THE APPLICATION AND POTENTIAL OF BAYESIAN NETWORK FUSION FOR AUTOMATIC CARTOGRAPHIC MAPPING

Karin Hedman

Technische Universitaet Muenchen Institute for Astronomical and Physical Geodesy Arcisstrasse 21, 80333 Muenchen Germany Stefan Hinz

KIT- Karlsruhe Institute of Technology Institute for Photogrammetry and Remote Sensing Englerstrasse 7, 76128 Karlsruhe Germany

ABSTRACT

The research into automatic cartographic mapping is a current topic due to today's availability of high resolution remote sensing data. In order to get as much reliable information as possible, it is recommendable to fuse different image data of the same scene. No matter if the images are acquired by different sensors, from different directions (i.e. multi-aspect data), or are multi-temporal, a careful fusion is required. In this paper we present a high-level decision fusion based on Bayesian network theory developed for automatic road extraction from multi-aspect SAR data. First, the Bayesian network theory is briefly introduced, followed by the process of developing the fusion for the road extraction: 1) Formulating the problem by means of a Bayesian network 2) Learning by estimating up conditional probabilities. Results of the fusion tested on TerraSAR-X data are presented. In the end the potential of the Bayesian network fusion for automatic mapping of cartographic features are discussed.

Index Terms— Bayesian networks, SAR, multi-aspect data, fusion, road extraction

1. INTRODUCTION

The availability of high resolution remote sensing data of urban areas has increased enormously in the recent years. Not only optical but also recently launched synthetic aperture radar (SAR) sensors have reached a new dimension in terms of resolution and offers new potential for urban area mapping. Detailed information, not only building structures, but even cars, traffic lights, and crash barriers can be extracted. In order to get as much information as possible out of urban scenes, it is advisable to utilize different images of the same scene. Previous research has shown that a combination of images acquired by different sensors [1], multi-temporal data sets [2], and fusion of images taken from different directions, so called multi-aspect images [3], can improve the results of urban mapping. Each image combination faces different

challenges. Often the extracted information from each single image is not only complementary and redundant but, depending on the level of complexity, also contradicting. Due to the high level of details in urban areas, it is often not recommendable to do direct fusion on pixel-level. Instead a careful fusion on decision-level, which is able to not only combine different information but also to estimate the uncertainty of the extracted information, can be applied. Bayesian network theory has here turned out to be especially useful.

In this paper, we present a fusion strategy based on Bayesian network theory which fully exploits the capabilities of multi-aspect SAR data [4][5]. Previous research has shown that multi-aspect SAR data indeed improves the result of road extraction [6][7]. The first part of this paper briefly explains the Bayesian network theory (see Sect. 2). The second part presents the fusion (see Sect. 3) itself, where the focus is on its structure and learning. Section 4 contains the results of the fusion tested on TerraSAR-X data. The fusion strategy was in the first place developed for multi-aspect SAR data, but we will also show that it can easily integrate additional information such as optical images or GIS (see Sect. 5).

2. BAYESIAN NETWORK THEORY

Bayesian network (BN) theory originates from the classical Bayesian inference theory. The BN U is a set of nodes $U = \{X_1, ..., X_n\}$, which are connected by a set of arrows $\mathbf{A} = \{(X_i, X_j) | X_i, X_j \in \mathbf{U}, i \neq j\}$ in a directed acyclic graph [8]. Nodes represent the variables (i.e. temperature of a device, gender of a patient or feature of an object) and the arrows symbolizes the causal dependencies among the variables. Variables may have discrete or conditional states. If there is an arrow from node Y to node X; this means that Y has an influence on X. Y is called the parental node and X is called the child node. X is assumed to have X states X_1, \ldots, X_n and X_n is the probability of each certain state X_n . Let X_n be the joint probability distribution of the state values X_n . For being a Bayesian net-

work, U has to satisfy the Markov condition, which means that a variable must be conditionally independent of its non-descendents given its parents. $P(x_1, x_2, ..., x_n)$ can therefore be defined as

$$P(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} P(x_i | pa(X_i))$$
 (1)

where $\mathbf{pa}(X_i)$ represent the parents states of node X_i . Often one is interested in which state a certain node is, the so called *hypothesis variable*. The estimation is done by means of known or partly known variables, *information variables*. Evidence can flow both upwards and downwards. Information coming from a parental node are called *causal* evidence while *diagnostic* evidence enter via a child node.

