FUSION OF MULTITEMPORAL ENVISAT ASAR AND HJ-1 DATA FOR OBJECT-BASED URBAN LAND COVER CLASSIFICATION

Alexander Jacob[1], Yifang Ban[2]

1. Royal Institute of Technology, Division of Geoinformatics and Geodesy, Stockholm, Sweden; aljacob@kth.se
2. Royal Institute of Technology, Division of Geoinformatics and Geodesy, Stockholm, Sweden; yifang@kth.se

ABSTRACT

The key goal of this work is to analyze the synergistic effects of multitemporal data fusion for urban land cover mapping. In particular this analysis is carried out using multitemporal ENVISAT ASAR images and one Chinese HJ-1 optical image acquired over Beijing in 2009. The major land cover classes are high-density built-up areas, low-density built-up areas, roads, airports, forests, parks, golf courses, grass/pasture, crops, bare fields and water. The methodology used in this research including orthorectification, SAR speckle filtering, and object-based classification. The segmentation is based on the newly developed algorithm KTH-SEG that utilizes an edge-aware region growing and merging approach. Fusion of the various combinations of multitemporal multi-angle SAR data and HJ-1 data were compared with SAR and optical data alone. The preliminary results show that the fusion of ENVISAT ASAR and HJ-1 data performed much better than optical data alone or SAR data alone. The fusion of 4-date SAR data and optical data can achieve similar classification accuracy as the fusion of 8-date SAR data and optical data if multi-angle, dual look direction SAR data with suitable temporal compositions are selected. Compared to eCognition, the KTH-SEG performed better in extracting linear features such as roads and rivers.

INTRODUCTION

The increase in urbanization that can be observed in China and other countries around globe causes drastic changes in the surroundings of the regions where mega cities are growing (1 & 2). This raises the need for good monitoring, which can be achieved by means of remote sensing and land cover classification.

Past studies have demonstrated that SAR and optical data provide complementary information thus improve the classification accuracy (3 – 5, 6). SAR can also overcome the problems such as cloud coverage and haziness when optical images are unavailable due to weather conditions (1, 3 – 5, 7). The benefits of object-based image analysis are also shown in various studies (5, 9 & 10), hence we decided to apply image segmentation prior to classification. Classic approaches to segmentation include thresholding techniques, edge detection and region-based approaches (11) and more recent approaches often combine several of those, e.g. edge-detection and region growing (12 & 13). Since those are based on simple examples of only a few objects further development is needed to adapt to very complex urban scenes from SAR and optical imagery. Thus, the overall objective of this research is to investigate multitemporal SAR and optical data for urban land cover mapping using an object-based approach.

STUDY AREA & DATA DESCRIPTION

In this study the capital of China, Beijing is selected as an example. The area has already undergone significant changes during the past decades (14). The 6th ring road around the city is the approximate boundary of the area analyzed of roughly 50km x 55km. The SAR data is from ENVISAT's ASAR sensor (C-Band) and the optical image is taken from HJ-1B (15). Both sensors

deliver images in 30m resolution. Most of the SAR images were acquired in dual polarization. Table 1 provides an overview of the SAR and optical data.

Table 1: Overview of SAR and optical images.

Satellite	*Sensor*	*Band*	*Beam Mode*	*Orbit*	*Observation Date*
Envisat	ASAR	HH/VV	IS2	A	2009-05-17
Envisat	ASAR	HH/VV	IS2	D	2009-05-27
Envisat	ASAR	HH/VV	IS4	A	2009-06-08
Envisat	ASAR	HH/VV	IS6	A	2009-06-11
Envisat	ASAR	VV	IS7	A	2009-08-04
Envisat	ASAR	VV	IS7	A	2009-09-08
HJ-1B	CCD2	B, G, R, NIR			2009-05-12

METHODS

The overview of the methodology used in this research in shown in *Figure 1*. The methodology is divided into three major phases, the pre-processing, the image segmentation, the classification of those segments.

