

# WEIGHTED PYRAMID LINKING FOR SEGMENTATION OF FULLY-POLARIMETRIC SAR DATA

Ronny Hänsch, Olaf Hellwich

Berlin Institute of Technology, Computer Vision and Remote Sensing  
Franklinstrasse 28/29, Office FR3-1, 10587 Berlin, Germany.  
Tel. +49 30-314 73 107, Fax. +49 30-314 21 104, E-mail: rhaensch@fpk.tu-berlin.de

**KEY WORDS:** SAR, Segmentation, Polarization, Algorithms, Multiresolution, Spatial

## ABSTRACT:

Image segmentation has the general goal to define regions within an image, in which all pixels have similar properties. For fully-polarimetric SAR data this is often done by spectral classification without any use of spatial information. On the contrary the proposed method aims to find homogenous segments in the image, which should be compact and connected if possible. A multiresolution image pyramid allows to calculate information based on regions of different size instead of single pixels or small neighbourhoods. Furthermore, a relaxation approach is used to defer the segmentation decisions until more accurate information is available.

## 1 INTRODUCTION

Image segmentation is an important preprocessing step in many applications. Numerous tasks such as classification, object detection and so forth can be achieved much more easily and accurately given an appropriate segmentation.

Due to the coherent nature of SAR sensors homogeneous areas are no longer homogeneous in the image, but contain strong multiplicative distortions. This speckle effect poses severe problems to spatial segmentation algorithms. A lot of work is done for radiometric classification without using any spatial information, e.g. (Lee et al., 1997, Ferro-Famil et al., 2001, Anfinsen et al., 2007, Hänsch et al., 2008) and there are very few approaches that try to combine spatial context and radiometric evidence as in (Reigber et al., 2007).

In this paper, the segmentation algorithm proposed in (Hong and Rosenfeld, 1984) is used to automatically derive the hierarchical structure of an image of fully-polarimetric SAR data. It is based on a multiresolution image pyramid with the original image at the base. Each higher level of the pyramid contains the image in a lower resolution. The different resolutions are obtained by simply averaging pixels in overlapping windows of certain size during the initialisation. Due to the overlap and window size each element in the pyramid has several parents (at the next higher level) and descendents (at the lower level).

Most algorithms for segmentation work with hard decisions: that means, each pixel is uniquely assigned to a certain cluster or segment. Other methods, which merge or split regions, have to decide for each region whether to split or to merge it. Because the true segments or clusters are a priori unknown, such hard decisions will be erroneous for some pixels. That is why the forced-choice aspect of segmentation has, in practice, a negative influence on the final segmentation result. Particularly, if it is difficult to undo wrong decisions made at the beginning. The algorithm presented avoids this by labelling links between each element and its parents with a certain link strength, representing the degree of association between node and parent. This association is based on a distance measure between the value of this pixel and the values of its parents. As this algorithm is applied to fully-polarimetric SAR data a distance measure is chosen, which respects the statistical properties of such data and is based on the Wishart distribu-

tion. Having established a set of weights, pixels at higher levels can be updated by the weighted average of the values of their descendents.

The entire process is then iteratively repeated until convergence, at which point a segmentation can be extracted. Some pixels in the pyramid will have small link strengths to all of their parents. They form independent subtrees in the pyramid and represent the searched segments.

## 2 THEORETICAL BACKGROUND

### 2.1 Fully-polarimetric SAR data

Fully-polarimetric SAR data measure amplitude and phase of the backscattered signal in four different transmit and receive polarisation combinations. However, a common assumption is that the cross polarisations are the same due to the reciprocity of natural targets. Therefore each data point is a three dimensional vector  $\vec{s}$ :

$$\vec{s} = (S_{HH}, \sqrt{2}S_{HV}, S_{VV}) \quad (1)$$

where  $S_{RT}$  is a complex component of the scattering matrix and  $R \in \{H, V\}$  is the receive and  $T \in \{H, V\}$  is the transmit polarisation.

