Computer Vision and Image Understanding 114 (2010) 210–219

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Generic self-calibration of central cameras

Srikumar Ramalingam [a,*], Peter Sturm [b], Suresh K. Lodha [c]

[a] Mitsubishi Electric Research Laboratories, Cambridge, USA
[b] INRIA Grenoble – Rhône-Alpes and Laboratoire Jean Kuntzmann, Grenoble, France
[c] University of California, Santa Cruz, USA

ARTICLE INFO

Article history:
Received 18 March 2008
Accepted 16 July 2009
Available online 31 August 2009

Keywords:
Generic imaging models
Self-calibration
Omnidirectional cameras

ABSTRACT

We consider the self-calibration problem for a generic imaging model that assigns projection rays to pixels without a parametric mapping. We consider the *central* variant of this model, which encompasses all camera models with a single effective viewpoint. Self-calibration refers to calibrating a camera's projection rays, purely from matches between images, i.e. without knowledge about the scene such as using a calibration grid. In order to do this we consider specific camera motions, concretely, pure translations and rotations, although without the knowledge of rotation and translation parameters (rotation angles, axis of rotation, translation vector). Knowledge of the type of motion, together with image matches, gives geometric constraints on the projection rays. We show for example that with translational motions alone, self-calibration can already be performed, but only up to an affine transformation of the set of projection rays. We then propose algorithms for full metric self-calibration, that use rotational and translational motions or just rotational motions.

1. Introduction

Many different types of cameras have been used in computer vision. Existing calibration and self-calibration procedures are often taylor-made for specific camera models, mostly for pinhole cameras (possibly including radial or decentering distortion), fisheyes, specific types of catadioptric cameras, etc.; see examples in [2,3,10,6,11,12].

A few works have proposed calibration methods for a highly generic camera model that encompasses the above mentioned models and others [7,4,8,19,18]: a camera acquires images consisting of pixels; each pixel captures light that travels along a projection ray in 3D. Projection rays may in principle be positioned arbitrarily, i.e. no functional relationship between projection rays and pixels, governed by a few intrinsic parameters, is assumed. Calibration is thus described by:

- the coordinates of these rays (given in some local coordinate frame).
- the mapping between rays and pixels; this is basically a simple indexing.

One motivation of the cited works is to provide flexible calibration methods that should work for many different camera types. The proposed methods rely on the use of a calibration grid and some of them on equipment to carry out precisely known motions.

The work presented in this paper aims at further flexibility, by addressing the problem of self-calibration for the above generic camera model. The fundamental questions are: can one calibrate the generic imaging model, without any other information than image correspondences, and how? This work presents a step in this direction, by presenting principles and methods for self-calibration using specific camera motions. Concretely, we consider how pure rotations and pure translations may enable self-calibration.

Further, we consider the *central* variant of the imaging model, i.e. the existence of an optical center through which all projection rays pass, is assumed. Besides this assumption, projection rays are unconstrained, although we do need some continuity (neighboring pixels should have "neighboring" projection rays), in order to match images.

The self-calibration problem has been addressed for a slightly more restricted model in [20,21,15]. Tardif et al. [20,21] introduced a generic radially symmetric model where images are modeled using a unique *distortion center* and concentric distortion circles centered about this point. Every distortion circle around the distortion center is mapped to a cone of rays. In [15] the self-calibration problem is transformed to a factorization requiring only a singular value decomposition of a matrix composed of dense image matches. Thirthala and Pollefeys [22] proposed a linear solution for recovering radial distortion which can also include non-central

* Corresponding author.
E-mail addresses: Srikumar.Ramalingam@merl.com (S. Ramalingam), Peter.Sturm@inrialpes.fr (P. Sturm), lodha@soe.ucsc.edu (S.K. Lodha).
URLs: http://www.merl.com/people/?user=ramalingam (S. Ramalingam), http://perception.inrialpes.fr/people/Sturm/ (P. Sturm), http://www.soe.ucsc.edu/~lodha (S.K. Lodha).

1077-3142/$ - see front matter
doi:10.1016/j.cviu.2009.07.007

cameras. Here, pixels on any line passing through the distortion center are mapped to coplanar rays.

This paper is an extended version of [16]. In addition to the methods proposed in [16], we study the self-calibration problem for two new scenarios. The first is to obtain a metric self-calibration from two pure rotations. Second we study the possibility of obtaining self-calibration up to an unknown focal length in the case of using one rotation and one translation. The same self-calibration problem has been studied independently in [14,9,5], where an algebraic approach is utilized for a differentiable imaging model and infinitesimal camera motion. In contrast to these works, we use a discrete imaging model and consider finite motions.

In this work we focus on restricted motions like pure translations and pure rotations. We compute dense matches over space and time, i.e. we assume that for any pixel $\mathbf{p}$, we have determined all pixels that match $\mathbf{p}$ at some stage during the rotational or translational motion. We call a complete such set of matching pixels, a *flowcurve*. Such flowcurves provide geometrical constraints on the projection rays. For example, a flowcurve in the case of a pure translation corresponds to a set of pixels whose projection rays are coplanar. In the case of pure rotation, the corresponding projection rays lie on a cone. These coplanarity and "co-cone" constraints are the basis of the self-calibration algorithms proposed in this paper.

