Computers & Geosciences 135 (2020) 104395

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

Superpixel generation for polarimetric SAR using Hierarchical Energy maximization☆

Shuai Yang [a,b], Xiaohui Yuan [b,*,1], Xiuguo Liu [a], Qihao Chen [a]

[a] Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China
[b] Department of Computer Science and Engineering, University of North Texas, Denton TX 76201, United States

ARTICLE INFO

Keywords:
Polarimetric SAR
Superpixel
Hierarchical structure
Speckles
Wishart energy

ABSTRACT

Superpixel provides a coherent division of an image that respects the integrity of objects and has been used in Polarimetric Synthetic Aperture Radar (PolSAR) image analysis. The generation of superpixels is sensitive to noise and cluttered objects, which is typical in SAR imagery. To address this issue, we proposed a novel PolSAR Hierarchical Energy Driven (PHED) method to handle PolSAR images. A hierarchical structure is built and an energy driven hill climbing strategy is employed. In the coarse level superpixel generation, histogram intersections of coherency matrix are used to divide the image into raw superpixels. In the fine level superpixel generation, our method evaluates each superpixel using Wishart energy. A boundary term is included in the pixel-level update to enable boundary choices. Experiments are conducted with UAVSAR, AIRSAR, and simulated datasets. Our results demonstrate that the proposed PHED outperforms Simple Linear Iterative Clustering (SLIC), Linear Spectral Clustering (LSC), TurboPixels (TP), and Superpixels Extracted via Energy-Driven Sampling (SEEDS) in terms of retaining fine boundaries of ground objects and accuracy. The robustness of PHED is also confirmed.

1. Introduction

Polarimetric Synthetic Aperture Radar (PolSAR) has become a popular tool to offer a comprehensive understanding of the Earth surface. However, due to the inner coherent imaging process, PolSAR data suffers from noise distortion, which limits the performance of further pixel-based applications such as decomposition and object recognition (Goodman, 1976; Foucher and Lopez-Martinez, 2014; Lee et al., 2014). Even though speckle filtering is able to reduce speckles to some extent. But the influence of speckles still exists in most pixel-based applications. To address this issue, Superpixel has been adopted in the process of PolSAR image analysis (Qi et al., 2012; Liu et al., 2013; Feng et al., 2014; Hou et al., 2016; Zhou et al., 2016; Pan et al., 2019).

Superpixel is a technique that assigns pixels into meaningful groups (Qin et al., 2015; Yang et al., 2017), it offers a new way to analyze the imagery. Superpixel generation methods are divided into three categories: graph-based methods, region growing, and hierarchical edge refining. A representative method of the graph-based methods is the Ncuts method (Shi and Malik, 2000). This method considers each pixel as a node, and the partition is achieved by globally minimizing a cost function that measures both the dissimilarity between different groups as well as the similarity within each group. However, Ncuts is computationally demanding (Van den Bergh et al., 2015).

The second category is based on region growing such as Watersheds (Vincent and Soille, 1991) and Turbopixels (Levinshtein et al., 2009). A simple and efficient method called Simple Linear Iterative Clustering (SLIC) was proposed (Achanta et al., 2012), which starts from a regular grid of centers or segments, and grows the superpixels by grouping pixels via K-means clustering. This method is easy to be applied in PolSAR by simply replacing the input channels with three polarimetric channels from PauliRGB image (Xu et al., 2016). Alternative distance measures have taken the statistical characteristics of the image into account. For instance, the revised Wishart distance was adopted by Qin et al. (2015), which also improved the cluster center initialization. The pixel intensity ratio was utilized in Xiang et al. (2013) for clustering, which modified the local similarity to refine the K-means strategy. Xiang et al. (2017) used polarimetric distances, textures, and spatial information for the local iterative clustering. This method is able to determine the trade-off automatically with

☆ The contributions of the authors are the following: Shuai Yang devised the method, implemented the code, and conducted experiments; Xiaohui Yuan devised the method, planned the experiments, and wrote the paper; Xiuguo Liu helped shape the idea and revised the paper; Qihao Chen helped shape the idea.

* Corresponding author.
E-mail address: xiaohui.yuan@unt.edu (X. Yuan).

[1] Equal first author.

https://doi.org/10.1016/j.cageo.2019.104395
Received 22 October 2018; Received in revised form 27 June 2019; Accepted 28 November 2019
Available online 30 November 2019
0098-3004/

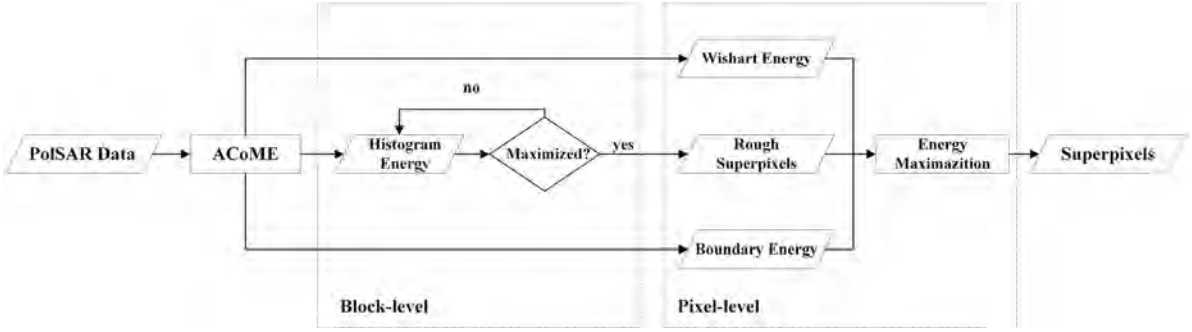

Fig. 1. The flowchart of proposed PHED.

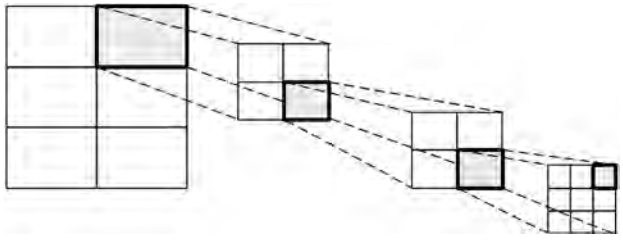

Fig. 2. The construction of hierarchical structure of PHED. The four rectangles, from left to right, are initial superpixels, second level, third level, and lowest level, respectively.

heterogeneity. However, SLIC adopts K-means as the clustering algorithm, which is a lack of stability and dependent on initialization.

