Registration of Very High Resolution SAR and Optical Images

Carlos Villamil-Lopez, German Aerospace Center (DLR), carlos.villamillopez@dlr.de, Germany
Lars Petersen, Airbus Defence and Space, lars.l.petersen@airbus.com, Germany
Rainer Speck, German Aerospace Center (DLR), rainer.speck@dlr.de, Germany
Dirk Frommholz, German Aerospace Center (DLR), dirk.frommholz@dlr.de, Germany

Abstract

The combined use of high resolution SAR and optical images is of great interest, especially in complex scenes like urban areas. However, this requires accurate registration of images with big radiometric and geometric differences. This paper describes an accurate and fast registration method that takes advantage of the geolocation accuracy of current sensors, instead of using more complex feature- or intensity-based methods. The proposed method can handle the ambiguities that arise in layover areas using a DEM or DSM. This is illustrated using several VHR SAR and optical images, including one urban scene in which a photogrammetric DSM was used.

1 Introduction

The interpretation of SAR images of complex scenes (e.g., urban areas) can be a difficult task due to radar specific imaging effects like speckle noise, multiple-bounce propagation, layover and shadowing. A useful approach for the analysis of such scenes is the combined use of SAR and optical images. Complex SAR signatures become much easier to interpret when using optical image data as an additional information source, and it has been shown how these two types of data can be used together for the retrieval of building information [1, 2].

In order to use SAR and optical images together, image registration must be performed. Because of the big radiometric and geometric differences between optical and SAR images, this was traditionally a complex task. Currently it is still an active research topic, with authors proposing and publishing new feature-based [3] and intensity-based [4] registration methods. However, when using images from current state-of-the-art spaceborne sensors like TerraSAR-X or Sentinel-1 representing SAR sensors, and Pleiades or WorldView representing optical sensors, these complex methods are no longer needed. Very accurate registration can be achieved nowadays just by using geometric registration, taking advantage of their high geolocation accuracy and using a digital elevation model (DEM) or a digital surface model (DSM). This approach is fully automatic and much faster than feature- and intensity-based registration methods. Also, when applied to images from sensors with tested geolocation accuracy, it is more robust.

In this paper, we will initially review the geolacation accuracy of current SAR and optical sensors, in order to prove that geometric registration is accurate enough and there is no need to use more complex registration methods. Then, an efficient and accurate implementation for the geometric image registration of SAR and optical images will be described. This implementation uses rational polynomial coefficients (RPC) for fast forward and backward geocoding of both SAR and optical images. The images can be in any arbitrary geometry (e.g., SAR images in slant-range), which means that there is no need to use orthorectified images. Also, the proposed method handles the ambiguities that arise in the layover areas of SAR images. Finally, results obtained with the described method will be shown for two different datasets: one with spaceborne optical data and a DEM, and another one with aerial optical data and a photogrammetric DSM. In both cases the SAR images were acquired by TerraSAR-X.

2 Geometric registration of SAR and optical images

2.1 Geolocation accuracy of current SAR and optical sensors

As already mentioned, this registration method is accurate and robust only when the sensors acquiring the images have a high geolocation accuracy. The accuracy of current sensors will be briefly reviewed here.

It is well known that SAR sensors are capable of very good geolocation accuracy due to their range-based imaging principle. In the case of TerraSAR-X, it is possible to achieve up to centimeter level accuracy after correcting different effects, like propagation delays and plate tectonics [5]. Just by using a simple atmospheric model, subpixel accuracy is easily obtained. In the case of Sentinel-1A, the commissioning phase also promises sub-pixel accuracy [6].

Optical spaceborne sensors traditionally had a much coarser geolocation accuracy [7]. However, this has changed with the current state of the art satellites: WorldView-1 has an accuracy of 4 m and WorldView-2 and 3 of 3.5 m [8]. These accuracy values are specified as the the circular error at 90% confidence (CE90). None of these sensors need any ground control points (GCP) for achieving the specified accuracies.

Regarding optical airborne sensors, its positional accuracy matches the accuracy of the GNSS receiver tied to the camera itself or its carrier. It can be as low as 0.6 m if operated in SBAS mode (satellite-based augmentation system) without a terrestrial reference station [9]. In practice, when the raw navigation data is refined by postprocessing or the GNSS is used in RTK mode (real-time kinematics) with a ground station, the geolocation accuracy will increase up to 0.07 m [10]. This is below one pixel for flight configurations and directly propagates to derived data products like aerial DSMs and true-ortho mosaics (TOMs).

