# Fusion of optical and radar data for the extraction of higher quality information

Gintautas Palubinskas*, Aliaksei Makarau, Junyi Tao
Remote Sensing Technology Institute, German Aerospace Center DLR, Oberpfaffenhofen, Germany

*Gintautas.Palubinskas@dlr.de, phone 49 8153 28-1490, fax 49 8153 28-1444, www.dlr.de/eoc

## ABSTRACT

Information extraction from multi-sensor remote sensing imagery is an important and challenging task for many applications such as urban area mapping and change detection. Especially for optical and radar data fusion a special acquisition (orthogonal) geometry is of great importance in order to minimize displacement effects due inaccuracy of Digital Elevation Model (DEM) used for data orthorectification and existence of unknown 3D structures in a scene. Final data spatial alignment is performed by recently proposed co-registration method based on a Mutual Information measure. For a combination of features originating from different sources, which are quite often non-commensurable, we propose an information fusion framework called INFOFUSE consisting of three main processing steps: feature fission (feature extraction aiming at complete description of a scene), unsupervised clustering (complexity reduction and feature conversion to a common dictionary) and supervised classification realized by Bayesian/Neural networks. An example of urban area classification is presented for the orthogonal acquisition of spaceborne very high resolution WorldView-2 and TerraSAR-X Spotlight imagery over Munich city, South Germany. Experimental results confirm our approach and show a great potential also for other applications such as change detection.
Additionally, simulation of optical and SAR data using high quality (Digital Surface Model) DSM in different acquisition geometries will be performed in order to make further interpretation and processing of SAR imagery easier. Learning of semantic relationships between objects in optical and radar data will help to enhance information extraction for various applications as classification, interpretation, change detection and scene reconstruction.

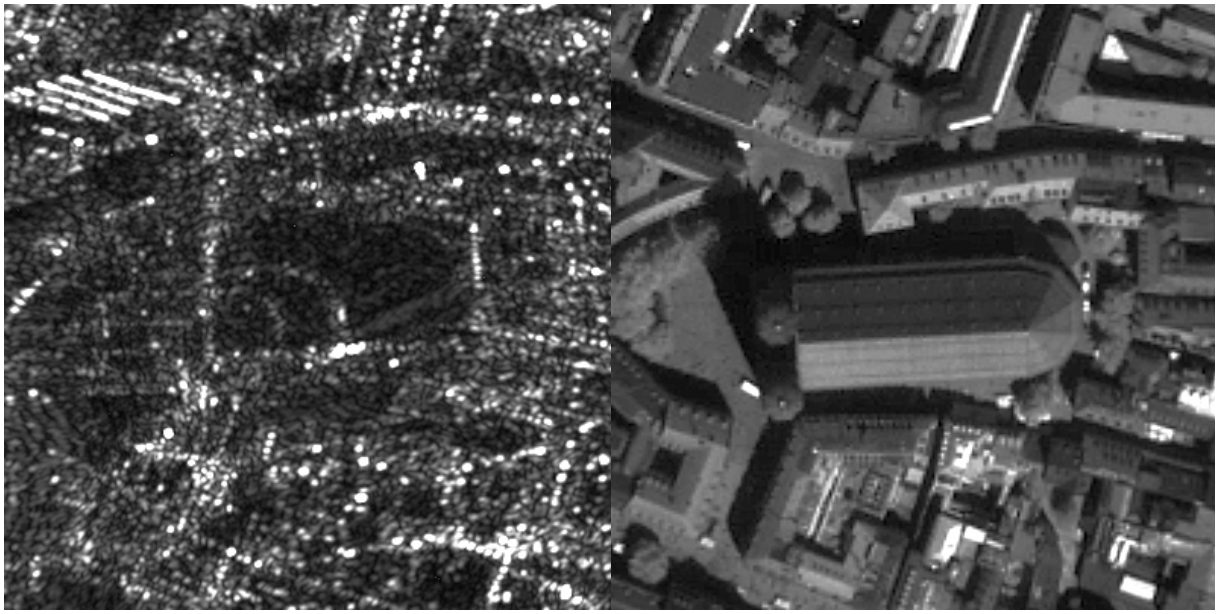
**Keywords:** multi-sensor, remote sensing, optical image, SAR, information fusion, classification, change detection, simulation

## 1. INTRODUCTION

Data fusion is a rapidly developing topic in various application areas during the last decades. Image fusion in remote sensing is one of them. However fusion of different sensor data such as optical and radar imagery is still a challenge. In this paper the term ‘radar’ is equivalent to Synthetic Aperture Radar (SAR).

### 1.1 Fusion problem

For the fusion of data from sensors exhibiting different acquisition geometries such as optical and radar systems it is important to understand their influence on the fusion process and to optimize it if necessary. For example, in Figure 1 it is practically impossible to recognize Frauenkirche (tourist attraction, Munich city) in SAR image even with the help of optical image. Special data acquisition geometry can help enormously as can be seen later in Section 2.