Often the data is represented as spatially averaged sample covariance matrix in order to reduce speckle and get more statistical information:

$$\mathbf{C} = \frac{1}{n} \sum_{i=1}^n \vec{s}_i \vec{s}_i^H \quad (2)$$

where  $H$  denotes the conjugate transpose and  $n$  is the number of samples used for averaging. If the distribution of  $\vec{s}$  is a multivariate complex Gaussian with zero mean, which is a standard assumption when dealing with fully-polarimetric SAR data, the sample covariance matrix  $\mathbf{C}$  of  $\vec{s}$  is complex Wishart distributed.

$$\vec{s} \sim N(0, \mathbf{\Sigma}) \Rightarrow \mathbf{C} \sim W(n, \mathbf{\Sigma}) \quad (3)$$

Hence, the density of  $\mathbf{C}$  given the covariance matrix  $\Sigma$  is defined by

$$p_n(\mathbf{C}|\Sigma) = \frac{n^{nq} |\mathbf{C}|^{n-q} \exp(-n \cdot \text{tr}(\Sigma^{-1} \mathbf{C}))}{|\Sigma|^n \cdot \pi^{q(q-1)/2} \prod_{k=1}^q \Gamma(n-k+1)} \quad (4)$$

where  $|\cdot|$  is the determinant and  $\text{tr}(\cdot)$  is the trace of a matrix,  $\Gamma(\cdot)$  is the standard gamma function and  $q$  is the dimensionality of  $\bar{s}$ .

## 2.2 Wishart-based distance measure

An often used distance measure for polarimetric SAR data is based on the Wishart distribution and is defined in (Lee et al., 1997):

$$d_W(\mathbf{C}, \Sigma) = -\frac{1}{n} \ln p(\mathbf{C}|\Sigma) \quad (5)$$

$$= \ln(|\Sigma|) + \text{tr}(\Sigma^{-1} \cdot \mathbf{C}) + c \quad (6)$$

where

$$c = -\frac{1}{n} \ln \left( \frac{n^{nq} |\mathbf{C}|^{n-q}}{\pi^{q(q-1)/2} \prod_{k=1}^q \Gamma(n-k+1)} \right) \quad (7)$$

The constant term  $c$  in (6) is class independent and can be omitted, if this distance is used as in (Lee et al., 1997) to decide if the data point  $\mathbf{C}$  more probably belongs to class  $c_1$  represented by the covariance matrix  $\Sigma_1$  instead of belonging to class  $c_2$  represented by  $\Sigma_2$ . The distance measure simplifies to:

$$d_W(\mathbf{C}, \Sigma) = \ln(|\Sigma|) + \text{tr}(\Sigma^{-1} \cdot \mathbf{C}) \quad (8)$$

Although this distance measure is not a metric, because it is neither homogeneous, nor symmetric and does not fulfill the triangle inequality, it is often used and has shown its effectiveness in practice. Because of this and its direct relation to the density function it will be used in a slightly modified version in this work.

## 3 WEIGHTED PYRAMID LINKING

### 3.1 Pyramid construction and initialisation

The basic structure used in this approach is a multiresolution image pyramid. While the original data (fully-polarimetric SAR data, multi-look complex covariance matrices) is forming the base, the higher levels are versions of the image with subsequent reduced resolutions. The height of the pyramid (the number of levels without the bottom level  $l=0$ ) shall be noted by  $L$ . Each pixel is represented by a node in this pyramid. The value  $v_1(n^l(x, y))$  of the node  $n$  at position  $(x, y)$  at level  $l$  is simply an average of a window with certain size  $s$  at the previous level  $l-1$ :

$$\forall l \in [1, L] : v_1(n^l) = v_2(n^l) = \frac{1}{s} \cdot \sum_{n' \in \text{des}(n^l)} v_1(n') \quad (9)$$

The difference between the covariance matrices  $v_1(n^l)$  and  $v_2(n^l)$  will be explained in section 3.3. Just note, that they are set to the same value during the initialisation. Each node  $n$  at level  $l$  ( $0 < l < L$ ) is therefore connected with a set of nodes at level  $l-1$ , called descendents  $\text{des}(n)$  and a set of nodes in level  $l+1$  called parents  $\text{par}(n)$ . Nodes at the bottom level  $l=0$  have only parents, while nodes at the top level  $l=L$  have only descendents. Only vertical connections between nodes at adjacent

levels and no horizontal relations between nodes of the same level are used.

