ISPRS Journal of Photogrammetry and Remote Sensing 138 (2018) 218–231

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Towards automatic SAR-optical stereogrammetry over urban areas using very high resolution imagery

Chunping Qiu [a], Michael Schmitt [a], Xiao Xiang Zhu [a,b,*]

[a] Signal Processing in Earth Observation, Technical University of Munich (TUM), Munich, Germany
[b] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany

ARTICLE INFO

Article history:
Received 29 August 2017
Received in revised form 9 December 2017
Accepted 11 December 2017

Keywords:
Synthetic Aperture Radar (SAR)
Optical images
Remote sensing
Data fusion
Stereogrammetry

ABSTRACT

In this paper we discuss the potential and challenges regarding SAR-optical stereogrammetry for urban areas, using very-high-resolution (VHR) remote sensing imagery. Since we do this mainly from a geometrical point of view, we first analyze the height reconstruction accuracy to be expected for different stereogrammetric configurations. Then, we propose a strategy for simultaneous tie point matching and 3D reconstruction, which exploits an epipolar-like search window constraint. To drive the matching and ensure some robustness, we combine different established hand-crafted similarity measures. For the experiments, we use real test data acquired by the Worldview-2, TerraSAR-X and MEMPHIS sensors. Our results show that SAR-optical stereogrammetry using VHR imagery is generally feasible with 3D positioning accuracies in the meter-domain, although the matching of these strongly hetereogeneous multi-sensor data remains very challenging.

1. Introduction

Currently, we are living in the "golden era of Earth observation", characterized by an abundance of airborne and spaceborne sensors providing a large variety of remote sensing data. In this situation, every sensor type possesses different peculiarities, designed for specific tasks. One prominent example is the German interferometric SAR mission TanDEM-X, whose task is the generation of a global Digital Elevation Model (DEM) (Krieger et al., 2007). In order to do so, for every region of interest highly coherent InSAR image pairs acquired by the two satellites of this mission are needed. The same holds for optical stereo sensors such as RapidEye, which additionally require cloudless weather and daylight during image acquisition (Tyc et al., 2005). Eventually, this means that there is a huge amount of data in the archives of which possibly a large potential remains unused, because information currently can only be extracted within those narrowly defined mission-specific configurations. If, e.g., a second coherent SAR acquisition is not (yet) available, or if one image of an optical stereo pair is obstructed by severe cloud coverage, the mission goal – topography reconstruction – can currently not be fulfilled. The solution to this problem is the development of methods for flexible multi-sensor data fusion (Schmitt and Zhu, 2016). This paper investigates the stereogrammetric fusion of SAR and optical imagery for a reconstruction of 3D information over urban areas less dependent on the type of the available remote sensing data.

The idea of SAR-optical stereogrammetry was first presented almost 30 years ago by Bloom et al. (1988), who investigated its general feasibility using low-resolution data provided by the SIR-B mission and the Landsat-4/5 satellites, focusing on the analysis of rural areas. Further investigations were carried out by the group of Raggam and Almer (1990), Raggam et al. (1993), Raggam et al. (1994), who combined low-resolution Seasat and SPOT/Landsat images for rural DEM generation. Some time later, similar experiments using ERS-2/Radarsat-1 and SPOT data were presented by Xing et al. (2008). All these studies have shown errors in the dekameter-domain, thus seemingly prohibiting an application of SAR-optical stereogrammetry to urban remote sensing. While in Wegner et al. (2014) building height estimation by fusing a single-pass interferometric SAR image pair and one aerial orthophoto has been shown to provide meter-accuracy, strict SAR-optical stereogrammetry was not studied by the authors. Only recently, Zhang et al. (2015) showed that TerraSAR-X and GeoEye-1 images can be used to carry out stereogrammetric 3D reconstruction with an error in the meter-domain – although the study only

* Corresponding author at: Signal Processing in Earth Observation, Technical University of Munich (TUM), Munich, Germany and Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany.
E-mail address: xiao.zhu@dlr.de (X.X. Zhu).

https://doi.org/10.1016/j.isprsjprs.2017.12.006
0924-2716/

used the manually measured corners of a simple-shaped building to proof the concept. In contrast, based on the preliminary considerations sketched in Schmitt and Zhu (2016) and Qiu et al. (2017), our work intends to provide a first step towards a solution for the non-trival problem of automatic stereo matching of VHR multi-sensor images of complex urban areas. In this context, it combines SAR-optical stereo intersection with a fully automatic selection of sparsely distributed tie points, i.e. it models both the matching and the reconstruction processes in a simultaneous manner. For this task, we first investigate the theoretically achievable accuracies of SAR-optical stereogrammetry, and how they depend on different intersection geometries. After that, we propose an epipolar-like constraint for an enhancement of the difficult search for homologue tie points. To drive the matching and ensure some robustness, we combine different hand-crafted image descriptors and evaluate our findings on test data comprised of VHR remote sensing imagery acquired by the spaceborne sensors Worldview-2 and TerraSAR-X as well as the airborne system MEMPHIS.