(a) (b)

Figure 1. Part of Munich center with Frauenkirche acquired by IKONOS panchromatic mode (a) and TerraSAR-X Spotlight mode (b) using the accidental satellite formation.

### 1.2 Fusion concept

We propose the following fusion concept shown in Figure 2.

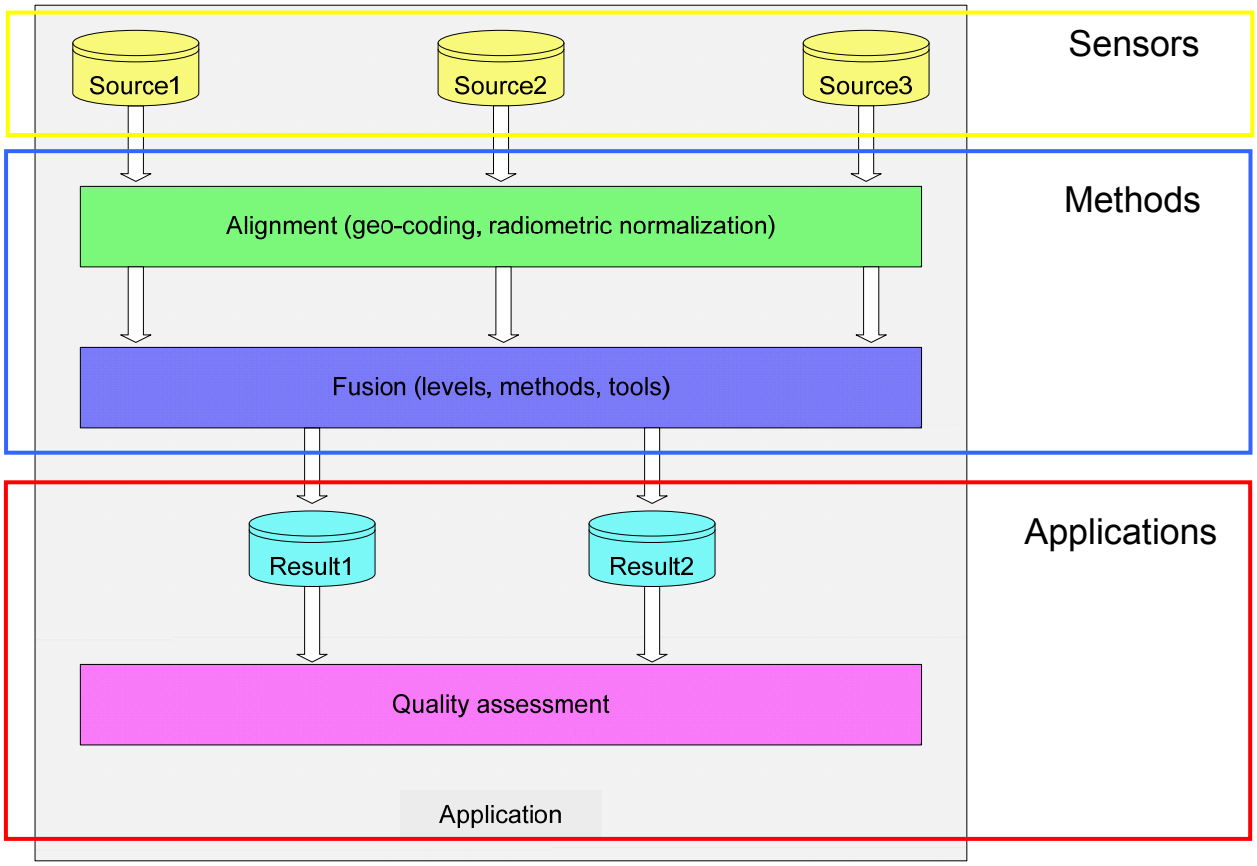



Figure 2. Proposed fusion concept.

According the proposed fusion concept already data preprocessing steps such orthorectification and coregistration introduced in the following Section belong to data fusion.

### 1.3 Preprocessing

For the data fusion on the lowest level (pixel or image-based) data preprocessing such as orthorectification and coregistration is an important prerequisite for a success. Here are the main tasks:

Orthogonal acquisition geometry (Palubinskas 2010)

to minimize displacement effects

Orthorectification

optical imagery (using TSX GCPs) (Reinartz 2011)

SAR imagery

Coregistration

automatical using Mutual Information (Suri 2010)

Pansharpening (multispectral + panchromatic)

General Fusion Framework (Palubinskas 2011)

Despeckling (SAR)

Nonlocal means filter (Deledalle 2009)

## 2. ORTHOGONAL ACQUISITION GEOMETRY

In this Section we propose an optimal optical and radar sensor formation for an image acquisition compensating/minimizing ground displacement effects of different sensors (Palubinskas 2010). A sum of look angles should give approximately 90° (Figure 3, left drawing). Flight directions should be as parallel as possible and perpendicular to look directions which are opposite for different sensors (Figure 3, right drawing). Same flight directions are not required in general e.g. airborne case. This sensor configuration allows e.g. a recovery of 3D object shadows during further data fusion, except a case when the Sun illumination direction is the same as for SAR look direction. Displayed left looking radar and right looking optical sensor formation can be preferable due to the Sun illumination direction which is from an optical sensor to the target on the Earth in order to see that side of a 3D object which is in shadow in the radar image and thus enable full reconstruction of a 3D object. Of course, the second possible sensor formation with a right looking radar and left looking optical sensor can be useful for data fusion too.

