On the Improvement of Structural Detection of Building Features in High-Resolution SAR Data by Edge Preserving Image Enhancement

E. Michaelsen, W. Middelmann, U. Soergel, U. Thoennessen

FGAN-FOM, Gutleuthausstrasse 1, 76275 Ettlingen, Germany
{mich,middelmann,soe,thoe}@fom.fgan.de

A production system approach is used for the detection of building structures in high-resolution X-band SAR-data. The productions code principles from perception psychology like good continuation and similarity for the use in computing machines. The process control of the system is assessment driven. Exhaustive search is not feasible and avoided. Instead, the search process is terminated after a predetermined effort. Thus the result is crucially dependent on the quality of the input data concerning noise and clutter. It is shown that an edge preserving image enhancement filter taking into account the special properties of such data can significantly improve the performance of the structural analysis.

1 Introduction

State-of-the-art airborne SAR systems can resolve objects of size even below a decimeter [2]. The recognition of e.g. urban buildings in such high-resolution SAR data requires the processing of large pictures. Objects of interest may measure several thousand pixels in length and width, while the features discriminating these objects from clutter best may well require full resolution. In these new X-band data very fine lines and salient point scatterer rows are obvious important hints for the presence of building structures (see Fig. 1).

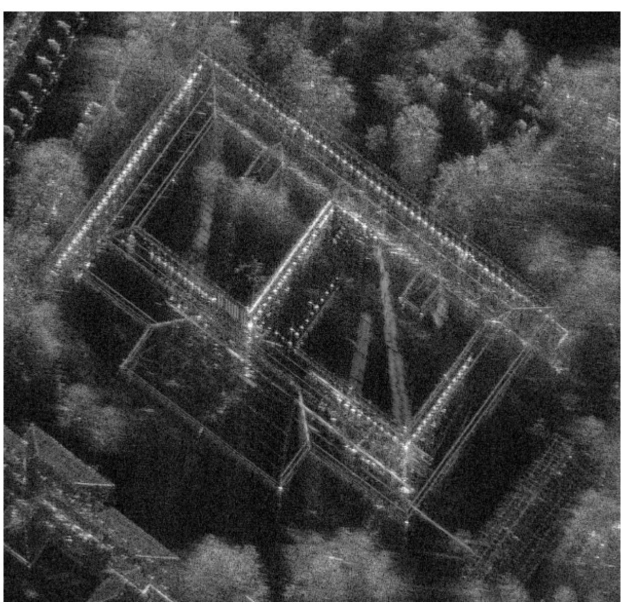


Fig. 1.High resolution X-band SAR-image containing a building block

Traditional standard image smoothing techniques or down-sampling will spoil many of these features. Hence, the iconic preprocessing has to be tailored to the particularities of the sensor and to the features demanded by the higher levels. Sect. 2 introduces such a filter. Its benefit will be discussed in the context of a segmentation and grouping of salient structures of a building block. For this purpose, the concepts of human perception [7] are integrated in a knowledge based system [6] for the automatic detection of man-made structures. Sections 2 to 5 briefly describe this approach while the last two sections present the experiment and result.

2 Efficient Edge Preserving Image Enhancement

The preprocessing with an image enhancing method aims to improve the performance of the structural analysis. Therefore an edge preserving method proposed by Cetin [1] has been used. Originally it is designed for a spot mode SAR image processing. But with respect to a strip mode SAR model the method is simplified here. It is defined by optimization of the following objective function that consists of a data fidelity term and additionally non-quadratic regularization terms:

$$J(f) = \| g - f \|_2^2 + \lambda_1 \| f \|_k^k + \lambda_2 \| \mathbf{D} | f | \|_k^k = \min_f \quad (1)$$

Here g is the SAR image and f is its enhanced version. The scalar parameters $\lambda_1,\lambda_2 \geq 0$ control the behavior of this enhancement $\|\cdot\|_k$ denotes the ℓ_k-norm, and $|f|$ is the column stacked vector of magnitudes. **D** is a discrete approximation to the 2-D gradient, i.e. all magnitude differences between right and downward neighbors are taken into consideration.
In general we chose *k<1*. Hence $\|\cdot\|_k$ is not a norm (violating the triangular in-equality). But for smoothing homogeneous regions and particularly for edge-preserving this choice yields improved behavior of the second regularization term. The first regularization term is used for point enhancement.

