Optik 127 (2016) 6183–6194

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Structure recovery from uncalibrated images using iterative perspective factorization

Shaohua Qiu*, Gongjian Wen

ATR Key Laboratory, National University of Defense Technology, Changsha 410073, P.R. China

ARTICLE INFO

Article history:
Received 10 January 2016
Accepted 6 April 2016

Keywords:
Eigen-decomposition
Perspective projection
Uncalibrated images
Factorization methods
Structure from motion

ABSTRACT

Structure recovery is a widely used data acquisition way for scene analysis. Factorization based methods have become the mainstream methods for structure recovery these years, which generally recover structures by solving the rectification matrix. However, these solutions will be invalid when eigen-decomposing a negative definite rectification matrix. Hence, we present an improved iterative perspective factorization method in which structure and motions are directly solved by iteratively imposing constraints rather than matrix decomposition. Experiments on both synthetic data and real images show that the proposed method can avoid the invalidation of general factorization-based solutions caused by failures and efficiently recover the geometric structure, camera motion parameters and camera focal lengths simultaneously from uncalibrated images with fast convergence rate, high accuracy and noise tolerance.

1. Introduction

With the growing demand for 3-dimensional (3D) data for virtual reality, 3D printing and archaeology, recovering 3D structure from multi-view image feature correspondences have been a fundamental problem in computer vision as well as photogrammetry, known as Structure from Motion (SFM). The factorization framework has been the most influential method in SFM at present due to its universality and robust noise tolerance.

There have been great efforts towards the factorization framework. Tomasi and Kanade [1] first introduced the factorization method under the orthographic projection model. Poelman and Kanade [2] then extended it to weak perspective and paraperspective projection models. These affine projection models are approximations of the real camera model, the perspective projection model, and the camera intrinsic parameters are not taken into account, thus a high accuracy can hardly be guaranteed. A perspective factorization method based on projective depth estimation was proposed by Triggs [3]; however, the constraints of projective matrix are not taken into account sufficiently. Wang et al. [4] proposed the quasi-perspective camera model which fills the gap between affine and perspective models. Han and Kanade [5,6] proposed a perspective factorization method fully considering the constraints of projective matrix, which broadens the horizons of projective factorization. Furthermore, bundle adjustment [7,8] from photogrammetry has made its way into SFM to estimate the optimal structure and camera parameters [9]. Lately, incremental or online SFM [10,11] were proposed by researchers, as well as the ones dealing with missing or uncertain input data [12–17]. Also, researchers [18] find that rigid SFM can be made robust to challenging intra-category variation for recognition-induced tasks.

* Corresponding author.
E-mail addresses: qiush125@163.com (S. Qiu), wengongjian@sina.com (G. Wen).

http://dx.doi.org/10.1016/j.ijleo.2016.04.007

More recently, the attention of the SFM community has been moved to non-rigid SFM (NRSFM) and great achievements have been made to robustly handle various problems in practical applications [19–24]. Xiao et al. [7] prove that enforcing both the basis and the rotation constraints leads to a closed-form solution of NRSFM. Akhter et al. [22] propose to treat the deformable shape as a collection of individual point trajectories. Based on these, Gotardo et al. [19] represent a smoothly deforming 3D shape as a smooth time-trajectory of a point within a linear shape space, while Hamsici et al. [20] considering deformations as spatial variations in shape space and enforcing spatial, rather than temporal, smoothness constraints. Additionally, great blossoms of SFM are also witnessed in vision related fields, such as remote sensing [25], augmented reality [26], recognition [27] and autonomous driving [28]. However, full matrix factorization for rigid objects is still an important issue to be considered [29] as a foundation of SFM.

There are two critical steps to obtain the structure under the perspective projection model [30]: (a) to recover a set of consistent perspective depths, and (b) to recover the camera intrinsic parameters. Up to now, many algorithms [3,5,30–32] have been presented to approximate projective depths directly from the measurement matrix. Meanwhile, due to the requirement of camera intrinsic parameters in the normalization process, there are researches recovering structure after self-calibrating the camera, which inevitably makes SFM more complex. Several methods [5,6,31,33,34] were presented to recover both structure and camera intrinsic parameters from uncalibrated image sequences following the idea of solving for rectification matrix by matrix decomposition. This procedure might go invalid in extreme situations when a negative definite matrix is decomposed as stated in Section 2, which inspired our research in this paper.

