Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computer Vision and Image Understanding 100 (2005) 416–441
Computer Vision and Image Understanding
www.elsevier.com/locate/cviu

Structure-from-motion using lines: Representation, triangulation, and bundle adjustment

Adrien Bartoli [a,*], Peter Sturm [b]

[a] *LASMEA, 24, Avenue des Landais, 63177 Aubière cedex, France*
[b] *INRIA, 655, Avenue de l'Europe, 38334 St Ismier cedex, France*

Received 12 March 2004; accepted 17 June 2005
Available online 11 August 2005

Abstract

We address the problem of camera motion and 3D structure reconstruction from line correspondences across multiple views, from initialization to final bundle adjustment. One of the main difficulties when dealing with line features is their algebraic representation. First, we consider the triangulation problem. Based on Plücker coordinates to represent the 3D lines, we propose a maximum likelihood algorithm, relying on linearizing the Plücker constraint and on a Plücker correction procedure, computing the closest Plücker coordinates to a given 6-vector. Second, we consider the bundle adjustment problem, which is essentially a nonlinear optimization process on camera motion and 3D line parameters. Previous overparameterizations of 3D lines induce gauge freedoms and/or internal consistency constraints. We propose the orthonormal representation, which allows handy nonlinear optimization of 3D lines using the minimum four parameters with an unconstrained optimization engine. We compare our algorithms to existing ones on simulated and real data. Results show that our triangulation algorithm outperforms standard linear and bias-corrected quasi-linear algorithms, and that bundle adjustment using our orthonormal representation yields results similar to the standard maximum likelihood trifocal tensor algorithm, while being usable for any number of views.

* Corresponding author. Fax: +33 473 407 262.
E-mail addresses: Adrien.Bartoli@gmail.com (A. Bartoli), Peter.Sturm@inria.fr (P. Sturm).

1077-3142/$ - see front matter
doi:10.1016/j.cviu.2005.06.001

Keywords: Structure-from-motion; Lines; Representation; Triangulation; Bundle adjustment

1. Introduction

The goal of this paper is to give methods for reconstruction of line features from image correspondences over multiple views, from initialization to final bundle adjustment. Reconstruction of line features is an important topic since it is used in areas such as scene modeling, augmented reality, and visual servoing. Bundle adjustment is the computation of an optimal visual reconstruction of camera motion and 3D scene structure, where optimal means maximum likelihood in terms of reprojected image error. We make no assumption about the calibration of the cameras. We assume that line correspondences over at least three views are available.[1]

While the multiple-view geometry of lines is well-understood, see, e.g. [5,11], there is still a need for practical structure and motion algorithms. The factorization algorithms [15,18,25] yield reliable results but requires all lines to be visible in all views. We focus on the common three-stage approach, see e.g. [11, Section 17.5] consisting in (i) computing camera motion using inter-image matching tensors, (ii) triangulating the features, and (iii) running bundle adjustment.

There exist reliable algorithms for step (i). In particular, it can be solved by computing trifocal tensors for triplets of consecutive images, using, e.g., the automatic computation algorithm described in [11, Section 15.6], and registering the triplets in a manner similar to [6]. Other integrated motion estimation systems are [20], based on Kalman filtering techniques and [26], registering each view in turn.

In steps (ii) and (iii), one of the main difficulties when dealing with line features arises: the algebraic representation. Indeed, there is no minimal, complete and globally nonsingular parameterization of the four-dimensional set of 3D lines, see, e.g. [11, Section 2.2]. Hence, they are often overparameterized, e.g., as the join of two points or as the meet of two planes (eight parameters), or by the six coefficients of their Plücker coordinates, which must satisfy the bilinear Plücker constraint. Another overparameterization is two images of the line (six parameters). The most appropriate representation depends upon the problem considered. For example, the algorithm in [11, Section 15.2] shows that the 'two image lines' representation is well-adapted to the computation of the trifocal tensor, while the sequential algorithm of [20] is based on Plücker coordinates.

Concerning step (ii), many of the previous works assume calibrated cameras, e.g. [14,21,23,27] and use specific Euclidean representations. The linear three view algorithm of [27] and the algorithm of [23] utilize a 'closest point + direction' representation, while [21] uses the projections of the line on the $x = 0$ and the $y = 0$ planes, which has obvious singularities. These algorithms yield sub-optimal results in that none of them maximizes the individual likelihood of the reconstructed lines.

Bundle adjustment, step (iii), is a nonlinear procedure involving camera and 3D line parameters, attempting to maximize the likelihood of the reconstruction, corresponding to minimizing the reprojection error when the noise on measured features has an

[1] Line correspondences over two views do not constrain the camera motion.