

Comparison of land cover classification using high-resolution TerraSAR-X and optical imagery

Jan Dirk Wegner, Shoaib Nezam, Sönke Müller, Uwe Sörgel, Institute of Photogrammetry and GeoInformation (IPI), Leibniz University Hanover, Germany

Abstract

TerraSAR-X is capable of acquiring imagery of one meter resolution. In data of such kind man-made objects become visible and typical land cover classification classes appear in high detail. Our aim is to find out if TerraSAR-X imagery may complement optical images for automatic land cover classification. Thus, we classify imagery of both data types into classes settlement, agriculture, streets, and forested areas and compare classification performances. We use Markov Random Fields (MRF) as learning based probabilistic framework to classify optical and SAR data. In case of the TerraSAR-X amplitude data we model the likelihood function with Fisher distributions, whereas texture measures are evaluated using Gibbs probability distributions for the optical images. First results show that high-resolution TerraSAR-X imagery may complement land cover classification.

1 Introduction

State-of-the-art space borne high-resolution SAR sensors like TerraSAR-X and Cosmo-SkyMed acquire imagery with a geometric resolution of one meter. Due to the sensor principle of actively emitting pulses in the microwave domain, they are nearly independent of cloud coverage and daylight. Additionally, they image different properties of objects on the ground compared to optical sensors due to the principle of measuring slant-ranges in side-looking viewing geometry and the signal in the microwave domain. Various approaches for land cover classification using SAR data [1], InSAR data [2], or a combination of SAR and optical data [3,4,5] have been proposed. Multiple probabilistic and non-probabilistic classification techniques may be applied; Lu and Weng [6] provide an overview.

Our aim is to investigate whether certain object classes may be better detected or distinguished in high-resolution SAR data than in optical imagery. We use Markov Random Fields (MRF) as classification framework for several reasons. First, we want to achieve probabilities instead of just decisions (as e.g., Support Vector Machines) in order to allow for later combination of multi-sensor data during post-processing in a probabilistic way. Second, we need to learn as many parameters as possible in order to avoid manual parameter settings for each new scene. Third, context-knowledge may support automatic scene analysis particularly if dealing with high-resolution data containing urban scenes. It may be integrated through the MRF prior term. Therefore, we use MRFs for classifying both the optical and the SAR data.

For the classification of the TerraSAR-X images we use an approach proposed by Tison et al. [7], which

models the likelihood term using Fisher distributions. The optical data is classified with a texture-based approach applying Gibbs distributions proposed by Gímel'farb [8]. They have the advantage of being particularly adapted to the statistics of high-resolution SAR data (intensity and amplitude).

2 Classification

For the classification of the high-resolution spotlight TerraSAR-X images we use an approach developed by Tison et al. [7], which models the likelihood distribution for intensity and amplitude SAR values with Fisher distributions. Due to the very high resolution of the data, the return signal is often governed by a single dominant scatterer, particularly in urban areas. Fisher distributions are capable of modelling heavy tails of distributions occurring particularly in urban areas with lots of bright scatterers. Additionally, they are highly flexible and may thus model various data distributions that had to be modelled with different functions beforehand. So far, this approach has only been applied to classify objects in urban areas. Nonetheless, we show that it may also be used more generally for land cover classification.

A texture classification approach described by Gímel'farb [8] is applied to classify the optical images. It models image textures with a Markov random field using Gibbs probability distributions. The approach models a local pixel neighbourhood only with pairwise pixel interactions and shows good results in modelling natural textures. To classify textures of different spatial resolutions, like appearing in aerial images, the approach is extended with a resolution pyramid.

2.1 TerraSAR-X

This section provides an overview of the classification applied to the TerraSAR-X data. MRFs have the advantage of assigning probabilities to the final labelling instead of only providing final decisions, like all graphical models. Those probabilities can be exploited for post-processing or decision making. MRFs are generative models and can be viewed as an extension of Naïve Bayes. They estimate the joint distribution $P(\mathbf{x}, \mathbf{y})$ of data \mathbf{x} and labels \mathbf{y} , which can be decomposed into a product of factors $P(\mathbf{x}/\mathbf{y})P(\mathbf{y})$. In the Bayesian context the first term in **Equation 1** can be viewed as the likelihood term, whereas the second one is a prior over labels \mathbf{y} .

