Int J Comput Vis (2013) 103:240–266
DOI 10.1007/s11263-012-0581-0

Multilinear Factorizations for Multi-Camera Rigid Structure from Motion Problems

Roland Angst · Marc Pollefeys

Received: 28 January 2011 / Accepted: 28 September 2012 / Published online: 13 November 2012

Abstract Camera networks have gained increased importance in recent years. Existing approaches mostly use point correspondences between different camera views to calibrate such systems. However, it is often difficult or even impossible to establish such correspondences. But even without feature point correspondences between different camera views, if the cameras are temporally synchronized then the data from the cameras are strongly linked together by the motion correspondence: all the cameras observe the same motion. The present article therefore develops the necessary theory to use this motion correspondence for general rigid as well as planar rigid motions. Given multiple static affine cameras which observe a rigidly moving object and track feature points located on this object, what can be said about the resulting point trajectories? Are there any useful algebraic constraints hidden in the data? Is a 3D reconstruction of the scene possible even if there are no point correspondences between the different cameras? And if so, how many points are sufficient? Is there an algorithm which warrants finding the correct solution to this highly non-convex problem? This article addresses these questions and thereby introduces the concept of low-dimensional motion subspaces. The constraints provided by these motion subspaces enable an algorithm which ensures finding the correct solution to this non-convex reconstruction problem. The algorithm is based on multilinear analysis, matrix and tensor factorizations. Our new approach can handle extreme configurations, e.g. a camera in a camera network tracking only one single point. Results on synthetic as well as on real data sequences act as a proof of concept for the presented insights.

Keywords Computer vision · 3D reconstruction · Structure from motion · Multilinear factorizations · Tensor algebra

R. Angst (✉) · M. Pollefeys
ETH Zürich, Universitätstrasse 6, 8092 Zürich, Switzerland
e-mail: rangst@inf.ethz.ch

M. Pollefeys
e-mail: marc.pollefeys@inf.ethz.ch

1 Introduction

1.1 Related Work and Motivation

Factorization-based solutions to the structure from motion (SfM) problem have been heavily investigated and extended ever since Tomasi's and Kanade's (1992) seminal work about rigid factorizations. Such factorization based approaches enjoy interesting properties: e.g.given an almost affine camera these techniques provide an optimal, closed-form[1] solution using only non-iterative techniques from linear algebra. The factorization approach, which is based on the singular value decomposition of a data matrix, has been further extended to multi-body motion segmentation (Tron and Vidal 2007), to perspective cameras (Sturm and Triggs 1996), non-rigid objects (Bregler et al. 2000; Torresani et al. 2001; Brand 2001, 2005; Wang et al. 2008), and articulated objects (Yan and Pollefeys 2008; Tresadern and Reid 2005). More flexible methods which can deal with missing data entries in the data matrix have been proposed in order to

[1] Throughout this paper, the term *closed-form solution* denotes a solution which is given by following a fixed number of prescribed, non-iterative steps. Solutions provided by algorithms which iteratively refine a current best guess are thus not closed-form solutions. Stretching the notion of closed-form solutions a little further, algorithms involving matrix factorization steps such as the singular value decomposition will still be considered as closed-form, even though nowadays such matrix decompositions are often implemented iteratively.