Road Extraction from SAR Multi-Aspect Data Supported by a Statistical Context-Based Fusion

K. Hedman[1], S. Hinz[2], U. Stilla[1]
[1]Photogrammetry and Remote Sensing, [2]Remote Sensing Technology
Technische Universitaet Muenchen, Arcisstrasse 21, 80333 Munich, Germany

***Abstract*— In this paper we describe a fusion approach for automatic object extraction from multi-aspect SAR images. The fusion is carried out by means of the Bayesian probability theory. The first step consists of a line extraction in each image, followed by attribute extraction. Based on these attributes the uncertainty of each line segment is estimated, followed by an iterative fusion of these uncertainties supported by context information and sensor geometry. On the basis of a resulting uncertainty vector each line obtains an estimation of the probability that the line really belongs to a road.**

I. INTRODUCTION

Road extraction is an important topic for the acquisition and updating of urban GIS. Synthetic aperture radar (SAR) holds some advantages against optical image acquisition. SAR is an active system, which can operate during day and night. It is also nearly weather-independent and, moreover, during bad weather conditions, SAR is the only operational system available today. This is central for several applications among others risk management. Road extraction from SAR images therefore offers a suitable complement or alternative to road extraction from optical images.

By the development of new, sophisticated SAR-system, automatic road extraction has reached a new dimension. Satellite SAR images up to 1 m resolution will be available in 2007 by the German satellite TerraSAR-X [1]. Airborne images already provide resolution up to 1 decimetre [2].

When working with road extraction from SAR images, we should keep in mind the inevitable consequences of the side-looking geometry of the SAR sensor; occlusions caused by shadow- and layover. These effects make road extraction complicated, especially in urban areas. In case of adjacent high buildings and narrow streets, the roads might not even be visible on the radar image. Furthermore in urban areas, the complexity arises through dominant scattering caused by building structures, traffic signs and metallic objects in cities. In order to compensate for eventual gaps additional information can be considered. One suggestion can be the introduction of context information. Bright features and their contextual relationships can be incorporated into the road extraction procedure. Detected vehicles and rows of building layover as well as metallic scattering caused by road signs are indicators of roads,[3][4].

Preliminary work has shown that the usage of SAR images illuminated from different directions (i.e. multi-aspect images) improves the road extraction results. This has been tested both for real and simulated SAR scenes [5],[6]. Multi-aspect SAR images contain different information, which is both redundant and complementary. A correct fusion step has the ability to combine information from different sensors, which in the end is more accurate and better than the information acquired from one sensor alone.

In this paper, we will present a fusion strategy carried out in a statistical framework, supported by global context and SAR sensor geometry.

II. ROAD EXTRACTION SYSTEM

The extraction of roads from SAR images is based on an already existing road extraction approach [7], which was originally designed for optical images with a ground pixel size of about 2m [8]. The first step consists of line extraction using Steger's differential geometry approach [9], which is followed by a smoothening and splitting step. By applying explicit knowledge about roads, the line segments are evaluated according to their attributes such as width, length, curvature, etc. The evaluation is performed within the fuzzy theory. A weighted graph of the evaluated road segments is constructed. For the extraction of the roads from the graph, supplementary road segments are introduced and seed points are defined. Best-valued road segments serve as seed points, which are connected by an optimal path search through the graph. The approach is illustrated in Fig. 1.

The novelty presented in this paper refers on one hand to the adoption of the fusion module to multi-aspect SAR images and on the other hand to a probabilistic formulation of the fusion problem instead of using fuzzy-functions (marked in gray in Fig. 1).

