AUTOMATIC ALIGNMENT OF HIGH RESOLUTION OPTICAL AND SAR IMAGES FOR URBAN AREAS

Stefan Auer[1], Michael Schmitt[2], Peter Reinartz[1]

[1] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen
[2] Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich

ABSTRACT

This paper presents the basics and functionality of SimGeoI, a simulation-based framework for the automated interpretation and alignment of optical and SAR remote sensing data. SimGeoI has been developed in order to align optical and SAR data based on given geometric information about objects represented by digital surface models. Thereby, the analysis of urban scenes is possible with independence of sensor type and perspective. After a brief introduction of the processor environment, possible applications of the framework are indicated with results of a case study for Istanbul (WorldView-2 and TerraSAR-X data). In this context, opportunities in the context of a joint analysis of high resolution optical and SAR data are addressed, i.e. concerning data fusion, change detection, and machine learning tasks.

Index Terms— Optical Data, SAR Data, Data Fusion, Simulation, Interpretation, Urban Areas, Ray Tracing

1. INTRODUCTION

The alignment of optical data and synthetic aperture radar (SAR) data from airborne or spaceborne sensors is a challenging task due to fundamental differences of the pertinent imaging concepts. The different imaging concepts of both sensor types lead to contrary geometric distortion effects in the images, in particular for elevated objects such as buildings [1]. That said, a pixel-to-pixel or patch-to-patch spatial matching of high-resolution optical images and SAR images is only achievable for flat scenes without elevated objects. However, typical urban scenes are composed by man-made structures, trees, and height changes of the ground (see e.g. [2] for a discussion of related challenges). SimGeoI (simulator of geo-referenced interpretation layers) tackles the problem with an object-related procedure, including geometric prior knowledge from digital surface models (DSMs).

2. INTRODUCTION OF FRAMEWORK SIMGEOI

SimGeoI comprises two components (see workflow in Fig. 1): GeoRaySAR [3] for simulating interpretation layers for SAR imagery and GeoRayOpt, a newly developed component for
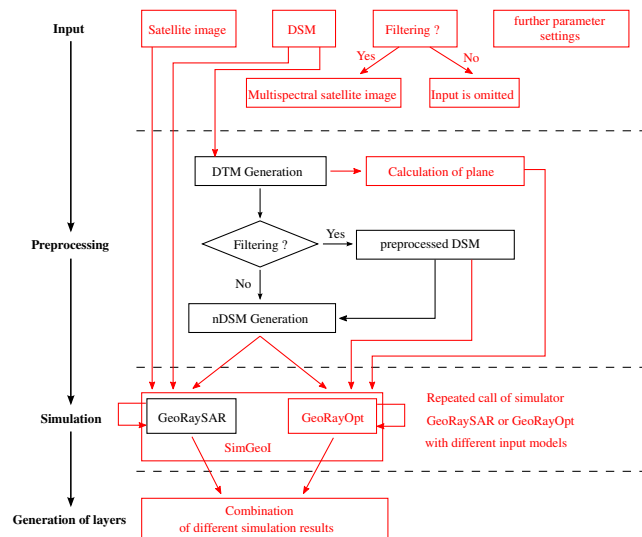

Fig. 1: Procedure of the SimGeoI environment (red: developed and implemented methodology; black: existing modules for simulation and DSM filtering/decomposition).

providing artificial images and interpretation layers for optical data (see details in [4]). The name SimGeoI (Simulator of Geo-referenced Interpretation Layers) refers to the integrated framework where the type of simulation and scene interpretation is triggered by input information. The first input is the image meta file, corresponding to a geo-referenced optical image or SAR image with UTM coordinates. The file contains the necessary information for defining the simulation set-up (sensor perspective, image properties, scene average height) and is interpreted automatically. The second input is a scene DSM in UTM coordinates which is required to generate the scene model for ray tracing.