2.2 A fast and simple implementation

The idea behind this registration approach is to compute the corresponding geographic coordinates for each pixel in the master image (forward geocoding), and then computing the correspondings pixel in the slave image from these coordinates (backward geocoding). Using this information the slave image is then resampled, obtaining a new image that matches the master image. However, the geographic coordinates for a given pixel usually cannot be computed directly, as the terrain height for that pixel is a priori unknown (and needed for performing forward geocoding). This height information should be obtained from a DEM or, if we also want to take into account man-made objects such as buildings, a DSM.

Because forward and backward geocoding has to be performed for every pixel and this can be a time consuming task, we need an efficient way to perform these calculations without losing accuracy. Optical image products usually provide the user with rational polynomial coefficients (RPC), sometimes also called rapid positioning capability. These are a set of polynomials that can be used for easy and accurate geocoding (i.e. errors in the order of 10^{-3} pixels) and that can be evaluated in a very fast way. RPCs can also be computed for a SAR image from the orbit information and other parameters in the image metadata, and also provide a similar accuracy [11]. An special case where no RPCs are needed would be an orthorectified image, which has already been processed and for which the geographic coordinates are known for every pixel.

Once the RPCs have been computed for both the SAR and the optical image, forward and backward geocoding can be easily performed. However, we still need the height of every pixel in the master image as an input for the forward geocoding. If we assume that each pixel only has one possible height, the registered image can be obtained very quickly. Although this assumption is not always true (e.g., layover areas in SAR images), it allows us to use a fast iterative solution. This iterative solution will be described here because it works well in most cases, and because of its simplicity and speed. The effects of the ambiguities in the layover areas will be shown in the next section, and we will explain how these should be handled.

For the mentioned iterative solution, we use an initial height value to compute the geographic coordinates of a pixel using the RPCs, and then this height value is updated using the DEM. Afterwards, new coordinates are computed, and the process is repeated iteratively until the height converges. Because terrain height varies smoothly, the height of a neighboring pixel can always be used as initial value, greatly reducing the number of iterations needed.

The runtime of this image registration algorithm depends only on the size of the master image. For a TerraSAR-X high resolution spotlight image, the registration takes approximately 5 seconds of runtime on a computer with an Intel i7 CPU and 16 GB of RAM.

2.3 Handling ambiguities in layover areas

As already mentioned, when registering an optical image to a slant-range SAR image, problems can arise in areas where layover effects are present. This occurs because multiple scatterers located at equal distance to the sensor and different heights appear superimposed at the same image pixel. When using the RPC and a DEM or DSM to find the geographic coordinates of this pixel, several solutions are possible (one for each of these scatterers). While all these solutions are valid and make sense for a SAR image, the co-registered optical image will have some artefacts unless the appropriate solution is chosen. The iterative approach described in the previous subsection will find one of these solutions, which may or may not be the appropriate one. In order to properly handle the ambiguities in layover areas we need to find all of these solutions, and then choose the one that provides a natural looking co-registered optical image.

To find all these solutions for a given pixel in the SAR image, we have to compute its geographic coordinates for different height values, and then intersect the resulting curve with the DEM or DSM. An example of this is shown in **Figure 1**, using the SAR image of a building and its corresponding DSM. This SAR image (**Figure 1a**) is in slant-range geometry, rotated (range axis directed downward), and has a pixel marked in red. This pixel corresponds to a layover area where backscattering from the ground, wall and roof appear mixed. All the possible locations for this pixel are computed and projected into the corresponding DSM. These are shown in **Figure 1b** as a blue line. The intersections of the DSM with the curve that defines the possible locations of the highlighted pixel can be seen in the plot from **Figure 1c**, with the intersections highlighted in red. In this plot, the three points from the DSM that appear in the highlighted pixel of the SAR image can be clearly identified.