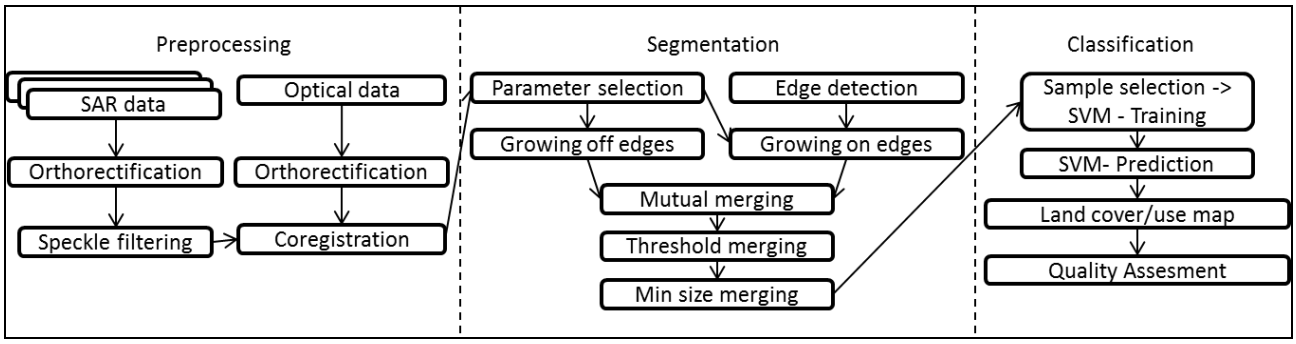

Figure 1: Overview of the methodology.

Image pre-processing

All SAR scenes are first of all orthorectified based on the satellite orbital model and a DEM from SRTM in 3" and then automatically co-registered using NEST 4B. Additionally a multi-temporal speckle filter is applied on the co-registered SAR image stack and the result is compressed to 8 bit by a 95% histogram clipped linear scaling function. The optical image is manually co-registered to the SAR image stack after orthorectification using PCI Geomatica 2012 SP2 Ortho-Engine.

Image segmentation

The segmentation is mainly carried out using our own edge-aware region growing and merging algorithm KTH-SEG which is implemented in JAVA. For comparison, the best data combinations were also processed using eCognition 8.7.2. KTH-SEG uses both region growing and merging and a simple edge-detection scheme. The edge detection is first performed on each input band individually by first enhancing the contrast in the band using histogram equalization and then applying a 5x5 Sobel-filter. The gradient magnitude images are then thresholded with a 90% margin to derive binary edge images. Those binary image layers are fused using a majority voting where only edges that appear in the majority of bands from one sensor are kept. The results of all sensors are then merged with a logical OR-operator.

With this information the region growing can be initiated. The growing is performed separately on edges and off edges. The homogeneity that is used for finding the best pixels and segments for

growing and merging is based on the weighted sum of two measures. Those are for the spectral part the change of mean due to a possible merge and for a more texture orientated feature the change of standard deviation. Both measures are averaged over all input bands. The growing is limited by the minimum segment size which is given as a parameter. The growing phase is followed by a two-step merging phase. The first step is merging mutually best fitting segments based on the homogeneity criterion and limited by the maximum segment size and the second step is merging segments on basis of threshold of the homogeneity value of two segments. The threshold is automatically derived from the average homogeneity of all neighboring segments. Finally those segments still below the minimum segment size are merged to their best fitting neighbor.

Image classification

The classification is performed using a support vector machine, which showed to be suitable for the task in e.g. (16 & 17), trained in a semi-supervised manner. It is called semi-supervised since it is only necessary to select one set of training areas. This initial set is then applied to all sets of segments from all the different combinations of SAR and optical data. This is achieved by overlapping the labeled training areas with the segments and using always the segment with the biggest overlap as the training sample for the individual classification.

The SVM used is based on a Radial Base Function Kernel and hence needs two parameters to be set, the general penalty value C and the RBF specific parameter gamma. These parameters are found automatically using a grid search approach. The performance of each pair is here evaluated by a five folded cross-validation within the training data. The best pair is used to train the machine and predict the whole dataset of segments using this machine. The one against one approach is used to adapt the SVM to the multi-class case. The java implementation of LibSVM 3.12 is utilized to integrate classification capability into KTH-SEG. The segments created from eCognition are also imported into KTH-SEG for classification purposes. This guaranties a fair comparison. The features used in the SVM are the mean and standard deviation within each segment of each band.