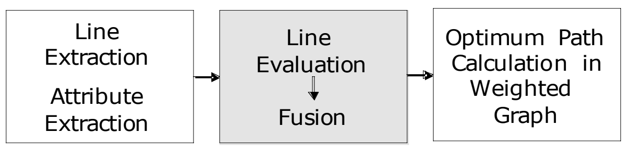

Figure 1. Automatic Road Extraction Process

III. FUSION APPROACH

A line extraction from SAR images often delivers partly fragmented and erroneous results. Especially in forestry and in urban areas over-segmentation occurs frequently. Attributes describing geometrical and radiometric properties of the line segments can be helpful in the selection and especially for sorting out the most probable false alarms. However, these attributes may be ambiguous and are not considered to be reliable enough when used alone. Furthermore, occlusion due to surrounding objects may cause gaps, which are hard to compensate. If line extraction fails to detect a road in one SAR view, it might succeed in another view illuminated from a more favorable direction. Multi-aspect images supply the interpreter with both complementary and redundant information. But due to the over-segmented line extraction, the information is often contradicting as well. Therefore multi-aspect SAR data requires a careful fusion module. The first task is to decide where in the road extraction process, the fusion shall be implemented and to decide what kind of input information shall be fused.

In general, better accuracy is obtained by fusing information closer to the source. But in contrary to multi-spectral optical images, a fusion of multi-aspect SAR data on pixel-level makes hardly any sense. SAR data is far too complex. Instead, of fusing pixel-information, features (line segments) shall be fused. Decision-level fusion means that an estimate (decision) is made based on the information from each sensor alone and these estimates are subsequently combined in a fusion process. If we put this into practice, the first step consists of a line extraction in each image, followed by attribute extraction. Based on these attributes the uncertainty of each line segment is estimated, followed by a fusion of these uncertainties supported by context information and sensor geometry. Based on an uncertainty vector a decision is made. The fusion module is illustrated in Fig. 2.

Techniques for decision-level fusion worth to mention are fuzzy-theory, Dempster-Shafer's method and Bayesian theory. In the following chapter we will discuss the application of a fusion process, which accommodates for these aspects.

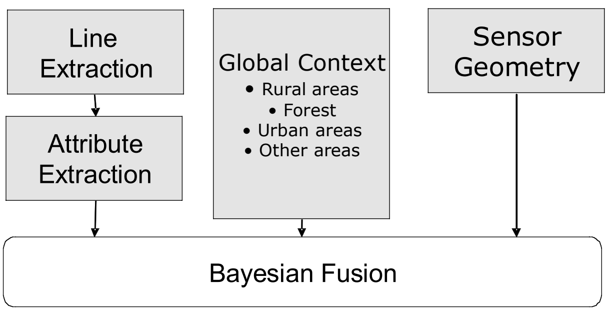

Figure 2. Fusion module and its input data

A. Theory

The underlying theory of the approach originates from Bayesian probability theory and can be drawn from the well-known Bayes' theorem;

$$p\left(Y|X,I\right)=\frac{p\left(X|Y,I\right)\cdot p\left(Y|I\right)}{p\left(X|I\right)} \tag{1}$$

Bayes' theorem follows directly from the product rule:

$$p\left(Y,X|I\right)=p\left(Y|X,I\right)\cdot p\left(X|I\right) \tag{2}$$

The strength of Bayes' theorem is that it relates the probability that the hypothesis Y is true given the data X to the probability that we have observed the measured data X if the hypothesis Y is true. The latter term is much easier to estimate. All probabilities are conditional on I, which is made to denote the relevant background information at hand.

The main feature involved in the road extraction process is the line segment, which can either be identified as a ROAD or as a FALSE_ALARM. Hence, we deal with the following hypotheses:

- Y_1 = an extracted line belongs to a ROAD
- Y_2 = an extracted line belongs to a FALSE ALARM

If relevant, the hypotheses above can be extended with more classes Y_3 ,..,Y_n (e.g. river, shadows).