Once all the intersections are found, we need to choose the solution that provides a natural looking co-registered optical image. In [12], the authors relate the viewing geometries of SAR and optical sensors, and explain how a properly co-registered optical image should look like an optical view with a line of sight is perpendicular to the SAR slant plane. This implies that among the possible solutions for a given pixel, the one we would expect to see in the co-registered optical image is the one

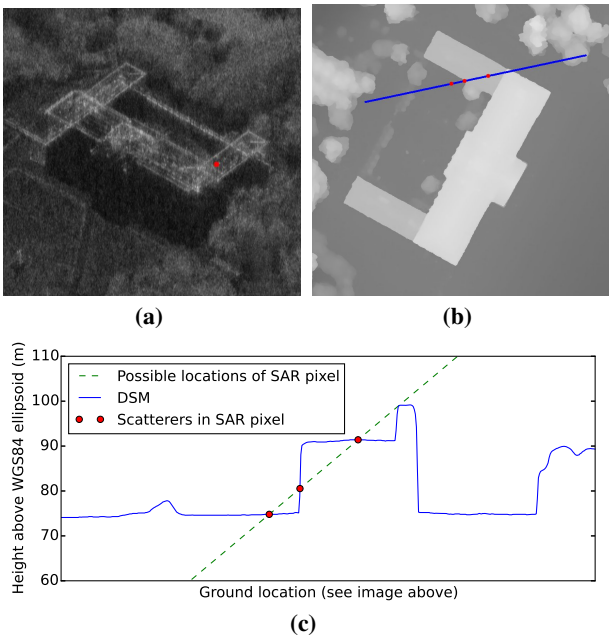

Figure 1: SAR image of a building with a pixel marked in red (a), DSM with all the possible locations of this pixel shown in blue (b), and plot showing the intersections between these locations and the DSM (c).

located at the highest altitude, as all the other solutions would appear hidden behind this one. For the example from **Figure 1**, this means that the point we should choose is the one from the roof.

An example of the effects that the ambiguities in the layover areas can have in the registration results is shown in **Figure 2**. A SAR image where a mountain appears superimposed over a lake due to layover can be seen in **Figure 2a**. This TerraSAR-X image is in slant-range geometry and rotated (range axis directed downward), and was acquired with a incidence angle of 22°. The co-registered optical image obtained with the previously described iterative approach is shown in **Figure 2b**, where some artefacts can be seen in the edge of the mountain. The correct registration results obtained applying the method described in this subsection can be seen in **Figure 2c**. The small differences between this last optical image and the SAR image are due to the coarse resolution of the DEM used for the registration. It is important to note that in this case the time needed for the registration will be higher. As a reference, in a test with a TerraSAR-X staring spotlight image and very high resolution DSM and optical images, the registration time was approx. 90 seconds.

3 Results

In this section, results obtained with the described method will be shown for two different scenes: one from Colorado Springs and another one from Oslo. In both cases, SAR images in slant-range geometry are used as the master images for the registration, i.e., the optical images are transformed to the SAR slant plane.

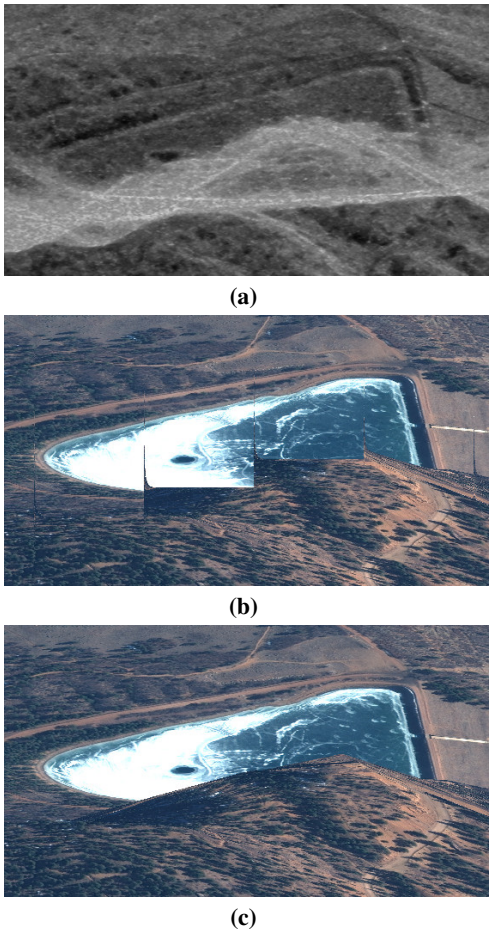

Figure 2: Layover area in SAR image (a), artefacts in co-registered optical image (b), and correct registration by selecting the appropriate scatterer for each pixel (c).