In this work we make the following specific contributions. First, we revisit the process of general factorization methods and analyze the extreme situation in which conventional methods might fail, which is the first attempt to our knowledge. Second, assuming that the camera focal length is the only unknown and varying intrinsic parameter, we present a simple but effective iterative perspective factorization method to simultaneously recover the structure, camera motions and camera focal lengths from uncalibrated images, while avoiding the camera calibration process, as well as the complex and unstable rectification matrix solution. Third, we present an explicit experimental study on both synthetic data and real images, which is evaluated from the aspects of accuracy, convergence rate and noise tolerance. The results are on par with the state-of-the-art.

2. Problem statement

The goal of perspective factorization is to recover the 3D structure, camera motion and intrinsic parameters from F uncalibrated perspective images of N 3D object points. The object structure matrix is denoted by$\boldsymbol{S} = [\boldsymbol{S}_1, \cdots, \boldsymbol{S}_j, \cdots, \boldsymbol{S}_N]$, $j = 1, \ldots, N$, where $\boldsymbol{S}_j$ is the unknown homogeneous 3D object point vector, $\boldsymbol{P}_i$ ($i = 1, 2, \ldots, F$) represents the unknown projective matrix of each image and $\boldsymbol{x}j\ I$ is the measured homogeneous image point vector. For the real camera projection model, the relation between object points and image points is

$$\boldsymbol{W} = \begin{bmatrix} \lambda_{11}\boldsymbol{x}_1^1 & \cdots & \lambda_{1N}\boldsymbol{x}_1^N \\ \vdots & \ddots & \vdots \\ \lambda_{F1}\boldsymbol{x}_F^1 & \cdots & \lambda_{FN}\boldsymbol{x}_F^N \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_1 \\ \vdots \\ \boldsymbol{P}_F \end{bmatrix} \begin{bmatrix} \boldsymbol{S}_1 & \cdots & \boldsymbol{S}_N \end{bmatrix} = \boldsymbol{PS} \tag{1}$$

where λ_{ij} is a non-zero scale factor, commonly called the projective depth, $\boldsymbol{W}$ is a $3F \times N$ scaled measurement matrix, $\boldsymbol{P}$ is a $3F \times 4$ projective matrix and $\boldsymbol{S}$ is a $4 \times N$ structure matrix.

Assuming that Projective Depth Recovery (PDR) has been done by certain algorithm and the scaled measurement matrix $\boldsymbol{W}$ is obtained, we can decompose it by rank-4 Singular Value Decomposition (SVD), which leads to a result up to a linear rectification matrix $\boldsymbol{Q}$ as shown in Eq. (2), where $\boldsymbol{Q}$ is an arbitrary 4×4 nonsingular matrix.

$$\boldsymbol{W} = \boldsymbol{P} \cdot \boldsymbol{S} = \boldsymbol{PQ} \cdot \boldsymbol{Q}^{-1}\boldsymbol{S} \tag{2}$$

The elements $\boldsymbol{M}$ of rotation matrix included in the rectified projective matrix $\boldsymbol{P}$ are orthonormal, and this leads to a linear equation set $\boldsymbol{MQ}(\boldsymbol{MQ})^{\mathrm{T}} = \boldsymbol{MGM}^{\mathrm{T}} = \boldsymbol{C}$, where $\boldsymbol{G} = \boldsymbol{QQ}^{\mathrm{T}}$ denotes the Gram matrix of $\boldsymbol{Q}$ and $\boldsymbol{C}$ is a diagonally coefficient matrix. $\boldsymbol{G}$ will be obtained by solving this equation set, and then the final rectification matrix $\boldsymbol{Q}$ is solved by eigen-decomposing $\boldsymbol{G}$.

Since $\boldsymbol{M}$ contains intrinsic parameters under the perspective model, the diagonal elements of $\boldsymbol{C}$ are unknown and they cannot be used as constraints. Then, the linear equation set above becomes homogeneous and cannot be uniquely determined without adding some assumptions. Thus, it makes the solution for the homogeneous problem possess a degree of randomness. Moreover, in applications affected by mismatching of some point pairs or strong noises, $\boldsymbol{G}$ cannot be eigen-decomposed to find the real matrix squared root as it is not positive semi-definite. To solve this problem, positive semi-definite programming [35,36] will be involved in the linear solving process and this will bring in more complexity and computation. Therefore, the attempts to avoid the process of solving the rectification matrix will be meaningful.

3. Methodology

In this section, an improved iterative perspective factorization method is proposed assuming that PDR has been done and the scaled measurement matrix is constructed.