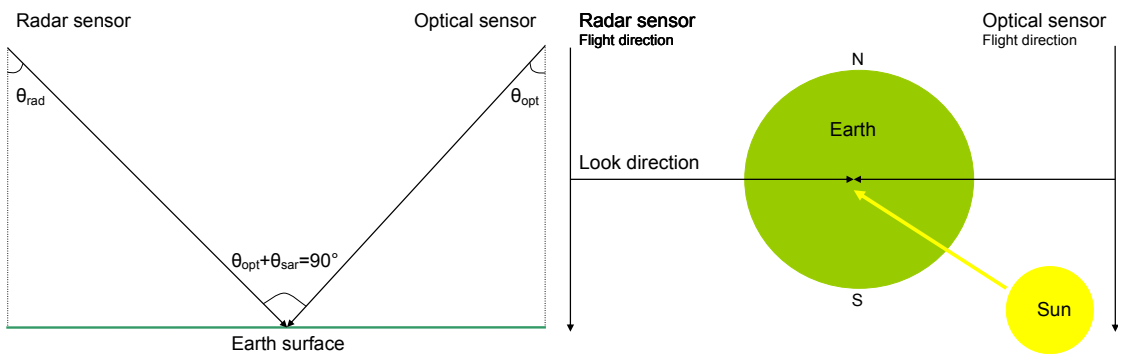
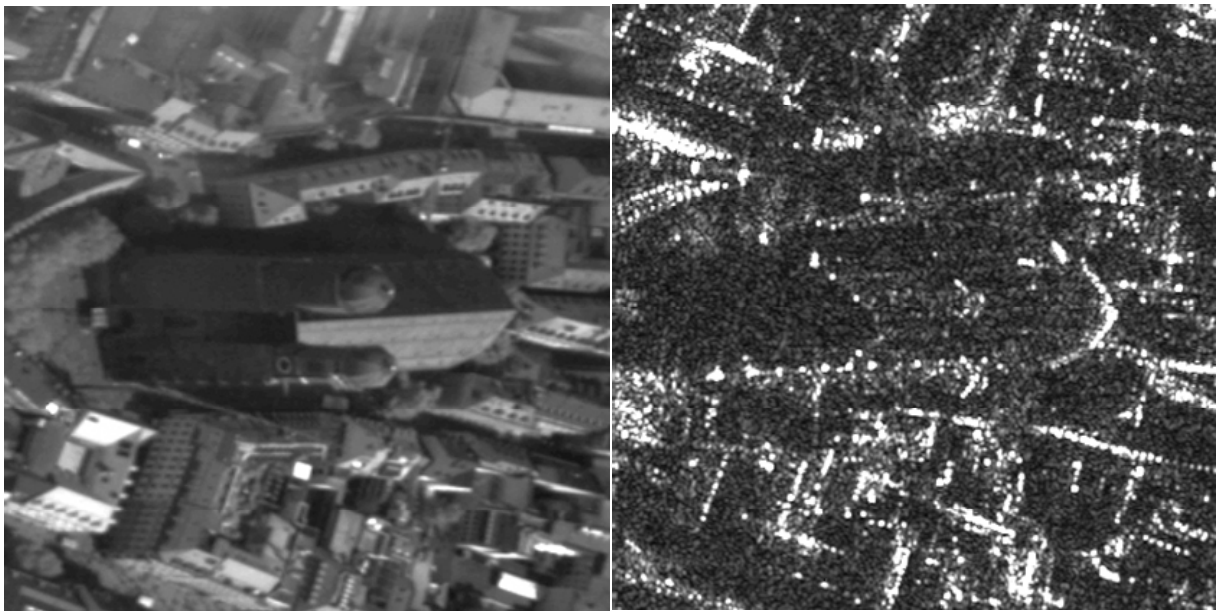



Figure 3. Proposed optical and radar sensor formation is illustrated. A sum of look angles should give 90° (left drawing). Flight directions should be parallel, in same direction and perpendicular to look directions which are opposite for different sensors (right drawing). Sun illumination direction is from an optical sensor to the target on the Earth.

Part of Munich center with Frauenkirche (tourist attraction) acquired by WV-1 (a) and TS-X (b) using the proposed satellite formation is shown in Figure 4. Ground objects like streets and plazas can be easily detected and found at the same geographical position in both images. Other structures: buildings (e.g. building block in the upper left corner of the image, church with two towers in the bottom left corner of the image) and trees can be easily indentified in both images. Only the feet of the buildings, which are differently projected in the radar image due to foreshortening in radar are found

at slightly different positions. So the roofs and tree crowns are well in place and can be overlaid correctly for any further processing.