The SAR images of both scenes are TerraSAR-X single look slant range complex (SSC) images. These were acquired in staring spotlight mode, and have a resolution of approx. 20 cm in azimuth and 60 cm in slant range. Initially, speckle filtering was initially performed to these images. Then, the images were resampled to a square pixel size of 30 cm, and finally they were rotated (range axis directed downward). This simple transformation gives the co-registered optical images a more natural look, as if an optical sensor was looking perpendicular to the SAR slant plane.
Regarding the optical and elevation data, the Colorado Spring dataset uses a spaceborne optical image from WorldView-2 and the SRTM DEM, while the Oslo dataset uses aerial optical images and a photogrammetric DSM. The characteristics of this data will be described in more detail in the two following subsections.

3.1 Spaceborne optical data and SRTM

For this first scene (Colorado Springs) we used a WorldView-2 RGB image, pan-sharpened and resampled to a 50 cm ground space distance (GSD). This is a Ortho Ready Standard (2A) product, which are radiometrically corrected, sensor corrected and projected

to a constant base elevation, with no terrain corrections applied. The provided RPC can be used for geocoding with effectively the same accuracy as the rigorous sensor model (which, as described before, is 3.5 m CE90 for this sensor). The TerraSAR-X image has the previously mentioned characteristics and an incidence angle of 39.26°. Finally, the DEM is SRTM, which has 90 m horizontal resolution. As it will be shown, in many cases even a coarse DEM such as this one is sufficient for achieving high registration accuracy.

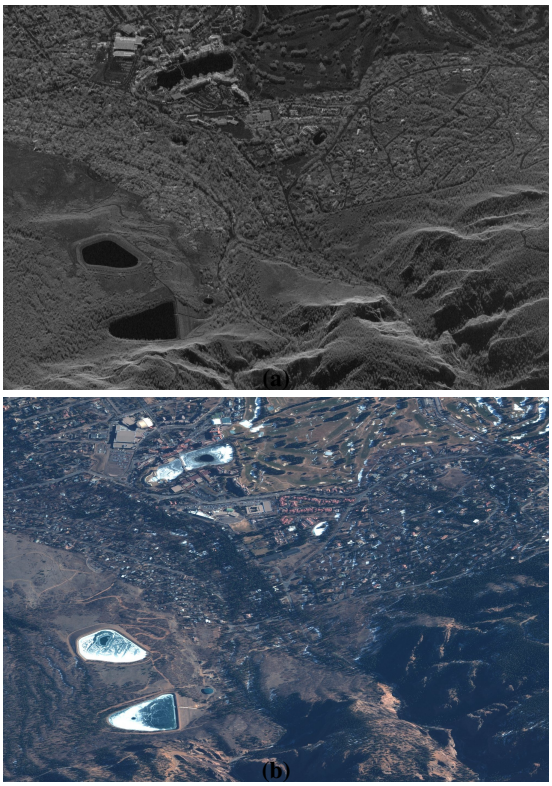

Figure 3: SAR image of Colorado Springs in slant-range geometry (a) and co-registered optical image (b).

Figure 4: Mosaic view showing the accuracy achieved using only geometric information and the SRTM DEM.

Figure 3 shows a part of the SAR image and the co-registered optical image. The registration accuracy achieved can be observed in the detailed mosaic view of **Figure 4**. Here, it can be seen how the roads and the large building align perfectly. Even though the building does not appear in the DEM, it is well aligned because of its small height (approx. 10 m) and the non-steep incidence angle with which the SAR image was acquired. If the building was higher and/or the incidence angle was steeper the layover effect would be more significant, and the registration results would not be so good. An example with tall buildings can be seen in the next subsection for the Oslo scene.

3.2 Aerial optical data and DSM

For this second scene (Oslo) we used aerial optical data and the corresponding photogrammetric DSM. In order to obtain this data a set of 276 UltraCam-X images with a GSD of 0.15 to 0.17 m was photogrammetrically processed. To refine the initial orientation from the GNSS and IMU tied to the sensor bundle adjustment with GCPs was applied. This yielded a positional RMS error of 0.03 x 0.04 x 0.08 m and an epipolar line deviation to selected corresponding points of 0.1 pixels. Subsequently a dense stereo matcher based on the SGM algorithm [13] with the census cost function was run to obtain the DSM. Having the height information the original bitmaps were orthogonally reprojected with enforced homogeneity constraints to generate the TOM to be transformed to the SAR slant-range image plane. The resolution of the derived data products which can be arbitrarily chosen was harmonized to 0.2 m to reflect the worst-case GSD of the input.