$$P(\mathbf{y} | \mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp \left(\sum_{i \in S} \log P_i(x_i | y_i) + \alpha \sum_{i \in S} \sum_{j \in N_i} \beta y_i y_j \right) \quad (1)$$

i is an image site out of the set of all image sites S of a single image, j is an image site of the 8-connectivity neighborhood N_i of site i . In our case a site i corresponds to a window centered on the pixel of interest. This window is shifted across the entire image. β is the weight of the Potts model [10]. Weighting parameter α balances the influence of the likelihood term and the prior term on the posterior probability $P(\mathbf{y}/\mathbf{x})$. $Z(\mathbf{x})$ is the partition function and can be interpreted as the distribution $P(\mathbf{x})$ of data \mathbf{x} in the Bayesian framework. It acts as a normalization constant (for a given data set) and can be expressed as sum over all possible label configurations of the product $P(\mathbf{x}/\mathbf{y})P(\mathbf{y})$. The data likelihood $P_i(x_i/y_i)$ uses data from a single site i , while the prior term introduces local interactions between adjacent labels y_j and y_i within a local neighborhood.

Fisher distributions well approximate SAR amplitude data. Thus they are used for the likelihood term $P_i(x_i/y_j)$ (**Equation 2**).

$$P_i(x_i | y_i) = \frac{\Gamma(L_{y_i} + M_{y_i})}{\Gamma(L_{y_i})\Gamma(M_{y_i})} \sqrt{\frac{L_{y_i}}{M_{y_i}}} \frac{2}{\mu_{y_i}} \cdot \quad (2)$$

$$\cdot \left(\frac{\sqrt{\frac{L_{y_i}}{M_{y_i}}} x_i}{\mu_{y_i}} \right)^{2L_{y_i}-1} \left/ \left(1 + \left(\frac{\sqrt{\frac{L_{y_i}}{M_{y_i}}} x_i}{\mu_{y_i}} \right)^2 \right)^{L_{y_i}+M_{y_i}} \right.$$

μ_y , L_y , and M_y are the parameters of the Fisher distribution of class y and Γ represents the gamma distribution. The Fisher parameters are estimated making use of the second-kind statistics based on the Mellin-Transformation (see more details in [7]). The

posterior probabilities $P(\mathbf{y}/\mathbf{x})$ are computed with the Metropolis algorithm in combination with simulated annealing.

2.2 Optical imagery

The texture classification approach is derived from the texture segmentation approach described in [8]. A detailed description of the used Markov random field model is given in [9]. Each different texture in the image represents an image class during classification. The classification approach is extended to use a multi-resolution technique in order to represent textures of different resolutions. This is done by first calculating a resolution pyramid of the optical input image and classifying each resolution level separately. Then, an over-all classification is conducted. The supervised approach learns the properties of the texture classes with training samples. The developed learning procedure determines the resolution level on which a texture class gains significant signatures. The resolution with the best separation characteristic may differ from one class to another; the classification of inhabited areas is, for example, significantly better in the lower resolutions and therefore preferably used. The learning step is a crucial part for the effectiveness and correctness of the derived results.

3 Results

We apply the previously described classification methods to two test data sets and show first results. The first set contains a multi-spectral orthophoto of 0.4 m resolution (**Figure 1(a)**) and a TerraSAR-X image (**Figure 1(b)**) of the Fuhrberger Feld, a rural area north of the city of Hannover, Germany. The TerraSAR-X image was acquired in single polarized (HH) spotlight mode with 1 m resolution in range and 2 m resolution in azimuth.

Our second test scene is located close to the airport of Algiers, Algeria. We took one IKONOS image (1 m

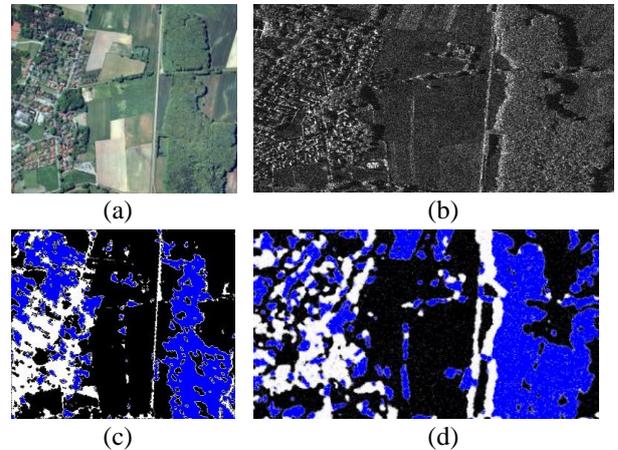


Figure 1 Test scene Fuhrberg (a) Optical aerial image, (b) TerraSAR-X image (slant-range geometry, range from left to right), (c) optical classification result, (d) SAR classification result.