The third category is the hierarchical edge refining. Superpixels Extracted via Energy-Driven Sampling (SEEDS) constructs a hierarchical structure to represent an image (Van den Bergh et al., 2015; Stutz et al., 2018). This method adjusts boundary blocks or pixels based on an energy function following a hill-climbing strategy. SEEDS exhibited superior performance in comparison to the state-of-the-art methods (Van den Bergh et al., 2015). It is straightforward to apply SEEDS with PauliRGB images to obtain superpixels in PolSAR. However, SEEDS adopts histogram intersection to calculate the energy, which is not a good choice in pixel-level for PolSAR data due to the presence of speckles. Besides, SEEDS using PauliRGB includes only partially discrete information, which is a partial representation of the PolSAR data.

This paper proposes a novel Polarimetric Hierarchical Energy Driven (PHED) superpixel generation method. A hierarchical structure is built based on the desired number of superpixels. The structure is divided into two parts: the block-level and the pixel-level. For the block-level part, histogram intersection is adopted to adjust the boundary to form a preliminary skeleton segmentation. For each level, assign the boundary sub-level blocks due to the energy maximization, and go to the lower level iteration until the job is complete. At the pixel-level, all the boundary pixels are evaluated as edge refining candidates via a Wishart energy metric combined with boundary term to take both the polarimetric and the spatial information into accounts.

The remainder of this paper is organized as follows. Section 2 gives the background and description of the proposed method in detail. Section 3 discusses our experimental results using simulated and real PolSAR imagery. Section 4 concludes this paper with a summary.

2. PolSAR Hierarchical Energy Driven method

For a reciprocal object illuminated by a monostatic, fully polarimetric SAR system, the polarimetric scattering information is described with a 2×2 complex scattering matrix $\boldsymbol{S}$ (Lee and Pottier, 2009):

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{HH} & \boldsymbol{S}_{HV} \\ \boldsymbol{S}_{HV} & \boldsymbol{S}_{VV} \end{bmatrix}, \tag{1}$$

where each element denotes a complex scattering coefficient. Subscripts H and V denote the horizontal and vertical polarization directions, respectively. In this case, the scattering matrix is symmetric and can be mapped to a 3-D single-look scattering vector based on the complex Pauli spin matrix basis set:

$$\boldsymbol{k} = \frac{1}{\sqrt{2}}[\boldsymbol{S}_{HH} + \boldsymbol{S}_{VV} \quad \boldsymbol{S}_{HH} - \boldsymbol{S}_{VV} \quad 2\boldsymbol{S}_{HV}]^T. \tag{2}$$

For multi-look PolSAR data, the polarimetric information is further represented by an average coherence matrix as follows, which is called Sample Covariance Matrix estimation (SCM):

$$T = \frac{1}{L}\sum_{i=1}^{L} \boldsymbol{k}_i \boldsymbol{k}_i^H, \tag{3}$$

where $\boldsymbol{k}_i^H$ is the conjugate transpose of $\boldsymbol{k}_i$, and L is the number of looks.

The main problem exists in the PolSAR data is the severe influence caused by speckle noises. By grouping homogeneous pixels together into a local region, superpixel generation offers a promising way to eliminate the influence of speckles to some extent. However, the current superpixels generation methods still face some problems. Fig. 1 illustrates the flowchart of our proposed PHED method, which includes three steps: a pre-process that suppresses noise, a block-level update that generates a skeleton segmentation, and a pixel-level update that produces the superpixels.

2.1. Superpixel generation as an energy maximization

Any superpixel result can be considered as a solution to oversegment the image. In most superpixel generation methods, the algorithm starts with an initial cut and then modify it through a metric evaluation. Following this idea, we consider the superpixel segmentation as an energy maximization problem, which is easy to obtain the results via the hill-climbing method.

We represent an image partitioning by referring to the set of pixels in a superpixel, denoted as A_k:

$$A_k = \{p_1, p_2, \ldots, p_n\}, \tag{4}$$

where p_i represents pixel which belongs to superpixel k. In this case, the whole partitioning of an image is represented with a set of superpixels, denoted with s, $s = \{A_1, A_2, \ldots, A_M\}$. Due to the fact that a pixel can only be assigned to a unique superpixel, all sets A_k are restricted to be disjoint, and thus, the intersection between any pair of superpixels is always the empty set.

Each solution can be evaluated by a so-called energy function, hence, the superpixel task is transformed into finding a solution to maximize the energy function, i.e.,

$$s^* = \arg\max E(s), \tag{5}$$

where s^* denotes the optimized solution and $E(s)$ is the energy function of solution s.

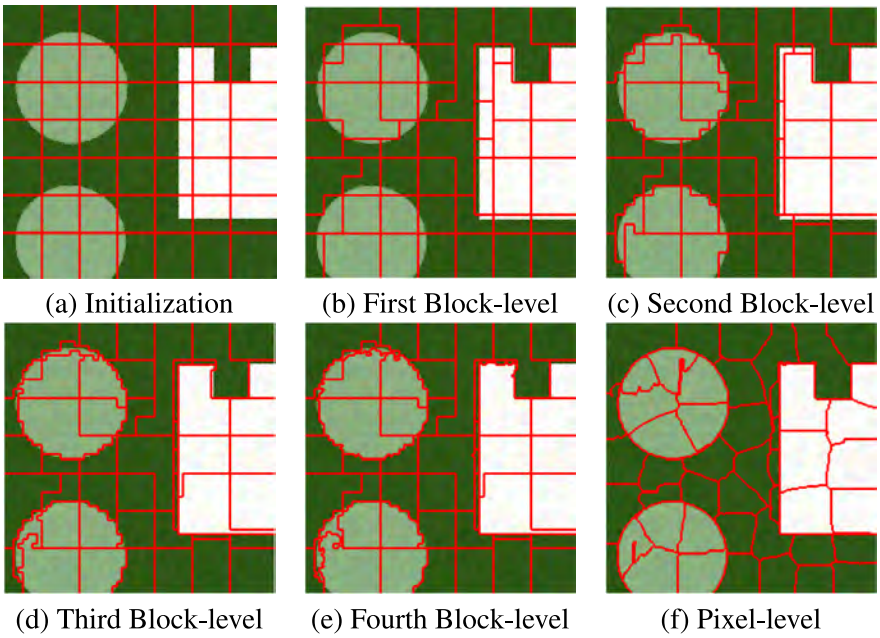

Fig. 3. Initialization, block-level, and pixel-level updates.