In our case the measured data X corresponds to geometric and radiometric attributes of the line segment – an attribute vector. Since we are mostly interested in the solution, which yields the greatest value for the probability of the observed data, usually referred to as the maximum likelihood estimate, we can write Bayes' theorem in a compact form:

$$p\left(Y|X,I\right)\propto p\left(X|Y,I\right)\cdot p\left(Y|I\right) \tag{3}$$

If two or more images are available, we shall combine data from two or more images. Then the hypotheses above will be extended to the assumptions whether a ROAD truly exist in the scene or not. We need to add a third term to our measured data X; the fact that a line has been extracted (L) or not extracted ($\overline{L}$) from one or more images. The probability that an object Y_j exist given the measurements $X_1,..X_n$, $L_1,...L_n$, can by means of Bayes' theorem be expressed as

$$p\left(Y_j|X_1,..,X_n,L_1,..,L_n,I\right)\propto p\left(X_1,...,X_n,L_1,..,L_n,|Y_j,I\right)\cdot \cdot p\left(Y_j|I\right) \tag{4}$$

The images can be regarded as independent observations. By means of this statement and by means of the product rule, the expression above can be written as;

1-4244-0712-5/07/$20.00 ©2007 IEEE.

$$p\left(Y_j \left| X_1, X_2, ..., X_n, L_1, L_2, ...L_n, I\right.\right) \propto p\left(X_1 \left| L_1, H, I\right.\right),. \\ .,p\left(X_n \left| L_n, H, I\right.\right) p\left(L_1 \left| Y_j, I\right.\right), ..., p\left(L_n \left| Y_j, I\right.\right) p\left(Y_j \left| I\right.\right) \quad (5)$$

where

$p(L_n|Y_1,I)$ = the posterior probability that a line is extracted from image *n* if a ROAD truly exist.

$p(X|L_n,Y_1,I)$ = the posterior probability that the data X is measured if a ROAD exist AND a line has been extracted.

$p(Y_1,I)$ = the prior or subjective probability that a road exist in the image.

The last one represents a subjective probability and can be defined by the user. Global context can here be especially useful.

The selection of attributes of the line segments is based on the knowledge about roads. Radiometric attributes such as mean and constant intensity, and contrast of a line as well as geometrical attributes like length and straightness are all good examples. It should be pointed out that more attributes does not necessarily mean better results, instead rather the opposite occur. A selection including a few, but significant attributes is recommended.

If there is no correlation between the attributes, the likelihood $p(X|Y_i)$ can be assumed equal to the product of the separate likelihoods for each attribute.

$$p\left(X \left| Y_j\right.\right) = p\left(x_1, x_2, ..x_n \left| Y_j\right.\right) = p\left(x_1 \left| Y_j\right.\right) p\left(x_2 \left| Y_j\right.\right) \ldots p\left(x_n \left| Y_j\right.\right). \quad (6)$$

Each separate likelihood $p(x_i|Y_j)$ can be approximated by a probability density function learned from training data as discussed in [10]. Learning from training data means that the extracted line segments are sorted manually into two groups, ROADS and FALSE_ALARMS. A fitting carried out in a histogram with one dimension is relatively uncomplicated, but as soon as the dimensions increase, the task of fitting becomes more complicated.

Please note that the estimated probability density functions should represent a degree of belief rather than a frequency of the behavior of the training data. The obtained probability assessment shall correspond to our knowledge about roads.

B. Context-based approach

Since even a very sophisticated feature extractor delivers generally results with ambiguous semantics, additional information of global and local context is helpful to support or reject certain hypotheses during fusion.

Global context plays an important role for the definition of the priori term. The frequency of roads is proportionately high in some context areas, for instance in urban regions. The a-priori probability must be different in these areas. In this work the user specifies the priors (see Tab. 1). Therefore the priors represent the belief of the user to a certain degree. In future work, these values will be compared with values learned from training data.

Global context has as well influence on the posterior probability ($p(L_n|Y_1,I)$ that a line is extracted if not conditioning on that a ROAD truly exist or not. It is much more likely to successfully detect a road surrounded by fields than a road in the middle of the forest. Furthermore, shadows (i.e. FALSE ALARMS) occur often in forest areas, which are likely to be extracted as lines from the SAR data.