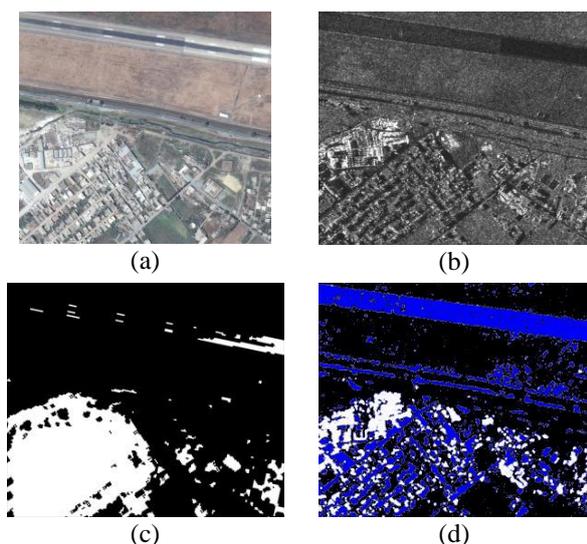


Figure 2 Test scene Algiers (a) IKONOS image, (b) TerraSAR-X image (slant-range geometry, range from left to right), (c) optical classification result, (d) SAR classification result.

panchromatic resolution, **Figure 2(a)** and one TerraSAR-X image (**Figure 2(b)**). The TerraSAR-X image was taken in single polarized (HH) high-resolution spotlight mode leading to resolutions of 0.6 m in range and 1.1 m in azimuth.

We initialize the MRFs of the TerraSAR-X images with a label image obtained from a pixel-wise classification with Support Vector Machines (SVM). Radial basis functions are used for kernel mapping. Three different features are computed as input to the SVM: one image filtered with the Frost operator [11], the coefficient of variation, and the probability image generated with the ratio line detector of Tupin et al. [12]. The latter is a characteristic hint of urban areas. Long bright lines often occur where building walls meet the ground due to double-bounce effects of the radar signal (see [13] for further details).

Classification results are shown in **Figures 1(c,d)** and **2(c,d)**. The overall true positive rates (TPR) and false positive rates (FPR) are summarized in **Table 1**. The overall TPRs of the optical data are slightly higher than those of the SAR data and FPRs are lower. The confusion matrices shown in **Table 2** allow for a more detailed interpretation. The left number always represents TerraSAR-X results while the right refers to classification results of the optical data. The elements on the matrix diagonals indicate how many pixels of a class have been correctly classified. For example, the central element of **Table 2(a)** shows that 73% of all TerraSAR-X pixels of the Fuhrberg scene belonging to class "agriculture" have been correctly classified as such. All non-diagonal elements indicate the percentage of pixels that have erroneously been assigned to another class. For instance, 15% of forest pixels in the TerraSAR-X image of Fuhrberg have erroneously been classified as "agriculture" (see bottom-centre element in **Table 2(a)**). TerraSAR-X imagery provides

much better results for the classification of smooth asphalt and concrete surfaces than the optical classification based on texture measures. Streets and the airport runway in the Algiers scene (**Figure 2(a,b)**) could not be detected at all in Algiers scene using the IKONOS image. Thus, this class was left away and the IKONOS image of the Algiers scene was classified into settlement and agricultural areas.

Table 1 Overall true positive rate (TPR) and false positive rate (FPR) on per-pixel level of the TerraSAR-X and of the optical classification results in percent

Sensor \ Scene	Fuhrberg		Algiers	
	TPR	FPR	TPR	FPR
TerraSAR-X	67	33	78	22
Optical sensor	72	28	93	7

Table 2 Confusion matrices of the classification results (a) test scene Fuhrberg, (b) test scene Algiers in percent; the left number in each cell refers to the TerraSAR-X data whereas the right one refers to the optical data.

Reference Class \ Classification	Settlement	Agriculture	Forest
	Settlement	59 54	18 22
Agriculture	12 8	73 87	15 5
Forest	32 11	15 34	53 55

(a)

Reference Class \ Classification	Settlement	Agriculture	Street
	Settlement	42 58	50 42
Agriculture	4 3	88 97	9 -
Street	6 -	22 -	72 -

(b)