(a) (b)

Figure 4. Part of Munich center with Frauenkirche acquired by WorldView 1 panchromatic mode (a) and TerraSAR-X Spotlight mode (b) using the proposed orthogonal satellite formation.

## 3. INFOFUSE MULTISENSOR CLASSIFICATION

INFOFUSE classification approach is based on a combination of both unsupervised clustering and supervised classification, thus allowing the usage of different features and scales for data and easy inclusion of the prior information using Bayesian/Neural Networks.

### 3.1 INFOFUSE framework

Data fusion framework consists of three main stages as shown in Figure 5 (Palubinskas 2008).

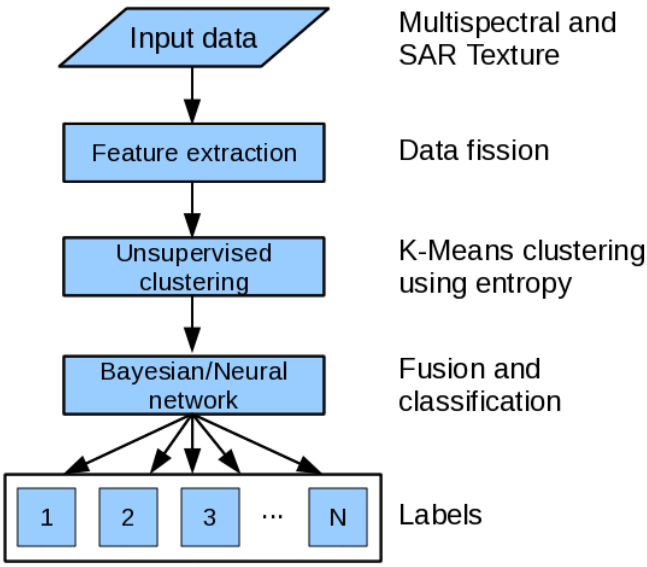



Figure 5. INFOFUSE framework for multi-sensor data fusion and classification.

**Feature extraction from input datasets**

Raw signal data are usually quite difficult to interpret or classify, so the preprocessing step—feature extraction—is often unavoidable. One tries to extract as much as possible various features from one source (information fission) or set of sources in order to produce a good classification. These features are expected to characterize different properties of structures and objects. After the feature extraction a large amount of redundant information is obtained.

**Dimensionality reduction using unsupervised clustering**

Since the aim of this step is to combine features with similar properties and to reduce the dimensionality of the calculated feature data, any unsupervised clustering method can be employed for this task. K-means clustering based on entropy (Palubinskas 1999) was applied on each extracted feature separately. The number of clusters for each feature can be different and defined individually according to the type of the feature.This step is performed to acquire the unique description of the data in terms of clusters and to reduce the dimensionality of the extracted features.

**Fusion of the clustered features**

A Bayes network or a Neural network is employed to fuse the extracted features and to produce the inference (i.e. classification through fusion). Bayesian or Neural network allows to combine information from different sources of measurement, therefore the fusion of incommensurable features (numerical, logical, semantical, etc.) can be performed. Supervised training of the network allows to estimate the network state and a classification is possible to perform.

### 3.2 Classification WV-2 VNIR and TS-X HS data (Munich center)

In order to investigate and illustrate the effectiveness of the proposed approach we have chosen the area of Munich city as a test scene. Munich contains variety of urban building types and structures, such as old town, residential area, low- , medium-, and high-rise buildings, rail road, water regions, bare soil, etc. Two very high resolution datasets were chosen: WorldView-2 multispectral imagery and TerraSAR-X Spotlight HS mode data. WorldView-2 multispectral data were obtained at the 12-th July 2010, 10:30:17 local time. Multispectral data contains 8 11-bit bands in 2m spatial resolution, the panchromatic data contain one 11-bit band in 0.5m spatial resolution. The spectral bands were pan-sharpened using an image fusion method based on high-frequency image data addition to low-resolution spectral image (Palubinskas 2011). This method provides minimal distortion of spectral and spatial characteristics of multispectral imagery (Makarau 2010). VNIR bands were especially used in our experiment since most of the very high resolution spaceborne multispectral sensors (e.g. IKONOS, GeoEye-1, Quickbird, etc.) acquire only VNIR data. TerraSAR-X data (Spotlight HS, EEC, VV polarization) were acquired at the 7-th June 2008, 06:17:48 local time. TerraSAR-X data were registered to the pansharpened multispectral image. Co-Occurence texture features (Haralick 1973) for SAR data were calculated (among the texture features are the Mean, Variance, Homogeneity, Contrast, Dissimilarity, Entropy, Second Moment, and Correlation).