3. BAYESIAN NETWORK FUSION

Bayesian networks express causal relationships and enable us to model data similar to a human's reasoning process. By applying Bayesian network to fusion of remote sensing data for object extraction, one need to start by analyzing the available information, i.e. searching for the information variables. That means to identify what features (pixels, image primitives, textural patterns, etc.) are extractable and estimate their reliability and which information is needed for estimating the state of the hypothesis variable. The development of a fusion contains the following steps; 1) Analysis of the problem in order to find information and hypothesis, 2) The set up of a Bayesian network structure, which means finding dependencies/independencies among the information and hypothesis variables and modeling the flow of information among the variables, and 3) The learning, which is carried out by estimation of conditional probability functions and conditional probability tables.

In this paper the process is exemplified by describing a fusion for road extraction from multi-aspect data in the context of cartographic mapping.

3.1. Fusion for road extraction from multi-aspect SAR images

The fusion presented in this paper was aimed to adapt an already existing road extraction system [9][10] to multi-aspect SAR data. The road extraction system is based on explicit modeling which makes it especially convenient to apply Bayesian network to the fusion. More information on how the fusion is integrated into the road extraction system can be found in [4].

3.1.1. Set-up of the Bayesian network structure

The main input to the fusion is not only the information extracted from the image (in this case a dark and bright line

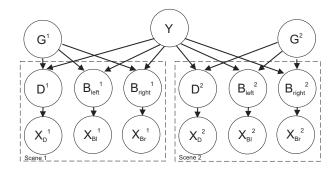


Fig. 1. Bayesian network for road extraction from two multi-aspect images. Superscript 1 and 2 refers to image 1 versus image 2. The nodes are the following: Y: The hypothesis variable, i.e. object in the real world, G: Sensor geometry, D: Extracted dark line primitive, B_{left} : Extracted bright line primitive on the left, B_{right} : Extracted bright line primitive on the right, X: Attributes of D and B

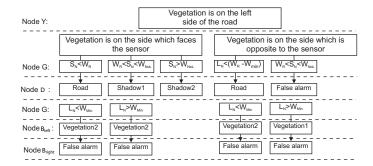


Fig. 2. For each state of Y, a detailed image model is required, from which the probability of expecting states of D, B_{left} and B_{right} can be modeled. S is the expected width of the shadow caused by the trees. W are thresholds representing maximum and minimum width which are related to the line extraction settings. The state vegetation is here divided into two groups, since the line extractor is assumed to behave differently for the two groups.

extraction) but also the image model, i.e. how roads usually appear in high resolution SAR data. A reasoning step is modeled, which describes the relation between the extracted object, nearby high objects and the sensor geometry. For instance an extracted road oriented in the looking angle of the SAR sensor (range) is considered more reliable than other detections closer to azimuth. Furthermore information about neighboring high objects (local context information) is integrated since these objects can be detected by a bright line extraction. Examples of neighboring high objects are trees and buildings. By incorporating this into the reasoning step, contradicting hypotheses (e.g. detection of a road in the first image, detection of parallel shadow and layover regions caused by neighboring high objects in the second image) can be solved. The reasoning step is modeled as a Bayesian network by setting up rules and defining a priori knowledge.

The line extraction is based on Steger's differential approach as also applied in [10] which is able to detect both

dark and bright linear features. In high resolution SAR data, not only roads are detected by the dark extraction, but also linear shadow regions and clutter (false alarms) due to vegetation. Many gaps occur as well as the over-segmentation is high, meaning that the intermediate results contain a high uncertainty. The incorporation of bright features enables us to include layover, which is a necessity for the reasoning step. Crucial is as well the sensor geometry information and relates to the assumption that if there are high trees next to the road, in some geometries the road can actually be detected, in other merely the shadow or bright scattering from the vegetation appears. The resulting Bayesian network with a list of the nodes can be seen in Fig. 1. The information nodes are the attributes of the line primitives, X, the sensor geometry, G, and to some extent the dark (D) and the bright line primitives B, i.e. detected or not. Hence information flow both up- and downwards in the system. All the three variables D, B_{left} , and B_{right} are treated as independent of each other. They are partly hidden nodes, since the states are not completely visible. Y and G are connected through incoming evidence from their child nodes.