The TerraSAR-X data however allows to detect smooth surfaces. Those areas that are smooth relative to the radar wavelength reflect the radar signal away from the sensor and thus appear dark in the SAR image. Thus, streets as one class of land cover could be detected with the support of high-resolution SAR data. Classification of forests and settlement areas may also be supported with high-resolution SAR data. Concerning the Fuhrberg scene the 59% of the settlement area was classified as such based on the TerraSAR-X image, whereas only 54% was detected based on the optical data. The Algiers scene however shows inverse results with the limitation that only two classes were classified in the optical data. Classification result of the classes forest and settlement could be further improved by post-processing steps or the integration of additional features. A closer look at **Table 2(a)** reveals that a large percentage of the misclassifications of those classes is due to forested areas being classified as settlements (32%). This effect is mainly caused by the layover of the forest (cf. **Figure 1(b,d)**) being erroneously classified as settlement due to its similar amplitude distribution. If we neglect those layover areas 62% of the forests and 64% of the settlements are detected (which outperforms the optical results).

4 Conclusions and Outlook

First results show that MRFs using Fisher distributions as likelihood may be used for land cover classification of high-resolution TerraSAR-X amplitude data. Some object classes that cannot be distinguished based on optical data are detectable in TerraSAR-X data (e.g., rural streets).

One drawback of the current state of the SAR classification approach is the limitation of the likelihood term to only one feature (i.e., the SAR amplitude in our case). Future work will comprise the extension to a multi-dimensional feature space in order to enable the incorporation of additional features directly (instead of taking the detour via SVMs), for example the coefficient of variation and various moments.

The source code of the TerraSAR-X classification using MRFs and the estimation of the Fisher distributions with second kind statistics will soon be available within the open source library Orfeo Toolbox (OTB). It may be downloaded at <http://www.orfeo-toolbox.org/otb/>.

Acknowledgements

We thank Jean-Marie Nicolas and Céline Tison for their support of the implementation of the Fisher distributions.

References

[1] Dobson, M.C.; Ulaby, F.T.; Pierce, L.E.: *Land-Cover Classification and Estimation of Terrain*

Attributes Using Synthetic Aperture Radar. Remote Sensing of Environment, 51(1): 199 - 214, 1995.

- [2] Engdahl, M.E.; Hyyppä, J.M.: *Land-Cover Classification Using Multitemporal ERS-1/2 InSAR Data*
- [3] Schistad Solberg, A.H.; Jain, A.K.; Taxt, T.: *Multisource Classification of Remotely Sensed Data: Fusion of Landsat TM and SAR Images*. IEEE Transactions on Geoscience and Remote Sensing, 32(4): 768 – 778, 1994.
- [4] Le Hégarat-Masclé, S.; Bloch, I.; Vidal-Madjar, D.: *Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing*. IEEE Transactions on Geoscience and Remote Sensing, 35(4): 1018 – 1031, 1997.
- [5] Waske, B.; Benediktsson, J.A.: *Fusion of Support Vector Machines for Classification of Multisensor Data*. IEEE Transactions on Geoscience and Remote Sensing, 45(12): 3858 – 3866, 2003.
- [6] Lu, D.; Weng, Q.: *A survey of image classification methods and techniques for improving classification performance*. International Journal of Remote Sensing, 28(5), pp. 823–870, 2007.
- [7] Tison, C.; Nicolas, J.-M.; Tupin, F.; Maître, H.: *A new statistical model of urban areas in high resolution SAR images for Markovian Segmentation*. IEEE Transactions on Geoscience and Remote Sensing, 42(10): 2046 – 2057, 2004.
- [8] Gimel'farb, G.L.: *Texture Modelling by Multiple Pairwise Pixel Interactions*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18(11), 1110–1114, 1996.
- [9] Gimel'farb, G. L.: *Gibbs Fields with Multiple Pairwise Pixel Interactions for Texture Simulation and Segmentation*. Rapport de recherche RR-3202, INRIA, Sophia Antipolis France, 1997.
- [10] Wu, F.: The Potts model. Rev. Modern Phys., vol. 54, pp. 235 – 268, 1982.
- [11] Frost, V.S.; Abbott Stiles, J.; Shanmugan, K.S.; Holtzman, J.C.: *A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 4(2): 157 - 166, 1982.
- [12] Tupin, F.; Maître, H.; Mangin, J.-F.; Nicolas, J.-M.; Pechersky, E.: *Detection of Linear Features in SAR Images: Application to Road Network Extraction*. IEEE Transactions on Geoscience and Remote Sensing, Vol. 36(2): 434 - 453, 1998.
- [13] Thiele, A.; Wegner, J.D.; Soergel, U.: *Building reconstruction from multi-aspect InSAR data*. In U. Soergel (Ed), Radar Remote Sensing of Urban Areas, Springer, 1st Edition, ISBN-13: 978-9048137500, 2010.