The windows overlap by a predefined amount  $o$  of pixels ( $0 < o < s$ ). The size  $s$  of the window and the overlap  $o$  define the decrease in resolution of the next level. The parameters used in this paper are a quadratic window size  $s$  of  $4 \times 4 = 16$  and an overlap  $o$  of two pixels in  $x$ - and  $y$ -direction. Given this setting of  $o$  and  $s$  there will not be enough pixels at the border of a level for a whole window, if the dimensions of this level are not even. In that case, the level is simply extended with as many pixels as needed. These additional pixels have the same value as the border pixels. Due to this manipulation the border pixels gain greater influence on the pixels at the next level. However, this effect is insignificant as experiments have shown.

### 3.2 Weight adjustment

The most important part of this approach is the introduction of link strengths  $\tilde{w}$  between nodes on adjacent levels of the pyramid. Instead of using only the descendent with the largest degree of association, all descendents contribute accordingly to their link strength to the node value at the next level. The link strength  $\tilde{w}(n, n')$  between node  $n$  and its descendent  $n' \in \text{des}(n)$  is based on proximity and similarity:

$$\begin{aligned} \tilde{w}(n, n') &= \exp(-d_{\text{spec}}(v_2(n'), v_1(n))) \quad (10) \\ &\cdot \exp(-\text{var}(n')) \\ &\cdot \exp(-d_{\text{spat}}(n', n)) \end{aligned}$$

The first factor has the most crucial role. It measures the spectral distance between two nodes at adjacent levels in the image pyramid. Any proper distance measure can be used here. As mentioned above an often used distance measure for polarimetric SAR data is (8), which is based on the Wishart distribution. Note, that the link strength is used to define the contribution of a descendent to the value of the current node in comparison to all other descendents of this node. Furthermore, the values of neighbouring nodes at one level are unlikely to be equal and the number of looks can be different, too. All nodes in the pyramid at the same level will have the same number of looks merely after the initialisation. That is why  $c$  in (6) cannot be omitted and (8) cannot be used here. Therefore  $d_{\text{spec}}(v_2(n'), v_1(n))$  is defined as:

$$d_{\text{spec}}(v_2(n'), v_1(n)) = -\ln p(v_2(n')|v_1(n)) \quad (11)$$

where  $p(v_2(n')|v_1(n))$  is the density of the Wishart distribution defined in (4).

Within the second factor the euclidian distance  $d_{\text{spat}}(n', n)$  between the spatial positions of the two nodes is used. The spatial distance within the  $4 \times 4$  neighbourhood is defined as:

$$\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} \sqrt{9} & \sqrt{5} & \sqrt{5} & \sqrt{9} \\ \sqrt{5} & \sqrt{1} & \sqrt{1} & \sqrt{5} \\ \sqrt{5} & \sqrt{1} & \sqrt{1} & \sqrt{5} \\ \sqrt{9} & \sqrt{5} & \sqrt{5} & \sqrt{9} \end{pmatrix} \quad (12)$$

As the link strength now depends on geometric closeness the regions tend to be more compact, whereas they would have more irregular shapes without this factor.

The third factor represents the variability of the descendent of node  $n$ . Since the goal is to segment the image into homogeneous regions, nodes that represent segments with high variability should get a lower link strength than nodes representing more

homogeneous regions.

$$var(n) = \sum_{n' \in des(n)} d_{spec}(v_2(n'), v_1(n))w(n, n') \quad (13)$$

$$w(n, n') = \frac{\tilde{w}(n, n') \cdot a(n')}{\sum_{n^* \in des(n)} \tilde{w}(n, n^*) \cdot a(n^*)} \quad (14)$$

The weight  $w(n, n')$  between a node  $n$  and its descendent  $n'$  have to be normalized, so that  $\sum_{n' \in des(n)} w(n, n') = 1$  holds. The variability  $var(n)$  of a node  $n$  becomes therefore the weighted average of the spectral distances  $d_{spec}$  to its descendents  $des(n)$ . The area  $a(n)$  of node  $n$  will be explained in more detail in the next section.

### 3.3 Node value recalculation

After the weights of each connection have been adjusted, the values of every parent have to be recalculated. Starting at the level  $l=1$  of the pyramid the values of all nodes in all levels have to be recomputed. The weights used have to depend on the link strength between the two nodes. However, they should also depend on the size of the image area  $a(n')$  represented by the descendent  $n'$ : Consider a node at a particular level of the pyramid, that has only one strong connection down the pyramid to only one image pixel. This node should have less influence than another node, that covers a large area within the image.

$$a(n) = \sum_{n' \in des(n)} \frac{\tilde{w}(n, n') \cdot a(n')}{\sum_{n^* \in par(n')} \tilde{w}(n^*, n')} \quad (15)$$

As the sum in the denominator is computed over the parents of  $n'$ , the area of a node is distributed among its parents in a normalized way. This ensures that the total area of all nodes at each level is the same as the area of the original image.