Accuracy assessment is performed against another set of testing areas. Confusion matrices are derived and key indicators of quality such as the kappa value and overall accuracy are calculated.

RESULTS

First, a large no. of different combinations of multitemporal, multi-angle SAR scenes were compared. The selection of SAR combinations were based on a good coverage of the vegetated season, and as many as possible different beam modes. Additionally the first two scenes were available in ascending and descending orbit and very close in time, which also gave possibility to make a few comparisons with different orbit configurations. The best two results were then selected for fusion with the optical image and re-processed. For validation purposes, the best result was also processed using eCognition for the segmentation as an additional benchmark for assessing the quality of the segmentation from KTH-SEG. The complete list of results can be found in *Table 2*. All KTH-SEG results were processed with 40 pixels as minimum segment size and 4000 pixels as maximum segment size. The weighting between spectral and texture criterion where equally set. For the eCognition processing, standard settings, as seen in many studies were applied. The shape criterion was set to 0.9 and the compactness 0.5. The scale was selected in a way that the no. of segments is comparable to the no of segments from KTH-SEG processing ~146k. Scale 25 has slightly less segments ~104k and 20 slightly more~156k.

Table 2: Combinations of prepared image stacks and classification accuracy achieved; Red – best result with all images used; Green – second best result with reduced no of images; Yellow – comparison with eCognition results using scale parameter 20 and 25.

May 17	*May 27*	*June 08*	*June 11*	*Aug. 04*	*Sept. 08*	*May 12*				
IS 2 A	*IS 2 D*	*IS 4*	*IS 6*	*IS 7*	*IS 7*	*HJ 1*	*Avg. Ac.*	*Over. Ac.*	*Kappa*	*Name*
x	x						42.59	47.29	0.41	
	x		x				44.33	48.32	0.42	
	x			x	x		46.16	50.55	0.44	
		x		x	x		50.07	53.40	0.48	
		x	x				50.61	53.90	0.48	
x	x		x	x			50.85	54.47	0.49	
			x	x	x		50.91	53.08	0.47	
	x	x					52.08	54.56	0.49	
x		x	x	x	x		52.11	54.43	0.49	Asc. Only
	x	x	x	x	x		53.38	56.85	0.52	Asc. Dsc.
	x	x	x	x	x		53.69	56.86	0.52	
	x		x	x	x		54.36	56.52	0.51	
	x	x		x	x		56.55	58.96	0.54	SAR Red.
x	x	x	x	x	x		63.83	65.14	0.61	All SAR
	x	x		x	x	x	72.08	74.12	0.71	Fusion Red.
x	x	x	x	x	x	x	75.30	75.45	0.72	Fusion All
x	x	x	x	x	x	x	66.75	67.92	0.64	Scale 25
x	x	x	x	x	x	x	69.79	70.51	0.67	Scale 20
						x	70.66	70.58	0.67	Opt only

The first thing we found from those results is that the more images used, the better the final outcome of the classification. This is valid both for the SAR data alone as well as when fused with the optical data. But when selecting carefully it is possible to reach almost as good results with less data available at least when fusing the SAR data with the optical data. However when just using SAR data the drop in accuracy is much bigger (compare the first green and red rows with the second green and red rows in *Table 2*). In this study SAR data could clearly complement the optical data, even though SAR alone yielded slightly less classification accuracy. The main reason for the lower SAR accuracy was that the training data for pasture/crops and bare fields were selected based on the May 12 HJ-1 data while in SAR images bare fields were converted to pasture/crops and vice versa throughout the growing season. Further, it should be noted that both SAR and optical data have very different characteristics and hence the inaccuracy of the results is very different as can be seen in **Error! Reference source not found.**. There we can see in the right encircled area an example of the new terminal on Beijing International Airport. This could easily be identified as a building using SAR data (B) but was found to be bare soil when using optical data (E). In the direct neighborhood of the building and in the encircled area in the lower left corner a problem of the SAR (B) classification becomes evident, the confusion of low backscatter classes like water, airport, golf courses and roads. These where much better separated, when in direct neighborhood, from the optical (E) result. When looking at these areas in the combined result (C) it is evident that both issues are mostly resolved and hence yielding a good synergy of the data fusion. In (F) the best classification result based on the segments from eCognition is shown. The accuracy of the eCognition result is around 5% below the best result from KTH-SEG. A possible reason for this could be that due to the edge detection the segments of KTH-SEG find

better boundaries especially when small almost linear features like roads and rivers are extracted (see river to golf course in the left encircled area in C and F).