2. Principle of SAR-optical stereogrammetry

2.1. Geometric interpretation

Fig. 1 sketches the basic principle of SAR-optical stereogrammetry, which can be described as an intersection of the range-Doppler projection circle defined by the principal SAR measurements time t and slant range R, and an optical projection ray defined by the optical image coordinates x and y. Conceptually, this yields a set of four equations, namely the range-Doppler equations (Leberl, 1990)

$$R = \sqrt{(X_s(t) - X)^2 + (Y_s(t) - Y)^2 + (Z_s(t) - Z)^2} \tag{1}$$

$$V_x(X - X_s(t)) + V_y(Y - Y_s(t)) + V_z(Z - Z_s(t)) = 0 \tag{2}$$

of the zero-Doppler processed SAR data, and the central projection equations (Egels and Kasser, 2003)

$$x = x_0 + c\frac{r_{11}(X - X_o) + r_{21}(Y - Y_o) + r_{31}(Z - Z_o)}{r_{13}(X - X_o) + r_{23}(Y - Y_o) + r_{33}(Z - Z_o)} \tag{3}$$

$$y = y_0 + c\frac{r_{12}(X - X_o) + r_{22}(Y - Y_o) + r_{32}(Z - Z_o)}{r_{13}(X - X_o) + r_{23}(Y - Y_o) + r_{33}(Z - Z_o)} \tag{4}$$

of the optical imagery. This overdetermined equation system can be solved for the unknown object coordinates X, Y and Z, if the orientation parameters, i.e. sensor position $\mathbf{S}(t) = [X_s(t), Y_s(t), Z_s(t)]^T$ and instantaneous velocity $\mathbf{V}(t) = [V_x(t), V_y(t), V_z(t)]^T$ at zero-Doppler time t of the SAR sensor, as well as projection center $\mathbf{P}_c = [X_o, Y_o, Z_o]$ of the optical sensor in addition to the elements of the rotation matrix $\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ depending on its orientation parameters ϕ, ω, κ, are known. A solution can be found, e.g. using least-squares estmation in the Gauss-Newton model.

2.2. Theoretical accuracy analysis

Obviously, the height accuracy that can be expected from SAR-optical stereogrammetry on the one hand depends on the accuracy of the observations, i.e. zero-Doppler time t and slant range R as measured by the SAR sensor and defined by the SAR image coordinates $(r, c)_s$, and the optical image coordinates $(r, c)_o$, which define the angular orientation of the optical projection ray. On the other hand, the accuracy also depends on the orbit and orientation parameters of the sensors. In this section, we derive some rules-of-thumb for the expectable accuracies depending on different acquisition configurations. Since the orbits of modern satellite missions are well-controlled, we focus on the modeling of the inaccuracies of the principal measurements, which define the intersection geometry given the sensor positions and orientations.

2.2.1. Opposite-side stereo

For sake of simplicity, Fig. 2 sketches SAR-optical stereogrammetry as a trigonometric, in-plane intersection problem: Using Z_s for the satellite height, the approximation error caused by ignoring the Earth curvature is $\left(\frac{Z_s - H_s}{H_s}\right)$, where Z_s can be calculated given the radius of the reference ellipsoid r and the radar viewing angle θ. For TerraSAR-X (TSX), $H_s = 515$ km and $\theta = 23^\circ$, given $r = 6371$ km, the approximation error is within 0.8%. As explained in Section 2.1, the optical projection ray defined by the viewing angle α intersects with the range-Doppler circle of the SAR sensor, whose radius is defined by the range measurement R, which can also be represented by the radar viewing angle $\theta = \arccos\left(\frac{Z_s - h}{R}\right)$.

Thus, in a generalized manner, we can denote the height reconstruction process as a nonlinear function of R and α:

$$h = f(R, \alpha) \tag{5}$$

If we then want to derive information about the height reconstruction accuracy σ_h, we can resort to variance–covariance propagation, which gives us

$$\sigma_h^2 = \left(\frac{\partial h}{\partial R}\right)^2 \sigma_R^2 + \left(\frac{\partial h}{\partial \alpha}\right)^2 \sigma_\alpha^2 \tag{6}$$

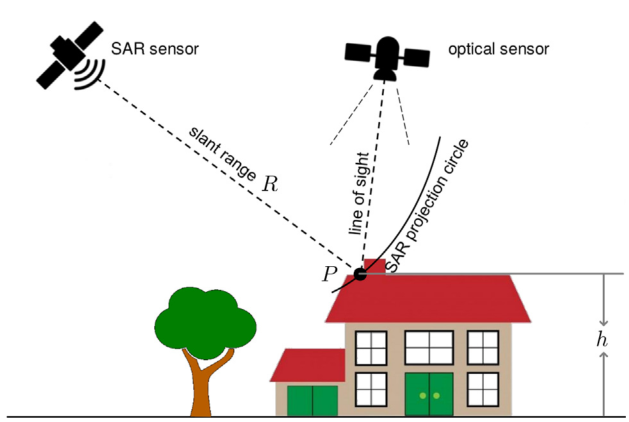

Fig. 1. The basic principle of SAR-optical stereogrammetry.

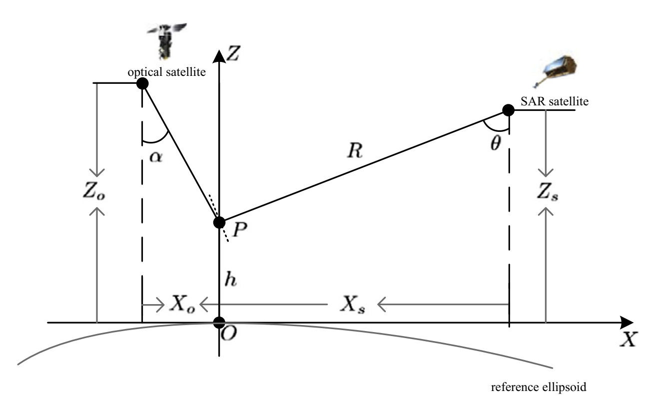

Fig. 2. SAR-optical intersection geometry for opposite-side stereo. The thin dotted line perpendicular to the look direction shows the SAR wave front, which can be considered as a tangent to the range-Doppler circle.