2.2. Hierarchical structure

To perform optimization effectively, hill-climbing (Russell and Peter, 2016) is adopted. Hill-climbing is a well-known algorithm that iteratively updates the solution by proposing small local changes at each iteration. Once the energy increases, the proposed partitioning is adopted to replace the current one. The whole optimization is achieved when all the energies from randomly proposed solution are less than the current one or meet the maximum iterations.

However, a potential matter is that a locally optimal solution may be obtained as a local region is used in the traditional hill-climbing algorithm. To avoid such a problem, a hierarchical structure is incorporated, which is also able to enhance efficiency. To build such hierarchical structure, the PolSAR image is divided into small pieces to form the lowest level, which is in size of $m \times n, (m, n \in \{2, 3, 4\})$. This level is considered as the smallest blocks level in the hierarchical structure, then the higher level is built by concatenating lower level blocks in a 2×2 fashion. The largest blocks are concatenated 2×2 to create the initial superpixels, as shown in Fig. 2. The desired number of superpixels is obtained by choosing the initial block size and number of block levels accordingly.

With the help of this hierarchical structure, optimization performs efficiently. It starts with the highest block update, which is a quarter of the target superpixel size. The algorithm moves to the lower level and performs the iterations. The lowest level is the pixel level.

Based on the hierarchical structure, the assignments are divided into two parts: the block-level and pixel-level updates. A diagram of hierarchical structure based superpixel generation is illustrated in Fig. 3. Fig. 3(a) depicts the grid initialization of the input image. The block-level update enables a large amount of pixels assignment during each iteration, and helps to build a skeleton for the superpixels as shown in Fig. 3(b). The algorithm keeps going down and performing updates until the smallest block-level update is done, as depicted in Fig. 3(c)–(e). The pixel-level update is performed to fit boundaries, as shown in Fig. 3(f).

2.2.1. Block-level update

The main objective of the block-level update is to form a valid and close to boundary solution rapidly, which means accuracy is the second priority comparing to the speed for the block-level. There are two major concerns regarding the block-update: the input feature extraction and the corresponding energy function.

For a nine-channel (six real and three imagery parts) polarimetric coherency matrix, l channels ($0 < l \leqslant 9$) are extracted and normalized as the input of the block-level update. Each channel is scaled as follows

$$b_n = 10 \log_{10} \boldsymbol{T}_i, \tag{6}$$

where n denotes the index of the block-level input channels ($n \in \{1, \ldots, 9\}$) and i denotes the index of polarimetric coherency matrix channels. It is straightforward that less channels will speed up the algorithm yet more channels result in fine performance. When l is nine, all the information contained in the PolSAR data are utilized. When l is three and $i = \{1, 6, 9\}$, the input is equivalent to PauliRGB of the Polarimetric SAR data except that the latter is discrete.

Another issue is to determine the energy function for the block-level update. The essence of the energy function is to score the current solution, which means it should be able to represent all the pixels included in the selected blocks and be sensitive to smaller block movements. We denote the energy function in the block-level update as

$$E(s) = H(s), \tag{7}$$

where $H(s)$ denotes the similarity between two blocks. A simple and efficient approach is histogram intersection, denoted as $\mathbb{H}$:

$$\mathbb{H}(c_{A_a}, c_{A_b}) = \sum_j \min\{c_{A_a}(j), c_{A_b}(j)\}, \tag{8}$$

where $c_{(\cdot)}$ denotes the histogram of a group of pixels.

Given the similarity measure, it is easy to evaluate the current assignment by validating if $H(s)+\tau > H(s_t)$. The τ is tolerance to reduce the influence of speckles. We have

$$\mathbb{H}(c_{A_a}, c_{A_b^l}) \geqslant \mathbb{H}(c_{A_b} \setminus c_{A_b^l}, c_{A_b^l}) \;\Leftrightarrow\; H(s) + \tau \geqslant H(s_t), \tag{9}$$

where A_b^l is the set of pixels that are candidates to be assigned from the superpixel A_b to A_a.

During each iteration, a boundary block is picked as candidate A_b^l, and the energy is calculated. We search the adjacent superpixels of the block candidate to find a potential solution. For each new solution, the energy is calculated and compared with the current one. If the energy increases, the proposed solution replaces the current one. A tolerance is added to the energy to reduce errors caused by speckle noises.

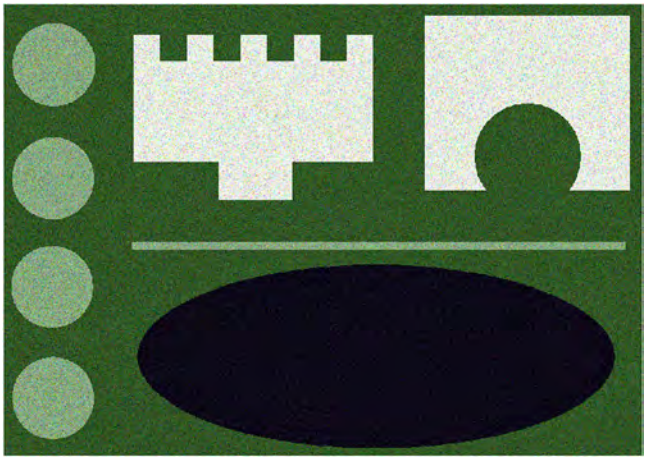

Fig. 4. Simulated PolSAR data #1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2.2. Pixel-level update

After block-level update, only boundary pixel movements are required from the block-level result. The energy function used in pixel-level update is as follows:

$$E(s) = (1-\gamma)E_W(s) + \gamma E_B(s), \tag{10}$$

where $E_W(s)$ represents the polarimetric term and $E_B(s)$ denotes the boundary term, γ is weight parameter. The range of $E_B(s)$ and $E_W(s)$ is $[0, 1]$.

