to be integers. Then, our method starts the process of determining the extrinsic parameters of all devices by following the estimated path. Finally, the three-dimensional points and extrinsic estimations are further refined using Sparse Bundle Adjustment (SBA) [16].

The rest of the paper is structured as follows. Section 2 reviews related work. Section 3 presents the basics behind structured light techniques. Section 4 details the proposed calibration method. Finally, Section 5 presents the experimentation carried out, and Section 6 draws some conclusions.

2. Related work

In general, a single-view structured light scanning cannot scan a whole scene or object since it can only provide one viewpoint. Some typical examples of these situations are 360 degrees scans or large scene reconstructions, such a whole room. To address these situations, some authors have opted for keeping a single-view scanner but employing some special equipment that permits the complete scan. An usual solution is a computer-controlled turntable [9,10] which rotates the object while a single-view scanner takes several scans on different angles, allowing a 360 degrees scan. The main problems are the extra cost of the equipment, the size limitation of the scanned object and the additional scanning time. Furthermore, the system is limited to 360 degrees scans, not allowing other types of multi-view scans. An alternative to turntables is the use of computer-controlled robotic devices [17] that can move the single-view scanner to different viewpoints while, at the same time, tracking the scanner positions. This approach can adapt to different ranges of multi-view scans, nevertheless, it involves a high cost of the precise equipment and it is still limited to the device scope.

Some approaches obtain various single-view reconstructions from different viewpoints and then integrate them. In general, this is a complicated task which usually requires manual intervention to provide an initial registration. An alternative is the use of synthetic markers to easily perform an automatic estimation of the scanner positions [18–20]. The main problems of these approaches are that the scene needs to be manipulated to attach the markers and, furthermore, the single-view scanner has still to be moved to obtain the reconstructions from different viewpoints.

In [21], a mirror system is proposed to obtain complete reconstructions of an object using a single-view scanner. The projected patterns are reflected by the set of mirrors, simulating additional virtual scanners that compose a multi-view scanner. The main benefit is that the complete reconstruction is performed using a single scan, reducing the scanning time. However, the method applicability is very restrictive, since the object size is limited to the working area surrounded by the set of mirrors, which also needs to be calibrated a priori. Furthermore, the method requires Fresnel lens which are considerably expensive.

As an alternative to the previous methods, some proposals employ a higher number of cameras and projectors, i.e. a MVSLS. These systems can also be expensive if using a high amount of devices but, contrary to previous alternatives, they can adapt to different types of scenes or objects by simply reorienting the devices to the desired areas. However, they also present some drawbacks. First, scanning with multiple projectors simultaneously is complex because of the interferences generated on the overlapped surfaces. As a consequence, many authors employ the projectors sequentially. Second, the calibration process is more complex due to the higher number of devices.

The simplest solution for extrinsic calibration is to use a calibration pattern, e.g. a chessboard pattern, which is visible by all devices, and perform the cameras and projectors calibration based on that pattern, such as [22]. However, this solution is not feasible in most cases, since each device view is usually covering different scene areas and, thus, a pattern cannot be observed by all devices simultaneously.

Some of the most basic proposals for multi-view structured light calibration require manual intervention or special equipment. In [11], a manual method for multi-view calibration is presented. First, the camera parameters are calibrated independently using the method in [23], which employs a led or laser pointer projected over the scene. Then, the calibration of the projectors is performed by projecting squared fiducial markers [24] while the user moves a white surface in front of them. The previously calibrated cameras triangulate the marker corners to obtain their 3D position which are finally used to calibrate the projectors. The main drawback is that the system requires manual operation, which in many cases can be tedious and inconvenient. Furthermore, if any of the devices is moved, the calibration process needs to be repeated.

In [12], a complete MVSLS is calibrated using a special translucent sheet. The system is limited to devices that surround the scene, i.e. 360 degrees scans. The patterns are projected over the sheet and, thanks to the translucent material, they are observed by all the cameras, independently of their position, allowing establishing correspondences between all the devices. The main drawback is that the translucent sheet needs to be moved manually