Every node in the pyramid plays two roles. On the one hand it is the sample covariance matrix of its descendents estimated by weighted averaging. On the other hand it is the descendent of a node on the next level. In order to apply the distance measure (11) it has to be Wishart distributed and the number of looks has to be known. One can show, that the sum of Wishart distributed random variables  $X_i$  is again Wishart distributed:

$$X_i \sim W(n_i, \Sigma), i = 1, \dots, k \Rightarrow \sum_{i=1}^k X_i \sim W\left(\sum_i n_i, \Sigma\right) \quad (16)$$

However, this holds only if all  $X_i$  have the same covariance matrix  $\Sigma$  and are independent. This assumption should be strongly violated at higher levels of the pyramid, because their nodes cover large regions of the image. Furthermore, a multiplication with a scalar changes the distribution:

$$X \sim W(n, \Sigma) \Rightarrow a \cdot X \sim W(n, a \cdot \Sigma) \quad (17)$$

The weighted average can therefore not be assumed to be Wishart distributed. That is why each node holds two values. The first one is simply the weighted average of the values of its descendents and therefore an estimation of the true covariance matrix:

$$v_1(n) = \sum_{n' \in des(n)} v_2(n') \cdot w(n, n') \quad (18)$$

where  $w(n, n')$  is defined by (14). The second one is the (unweighted) average of the descendents with the strongest connections. This ensures, that only descendents which are very likely

to have the same distribution contribute to this value:

$$v_2(n) = \frac{1}{Z} \sum_{n' \in des(n)} v_2(n') \cdot \delta(w(n, n')) \quad (19)$$

$$\delta(w(n, n')) = \begin{cases} 1, & \text{if } w(n, n') > \theta \\ 0, & \text{else} \end{cases} \quad (20)$$

$$Z = \sum_{n' \in des(n)} \delta(w(n, n')) \quad (21)$$

### 3.4 Iterative processing

All calculations, in particular adjustments of link strengths and recalculations of the values of each node are done iteratively. The following gives an overview of the algorithm:

0: INIT: Construct pyramid

1: While: levels not converged

1.1: FOR  $l=1$  TO  $L$

1.1.1: Adjust weights  $\tilde{w}(n^l, n^{l-1})$

1.1.2: Recalculate area  $a(n^l)$

1.1.3: Recalculate values  $v_1(n^l)$  and  $v_2(n^l)$

1.1.4: Recalculate variability  $var(n^l)$

2: Construct tree  $\Rightarrow$  Extract segments

After a few iterations the link strengths will stabilise and not change anymore. At first the level  $l = 1$  of the pyramid converges. At this time each node at this level will have strong links only to that subset of pixels in the set of descendents, which are very likely to have the same distribution governed by covariance matrix  $\Sigma$  of which the node value is an estimation. The second value of a node at level  $l = 1$  will now be an (unweighted) average of Wishart distributed random variables of the same distribution. That is why its own distribution can be assumed to be Wishart, too. However, the averaged variables cannot be assumed to be independent, because of the possible overlap of their areas in the image. Therefore the number of looks is estimated by the area the node covers in the original image and not by the number of looks of its descendents. All levels will converge in ascending order after a few iterations.

### 3.5 Tree construction

If the whole pyramid has converged, meaning that link strengths and, therefore, values of all nodes do not change anymore, there are two general types of nodes in the pyramid. On the one hand, nodes that have strong connections to one or more parents and, on the other hand, nodes which have no strong connection to any parent at all. Latter ones define roots of independent subtrees within the pyramid and represent homogeneous regions in the image. All nodes at the top level of the pyramid or nodes whose link strengths to all their parents are below a certain threshold are considered as such roots. However, because of the different statistics at each level, e.g. the mean number of looks decreases with decreasing height, one cannot use a global threshold. But there is in each level an abrupt rise in the number of roots for a certain value  $t_r$ . This value is used as threshold to define the roots in each level. Figure 1 shows an example of the relationship between the number of roots at a certain level and the threshold to define them.