The boundary term E_B evaluates the shape of a superpixel, which penalizes local irregularity. It can be replaced with other metrics to favor either compactness or other specific shapes. We denote the boundary term as follows:

$$E_B = \frac{\sqrt{({d_x}^2 + {d_y}^2)}}{N_B}. \tag{11}$$

Note that d_x and d_y are the horizontal and vertical distances between a superpixel and the candidate pixel, respectively. N_B is the normalization term used to scale the boundary term to the range of $[0, 1]$.

The choices of E_W is plentiful, any metric that reflects the similarity between two coherency matrices is feasible, for instance, the polarimetric similarity parameter (An et al., 2009) and polarimetric distribution distances (Lee and Pottier, 2009). We adopt a metric derived from symmetric revised Wishart distance D_{SRW} (Anfinsen et al., 2007):

$$D_{SRW} = \frac{1}{2}(Tr(\bar{\boldsymbol{T}}_j^{-1}\bar{\boldsymbol{T}}_i) + Tr(\bar{\boldsymbol{T}}_i^{-1}\bar{\boldsymbol{T}}_j)) - d. \tag{12}$$

Note that T_i and T_j are two coherency/covariance matrices, d is the dimension of coherency matrix, i.e., 3 for monostatic case and 4 for bistatic case. This non-negative distance equals to 0 when T_i and T_j are identical. Higher value of D_{SRW} indicates less similarity between T_i and T_j. Hence, the polarimetric similarity term is computed as follows:

$$E_W = \frac{\arctan(D_{SRW})}{\pi/2}. \tag{13}$$

2.3. Post-processing

Due to the mechanism of the hierarchical structure, the numbers of superpixels remains the same after the pixel-level update. This fact may cause some problems, for instance, very small superpixels. To avoid such problems, our method includes an additional reassignment of small superpixels. The idea is to identify the superpixels that are smaller than a given threshold and calculate the energy between these small superpixels and their adjacent superpixels. The merge of superpixels is performed to the ones with the highest energy. The threshold is calculated as follows:

$$Th = \lambda\frac{N_{image}}{N_{sp}}, \tag{14}$$

where N_{image} and N_{sp} are numbers of image pixels and superpixels, respectively. λ is the weight, which is adjustable and set to 0.1 in our experiments.

3. Experiments and discussion

We use several polarimetric datasets to validate the performance of the proposed method including UAVSAR, AIRSAR, and two simulated datasets. Four popular superpixel generation methods, SLIC, LSC(Linear Spectral Clustering) (Li and Chen, 2015), TP(TurboPixels) (Levinshtein et al., 2009), and SEEDS are used for comparison. All data used in this section are pre-processed via ACoME(Adaptive Coherency Matrix Estimation) to suppress speckles (Yang et al., 2016).

To evaluate the results quantitatively, we adopt two metrics commonly used for evaluating the quality of superpixels. We use $g = \{g_1, g_2, \ldots, g_i\}$ to denote a ground truth segmentation with i segments, $s = \{s_1, s_2, \ldots, s_j\}$ to represent corresponding superpixels with j segments, and $|\cdot|$ denotes the segment size.

The first metric is Achievable Segmentation Accuracy (ASA), which is a performance upper bound measure. ASA offers an achievable highest accuracy when taking superpixels as units for object segmentation. All the superpixels are labeled with the label of ground truth segment that has the largest overlap. The fraction of correctly labeled pixels out of the whole image is the computed ASA, as shown in Eq. (15). A higher ASA indicates better performance.

$$ASA(s) = \frac{\sum_j \max_i |s_j \cap g_i|}{\sum_i |g_i|}. \tag{15}$$

In addition to accuracy, the adherence of boundaries is also a valuable way to evaluate the results quantitatively. The Boundary Recall (BR) is formulated as

$$BR(s) = \frac{\sum_{p\in\mathfrak{B}_g} \mathbb{I}[(\min_{q\in\mathfrak{B}_s} \|p-q\|) < \epsilon]}{|\mathfrak{B}_g|}, \tag{16}$$

where $\mathfrak{B}_g$ and $\mathfrak{B}_s$ are union sets of ground truth boundaries and the computed superpixels boundaries, respectively. The indicator function $\mathbb{I}[\cdot]$ checks if a boundary pixel of superpixel is within a tolerance distance, ϵ, of the ground truth boundaries. Higher BR indicates better boundary adherence.

Besides quantitative metrics, we also use the boundary map and mean image map to evaluate the superpixel generation qualitatively. Boundary map is to mark boundaries of superpixels in red color based on original PauliRGB data, which is an intuitive way to show the results. The mean image map is calculated based on the averaged color of each superpixel of PauliRGB data. All superpixels are colored with averaged colors instead of original ones. Hence the difference between PauliRGB data and mean image can be used to evaluate the quality of superpixel intuitively.

3.1. Evaluation of wishart energy

We simulated a PolSAR data (note as simulated data #1) of a size of 515 × 728 (Lee et al., 1994). The image consists of four types of ground objects: pond (black), vegetation (dark green), forest (light green), and buildings (white), as shown in Fig. 4(a). The pond is simulated with Wishart distribution, while the others are simulated with K distribution. The structure of the simulated data is meticulously designed to contain different shapes including rectangle, circle, thin bar, sawtooth, and ellipse. Hence, it is explicit to evaluate the ability to preserve boundaries of superpixel generation methods using this