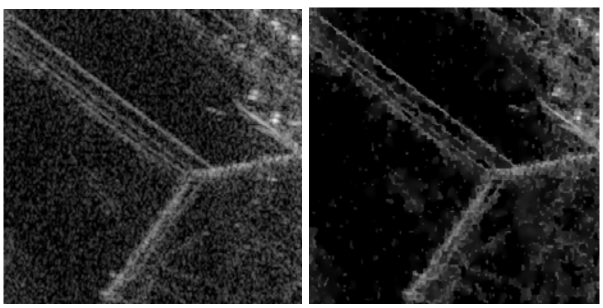


Fig. 2.Left original, right filtered image with λ_1=0.5, λ_2=0.9, k=0.2, ε=10^{-9}

In contrast to the original method [1] g is not complex valued here, i.e. similar results have been computed using the magnitude SAR image only. Restriction to ordinary float operations also helps reducing the computational effort.
Still, the enhancement of large images requires an efficient solving technique for the non-quadratic optimization problem. Following [1] J is approximated by a differentiable objective function J_ε with a small constant $\varepsilon \geq 0$. This is done by modifying the ℓ_k-norm in (1) by $\| x \|_k^k \approx \sum_i (x_i^2 + \varepsilon)^{k/2}$. The solution of the approximated version of (1) is done by solving the nonlinear system $\nabla J_\varepsilon(f) = 0$. This can be written as

$$\nabla J_\varepsilon(f) = H(f)f - 2g = 0 \qquad (2)$$

Here f and g are column stacked vectors. The symmetric system matrix H has five non-zero diagonals only. The matrix H produced by the original method does not have this advantageous sparse structure.

The system (2) is solved by a novel two-stage Jacobian method that works on multiplicatively decomposable sub-systems of H. From the sparsity of H it follows each iteration step is of only linear complexity related to the number of pixels. Hence this is particularly efficient for the enhancement of large images. In Fig. 2 the outcome is compared to the original using a small section of the image shown in Fig.1.

3 Feature Extraction

In order to apply assessment driven symbolic reasoning to the SAR-data the information has to be transformed from the iconic image matrix into an unordered set of attributed objects.
Here we use the squared averaged gradient method [3] to extract **L**-objects (lines) where an edge- or line-like structure is likely to be present in the image. Also these objects are fairly short – being located at just one pixel – they already contain an orientation attribute inherited from the gradient.
In high resolution SAR-data salient scatterers carry important information. To extract these we use a simple spot-operator developed for data from the thermal spectral domain [5]. Only a small portion of the pixels of an image will be non-zero after applying this filter. The edge preserving enhancement further reduces the number. Moreover, an appropriate threshold is set for the minimal value of a pixel in order to be accepted as **P**-object.

4 Production Net

Three structures are defined for the automation of assessment driven symbolic reasoning: 1) The set of objects, 2) the set of productions and 3) the control rationale. The interrelation between objects and productions is best understood displaying a production net. Fig. 3 shows the system used in this contribution for building detection.
Production nets are best understood discussing such an example production by production:
`P1` is constructing an **S**-object (“scatterer” or “spot”) from a set of **P**-objects (pixels trespassing the hot-spot feature extraction threshold). The condition part of this

production is proximity. The dots "…" indicate that the size of the set of **P**s is not fixed. The **S**-object produced is assessed according to its mass (sum of grey-values of the preceding **P**-objects).

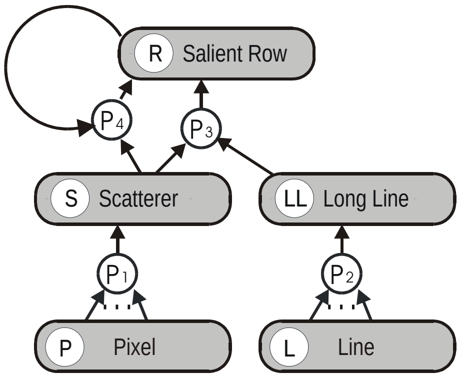


Fig. 3. Production Net

P2 is constructing an **LL**-object (prolonged line) from a set of **L**-objects (lines extracted by the gradient-based feature extraction filter). The condition part of this production is co-linearity. The **LL**-object produced is assessed according to its length.