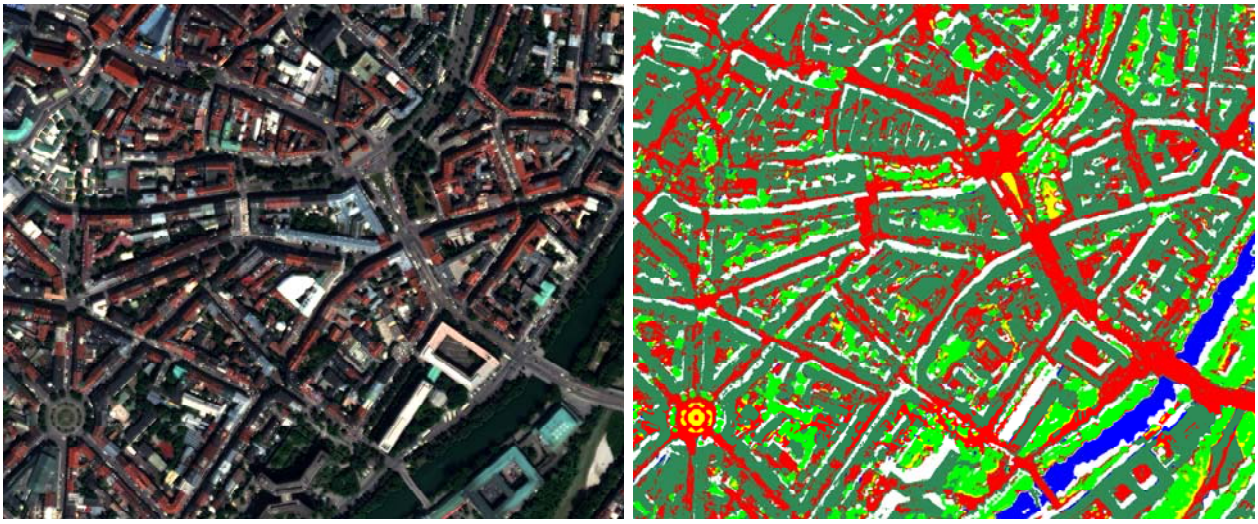
Table 1. List of classes/sub-classed used for classification.

| Label | Classes/subclasses |
|---|---|
| Buildings | 8 subclasses |
| Roads | 2 subclasses |
| Water | 1 class |
| Forest/Trees | 1 class |
| Grass | 1 class |
| Shadows | 1 class |

In this experiment we have selected 6 main classes for the urban scene: 1) Building; 2) Road; 3) Water; 4) Forest/Tree; 5) Grass; 6) Shadow (see Table 1). It should be noted that, for example, buildings have different material of the roofs, therefore highly varying spectral characteristics of the material (tiles, concrete, highly reflecting metal, etc.) make difficult to classify such inhomogeneous objects into one class of interest. In this experiment the building class contains the following types of building roofs: tiles roof, concrete roof, dark color roof, green color metal roof, blue color metal roof, glass roof, highly reflecting roof, grass on the roof. Roads class contains the following types of pavements: asphalt pavement and concrete pavement. The ground truth for the area under investigation was proofed by the ATKIS vector map provided by Bavarian State Agency for Surveying and Geoinformation (Landesamt für Vermessung und Geoinformation). The number of clusters in the unsupervised clustering usually has value between 40 and 100.

Experimental search found that the value of 80 provides significant dimensionality reduction with high accuracy of the land cover classification. A Multilayer perceptron (MLP) was employed for the data fusion and classification implemented in the IDL. A feed-forward neural network based was employed. The network contains two hidden layers with 16 neurons in each layer. Training of the MLP made 1000 training epochs.

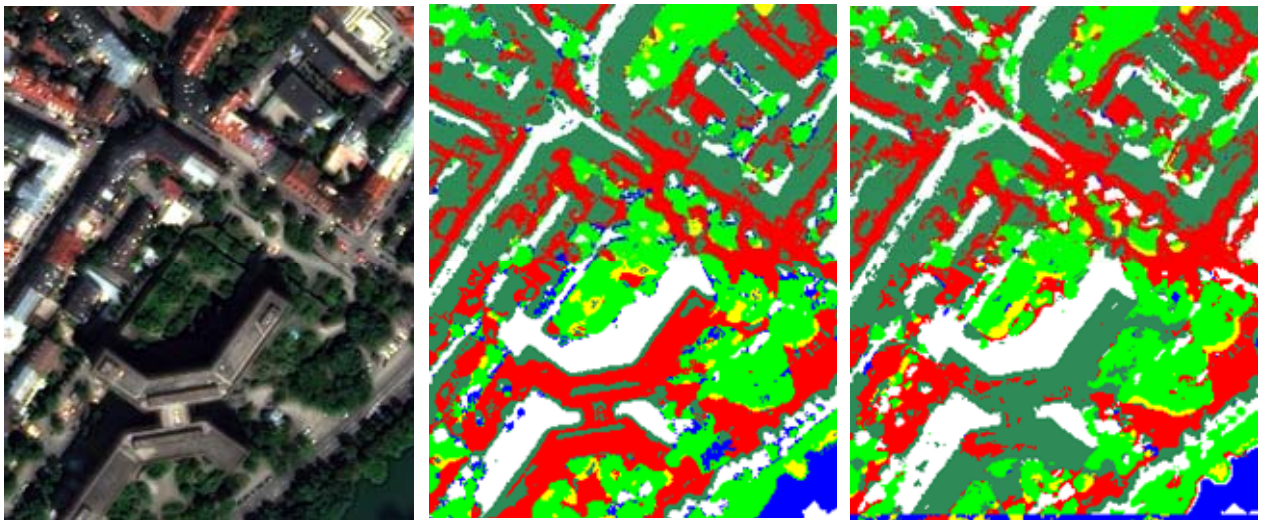
INFOFUSE classification maps for Munich city center are presented in Figure 6: WV-2 RGB composite (a), joint classification of WV-2 VNIR bands and TS-X features (b).