1.1. Overview of the paper

We formulate the generic self-calibration problem for central cameras in Section 2. In Section 3 we describe the geometrical constraints that can be obtained from pure translation and pure rotation. In Section 4 we show that with translational motions alone, self-calibration can already be performed, but only up to an affine transformation of the set of projection rays. Our main contribution is given in Section 5 where we show different self-calibration approaches using combinations of pure rotations and pure translations. Finally in Section 6 we show results for fisheye images using a self-calibration method that uses two rotations and one translation.

2. Problem formulation

We want to calibrate a central camera with n pixels. To do so, we have to recover the directions of the associated projection rays, in some common coordinate frame. Rays need only be recovered up to a euclidean transformation, i.e. ray *directions* need only be computed up to rotation. Let us denote by $\mathbf{D}_i$ the 3-vector describing the direction of the ray associated with the ith pixel $\mathbf{p}$.

Input for computing ray directions are pixel correspondences between images and the knowledge that the motion between images is a pure rotation (with unknown angle and axis) or a pure translation (with unknown direction and length). For simplicity of presentation, we assume that we have dense matches over space and time, i.e. we assume that for any pixel $\mathbf{p}$, we have determined all pixels that match $\mathbf{p}$ at some stage during the rotational or translational motion. Let us call a complete such set of matching pixels, a *flowcurve*. For ease of expression we sometimes call flowcurves arising from translational, respectively, rotational motion, *t-curves*, respectively, *r-curves*. Flowcurves can be obtained from multiple images undergoing the same motion (rotations about same axis but not necessarily by the same angle; translation in same direction but not necessarily with constant speed) or from just a pair of images I and I', as shown further below.

In Figs. 1 and 2, we show flowcurves obtained from a single image pair each for a pure translation and a pure rotation (rotation about an axis passing through the optical center). Let $\mathbf{p}$ and $\mathbf{p}'$ refer to two matching pixels, i.e. pixels observing the same 3D point in I and I'. Let $\mathbf{p}''$ refer to the pixel that in I' matches to pixel $\mathbf{p}'$ in I. Similarly let $\mathbf{p}'''$ be the pixel that in I' matches to pixel $\mathbf{p}''$ in I, and so forth. The sequence of pixels $\mathbf{p}, \mathbf{p}', \mathbf{p}'', \mathbf{p}''', \ldots$ gives a subset of a flowcurve. A dense flowcurve can be obtained in several ways: by interpolation or fusion of such subsets of matching pixels or by fusing the matches obtained from multiple images for the same

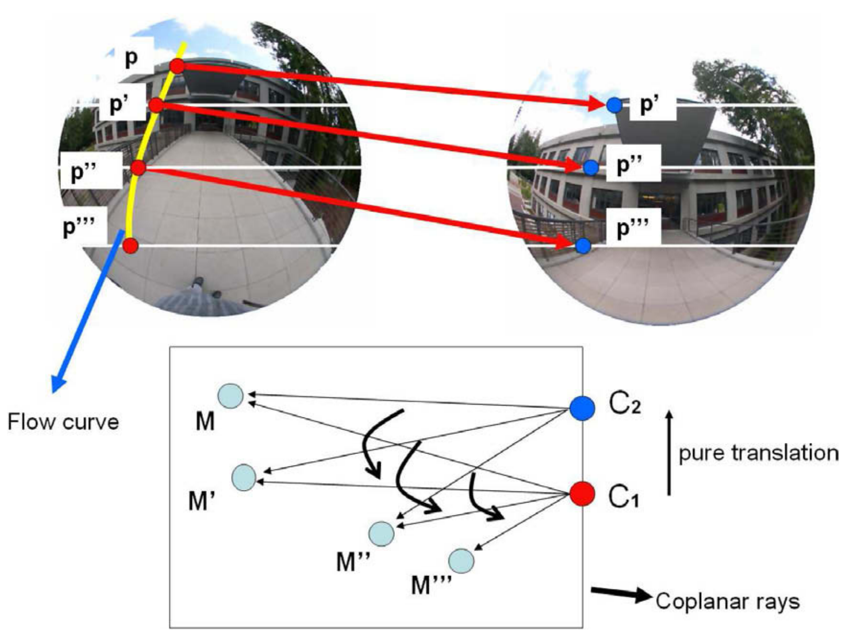

Fig. 1. Illustration of flowcurves from translation motions (t-curves). On the top we show two images related by a pure translation. Here the camera moves towards the building. Let $\mathbf{p}$ and $\mathbf{p}'$ be two matching pixels in the left and right images, respectively. Now we consider the pixel in the left image at the same location as $\mathbf{p}'$ in the right image. Let the matching pixel to this one in the right image be $\mathbf{p}''$. Doing this iteratively we obtain a set of pixels $\mathbf{p}, \mathbf{p}', \mathbf{p}'', \ldots$ which lie on the flowcurve. In the bottom we show the projection rays corresponding to pixels in the flowcurve. Let the optical center move from $\mathbf{C}_1$ to $\mathbf{C}_2$ and the projection rays corresponding to $\mathbf{p}$ be $\mathbf{C}_1\mathbf{M}$, to $\mathbf{p}'$ be $\mathbf{C}_1\mathbf{M}'$ and so on. It can be easily seen that the projection rays $\mathbf{C}_1\mathbf{M}, \mathbf{C}_1\mathbf{M}', \ldots$ as well as $\mathbf{C}_2\mathbf{M} \ldots$, are coplanar.