In advance to the simulation step, the DSM is filtered (removal of of trees, suppression of noise). Thereafter, a DTM describing terrain (see method in [5]) and an nDSM including elevated objects are generated. The combination of simulated images from the DSM, the nDSM and the DTM leads to interpretation layers for optical and SAR images. Tables 1 and 2 summarize simulation combinations used for generating the

case study results presented in the paper. To separate object from background, strong diffuse backscattering is assigned to model surfaces. Based on that, binary layers (background: 0, foreground: 1) are easily derived and combined which are used to extract image sections related to basic scene classes. In the optical case, the surface models (DSM, DTM, nDSM) are combined with different illumination settings (L1: signal source at camera position for simulating full object extent in image, L2: signal source representing sun light for shadow simulation). In the SAR case, the signal source is defined at the antenna position for identifying signal response from buildings and ground as well as double reflection regions and shadow areas. The analysis can be continued on the building level by extracting building block models from the nDSM. Therefore, a minimum size for isolated segments in the nDSM has to be defined (2000 pixels for case study below).

Table 1: Combination of simulation results to interpretation layers for optical images; L1: signal source at sensor position; L2: signal source representing sun illumination. Lay: layers, Sim: simulation.

Sim. / Lay.	L1 DSM	L1 nDSM	L1 DTM	L2 DSM	L2 nDSM
Building (full extent)		x			
Ground (full extent)	x	x			
Building (illuminated)					x
Ground (illuminated)				x	x
Sun shadow			x	x	
Building (shadowed)	combination of "Building" layers				
Ground (shadowed)	combination of "Ground" layers				

Table 2: Combination of simulation results to interpretation layers for SAR image; L1: signal source at sensor position.

Simulation / Layers	L1 DSM	L1 nDSM	L1 DTM
Building layover		x	
Ground		x	
Signal double reflection	x		
Shadow	x		x

coordinates) and orthophoto, as well as a TerraSAR-X spotlight image. The properties of the image data to be interpreted (WorldView-2 image and TerraSAR-X image, both geocoded in UTM coordinates) are summarized in Tables 3 and 4. The sensor perspectives w.r.t. the urban area are not the same.

**Table 3**: WorldView-2 data properties for Istanbul test site (geo-referenced, level 2A).

Pixel spacing (east, north)	0.5 m
Off-nadir angle	12.8°
Scene azimuth angle	235.6°
Sun azimuth angle	145.2°
Sun elevation angle	65°
Acquisition date	2015-07-27

Table 4: TerraSAR-X data properties for Istanbul test site (geo-referenced, level 1B).

Azimuth resolution	1.15 m
Ground range resolution	0.99 m
Pixel spacing (east, north)	0.5 m
Signal incidence angle	41.0°
Orbit	ascending
Acquisition date	2008-05-05

3. CASE STUDIES

3.1. Data pool

The data pool for the Istanbul scene includes a small set of WorldView-2 images (multispectral + pan, the latter with 0.5 m spatial resolution), the thereof derived digital surface model [6] (horizontal and vertical resolution: 0.5 m and 1 m; UTM

3.2. Scene Interpretation

As described in Section 2, interpretation layers are generated based on simulation combinations. Figs. 2 and 3 present examples of extracted image parts from the WorldView-2 and TerraSAR-X images using the interpretation layers. Despite the difference of sensor type and imaging perspective, image parts can be related (building, ground). Besides, sensor-specific interpretation layers can be generated which help to analyze the scene (examples: sun shadow in Fig. 2c, radar
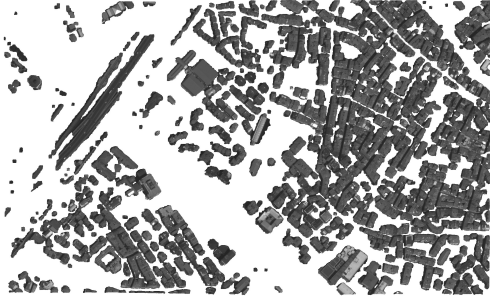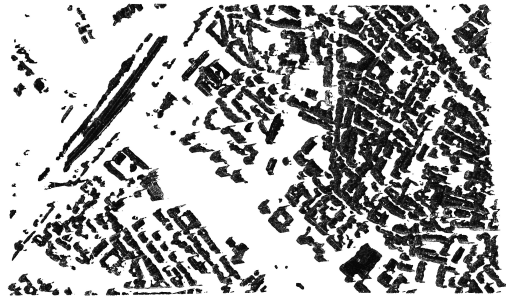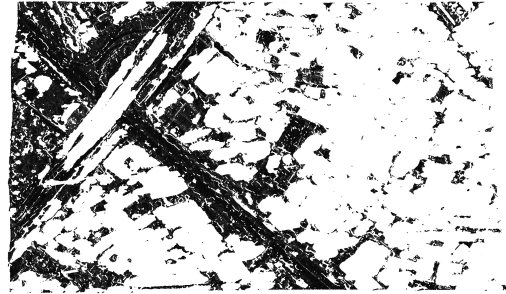
(a) WorldView-2: buildings (full extent)
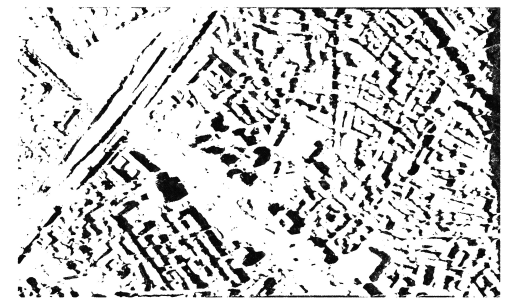