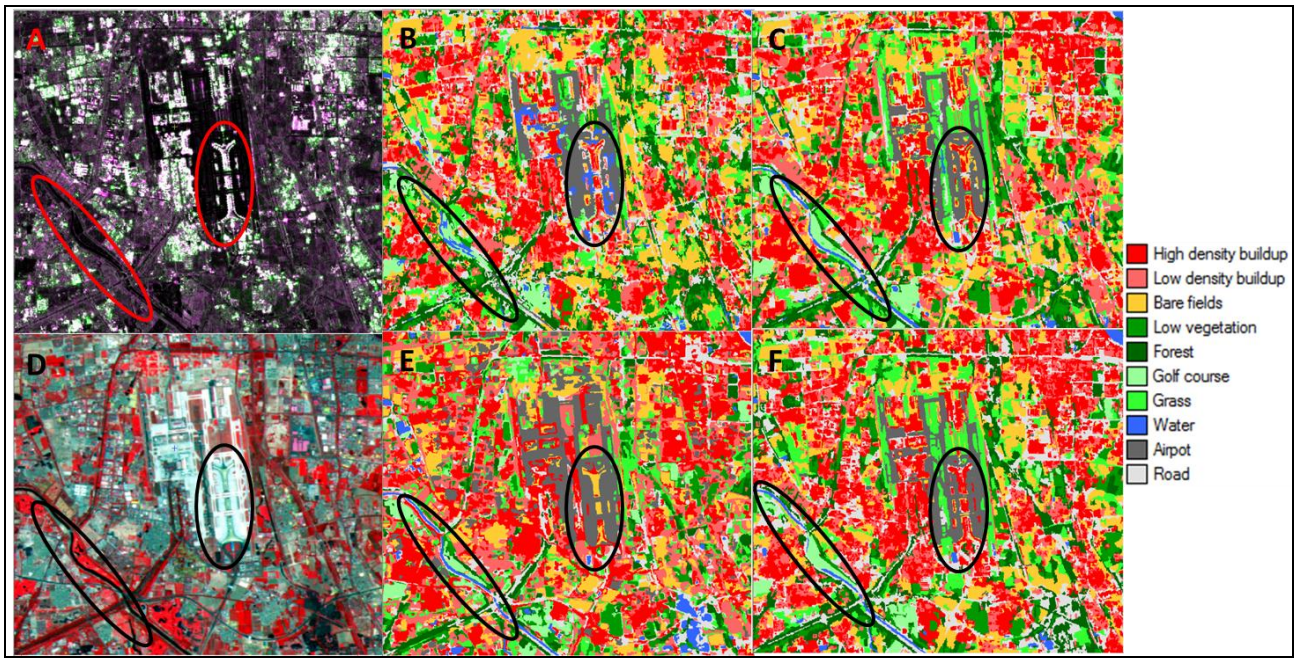

*Figure 2: Beijing international airport – classification comparison; A – SAR sample; B – All SAR; C – Fusion Red; D – optical sample; E – Opt only; F – eCognition Scale 20 (for naming see Name column in **Table 2**)*

In general the road segments in (F) often have a bigger width than in (C). Another more general impression was that the segments in eCognition could find large homogeneous areas like crops better but often failed to find good boundaries between the urban classes of high and low density as well as roads. From the results in *Table 2*marked in yellow, it can also be seen, that a higher no. of segments seems to yield better classification results when working with images in this resolution.

CONCLUSIONS

The research shows that the combinations of multitemporal multi-angle SAR data are suitable for urban land cover classification, especially when many scenes over the vegetation season can be acquired. Even when reducing the number of scenes still good results can be achieved as long as the remaining scenes still cover the season well and show a good diversity in terms of incidence angles and look directions. It is also evident, that the fusion of SAR and optical data is superior to individual classification of those data alone. Finally the results show that the segmentation algorithm KTH-SEG works very well on the given dataset and outperforms the leading commercial software for image segmentation eCognition in classifying linear features.