In this case the TerraSAR-X image has an incidence angle of 40.62°and all the other previously mentioned characteristics. The part of this SAR image corresponding to Oslo city center can be seen in **Figure 5**, together with the co-registered aerial optical image. The achieved registration accuracy can be observed in **Figure 6**, which shows a detailed mosaic view of some tall buildings. Here, it can be seen how all this tall buildings are perfectly aligned, proving how the proposed method can be used for achieving very accurate registration in complex scenes like urban areas if an accurate DSM is available.

Finally, a comparison of different images is shown in **Figure 7**. The SAR image is shown in **Figure 7a**. As explained before, most pixels in this SAR image correspond to more than one point in the DSM due to the layover effect. Of all the possible heights for a given pixel, the highest should be selected for the registration. These height values are shown in **Figure 7b**. As a reference, the tallest building in the image is approximately 100 m above the street level. The co-registered optical image obtained using the DSM is shown in **Figure 7c**. For comparison, the one obtained using SRTM is shown in **Figure 7d**. Here, two interesting details can be observed. The first one is that a DEM is no longer sufficient for achieving a good registration of elevated man-made objects as these buildings when using very high resolution images. The second

one is the lack of information on the walls of the buildings in **Figure 7c**, which are shown as constant surfaces of the color of the roof borders. This occurs because only the roofs of the buildings could be seen in the aerial optical image.

Figure 5: SAR image of the Oslo city center in slant-range geometry (a) and co-registered optical image (b).

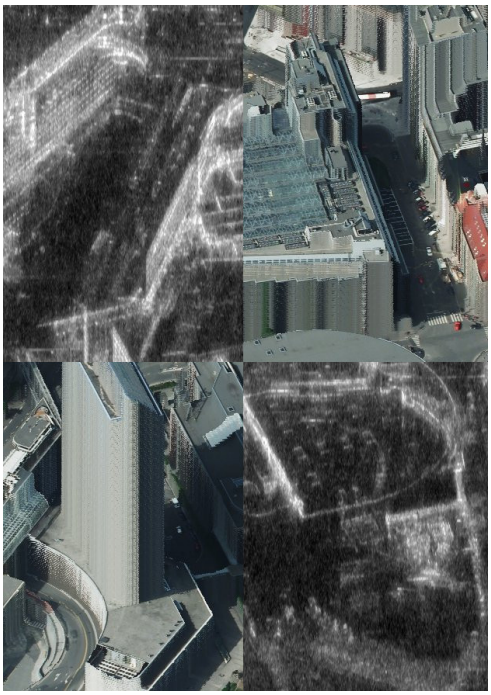

Figure 6: Mosaic view showing the accuracy achieved in urban areas with tall buildings when using a DSM.

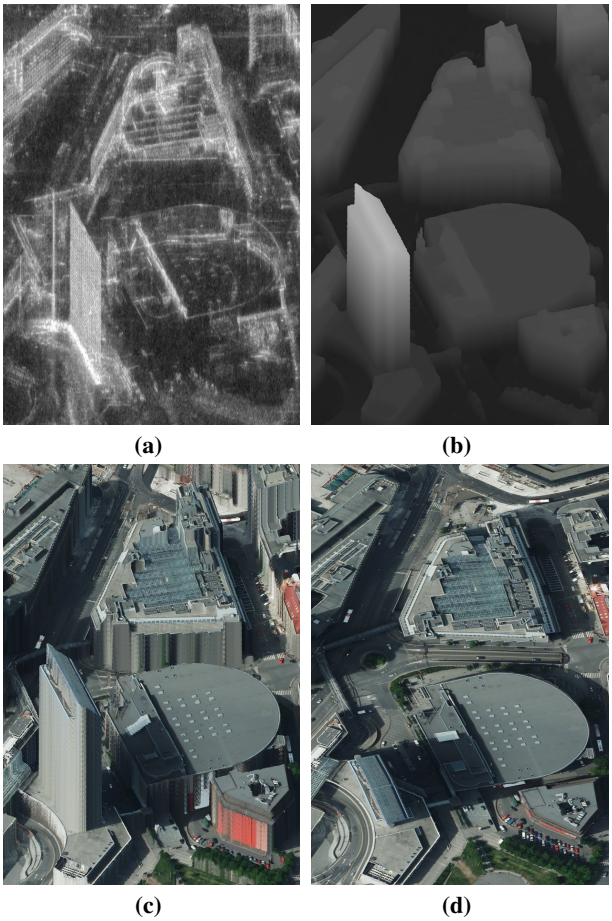

Figure 7: SAR image of some tall buildings in slant-range geometry (a), pixel height obtained from the DSM (b), co-registered optical image obtained with the photogrammetric DSM (c), and with the SRTM DEM (d).