(b) WorldView-2: ground
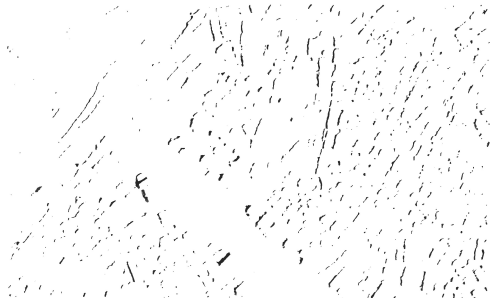
(c) WorldView-2: sun shadow

Fig. 2: Istanbul test site: extracted image parts from WorldView-2 image. Background marked by white color.

(a) TerraSAR-X: buildings

(b) TerraSAR-X: ground

(c) TerraSAR-X: shadow

Fig. 3: Istanbul test site: extracted image parts from TerraSAR-X image. Background marked by white color.

signal related shadow in Fig. 3c). Note that non-white pixels in Figs. 2 and 3 represent original image values.

The analysis is continued on the building level where 88 individual building models are identified from the nDSM. Again, interpretation layers are generated for the input building nDSMs. Fig. 4 presents the result for different building types, which clearly denies the impression of similarity in Figs. 2a and 3a. Both the shape and the position of the building-related image parts vary. Looking at the first row, it is seen that that the off-nadir view of WorldView-2 is considered by SimGeoI which extracts the full building extent. The second row of Fig. 4 indicates the impact of sun shadow on building-related parts (here: moderate effect due to big sun elevation angle). The third and fourth row of Fig. 4 complement extracted image parts from the TerraSAR-X image: building layover and shadow. Here, the simulated shapes are

also helpful for visual interpretation as signal responses are not obtained for all parts of the buildings and mostly linked to specular reflections at building corners.

4. DISCUSSION

The case study results above visualize the different imaging effects inherent to the interpretation layers (e.g. the impact of sun light in the optical case, different geometric distortions). From a general viewpoint, three main aspects can be added:

- Buildings: Image parts of optical and SAR data related to buildings do not directly correspond (individual pixels do not match) as the simulated information is generated for the original image plane. In contrast, the extracted image parts form a complementary source of in-
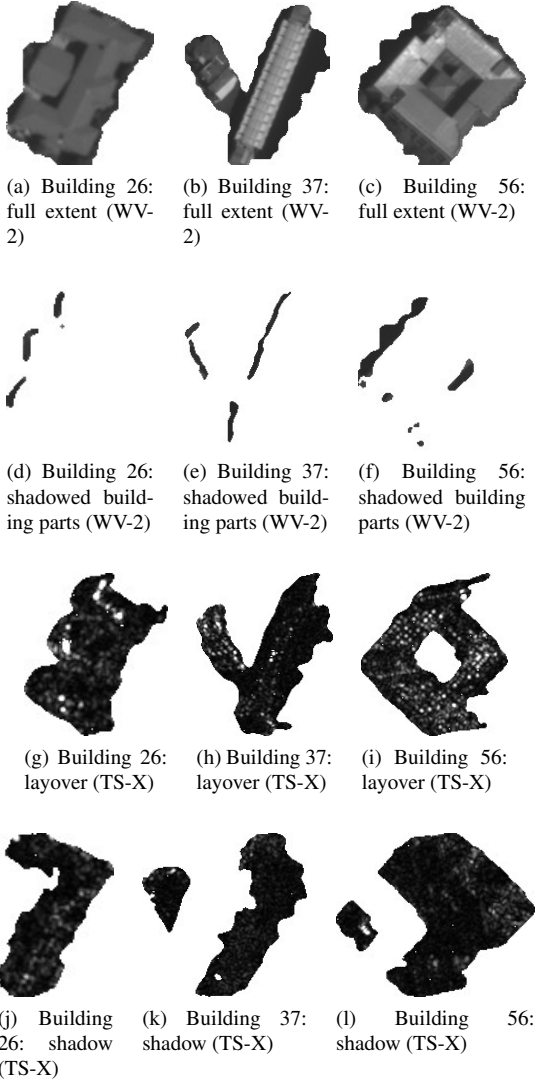
(a) Building 26: full extent (WV-2)