P3 initializes the recursive construction of **R**-objects (salient rows) by defining a single **S**-object to be an **R**-object with one member provided there is an **LL**-object close to it. The rows direction attribute is taken from the **LL**-object, while its spacing attribute is set to a default a-priori value. The new **R**-object inherits the assessment from the preceding **S**-object.

P4 defines the recursion step appending additional **S**-objects to an existing **R**-object whenever they fit sufficiently well into its direction and spacing attributes. The new **R**-object is assessed according to the deviation from the expected fit. The new attributes are obtained by averaging. The expected error (i.e. the size of the search region for the next possible members) is becoming smaller with rising number of members.

Examples of other production nets for other tasks may be found in [4,6].

5 Control

Our control strategy works on the set of processing elements *(i,p)*, where *i* is an object and *p* from $\{P_1, \ldots, P_n, nil\}$. Each new object also defines a new element with processing index *nil*. The elements are constantly being sorted according to the assessment attribute of their object part. Best elements are processed first.

Processing an element *(i,nil)* means replacing it by $\{(i, P_{n1}), \ldots, (i, P_{nj})\}$ with appropriate productions $P_{n1} \ldots P_{nj}$ according to the connectivity of the production net (which is known to the control module but not to the modules coding the productions).

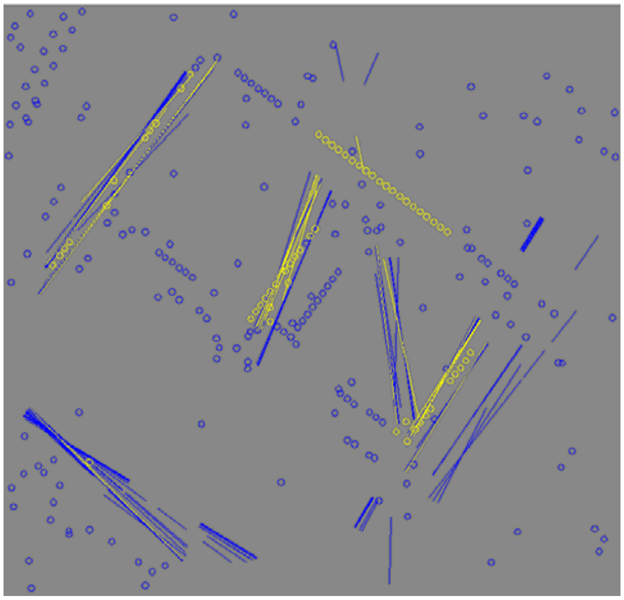
Processing an element (i, P_n) means calling the module coding the production P_n with the attribute values of object *i* as parameter. The module will search for partner objects that fulfill the productions condition together with object *i*. It may create none, one or many new objects – and thus also new processing elements. After such processing the element is deleted while its object and also the partner objects encountered remain in the object set.

The process may start if all or some objects are constructed by the feature extraction procedures. New objects may be added by external processes at any time. The process will end if all processing elements disappeared. Often this may not be feasible. In this case the process may be terminated by any external criterion. After termination the set of the best objects of highest rank in the net is output as result. A detailed description of the control structure may be found in [6].

