Stable least-squares matching for oblique images using bound constrained optimization and a robust loss function

CrossMark

Han Hu [a,b], Yulin Ding [a,*], Qing Zhu [a,c,d], Bo Wu [b], Linfu Xie [d], Min Chen [a]

[a] Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, PR China
[b] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[c] National-local Joint Engineering Laboratory of Spatial Information Technology for High-speed Railway Running Safety, Southwest Jiaotong University, PR China
[d] Collaborative Innovation Center for Geospatial Technology, Wuhan, PR China

ARTICLE INFO

ABSTRACT

Least-squares matching is a standard procedure in photogrammetric applications for obtaining sub-pixel accuracies of image correspondences. However, least-squares matching has also been criticized for its instability, which is primarily reflected by the requests for the initial correspondence and favorable image quality. In image matching between oblique images, due to the blur, illumination differences and other effects, the image attributes of different views are notably different, which results in a more severe convergence problem. Aiming at improving the convergence rate and robustness of least-squares matching of oblique images, we incorporated prior geometric knowledge in the optimization process, which is reflected as the bounded constraints on the optimizing parameters that constrain the search for a solution to a reasonable region. Furthermore, to be resilient to outliers, we substituted the square loss with a robust loss function. To solve the composite problem, we reformulated the least-squares matching problem as a bound constrained optimization problem, which can be solved with bounds constrained Levenberg–Marquardt solver. Experimental results consisting of images from two different penta-view oblique camera systems confirmed that the proposed method shows guaranteed final convergences in various scenarios compared to the approximately 20–50% convergence rate of classical least-squares matching.

© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Currently, aerial oblique images are becoming the new mapping standard for digital city modeling due to their capabilities for viewing building facades from different angles, such as the penta-view camera system with "Maltese Cross" configuration (Petrie, 2009; Xiao et al., 2012; Rupnik et al., 2014; Xiong et al., 2014; Lemmens, 2014a). The fundamental step in 3D applications with aerial oblique images is determining the exterior orientation (EO) parameters for all the camera heads. Due to possible synchronization problems among the camera heads and the rigidity of camera platform (Wiedemann and Moré, 2012; Lemmens, 2014b; Hu et al., 2015; Rupnik et al., 2015), all images should be oriented through a combined block adjustment; otherwise, if only nadir images are oriented, the inaccurate EO parameters for oblique images obtained from calibrated platform parameters will lead to obvious systematic errors in the object space. This has triggered new challenges for classical photogrammetric software, due to the significant geometrical and radiometric differences between oblique images. The problems are especially severe for image matching (Hu et al., 2015), which are explicitly affected by the geometrical deformations and radiometric changes between oblique images.

Least-squares matching (LSM) is a well-established image matching method providing feature correspondences in sub-pixel level (Förstner, 1982; Ackermann, 1984; Gruen, 1985). LSM minimizes the gray level differences between the template and the matching window whereby the position and the shape of the matching window are parameters to be determined in the adjustment process. LSM is a nonlinear optimization problem that needs to be linearized and solved iteratively to a convergent solution. To obtain the final convergence for LSM, two criteria should be satisfied (Gruen, 2012): (a) accurate initial correspondence must be established and (b) the image quality in both of the images in the LSM window should be similar. For traditional nadir images, the degree of geometrical deformations and radiometric

* Corresponding author. Tel.: +86 13437119857.
E-mail address: rainforests@126.com (Y. Ding).

differences are guaranteed by the systematic aerial flight and small varying viewing directions; therefore, the problem relies only on the initial match position. However, when processing oblique images, we have occasionally encountered significant blur and light condition differences due to lack of forward motion compensation in some oblique camera systems, as shown in Fig. 1. Furthermore, the tilting angles between oblique and nadir images will also cause obvious differences in appearance. In this case, accurate initial values for the unknown parameters sometimes cannot ensure that LSM will be convergent.

In each iteration in LSM, the method will generate a shift vector, which consists of the incremental values for the provisional parameters. Because the shift vector is the result of a first order approximation using the Taylor series, it is directly related to the grayscale values and gradients of the images in the LSM window. Incorrect shift vectors due to significant differences between image qualities may cause the problem of zigzagging (Moré and Thuente, 1994) and reach the maximum number of iterations set by the algorithm. Furthermore, in the case of similar textures or patterns, LSM may converge to another false local optimal value, which also causes incorrect results.

It should be noted that all of the parameters of LSM have physical significance, including affine transformation and image illumination differences; the ranges of the parameters will absolutely lie in a reasonable region, e.g., the lower and upper bounds. For example, the initial EO parameters for all images can be obtained from the GNSS/IMU system and the pre-calibrated installation rotation and translation parameters in a laboratory environment. Although large perspective deformations exist between vertical and oblique images, with the aid of the initial EO parameters, a homography transformation (Hartley and Zisserman, 2004) can be estimated, which will quantitatively portray the deformation; using the transformation, all of the images can be warped geometrically to alleviate the deformations, and the warped images will present almost no scale and rotation differences, as detailed in our previous work (Hu et al., 2015). Furthermore, the contrast and brightness differences between the aerial images will lie in a reasonable region.