(b) Building 37: full extent (WV-2)

(c) Building 56: full extent (WV-2)

(d) Building 26: shadowed building parts (WV-2)

(e) Building 37: shadowed building parts (WV-2)

(f) Building 56: shadowed building parts (WV-2)

(g) Building 26: layover (TS-X)

(h) Building 37: layover (TS-X)

(i) Building 56: layover (TS-X)

(j) Building 26: shadow (TS-X)

(k) Building 37: shadow (TS-X)

(l) Building 56: shadow (TS-X)

Fig. 4: Extracted image parts from WorldView-2 and TerraSAR-X images for individual buildings. WV-2: WorldView-2; TS-X: TerraSAR-X.

formation. The full building extent is described for the respective data source.

- Geo-referencing: The interpretation layers are generated with UTM coordinates and, hence, can be directly imposed on the images. The extracted image parts of for WorldView-2 and TerraSAR-X data indicate the spatial accuracy of the layers with respect to the satellite images. The result is promising as the underlying DSM is derived from four WorldView-2 images whose localization accuracy is lower than for TerraSAR-X.

- Applications: The extraction of related image parts enables an object-related analysis of satellite images. As an example, the extracted image parts for the full scene can be used to provide training data for basic classes or

extract characteristic parameters. The identification of image parts corresponding to individual objects support scene monitoring over time using multi-modal satellite data. In the extreme case of only two data sets, change detection methods can be supported with prior knowledge about scene objects. Thereby, images become comparable for different sensor viewing perspectives, which is crucial for urgent situations.

5. CONCLUSION

This paper has introduced the functionality of SimGeoI, a simulation-based method for the automatic alignment of optical and SAR satellite data. In this context, scene interpretation is conducted automatically and in absolute coordinates using a digital surface model (DSMs) and meta information of the images. As a core element of SimGeoI, ray tracing techniques are applied in order to generate interpretation layers which geometrically describe basic scene classes of urban areas. Case study results for a scene in Istanbul, based on WorldView-2 and TerraSAR-X data, indicate that SimGeoI can open the door to the object-related analysis of satellite images with independence of the sensor type and perspective.

6. REFERENCES

[1] G. Palubinskas, P. Reinartz, and R. Bamler, "Image acquisition geometry analysis for the fusion of optical and radar remote sensing data," *Int. J. Image Data Fusion*, vol. 1, no. 3, pp. 271–282, 2010.

[2] M. Schmitt and X. Zhu, "On the challenges in stereogrammetric fusion of SAR and optical imagery for urban areas," in *The Int. Arch. Photogram. Rem. Sens. Spatial. Inform. Sci.*, 2016, vol. 41, pp. 719–722.

[3] J. Tao, S. Auer, G. Palubinskas, P. Reinartz, and R. Bamler, "Automatic SAR simulation technique for object identification in complex urban scenarios," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 7, no. 3, pp. 994–1003, 2014.

[4] S. Auer, I. Hornig, M. Schmitt, and P. Reinartz, "Simulation-based interpretation and alignment of high-resolution optical and SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, submitted.

[5] R. Perko, H. Raggam, K.H. Gutjahr, and M. Schardt, "Advanced DTM generation from very high resolution satellite stereo images," in *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2015, vol. II-3/W4, pp. 165–172.

[6] P. d'Angelo and G. Kuschk, "Dense multi-view stereo from satellite imagery," in *Proceedings of IGARSS Conference*, 2012, pp. 6944–6947.