But the posterior probability ($p(L_n|Y_1,I)$ is as well dependent on the relation between the sensor geometry and the extracted road. A road, whose direction approaches range, is more likely to be extracted at its true position.

Incorporating the context information (global context and sensor geometry), C, in (5) gives the following expression;

$$p\left(Y_j \left| X_1, X_2, ..., X_n, C_1, C_2, ...C_n, L_1, L_2, ...L_n, I\right.\right) \propto \\ p\left(X_1 \left| L_1, C_1, Y_j, I\right.\right), .., p\left(X_n \left| L_n, C_n, Y_j, I\right.\right) p\left(L_1 \left| C_1, Y_j I\right.\right),. \\ .., p\left(L_n \left| Y_j, C_j, I\right.\right) p(C_1 | Y_j), ..., p(C_n | Y_j) p\left(Y_j \left| I\right.\right) \quad (7)$$

where

$p(L_n|C_n,Y_1,I)$ = the posterior probability that a line is extracted from image *n* if a ROAD truly exist AND is surrounded by the global context C_n In this work, this term is treated as a subjective probability (Tabs. 2 and 3). Please notice that this term is varied due to both context area and the relationship between the direction of the road and the SAR sensor geometry.

$p(C_n|Y_1,I)$ = the posterior probability that the context C_n occur if a ROAD exist. This term might be hard to define, but can be of significance in urban areas, if the main directions of the road are known in advance. Especially in modern cities, the road network tends to consist of parallel roads and perpendicular intersections. In this work, this probability is set equal to all occasions.

$p(X|C_nL_n,Y_1,I)$ = the posterior probability that the data X is measured if a ROAD exist AND surrounded by the context C_n AND a line has been extracted. Most probably the attributes of the line depends on the surrounding context. Roads tend to be relatively shorter in urban areas than in other global context areas. Due to limited amount of training data, probability density functions conditioned on the context information are pretty hard to define.

Since the probability density functions of the attributes are defined by training data in all context areas, we apply the theorem of marginalization

$$p\left(X \left| L, Y, I\right.\right) = \int_{-\infty}^{+\infty} p\left(X \left| C, L, Y, I\right.\right) dC. \quad (8)$$

C. *Iterative fusion of two images*

The roads extracted in each single image are fused together by the following iterative fusion strategy. All segments are sorted according to its discriminant value

$$g(x) = \ln\left(p\left(x \middle| L, Y_1, I\right)\right) - \ln\left(p\left(x \middle| L, Y_2, I\right)\right). \quad (9)$$

The line segment with the highest discriminant value is chosen first. Then, all neighbouring segments are searched for. Those parts of the neighbouring segments, which satisfy overlap and collinearity criteria (i.e. buffer width and direction difference) are assumed to be redundant extractions and are removed. If only a part of the neighbouring segment is fused, the segment is clipped and the non-fused part remains in the search. Also, lines with an all too deviant direction according to the best-evaluated line remain. The best-evaluated segment obtains a probability based on (5) or (7) depending on integrating context information or not.

Then, the segment yielding the second highest maximum likelihood of being ROAD is chosen and processed with the same algorithm. The whole fusion process ends after all segments have been processed.

TABLE I. PRIOR PROBABILITIES

Global context	**P(Y,I)**	
	Y=ROAD	***Y=FALSE ALARM***
FIELD	0.2	0.8
URBAN	0.3	0.7
FOREST	0.05	0.95
OTHER	0.2	0.8

TABLE II. POSTERIOR PROBABILITIES CONDITIONAL ON CONTEXT I

Global context C	**P(L|C,Y,I)**			
	Y=ROAD		***Y=FALSE ALARM***	
	L	$\bar{L}$	***L***	$\bar{L}$
FIELD	0.7	0.3	0.3	0.7
URBAN	0.5	0.5	0.5	0.5
FOREST	0.5	0.5	0.5	0.5
OTHER	0.6	0.4	0.6	0.4
MIXED	0.5	0.5	0.5	0.5