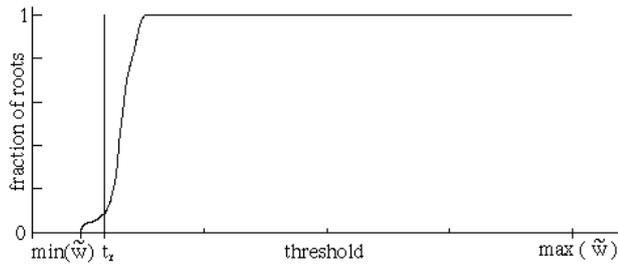


Figure 1: fraction of roots over threshold

Beginning at the top of the pyramid each node is labelled as root or to which root it belongs. When the label propagation has reached the bottom of the pyramid, each pixel of the original image data is labelled by a number indicating to which root, i.e. to which segment it belongs.

#### 4 MERGING

The described approach results in a very fine segmentation, which is very sensitive to changes in the image. This ensures that all real-world boundaries are contained in the segmentation, but results in an image, which is oversegmented to a certain degree. However, the obtained segments cover only strong homogeneous regions and have therefore good spectral properties. A simple merging algorithm is sufficient to reduce the number of segments and to obtain a good segmentation result. Beginning with the largest segment each of its neighbours is investigated. If the similarity  $s(R_1, R_2)$  between segment  $R_1$  and its currently investigated neighbour  $R_2$  is greater than a certain threshold both segments are merged. This is done until all similar neighbouring segments are merged. Afterwards the next region is investigated in the same way.

$$s(R_1, R_2) = \frac{1}{|R_2|} \sum_{\mathbf{X} \in R_2} \delta(\mathbf{X}) \quad (22)$$

$$\delta(\mathbf{X}) = \begin{cases} 1, & \text{if } d_W(\mathbf{X}, \mathbf{Y}_1) \leq \gamma d_W(\mathbf{X}, \mathbf{Y}_2) \\ 0, & \text{else} \end{cases} \quad (23)$$

where  $\mathbf{Y}_i$  is the sample covariance matrix of region  $R_i$  and  $d_W(\mathbf{X}, \mathbf{Y}_i)$  the distance measure defined by (8). The constant  $\gamma = (1 + \epsilon)$ , where  $\epsilon$  is a small number, relaxes this constraint a little bit. In our experiments  $\gamma$  was set to  $\gamma = 1.01$ .

#### 5 RESULTS

In Figure 2 the segmentation result of a  $727 \times 1047$  image of fully-polarimetric SAR data is shown.

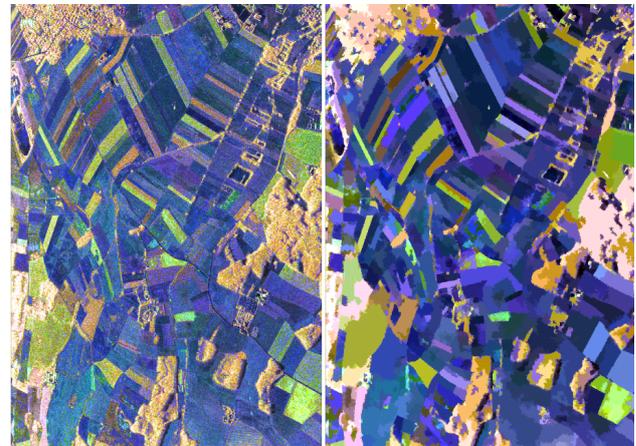


Figure 2: left: original; right: segmentation

Figure 3 shows a magnification of a part of Figure 2. The homogeneous regions have been extracted successfully. Regions with different properties have been separated and are considered as independent segments.

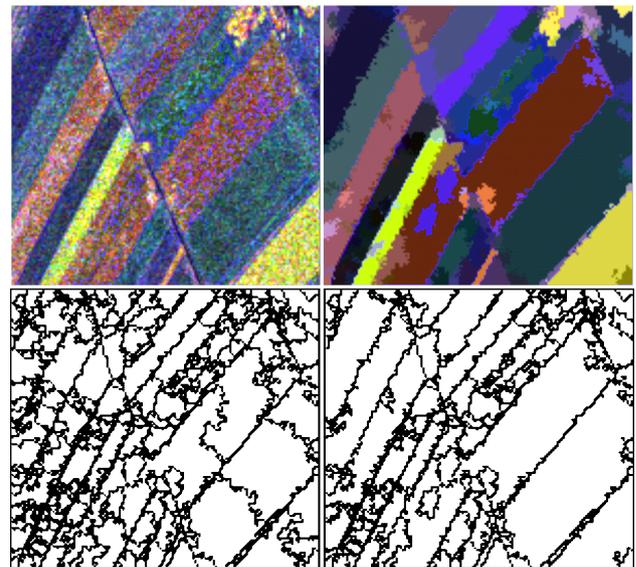


Figure 3: The figure shows the segmentation in more detail. At the top left side the original image data is shown and at the top right side the segmentation result. The bottom line shows the segment borders without (left) and with (right) merging.