Aiming at improving the quality and robustness of LSM for oblique images with regard to the significant differences in appearance caused by blur and diverse illumination conditions, this paper proposes to impose constraints on parameters to force LSM to iterate with unknowns in reasonable ranges using *a priori* information. Explicitly, we reformulated the LSM problem into a nonlinear optimization form subjected to bounded constraints (*c.f.* Section 3.2), which defined the lower and upper bounds of the parameters.

Then, the bounded problem could be solved efficiently and simply using publicly available solvers as long as the cost function and gradient for the parameters were evaluated in each iteration (*c.f.* Section 3.4). By assigning limits of the parameters, the shift vectors incremented to the unknowns would not stray from the final convergence, and the search space for the unknowns was significantly reduced. Furthermore robust loss function, e.g., Huber loss, was proposed to handle the potential outliers of pixel values (*c.f.* Section 3.3). Huber loss (Hastie et al., 2009) would behave the same as the square loss in the inlier region, but the influence would only grow linearly rather than quadratically in the outlier region. Thus, robust and accurate sub-pixel locations could be achieved. Due to the limited search space for convergent solutions and the robust loss functions, we could guarantee final convergence of LSM.

The remaining parts of this paper are organized as follows. Section 2 summarizes previous work on matching and locating feature points with sub-pixel accuracies. Section 3 describes the reformulated LSM problem and present solutions for the constrained problem. Then, experimental evaluations and analyses are demonstrated in Section 4. Finally, the concluding remarks are presented.

2. Related works

In the community of photogrammetry, where accuracy and quality control sometimes receive first priority, establishing correspondences at sub-pixel locations has been extensively studied (Förstner, 1982), and numerous methods have been proposed for this process. Generally, two strategies exist for achieving sub-pixel correspondences: localization of feature points or matching two images in a small window. The first strategy seeks to obtain the maximum responses of interest points in the two images independently; the latter strategy commonly fixes one point as the reference and optimizes the location of the other image at the sub-pixel level.

In close range photogrammetry, precise and automatic methods to orient the image blocks have been developed, which matured approximately two decades ago, using the artificial coded targets (Fraser, 1997). The core technique for developing such automatic systems generally involves two independent steps: recognition and localization of the target point (Wong et al., 1988). Because the targets are designed to be the peak or trough responses in images, they can be easily detected through dynamically and adaptively thresholding. Then, the location of the targets can be precisely obtained by several methods, including eclipse fitting with the edges and centroiding (and its variants) or Gaussian distribution fitting with the pixel/ binary values in the target area
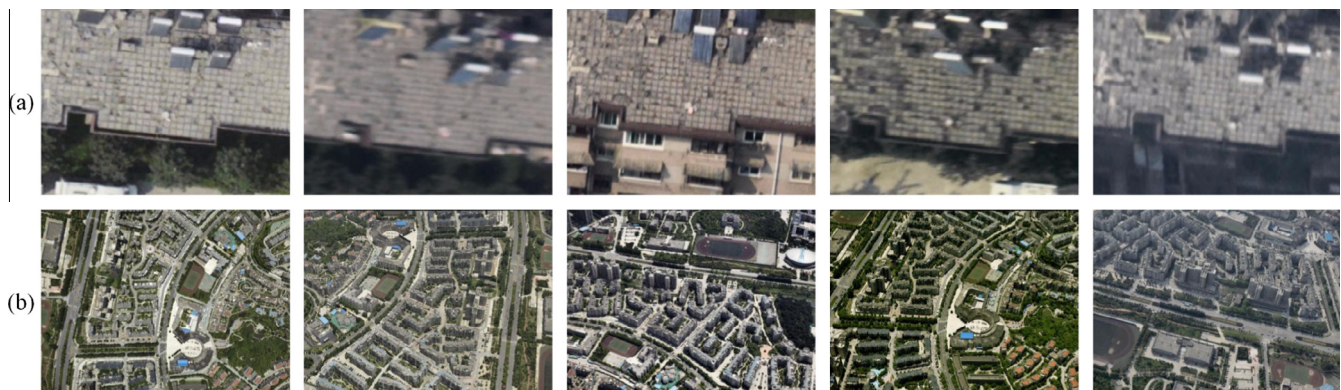

Fig. 1. Differences of appearance caused by diverse angles of incidence of the nadir cameras (first column) and four oblique cameras (other columns), including: (a) blur and (b) illumination differences. It should be noted that we pre-rectified the images in the first row for a more clear presentation. Lack of forward motion compensation in some of the aerial oblique camera systems and variation of ground sample distance (GSD) in the oblique views may lead to the blur effect. And the illumination differences may come from different sun-light directions and tuning of different camera sensors and lenses.