The node D contain the following states; (d_1 :ROAD, d_2 :SHADOW, and d_3 :FALSE ALARMS). The states can be distinguished by their significant attributes (such as intensity, length and straightness). The bright line extraction is able to detect local context information such as bright scattering and layover effects caused by buildings (b_1 :BUILDINGS), vegetation (b2:VEGETATION), or crash barriers or vegetation nearby the highway lanes (b_3 :HIGHWAYS). Also these three classes can be separated from each other by means of significant attributes. Node Y has in total ten different states, consisting of both road classes (e.g. y_1 :OPEN ROAD, y_2 :HIGHWAY, y_3 :ROAD WITH VEGETATION NEARBY - LEFT SIDE, etc.) and non-road classes (e.g. y_6 :HIGH VEGETATION, y_8 :BUILDING, y_{10} :CLUTTER), etc.). The non-road classes are included in order to sort out falsely extracted dark primitives that belong to shadows caused by high objects, or belong to simply clutter.

3.1.2. Conditional probabilities

The learning is carried out by estimating conditional probabilities between the nodes. The link between the detected line primitive and their attributes P(D|X) and P(B|X), was quantified by setting up a "classifier". This was done by estimating each separate likelihood by a probability density function fitted to histograms obtained by a training data set ($\tilde{1}000$ dark, 100 bright linear features) [5]. There was no significant correlation between the attributes meaning that the likelihoods $P(X|d_j)$ and $P(X|b_j)$ is the product of each separate likelihood for each attribute. Optical data and maps were used as reference.

The links expressing the relation between the existing object Y, the sensor geometry G and the extracted line primi-

tives D and B, (P(D|G,Y), P(B|G,Y)) are far more complex. It needs to consider the image model considering different sensor geometries, which practically mean that for each state of Y, one need to analyze how this object will appear in the SAR image for all kind of different sensor geometries [5]. The complexity of the task is exemplified by Fig. 2 and it gets clear why it must be defined whether the bright feature is detected on the left (B_{left}) or on the right side (B_{right}) of the dark line segment. Furthermore the link must also consider the performance of the line extractor as well as the performance of the "'classifier", i.e. what is the probability that a possible state of Y, acquired by a certain sensor geometry, is detected by the line extractor, and if detected, what is the probability that the line primitive is assigned to a certain state of D and B. The performance of the line extraction was carried out by comparing the results of the line extractor with reference data, and thereby the knowledge that the line extraction sometimes fail (i.e. many gaps) is included. The training data was randomly separated into one training set and one testing set. The first one was used for the definition of the continuous probability density functions (i.e. building the "classifier"") and the second one was used for a classification error matrix (i.e. evaluating the "'classifier""). These error matrices give information about how well the classification works for each state of D and B.

4. RESULTS

The fusion was tested on TerraSAR-X data (high resolution spotlight mode, multi-look ground range detected) acquired in ascending and descending orbits in March 2009. The antiparallel scenes show an area in Garching, close to Munich, Germany. The result of the classifier can be seen in Fig.3(a)-(b) and the final result excluding the non-road states after fusion in Fig. 3(d). The results were compared to manually digitized reference data. The fusion turned out to be especially successful by identifying highways (user's accuracy: 77%), but had some problems by roads with high vegetation nearby (user's accuracy:9%). This seems to be a consequence of a mis-classification of the bright features into highways, which is an indication of improperly learning. In contrary to the scene from where the training data was collected, the vegetation nearby the roads in this scene have similar characteristics as vegetation close to highways (long and regular shape). The overall completeness of finding roads was 77%. The overall correctness of 44% is expected to rise as soon as the extracted line primitives and their assessment through the fusion is included in a network generation as presented in [9]. Isolated falsely classified line primitives will then be sorted out. Considering the high error rate by the line extraction and that the fusion only delivers an intermediate result, the correctness and completeness are indeed reasonably good values.