#### 6 CONCLUSIONS

The proposed algorithm has several advantages with respect to other segmentation approaches. There is no need for any kind of handmade initialisation. Furthermore, not only a segmentation is obtained, but a hierarchical structure of the image, which contains more information than a simple collection of disjoint regions.

Although the segments obtained will tend to stay compact (contrary to regions gained e.g. by spectral clustering where no spatial information is used) they do not have to be connected. The root of a subtree at a high level covers a large area within the image and can thus connect regions, which are neither close to each other nor connected, but have similar spectral properties. Another important feature is that the number of segments is exclusively based on the given data. Neither an exact number has to be set, nor a maximum number, because the number of segments is a direct result of the algorithm.

As the proposed method was designed to segment homogeneous areas, one of its limits is shown in Figure 4. It is not able to consider a heterogeneous image area like regions with strong texture as one segment. Such a region (e.g. cities or forests with great fluctuations in height and/or changes in backscatter properties due to different vegetation) will be segmented in many small regions. Future work will include the analysis of all regions in a more rigorous way than the above mentioned simple merging algorithm to overcome this disadvantage.

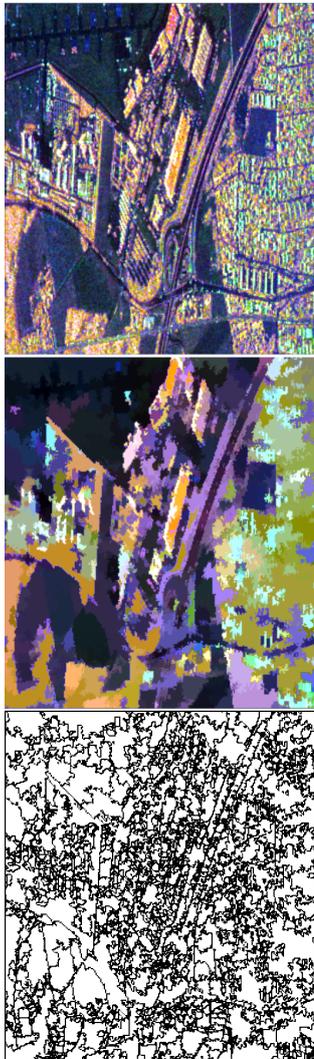


Figure 4: top: original; middle: segmentation; bottom: segment borders

## REFERENCES

- Lee, J.-S. and Grunes, M.R. and Kwok, R., 1994. Classification of multilook polarimetric SAR imagery based on complex Wishart distribution. *Int. J. Remote Sensing*, Volume 15, No.11, pp. 2299–2312.
- Ferro-Famil, L. and Pottier, E. and Lee, J.-S., 2001. Unsupervised classification of multifrequency and fullypolarimetric SAR images based on the H/A/Alpha-Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, Volume 39, Issue 11, pp. 2332–2342.
- Anfinsen, S. N. and Jenssen, R. and Eltoft, T., 2007. Spectral Clustering of Polarimetric SAR Data with Wishart-Derived Distance Measures. *Proc. Intl. Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR2007)*.
- Hänsch, R. and Jäger, M. and Hellwich, O., 2008. Clustering by deterministic annealing and Wishart based distance measures for fully-polarimetric SAR-data. *Proc. EUSAR2008*.
- Reigber, A. and Jäger, M. and Neumann, M. and Ferro-Famil, L., 2007. Polarimetric Fuzzy K-Means Classification with Consideration of Spatial Context. *Proceedings of POLINSAR'07*.
- Hong, T.H. and Rosenfeld, A., 1984. Compact Region Extraction Using Weighted Pixel Linking in a Pyramid. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol.PAMI-6,No.2, pp. 222–229.