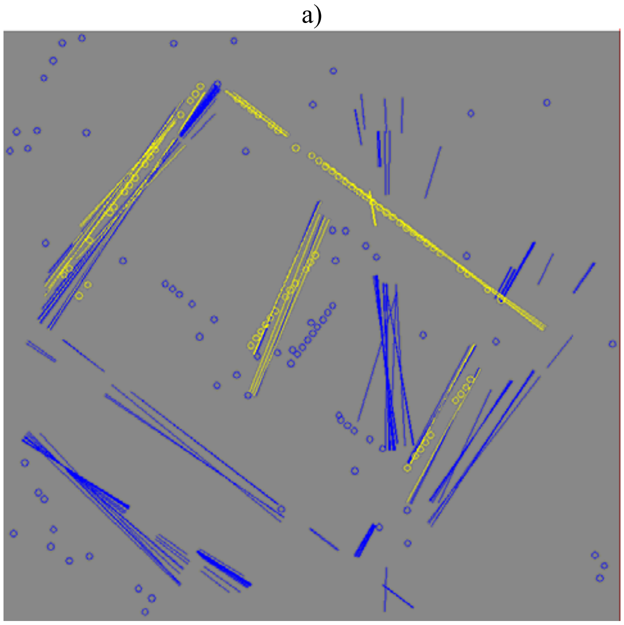
6 Experiment

A high resolution SAR-image of more than 2000 Pixel size in both directions has been filtered by the efficient edge preserving image enhancement mentioned above (Figs. 1 and 2). From both, the filtered version as well as the original primitive **P**-objects and **L**-objects were extracted using the same parameter setting. Then the search procedure described above was applied to these sets using the production system displayed at Fig. 3 and terminated after 5000 search cycles. Results are shown in Fig. 4 (all **S**-objects and all **LL**-object in blue, all **R**-objects with more than one member in yellow).

Without the preprocessing many more clutter objects burden the search.



a)



b)

Fig. 4. Result after 5000 search cycles: a) without pre-processing; b) with pre-processing

7 Conclusion

Termination of the symbolic reasoning process after only 5000 control cycles is far from exhaustive. Recall that a cyclic production net may cause exponential rising effort (in the number of primitives). Both runs benefit from allowing more searching. But the example shows that important structure will appear earlier if an enhancement filter reduces clutter and noise in the image. The original SAR image enhancement following [1] leads to high computational effort on such large images. However, simplifying this enhancement step in the proposed way leads to computational effort that grows only linear with the number of pixels while the results remain good enough to significantly improve the performance of the structural analysis.

Yet in the result displayed in Fig. 4b there is more computation involved than in the result displayed in Fig. 4a (namely the filter processing). It remains to be studied what partition of the overall computational resources and time between the iconic filter process and the symbolic reasoning process yields optimal performance.

References

1. M. Çetin, W.Clem. Karl, and D.A. Castañon, Analysis of the Impact of Feature-Enhanced SAR Imaging on ATR Performance. // E.G. Zelnio (ed.): Proceedings of SPIE - Algorithms for SAR Imagery IX, Orlando, Florida. – 2002. - Vol. 4727. – P. 134-145.
2. J. H. G. Ender, A. R. Brenner: PAMIR - a wideband phased array SAR/MTI system. // IEE Proceedings - Radar, Sonar, Navigation. - 2003. - Vol. 150, No. 3. - P. 165-172.
3. W. Foerstner: A Framework for Low Level Feature Extraction // J.-O. Eklundh (ed). Computer Vision – ECCV 94. -1994. - Vol. II, B1. – P. 383-394.
4. E. Michaelsen, U. Stilla, K. Lütjen: Which Mathematical Foundations Support Model Based Vehicle Recognition? // Pattern Recognition and Image Analysis: Representation, Processing, Analysis and Understanding of Images. - 2000. - Vol. 33, No.1. - P. 328-335.
5. Michaelsen, E., Stilla, U. 2001. Estimating urban activity on high- resolution thermal image sequences aided by large scale vector maps. // IEEE/ISPRRS Joint workshop on remote sensing and data fusion over urban areas, Urban 2001. IEEE, ISBN 0-7803-7059-7, 25-29U.
6. Stilla: Map-Aided Structural Analysis of Aerial Images. // ISPRS Journal of Photogrammetry and Remote Sensing. – 1995. – Vol. 50, No. 4. – P. 3-10.
7. M. Wertheimer: Untersuchungen zur Lehre von der Gestalt II. // Psychologische Forschung. – 1923. - Vol. 4. // Translated as ‘Principles of Perceptual Organization’ in: D. Beardslee, M. Wertheimer (eds.), Princeton, N. J. – 1958. – P. 115-135.