(a) (b)

Figure 6. WV-2 RGB composite (a), INFOFUSE classification WV-2 VNIR+TS-X (b) for Munich city center.

Zoomed part of INFOFUSE classification for Munich city center is presented in Figure 7: WV-2 RGB composite (a), classification of WV-2 VNIR bands (b), classification of WV-2 VNIR bands and TS-X features (c). We see how the addition of SAR information helps to extract correctly the building seen in lower part of an image.



(a) (b) (c)

Figure 7. WV-2 RGB composite (a), INFOFUSE classification WV-2 VNIR (b), INFOFUSE classification WV-2 VNIR+TS-X (c) for Munich city center zoom.

For the quantitative analysis of the quality of INFOFUSE classification confusion matrices are presented for WV-2 VNIR classification in Table 1 and for joint WV-2 VNIR and TS-X features classification in Table 2. When comparing

the Tables we see that the addition of SAR information allows to decrease confusion between road and building classes significantly.

Table 2. Confusion matrix of WV-2 VNIR classification.

| *Class* | *Water* | *Grass* | *Trees* | *Buildings* | *Road* | *Shadow* |
|---|---|---|---|---|---|---|
| *Water* | 100.00 | 0 | 0 | 0 | 0 | 0 |
| *Grass* | 0 | 85.47 | 0 | 0 | 0 | 0 |
| *Trees* | 0 | 14.53 | 100.00 | 0 | 0 | 0 |
| *Buildings* | 0 | 0 | 0 | **60.05** | **42.25** | 4.72 |
| *Road* | 0 | 0 | 0 | **39.95** | **57.75** | 0 |
| *Shadow* | 0 | 0 | 0 | 0 | 0 | 95.28 |
| *Total* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

Table 3. Confusion matrix of WV-2 VNIR + TS-X features classification.

| *Class* | *Water* | *Grass* | *Trees* | *Buildings* | *Road* | *Shadow* |
|---|---|---|---|---|---|---|
| *Water* | 100.00 | 0 | 0 | 0 | 0 | 0 |
| *Grass* | 0 | 98.20 | 0 | 0 | 0 | 0 |
| *Trees* | 0 | 1.80 | 98.90 | 0 | 0 | 0 |
| *Buildings* | 0 | 0 | 0 | **96.36** | **3.91** | 0.47 |
| *Road* | 0 | 0 | 1.10 | **3.64** | **96.06** | 0 |
| *Shadow* | 0 | 0 | 0 | 0 | 0 | 99.53 |
| *Total* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

## 4. IMAGE SIMULATION USING DSM

The method described in section 2 for the interpretation of SAR images with optical images has two main problems: (1) the prerequisite is too restricted, because of the fixed azimuth and incidence angle, also the condition of weather and time, (2) because of the perspective projection of the optical sensor data, the displacements of objects in the two images should theoretically only be the same in one line.

Based on this idea, we propose to provide a simulated optical image based on an orthographic projection instead of perspective projection. If the viewing directions of the virtual camera and the SAR sensor are orthogonal, the displacement of objects will be the same. We use the open-source ray tracing software POV-Ray (POV-Ray 2011) for optical data simulation.

For SAR simulation, RaySAR is used which has been developed at Remote Sensing Technology, Technische Universität München (TUM). It is based on POV-Ray and provides output data for generating images in SAR geometry. The detailed process of SAR simulation is shown in (Tao 2011).

For the simulation we use a DSM of Munich city center with a horizontal and vertical resolution of 1 meter and 0.1 meter, respectively. The DSM is derived from LiDAR data and is shown in Figure 8.

Figure 9 shows the simulation result of the optical simulation (a) and SAR simulation (d). Compared to the TerraSAR-X image (c), the similarities of the building structures are clearly visible in the simulation results. The shape of shadow areas is also similar in the simulated images. Vegetation (e.g. in the upper part of the scene) is not represented in the DSM. Hence, it can not be seen in the simulated images.