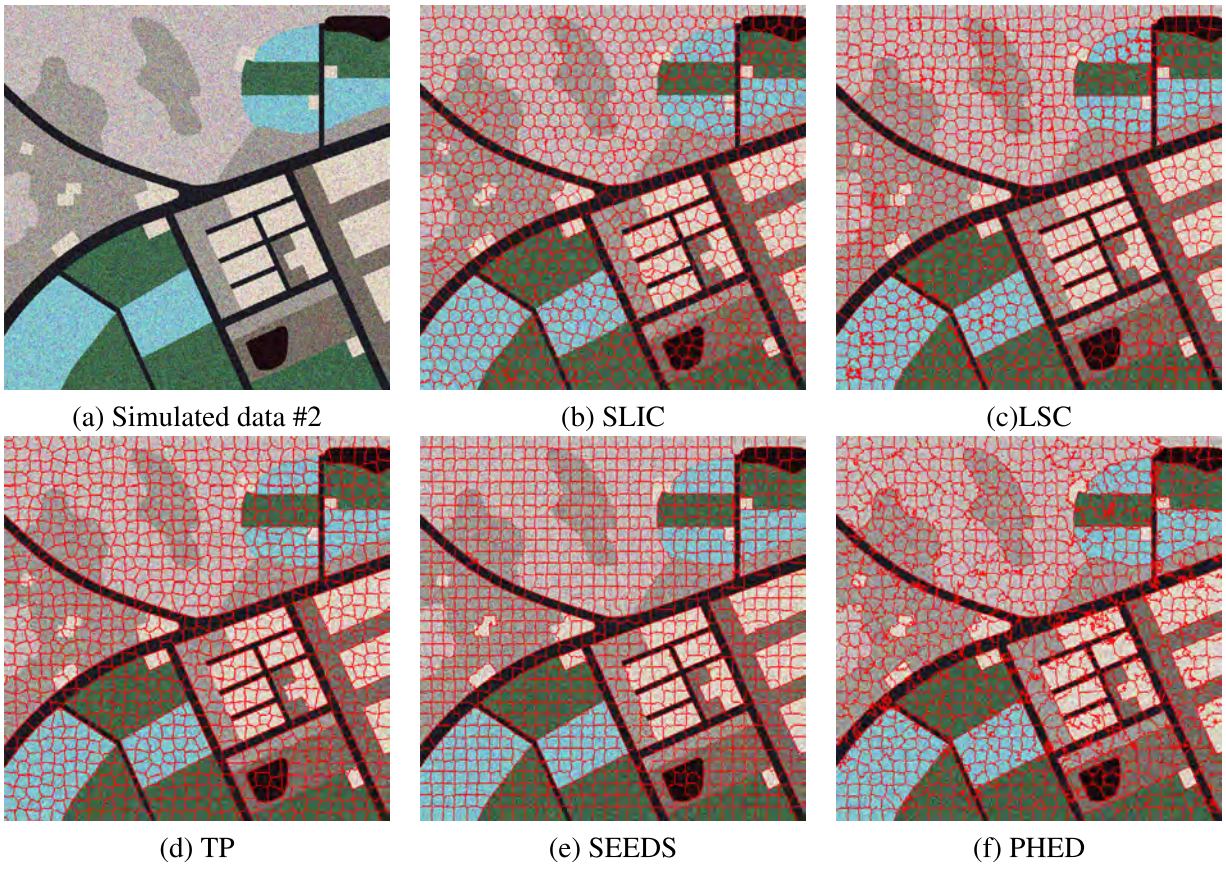

Fig. 5. Simulated data #2 PauliRGB image and superpixel results of different methods.

Table 1
Wishart energy between pixels and superpixels from different ground categories.

	sp_1	sp_2	sp_3	sp_4	Diff
p_1	**0.9474**	0.0060	0.0548	0.0704	0.8770
p_2	0.0051	**0.9655**	0.0006	0.0497	0.9158
p_3	0.0594	0.0007	**0.9276**	0.0061	0.8682
p_4	0.0661	0.0593	0.0060	**0.9677**	0.9084

dataset, especially the ability to handle different kinds of shapes of ground objects.

Recall that high energy implies a high similarity, which shows a better solution to be updated. We choose 10,000 pixels for each of the four categories from the simulated data #1 as the superpixels ($sp_i,\ i \in \{1,2,3,4\}$), and pick one pixel randomly ($p_j,\ j \in \{1,2,3,4\}$) as the candidate and calculate the Wishart energy. The four categories represent forests, pond, buildings, and grassland. The Wishart energy is shown in Table 1, which presents the average of 1000 repetitions. The term in row j, column i represents the mean energy between p_j and sp_i, where i, j are the category indices. As shown in this table, only when $i = j$, i.e., self-energy, high energy is achieved (in bold font). This indicates that Wishart energy characterizes the similarity very well. The difference between the self-energy and the second highest term (underlined) is significant, which indicates that Wishart energy is able to differentiate categories with success.

3.2. Evaluation with simulated data

PolSAR data simulation is a good way to evaluate the methods with known parameters, i.e., the accurate boundary in our case. To better compare the superpixel generation methods, a simulated dataset #2 of 800 × 800 is generated, which contains more complicated types and shapes. In general, eight different classes, including road, water body, grass-land, different types of crops, bushland, forest, and urban areas are simulated based on real-world data. Waterbody was generated by the Wishart distribution and the others were by the K distribution, as shown in Fig. 5(a).

Fig. 5(b)–(f) list the superpixel results of SLIC, LSC, TP, SEEDS, and PHED. All five methods obtained fine results in general, but the details still need further investigation. Fig. 6 shows the zoom-in view comparison. In the first two rows, all the five methods recognize the black road in the lower-left corner, but only PHED can adhere to the boundaries of gray crops. The last two rows of Fig. 6 show the comparison of urban areas, LSC, SEEDS, and PHED obtained the fine structure of roads, which proves their efficiency of preserving the linear structure. The blue box in Fig depicts a confusing area consist of two types of crops, it is clear that only PHED can differentiate these two types, the other four methods failed in the area due to the lack of polarimetric information usage.

Due to the serve influence by speckle noise, the threshold of BR is raised to offer a loose comparison (thresholds are to 3 and 4) in Table 2. It is obvious that TP, SEEDS, and PHED obtained accuracy higher than 95%, which shows their advantages in segmentation accuracy. However, the boundary recall of TP and SEEDS decrease rapidly due to the lack of polarimetric information usage. PHED achieved the best performance in terms of segmentation overlap and boundary adherence.