ACKNOWLEDGEMENTS

This research was supported by the Swedish National Space Board. The research is also part of the project ‘Satellite Monitoring of Urbanization for Sustainable Urban Development’ within the European Space Agency (ESA) and Chinese Ministry of Science and Technology (MOST)’s Dragon II program. The authors would like to thank ESA for the ENVISAT ASAR data and MOST for the HJ-1 data.

REFERENCES

1. Ban Y., Gamba P., Gong P., Du P, 2011. Satellite Monitoring of Urbanization in China for Sustainable Development, The Dragon 'Urbanization' Project. Earthzine.
2. Seto K. C., Fragkia M., Güneralp B., Reilly M. K., 2011. A Meta-Analysis of Global Urban Land Expansion. PLoS ONE, 6(8): e23777.
3. Ban, Y. 2003. Multitemporal ERS-1 SAR and Landsat TM data for agricultural crop classification: comparison and synergy", Canadian Journal of Remote Sensing. 29(4): 518-526.
4. Ban Y., Yousif O. A., 2012. Multitemporal Spaceborne SAR Data for Urban Change Detection in China. IEEE Journal on of Selected Topics in Applied Earth Observations and Remote Sensing (In Press).
5. Ban Y., Hu H., Rangel. I. M., 2010. Fusion of Quickbird MS and RADARSAT SAR for Urban Land-Cover Mapping: Object-based and Knowledge-based Approach. International Journal of Remote Sensing, 31(6): 1391-1410.
6. Waske B, Linden v. d. S., 2008. Classifying Multilevel Imagery From SAR and Optical Sensors by Decision Fusion. IEEE Transactions on Geoscience and Remote Sensing. 46(5): 1457-1466.
7. Henderson M. F., Lewis A. J. 1998. Principles & Applications in Radar Remote Sensing, Manual of Remote Sensing Volume 2. Chapter 13, p677.
8. Nezry E., Mougin E., Lopes A., Gastellu-Etchegorry J. P., Laumonier Y., 1993. Tropical Vegetation mapping with Combined Visible and SAR Spaceborne data, International Journal of Remote Sensing. 14(11).
9. Blaschke T., 2010. Object Based Image Analysis for Remote Sensing, ISPRS Journal of Photogrammetry and Remote Sensing. 65: 2-16.
10. Hu, H. and Y. Ban. 2012. Multitemporal RADARSAT-2 Ultra-Fine-Beam SAR Data for Urban Land Cover Classification. Canadian Journal of Remote Sensing. 29(4): 518-526, 10.5589/m03-014.
11. Pal N. R., Pal S. K., 1993, A Review on Image Segmentation, Pattern Recognition, 26(9):1277-1294.
12. Yu Q., Clausi D. A., 2008. IRGS: Image Segmentation Using Edge Penalties and Region Growing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(12): 2126-2139.
13. Qin A. K., Clausi D. A., 2010. Multivariate Image Segmentation Using Semantic Region Growing With Adaptive Edge Penalty. IEEE Transactions on Image Processing.19(8).
14. Tan M., Li X., Xie H., Lu C., 2005. Urban Land Expansion and Arable Land Loss in China – Case Study of Beijing-Tianjin-Hebei Region. Land Use Policy, 22(3): 187-196.
15. CEOS, CEOS EO HANDBOOK – MISSION SUMMARY - HJ-1B. 2011. http://database.eohandbook.com/database/missionsummary.aspx?missionID=464.
16. Niu X. ,Ban Y., Multitemporal RADARSAT-2 Polarimetric SAR Data for Urban Land Cover Classification using Object-based Support Vector Machine and Rule-based Approach, International Journal of Remote Sensing, (In Press).
17. Waske B., Benediktsson J. A., 2008. Fusion of Support Vector Machines for Classification of Multisensor Data. IEEE Transactions on Geoscience and Remote Sensing, 45(12): 3858-3866.