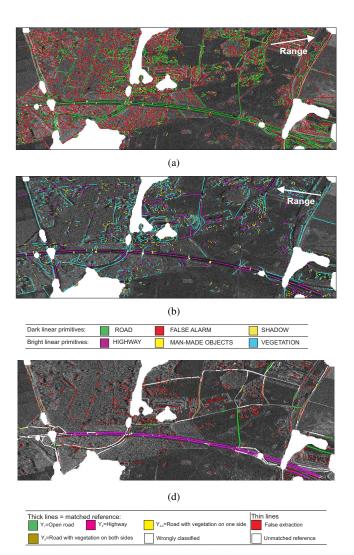


Fig. 3. Uncertainty assessment of dark linear features (a) and bright linear features (b) extracted in each of the two SAR sub-scenes of Garching. The linear features and their uncertainty assessment is the input to the subsequent fusion (c). The results after fusion based on the Bayesian network was matched with the reference. Masked built-up areas are white.

5. DISCUSSION AND FUTURE CONCEPTS

Despite the complexity of learning, the Bayesian network theory is optimal for this kind of work. The decisive point is that Bayesian network enables us to realize a complex image model within a probabilistic framework. Furthermore we are able to include the uncertainty of the line extraction as well as the "classifier" as a learning step. The proposed Bayesian network offers us the ability to easily integrate additional data, for instance GIS information or remote sensing data from other sensors, such as optical data. In case of GIS data, the information should be integrated as prior information, (e.g. "What is the rate of constructing new roads per year?" or "X% of existing road network remain from one year to an other"). Any information extracted from optical

data is able to be included through child nodes. The learning can be carried out in a similar way as presented in this work. An additional advantage of formulating the fusion in the frame of Bayesian network is that the output is not a definite classification, merely each line feature obtain an uncertainty value of being a road or not. The output of the fusion has therefore a high potential to be integrated in a network generation step based on Markov random field as in [6] or in the framework of conditional random fields [11].

A Bayesian network requires a complete probabilistic model and dependent on the amount of variables and states, the set-up of the system as well as the learning can be very time-consuming. But as long as the task is not too complex, Bayesian network theory has indeed due to its flexibility the potential of being the framework for future urban object extractions from remote sensing data.

6. REFERENCES

- [1] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and sar spaceborne images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 10, pp. 3932 3946, oct. 2011.
- [2] F. Dell'Acqua, P. Gamba, and G. Lisini, "Improvements to urban area characterization using multitemporal and multiangle sar images," *Geo*science and Remote Sensing, IEEE Transactions on, vol. 41, no. 9, pp. 1996–2004, Sept. 2003.
- [3] A. Thiele, E. Cadario, K. Schulz, U. Thonnessen, and U. Soergel, "Building recognition from multi-aspect high-resolution insar data in urban areas," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 11, pp. 3583–3593, Nov. 2007.
- [4] U. Stilla and K. Hedman, Feature fusion based on Bayesian network theory for automatic road extraction, vol. 15, chapter 3, pp. 69–86, Springer, 04 2010.
- [5] K. Hedman, Statistical fusion of multi-aspect Synthetic Aperture Radar data for automatic road extraction, Ph.D. thesis, Karlsruher Institute for Technology, 07 2010.
- [6] F. Tupin, B. Houshmand, and M. Datcu, "Road detection in dense urban areas using sar imagery and the usefulness of multiple views," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 11, pp. 2405–2414, Nov 2002.
- [7] K. Hedman, U. Stilla, G. Lisini, and P. Gamba, "Road network extraction in vhr sar images of urban and suburban areas by means of class-aided feature-level fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1294–1296, 02 2010.
- [8] Judea Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, September 1988.
- [9] C. Wiedemann and S. Hinz, "Automatic extraction and evaluation of road networks from satellite imagery," in *International Archives of Photogrammetry and Remote Sensing*, 1999, pp. 95–100.
- [10] B. Wessel and C. Wiedemann, "Analysis of automatic road extraction results from airborne sar imagery," in *Proceedings of the ISPRS Con*ference Photogrammetric Image Analysis, 2005, vol. 34, pp. 105–110.
- [11] J.D. Wegner, R. Ha? andnsch, A. Thiele, and U. Soergel, "Building detection from one orthophoto and high-resolution insar data using conditional random fields," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, vol. 4, no. 1, pp. 83 –91, march 2011