Available online at www.sciencedirect.com
ScienceDirect
Computer Vision and Image Understanding 103 (2006) 170–183
Computer Vision and Image Understanding
www.elsevier.com/locate/cviu

Pan–tilt–zoom camera calibration and high-resolution mosaic generation

Sudipta N. Sinha *, Marc Pollefeys

Department of Computer Science, University of North Carolina at Chapel Hill, CB # 3175, Sitterson Hall, Chapel Hill, NC 27599, USA

Received 8 April 2005; accepted 7 June 2006
Available online 21 July 2006
Communicated by Seth Teller

Abstract

In this paper, we discuss the problem of estimating parameters of a calibration model for active pan–tilt–zoom cameras. The variation of the intrinsic parameters of each camera over its full range of zoom settings is estimated through a two step procedure. We first determine the intrinsic parameters at the camera's lowest zoom setting very accurately by capturing an extended panorama. The camera intrinsics and radial distortion parameters are then determined at discrete steps in a monotonically increasing zoom sequence that spans the full zoom range of the camera. Our model incorporates the variation of radial distortion with camera zoom. Both calibration phases are fully automatic and do not assume any knowledge of the scene structure. High-resolution calibrated panoramic mosaics are also computed during this process. These fully calibrated panoramas are represented as multi-resolution pyramids of cube-maps. We describe a hierarchical approach for building multiple levels of detail in panoramas, by aligning hundreds of images captured within a 1–12× zoom range. Results are shown from datasets captured from two types of pan–tilt–zoom cameras placed in an uncontrolled outdoor environment. The estimated camera intrinsics model along with the cube-maps provides a calibration reference for images captured on the fly by the active pan–tilt–zoom camera under operation making our approach promising for active camera network calibration.

Keywords: Pan–tilt–zoom camera calibration; Active camera networks; Image mosaicing; Zoom calibration; Radial distortion; Radiometric alignment

1. Introduction

The use of active pan–tilt–zoom (PTZ) cameras in wide-area surveillance systems, reduces the actual number of cameras required for monitoring a certain environment. During operation, each PTZ camera can act like a high-resolution omnidirectional sensor, which can potentially track activities over a large area and capture high-resolution imagery around the tracked objects. While omnidirectional cameras simultaneously observe a scene with a large field of view (FOV) often from a single viewpoint, they typically capture low-resolution images and have a limited range of scale. High-resolution panoramic cameras need specialized hardware and can be extremely expensive. However, environments that are static or where events happen around a small region, do not require simultaneous imaging. For static scenes, multiple images captured over time can be aligned and composited into a complete panorama using image mosaicing algorithms [1,12,13]. PTZ cameras by virtue of their large zoom range can view a scene at a greater range of scale compared to an omnidirectional camera. At its finest scale, it can capture high-resolution imagery whereas a large range of pan and tilt gives it a large virtual FOV. Hence the PTZ camera combines the best of both worlds at an affordable cost.

A network of such active cameras could be used for 3D modeling of large scenes and reconstruction of events and activities within a large area, provided pixels captured from

* Corresponding author.
E-mail addresses: ssinha@cs.unc.edu (S.N. Sinha), marc@cs.unc.edu (M. Pollefeys).

1077-3142/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2006.06.002

an active camera under operation could be accurately mapped to visual rays in 3D space. This paper describes a fully automatic method for calibrating such a model for pan–tilt–zoom cameras that does not require physical access to the cameras or the observed space. A model for the camera's intrinsic parameters is estimated from images captured within its full range of pan, tilt and zoom configurations. Our method is inherently feature-based, but does not require a calibration object or specific structures in the scene.