3.3. Evaluation with real-world data

The first real-world dataset we used is single look L-band ESAR data of Oberpfaffenhofen, Germany with a spatial resolution of 1.5 m × 0.89 m (range × azimuth), as shown in Fig. 7(a), the size of the image is 800 × 800. The number of superpixels is 1600. Fig. 7(b)–(f)

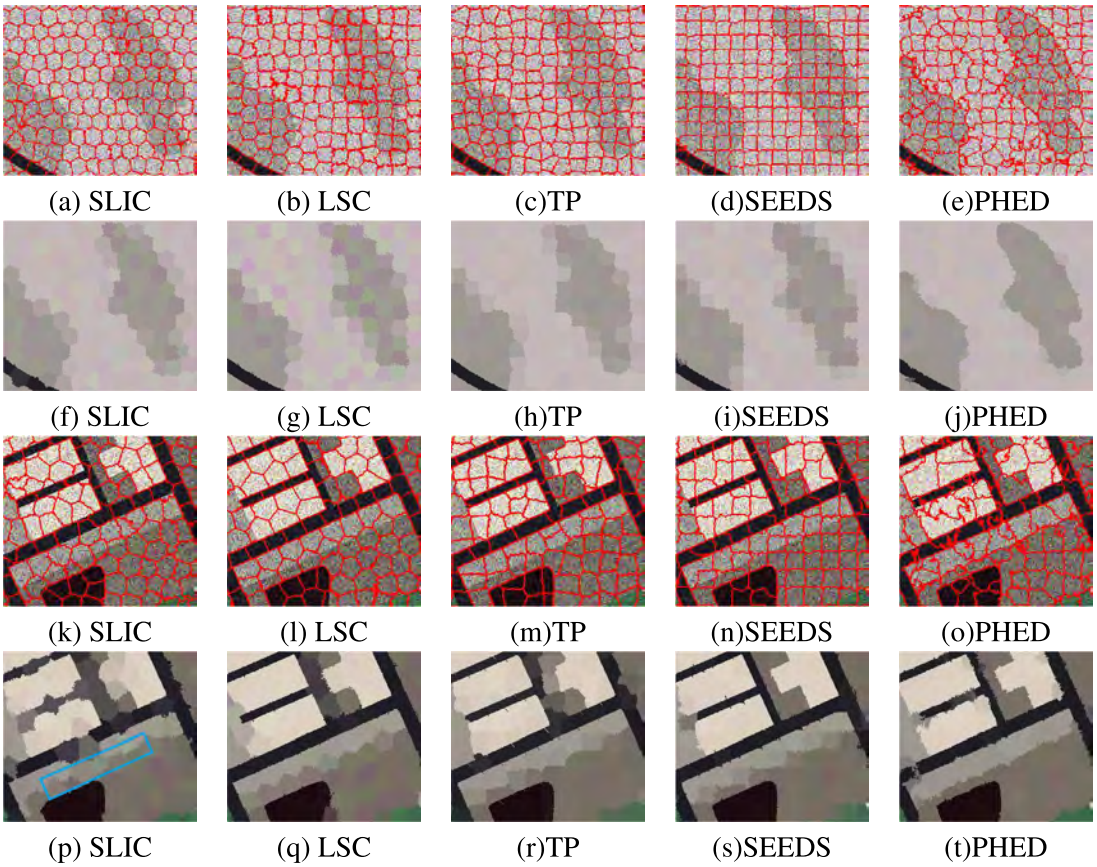

Fig. 6. Detailed comparison of Simulated data #2.

Table 2
Quantitative comparison results of the simulated data #2 using different methods.

	SLIC	LSC	TP	SEEDS	PHED
ASA	93.284%	87.635%	95.111%	96.586%	97.540%
BR3	80.465%	60.095%	86.718%	92.260%	98.929%
BR4	85.689%	70.973%	90.481%	94.434%	99.027%

Table 3
Quantitative comparison results of ESAR data using different methods.

	SLIC	LSC	TP	SEEDS	PHED
ASA	86.318%	82.297%	86.168%	87.928%	88.940%
BR2	69.769%	61.797%	67.683%	82.648%	83.515%
BR3	81.583%	75.675%	80.547%	91.340%	92.465%

Table 4
Quantitative comparison results of AIRSAR data using different methods.

	SLIC	LSC	TP	SEEDS	PHED
ASA	93.548%	89.908%	92.657%	94.572%	95.703%
BR2	79.072%	72.618%	73.659%	88.159%	90.430%
BR3	93.378%	92.136%	93.034%	96.031%	98.290%

demonstrate the superpixel results generated via SLIC, LSC, TP, SEEDS, and PHED. It is clear that all the methods are able to obtain rational superpixel results, which are close to the boundaries.

Fig. 8 illustrates the zoom-in views of two characteristic targets: buildings surrounding by grassland and farmland. All five methods are able to identify the shape of buildings. However, SLIC, LSC, and TP failed to adhere the boundaries, while SEEDS and PHED obtained fine results, except the bottom part of SEEDS, as demonstrated in the blue box in Fig. 8(d). In the second row of Fig. 8. The right pink object is diffused with the surroundings on the left boundary of the river for SLIC, LSC, TP, and SEEDS. The most integrate results are obtained by PHED.

Table 3 lists ASA and BR results of the five methods. For accuracy, PHED achieves the best performance, which is 3.04%, 8.07%, 3.22%, and 1.15% higher compared with SLIC, LSC, TP, and SEEDS, respectively. As for the boundary recall, PHED achieves 92.465%, which outperforms the rest four methods when tolerance is 2, where the LSC only obtains 61.797%. When we set the tolerance factor to 3, the recall is improved. PHED consistently achieves the best result.

Scattering information is important in PolSAR data representation, the ESAR experiments reconfirm the efficiency of PHED's utilization of scattering information. However, it is valuable to discuss the performance of PHED regarding weak contrast condition of scattering characteristics, i.e., using another dataset with similar scattering. Another experiment is carried out with L-band AIRSAR data over Flevoland, The Netherlands in August 1989. The size of this data is 380 × 424 pixels with a resolution of 8 m × 8 m (range × azimuth). The number of superpixels is 1000. The image and corresponding superpixel results are demonstrated in Fig. 9. The whole image consists of several types of crops, which means most of the image is dominated by surface scattering. It is clear that all the methods are able to generate fine superpixels, which indicates that they are all able to identify the scattering characteristics with success. Fig. 10 shows a detailed comparison of five methods. There is a T-shaped dark object in the area, it is obvious that PHED obtained the most integrate results compared to the rest four methods.

Table 4 lists the accuracy and boundary recall rate comparison results. PHED resulted in the highest ASA as well as the highest BR, which reconfirms the proficiency. The accuracy of the rest shows that these methods are able to obtain superpixels with fine results, but all of them are lack of the ability to handle the boundaries with success.