4 Conclusions

The combined use of SAR and optical images is of great interest, especially in complex scenes such as urban areas. Traditionally, the registration of SAR and optical images was performed using complex feature- or intensity-based registration methods. In this paper we have shown that these methods are no longer required when using current state-of-the-art sensors. Because of their high geolocation accuracy, geometric registration using a DEM or a DSM will provide very good results. A fast and accurate implementation of this registration method was described in this paper. The proposed method is able to handle the ambiguities that appear in the layover areas of the SAR images, achieving good registration results even in these areas. The effects of these ambiguities are shown in **Figure 2**, together with the correct registration results.

Results obtained using the described registration method were shown for two different scenes: one from Colorado Springs and another one from Oslo. For both scenes, TerraSAR-X staring spotlight images were used. However, we used different types of optical and elevation

data: spaceborne optical data a coarse DEM for Colorado Springs, and aerial optical data and a very accurate photogrammetric DSM for Oslo. The results obtained for both scenes can be seen in **Figure 3** and **Figure 5**. Detailed views showing the accuracy achieved in both cases can be seen in **Figure 4** and **Figure 6**. Finally, a comparison of the registration results obtained with a DEM and a DSM in an area with tall buildings in Oslo city center can be seen in **Figure 7**. From these results we can conclude that while the registration using a coarse DEM will provide good registration results in most cases, a DSM should be used in urban areas if possible.

Acknowledgement

The authors would like to thank Frank Lehmann from the Institute of Optical Sensor Systems of the German Aerospace Center (DLR) for his contribution to the processing of the aerial optical image and digital surface model shown in this paper.

References

[1] V. Poulain, J. Inglada, M. Spigai, J. Y. Tourneret, and P. Marthon, "High-Resolution Optical and SAR Image Fusion for Building Database Updating," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 2900–2910, Aug. 2011.

[2] T. L. Wang and Y. Q. Jin, "Postearthquake Building Damage Assessment Using Multi-Mutual Information From Pre-Event Optical Image and Postevent SAR Image," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 452–456, May 2012.

[3] H. Sui, C. Xu, J. Liu, and F. Hua, "Automatic Optical-to-SAR Image Registration by Iterative Line Extraction and Voronoi Integrated Spectral Point Matching," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6058–6072, Nov. 2015.

[4] S. Suri and P. Reinartz, "Mutual-Information-Based Registration of TerraSAR-X and Ikonos Imagery in Urban Areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 939–949, Feb. 2010.

[5] U. Balss, C. Gisinger, X. Y. Cong, R. Brcic, S. Hackel, and M. Eineder, "Precise Measurements on the Absolute Localization Accuracy of TerraSAR-X on the Base of Far-Distributed Test Sites," in *EUSAR*, Jun. 2014, pp. 1–4.

[6] A. Schubert, D. Small, N. Miranda, D. Geudtner, and E. Meier, "Sentinel-1a Product Geolocation Accuracy: Commissioning Phase Results," *Remote Sensing*, vol. 7, no. 7, pp. 9431–9449, Jul. 2015.

[7] P. Reinartz, R. Müller, S. Suri, M. Schneider, P. Schwind, and R. Bamler, "Using geometric accuracy of TerraSAR-X data for improvement of direct sensor orientation and ortho-rectification of optical satellite data," in *IGARSS*, Jul. 2009.

[8] DigitalGlobe, "Geolocation Accuracy of WorldView Products," Tech. Rep.

[9] Novatel, "Novatel OEM615 Dual-Frequency GNSS Receiver Product Sheet," Tech. Rep.

[10] W. Rüther-Kindel and J. Brauchle, "The Salsa Project - High-End Aerial 3D Camera," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-1/W2, pp. 343–348, Aug. 2013.

[11] G. Zhang, Z. Li, H. b. Pan, Q. Qiang, and L. Zhai, "Orientation of Spaceborne SAR Stereo Pairs Employing the RPC Adjustment Model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2782–2792, Jul. 2011.

[12] H. Anglberger, R. Speck, and H. Suess, "Transforming optical image data into a SAR system's range-based image space," in *Proc. of SPIE*, vol. 8714, 2013.

[13] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328–341, Feb. 2008.