TABLE III. POSTERIOR PROBABILITIES CONDITIONAL ON CONTEXT II

Global context C	**P(L|C,Y,I)**			
	Y=ROAD[a]		***Y=FALSE ALARM***[a]	
	L	$\bar{L}$	***L***	$\bar{L}$
FIELD	0.85	0.15	0.15	0.85
URBAN	0.7	0.3	0.7	0.3
FOREST	0.5	0.5	0.5	0.5
OTHER	0.7	0.3	0.3	0.7
MIXED	0.7	0.3	0.3	0.7

a. The direction of ROAD or FALSE ALARM approaches range of the SAR sensor

IV. RESULTS AND DISCUSSION

The fusion approach was tested on two multi-aspect SAR images (X-band, multi-looked, ground range SAR data) of a sub-urban scene near the airport of DLR in Oberpfaffenhofen, southern Germany. One image was illuminated from the south (Fig. 3) and one from the south-east (i.e. with roughly 45° difference). Global context regions can be derived from maps or GIS before road extraction, or can be segmented automatically by a texture analysis. As a start, global context (URBAN, FOREST, FIELDS and OTHER) is extracted manually (see Fig. 4).

The iterative fusion step is illustrated in Fig. 5. The reader can differentiate between best-evaluated line segments and their neighboring line segment, due to their different colors.

A fact that comes clear from the comparison of Figs. 6 and 7 is the importance of using global context for the evaluation, in particular for determining the Bayesian priors. Incorporating global context reduces the number of false alarms in forest regions (marked black in Fig. 4).

Unfortunately the main directions of the roads in Fig. 3 do not coincide with the range directions of the SAR sensors. But the parking lanes in the lower right corner are an exception. As a consequence, the parking lanes in the lower image obtain higher probabilities.

As can also be seen from Figs. 6 and especially 7, most line segments that correspond to roads still got a good evaluation. On the other hand, many of the false alarms in the urban and forest area are rated worse, even though also some correct segments got a bad rating. However, keeping in mind that this evaluation and the following fusion are intermediate steps before the network-based grouping (see flow charts in Figs. 1 and 2) the Bayesian fusion seem indeed being robust enough to be applied.

The results achieved so far are promising in terms that the fusion of the lines is on one hand statistically sound and, on the other hand, it closely matches the assumptions on the significance of different attributes with respect to their distinctiveness. However, the fusion of evaluated lines from different views still has to be connected with the following network-based grouping and above all be analyzed in depth.