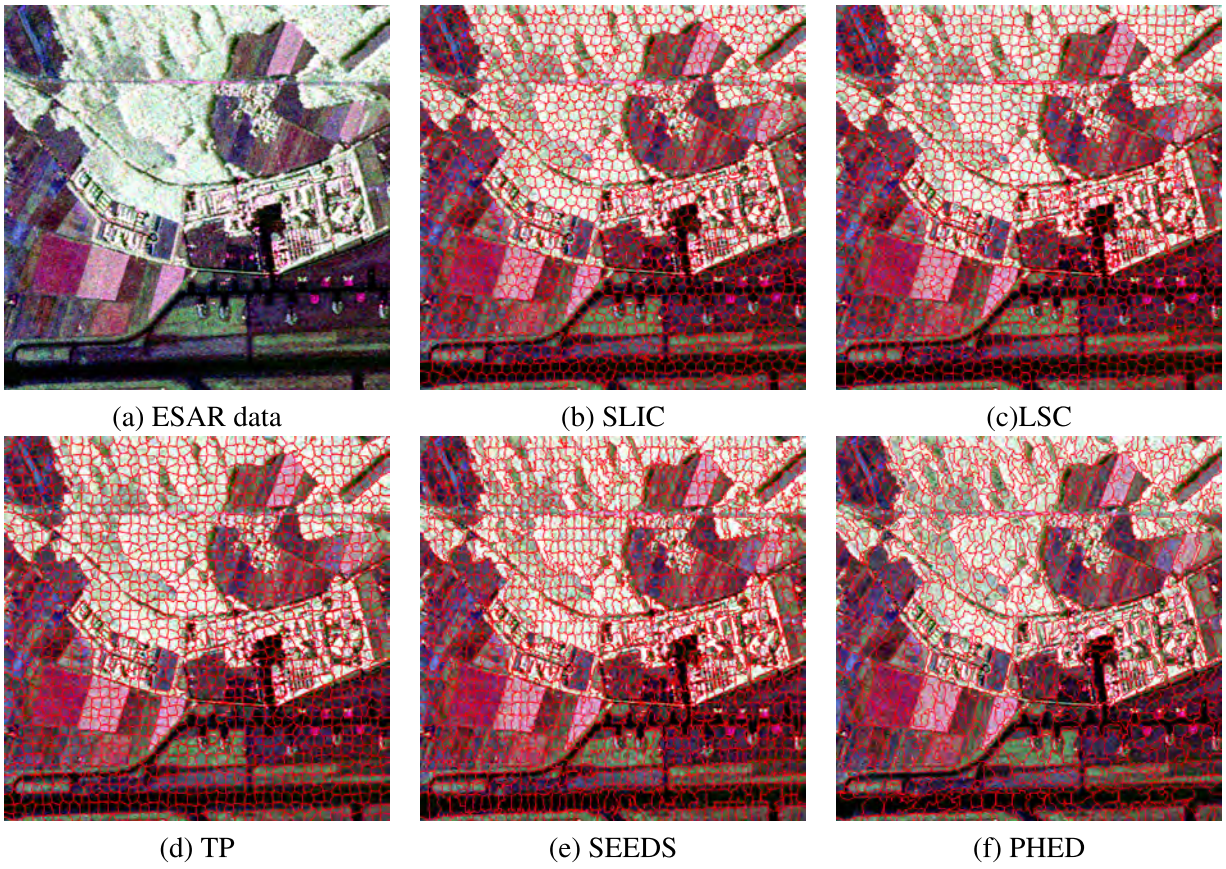

Fig. 7. ESAR PauliRGB image and superpixel results of different methods.

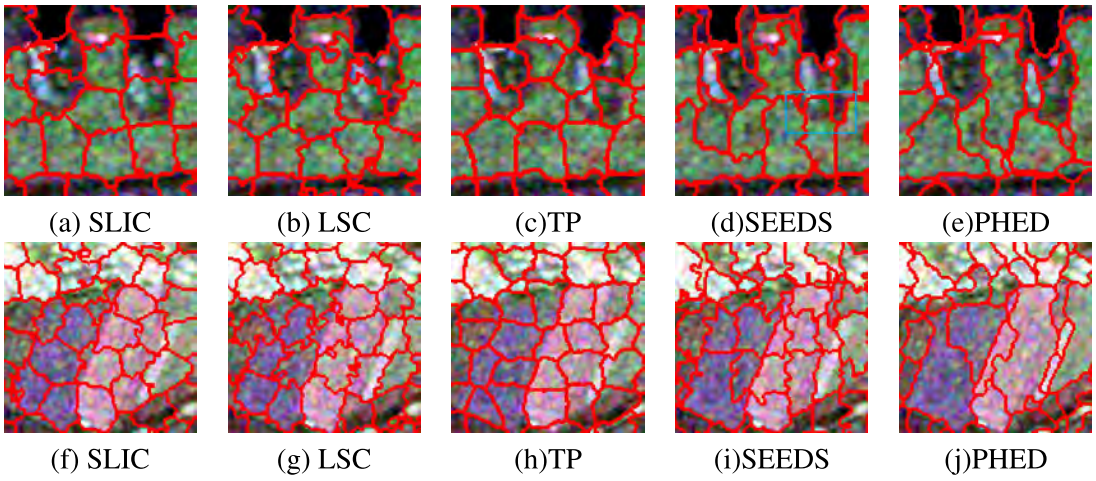

Fig. 8. Detailed comparison of ESAR data.

3.4. Sensitivity and uncertainty

This section evaluates the proposed method via time efficiency and sensitivity. Table 5 lists the running time comparison of ESAR data. All the methods were implemented using C++, and run on a laptop with a 64-bit Windows 10 operating system, a quad-core Intel Core i5 8250U CPU, 1.60 GHz and 16 GB memory. It is straightforward that LSC is the fastest method compared to others. TP, SEEDS, and SLIC are the second, third, and fourth place, respectively. PHED run for 115.67 s with ESAR data, which is the slowest method in this experiment since it contains the computation of Wishart energy. However, LSC, TP, and SLIC have disadvantages regarding boundary adherence and overlap accuracy. At this point, accuracy is the priority when compares to time efficiency. SEEDS performs better than PHED in running time. However, SEEDS is inferior to PHED in superpixel segmentation. PHED achieved the best performance in two minutes, which is competitive in terms of time complexity. In addition, computational time can be reduced via parallel computing.