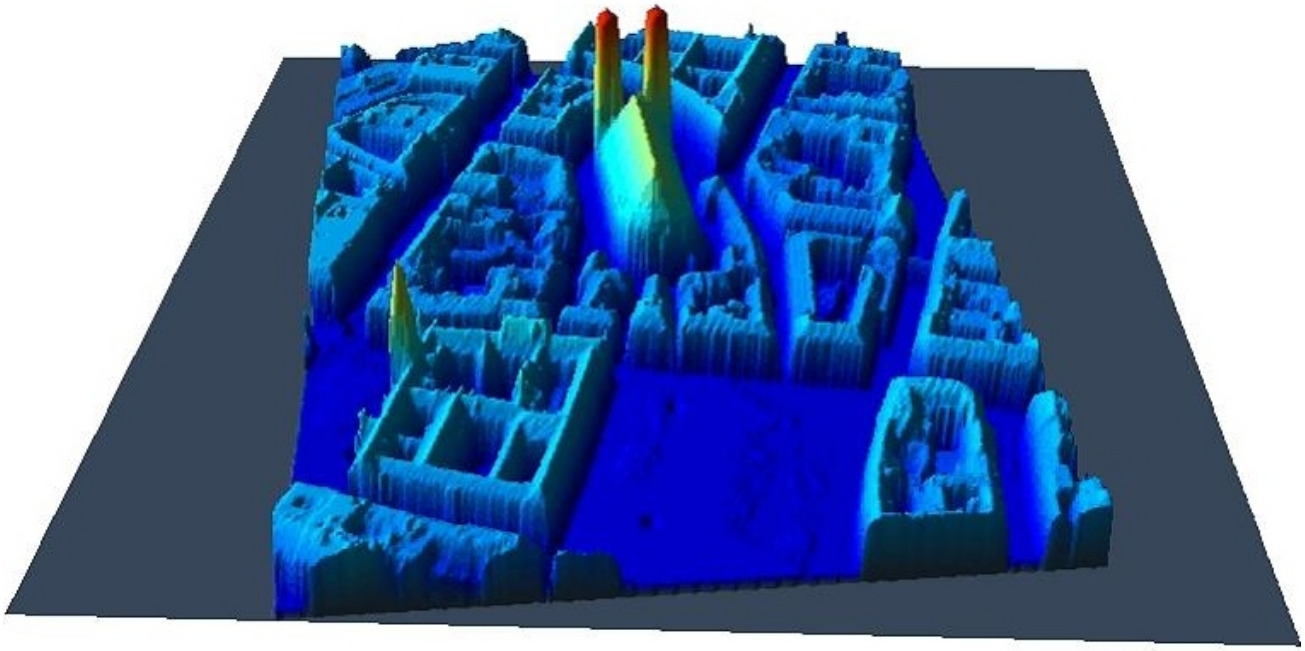
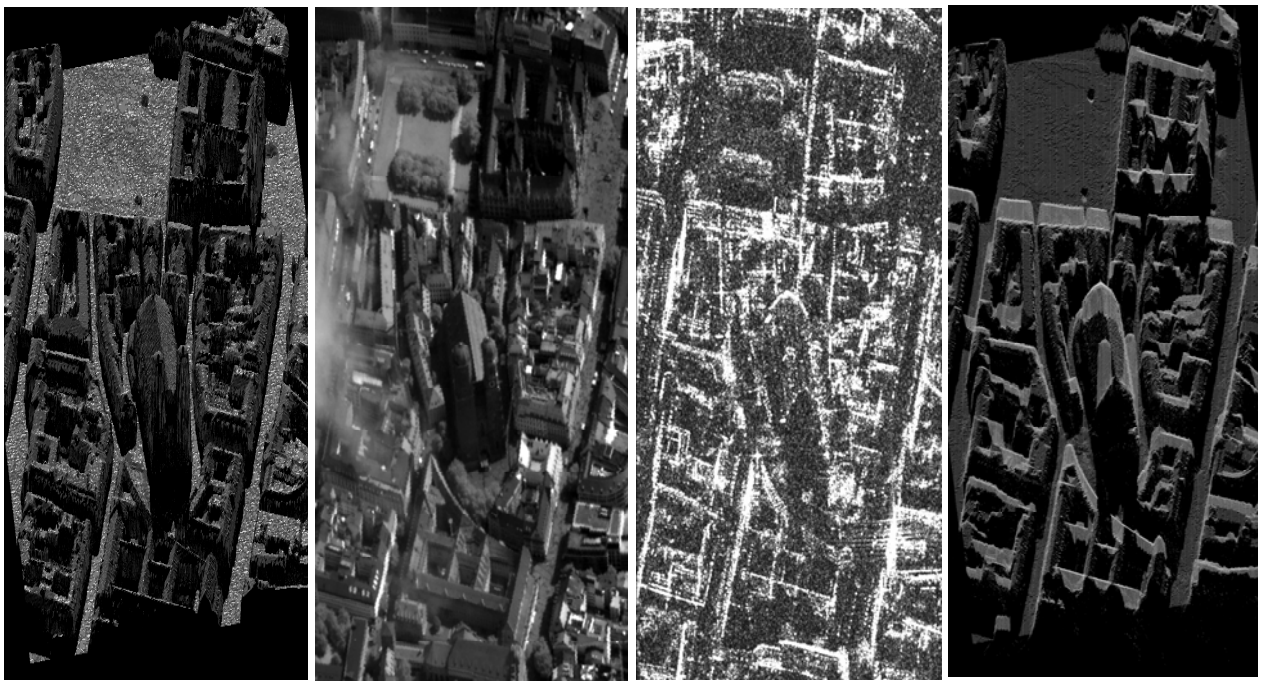


Figure 8. Digital elevation model of Munich center.