Figure 3. One of the two SAR images analysed in this work

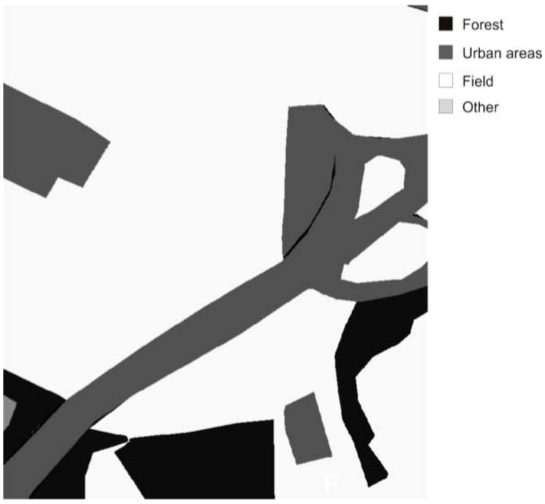

Figure 4. Manually extracted global context

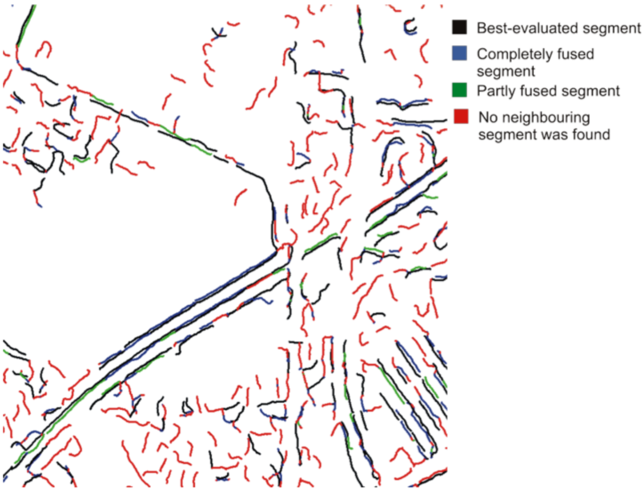

Figure 5. Iterative fusion

Figure 6. After fusion ignoring context, Prior: P(Y)=0.3, P(Y)=0.7, $P(L|Y_1,I)=0.6$, $P(L|Y_2,I)=0.4$, $P(\overline{L}|Y_1,I)=0.4$, $P(\overline{L}|Y_2,I)=0.6$

Figure 7. After fusion incorporating context information. Prior and posterior probabilities used can be seen in Tab. 2 and 3.

ACKNOWLEDGMENT

The authors would like to thank the Microwaves and Radar Institute, German Aerospace Center (DLR) for providing SAR data. This work was done within the TUM-DLR Joint Research Lab (JRL) [http://www.ipk.bv.tum.de/jrl], which is funded by Helmholtz Gemeinschaft.

[1] A. Roth, "TerraSAR-X: A new perspective for scientific use of high resolution spaceborne SAR data," 2nd GRSS/ISPRS Joint workshop on remote sensing and data fusion on urban areas, URBAN 2003. IEEE, pp. 4-7.

[2] J.H.G. Ender, A.R. Brenner, "PAMIR - a wideband phased array SAR/MTI system," IEEE Proceedings - Radar, Sensor, Navigation, 2003, vol 150(3): pp. 165-172.

[3] Wessel, B., Hinz, S., 2004. Context-supported road extraction from SAR imagery: transition from rural to built-up areas. In: *Proc. EUSAR 2004*, Ulm, Germany, pp. 399-402.

[4] Amberg, V., Coulon M., Marthon P., Spigai M., 2005. Improvement of road extraction in hihg.resolution SAR data by a context-based approach, Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Vol. 1, pp. 490-493.

[5] F. Tupin, B. Houshmand, M. Datcu, "Road Detection in Dense Urban Areas Using SAR Imagery and the Usefulness of Multiple Views", IEEE Transactions on Geoscience and Remote Sensing. Vol. 40, No 11, pp. 2405-2414, Nov. 2002.

[6] F. Dell'Acqua, P. Gamba, G. Lisini, "Improvements to Urban Area Characterization Using Multitemporal and Multiangle SAR Images", IEEE Transactions on Geoscience and Remote Sensing. Vol. 4, No. 9, pp. 1996-2004, Sep. 2003.

[7] B. Wessel, C. Wiedemann, "Analysis of Automatic Road Extraction Results from Airborne SAR Imagery", In: Proceedings of the ISPRS Conference "PIA'03", International Archieves of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich 2003, 34(3/W8), pp. 105-110.

[8] C. Wiedemann, S. Hinz, "Automatic extraction and evaluation of road networks from satellite imagery", International Archives of Photogrammetry and Remote Sensing. 32(3-2W5), Sep. 1999, pp. 95-100.

[9] C. Steger, "An unbiased detector of curvilinear structures", IEEE Trans. Pattern Anal. Machine Intell., 20(2), pp. 549-556, 1998.

[10] K. Hedman, S. Hinz, U. Stilla, "A Probabilistic Fusion Strategy Applied to Road Extraction from Multi-Aspect Sar Data", International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, Vol. 36-3, 55-60, 2006.