Past work on active camera calibration has mostly been done in a laboratory setup using calibration targets and LEDs or at least in a controlled environment. Some of these include active zoom lens calibration by Willson et al. [14,9,15], self-calibration from purely rotating cameras by deAgapito [4], and more recently pan–tilt camera calibration by Davis et al. [6]. Our approach towards zoom calibration is simpler than that of Wilson [15] who computed both focal length and radial distortion at many different zoom settings [15] and is similar to that of Collins et. al. [5], who calibrated a pan–tilt–zoom active camera system in an outdoor environment. However, we extend the lens distortion model proposed by Collins [5] who assumed constant radial distortion, estimated it only at a particular zoom level and modeled its variation using a magnification factor. We actually estimate the radial distortion caused by optical zoom of the camera, the effect of which varies with camera zoom. Our method computes the intrinsic parameters of a PTZ camera from images taken by the rotating and zooming camera in an unknown scene. The intrinsic parameters at the lowest zoom are first computed by estimating homographies between multiple images acquired by a rotating camera. Using bundle adjustment [11], the homography model is extended to take radial distortion into account and obtain a complete calibrated panorama of the scene with sub-pixel alignment error (see Fig. 1). We next use an image sequence from the full zoom range of the camera to estimate the variation of its intrinsics with zoom. In our calibration model, an active PTZ camera is a virtual, static omnidirectional sensor. Next multiple images captured at increasing zoom levels are aligned to the calibrated panorama to generate a multi-resolution cube-map panorama.

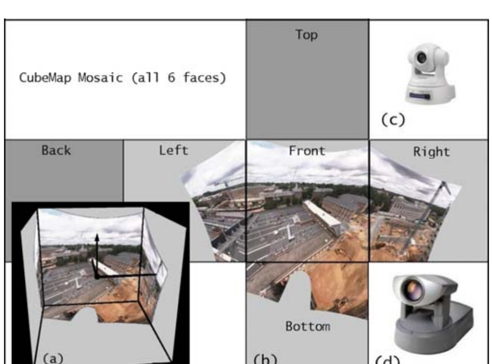
We presented preliminary work on PTZ camera calibration and multi-resolution cube-map generation in [2] and [3], respectively; this paper contains a comprehensive description of the work and shows its importance for active PTZ camera calibration. We do not address the extrinsic calibration of a PTZ camera network; the methods for conventional camera network calibration as proposed in [16,17], could be extended to PTZ cameras. The paper is organised as follows. Section 2 introduces the camera model while Section 3 explains the calibration procedure in detail followed by experimental results. Section 4 addresses the construction of multi-resolution panoramas using a method that overlaps with the calibration algorithm. We conclude with discussions and scope for future work in Section 5.

2. Theory and background

2.1. Camera model

We chose to use a simple pan–tilt–zoom (PTZ) camera model and make a tradeoff for simplicity over exactness in our choice, similar to [4,5]. Our model assumes that the center of rotation of the camera is fixed and coincides with the camera's center of projection. More general camera models have been proposed [6,14] for PTZ cameras that violate this assumption. Davis and Chen [6] model the pan and tilt rotations to be occurring about arbitrary axes in space; they estimate them in their calibration method. However, when cameras are used outdoors or in large environments, the deviation of the center is negligible compared to the average distance of the observed features, which are typically distant. Our experiments with the Canon VB-C10 and Sony SNC-RZ30 surveillance cameras (refer to the table in Fig. 1 for relevant specifications) have shown this assumption to be reasonably accurate.

In the pin-hole camera model (see Fig. 2(a)) for the perspective camera, a point $\mathbf{X}$ in 3D projective space $\mathbf{P}^3$ projects to a point $\mathbf{x}$ on the 2D projective plane $\mathbf{P}^2$



Feature	VB-C10	SNC-RZ30
Pan Angle Range	200.0^o	340.0^o
Tilt Angle Range	120.0^o	115.0^o
Zoom Range / Steps	$16X$ / 4450	$25X$ / 16000
fov at z_0	47.5^o	45.0^o
fov at z_{max}	3.0^o	2.0^o
Potential # pixels	1 Gpixels	3 Gpixels

Fig. 1. Cube-map Mosaic of 84 images computed from a PTZ camera mounted on a building-top. (a) Mapped on a cube. (b) Six faces unfolded on a plane. PTZ cameras used in our experiments. (c) Sony SNC-RZ30. (d) Canon VB-C10. See table for specifications.