(a) (b) (c) (d)

Figure 9. Simulated optical image (a), WV-1 image (b), TS-X image (c), simulated SAR image (d).

## CONCLUSIONS

This paper presents a multi-sensor data fusion method for urban area classification. The fusion model is based on information fission, dimensionality reduction, and information aggregation and employs relevant ways of multisource data combination. Utilized multi-sensor data (multispectral and SAR) allow to increase the number of classes and to

boost the accuracy of the classification. The results of the classification are used for common land cover and urban classes. Multi-sensor data are processed separately and the fusion and classification method follows consensus rules of multisource data classification. The data classification is not influenced by the limitations of dimensionality and the calculation complexity primarily depends on the step of dimensionality reduction. The shown method has also a high potential for the task of change detection, which is a matter for future research.

Additionally in this paper, a method for supporting the visual interpretation of SAR images with simulated optical and SAR images using LiDAR DEM has been presented. Since the location and shape of the objects are similarly represented in the simulated images, acquiring the semantic on a SAR image is easier. The simulated optical image can be used for direct and quick identification of objects in the SAR image. The simulated SAR image has a similar signal reflectivity as the TerraSAR-X image, and it can also separately present single or multiple scattering in the SAR images, which is very useful for building recognition and reconstruction. Both the simulated optical and SAR images are automatically geocoded and enable a direct comparison with the SAR images. Future work will concentrate on learning the semantic relationship between objects in optical and SAR images in order to improve, for instance, methods for change detection.

## ACKNOWLEDGEMENT

We would like to thank DigitalGlobe and European Space Imaging (EUSI) for the collection and provision of IKONOS, WorldView-1 and WorldView-2 scenes over Munich city. TerraSAR-X data were provided by DLR through the Science Projects MTH0505 and MTH0948. The work of Aliaksei Makarau and Junyi Tao was supported by the DLR-DAAD research grants A/09/95629 and A/09/95888, respectively.

## REFERENCES

[1] Auer, S., Hinz, S., Bamler, R., “Ray-Tracing Simulation Techniques for Understanding High-Resolution SAR Images”, IEEE Transactions on Geoscience and Remoe Sensing, 48(3), 1445-1456 (2010).

[2] Deledalle, C.A., Denis, L. and Tupin, F., “Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights,” IEEE Transactions on Image Processing, 18(12), 2661-2672 (2009).

[3] Haralick, R., Shanmugam, K., and Dinstein, I., “Textural features for image classification,” IEEE Transactions on Systems, Man, and Cybernetics, 3 (6), 610-621 (1973).

[4] Makarau, A., Palubinskas, G. and Reinartz, P., “Multiresolution Image Fusion: Phase Congruency for Spatial Consistency Assessment,” In: ISPRS Technical Commision VII Symposium - 100 Years ISPRS, 5-7 July, 2010, Vienna, Austria, Wagner, W., Szekely, B. (eds.), IAPRS, XXXVIII(7B), 383-388 (2010).

[5] Makarau, A., Palubinskas, G. and Reinartz, P., “Multi-sensor data fusion for urban area classification,” Proc. of Joint Urban Remote Sensing Event JURSE, 11-13 April, 2011, Munich, Germany (2011).

[6] Palubinskas, G., “An unsupervised clustering method by entropy minimization,” In: Maximum Entropy and Bayesian Methods, W. von der Linden et al., Eds., Kluwer Academic Publishers, Netherlands, 327-334 (1999).

[7] Palubinskas, G. and Datcu, M., “Information fusion approach for the data classification: an example for ERS-1/2 InSAR data”, International Journal of Remote Sensing 29(16), 4689-4703 (2008).

[8] Palubinskas, G., Reinartz, P. and Bamler, R., “Image acquisition geometry analysis for the fusion of optical and radar remote sensing data,” International Journal of Image and Data Fusion 1(3), 271-282 (2010).

[9] Palubinskas, G. and Reinartz, P., “Multi-resolution, multi-sensor image fusion: general fusion framework,” Proc. of Joint Urban Remote Sensing Event JURSE, 11-13 April, 2011, Munich, Germany (2011).

[10] Persistence of Vision Raytracer Propriety Limited. Available: www.povray.org (last access: 11.03.2011).

[11] Reinartz, P., Müller, R., Schwind, P., Suri, S. and Bamler, R., “Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data,” ISPRS Journal of Photogrammetry and Remote Sensing, 66, 124-132, (2011).

[12] Suri, S. and Reinartz, P., “Mutual-Information-Based Registration of TerraSAR-X and Ikonos Imagery in Urban Areas,” IEEE Transactions on Geoscience and Remote Sensing 48(2), 939-949 (2010).

[13] Tao, J., Palubinskas, G., Auer, S. and Reinartz, P., “Interpretation of SAR images in urban areas using simulated optical and radar images,” Proc. of Joint Urban Remote Sensing Event JURSE, 11-13 April, 2011, Munich, Germany, (2011).