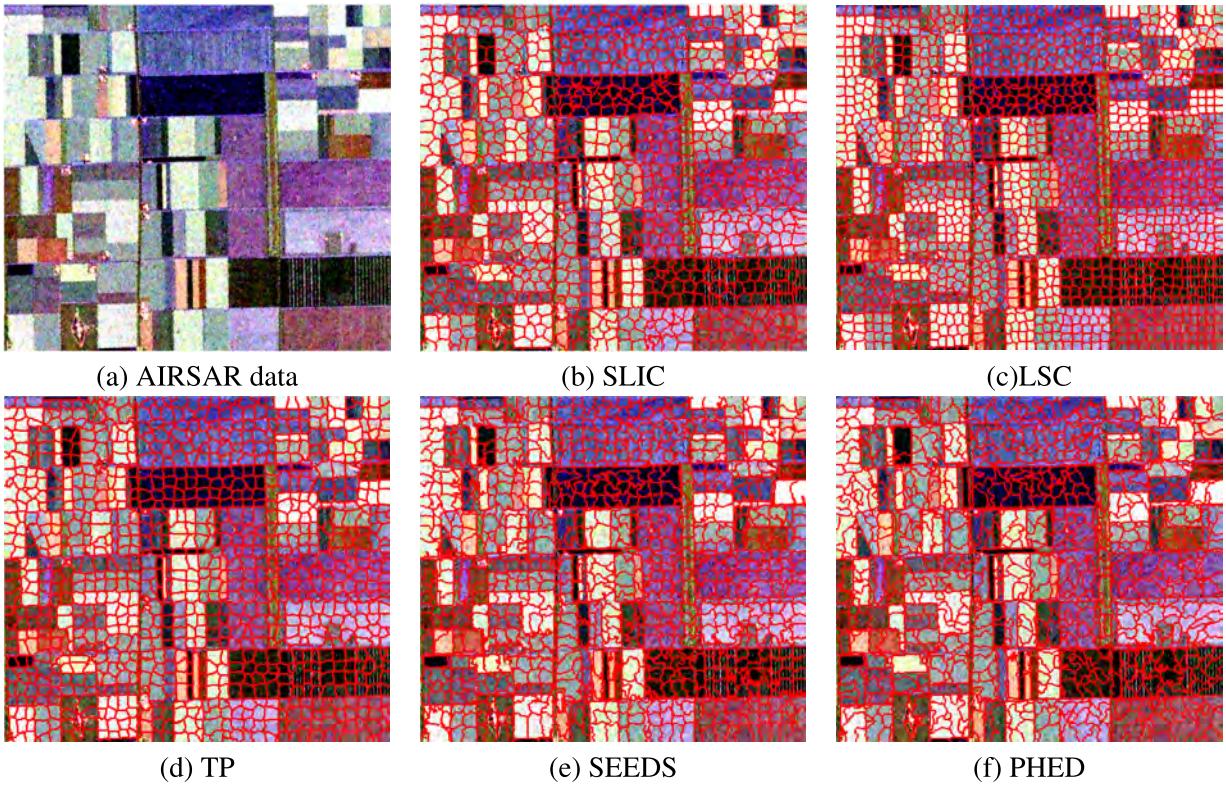

Fig. 9. AIRSAR PauliRGB image and superpixel results of different methods.

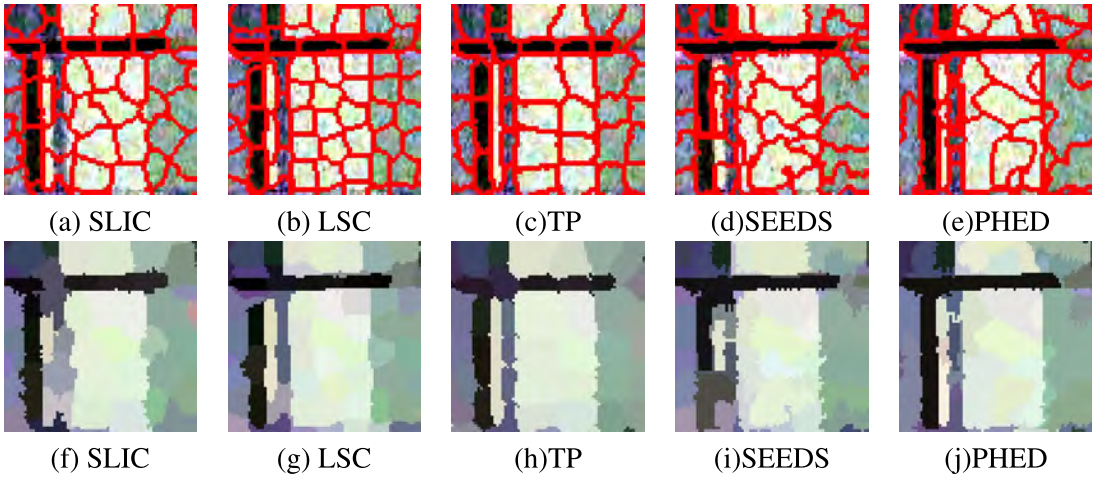

Fig. 10. Detailed comparison of AIRSAR data.

Table 5
Running time comparison of superpixels.

	SLIC	LSC	TP	SEEDS	PHED
Mean	93.67 s	4.33 s	25.33 s	54.67 s	115.67 s
Std	5.51 s	1.53 s	2.52 s	2.08 s	1.53 s

As a parameter based method, it is necessary to discuss the influence of changing parameters on the superpixel generation. In this section, several important parameters of PHED, including the desired number of superpixels, the boundary factor γ, and the minimum confidence τ used for block updates based on simulated PolSAR data #1.

Additional to Fig. 4, We modified the numbers of superpixels for all three methods to figure out the robustness. Recall that 1000 was adopted for the simulated data. We set numbers of superpixels as 100, 50, and 20, which are extremely low choices, and repeat this experiment. The results are shown in Fig. 11. The SLIC failed to retain boundaries of circles and thin bar object in all three cases, while the building with sawtooth is well retained. SEEDS is able to obtain accurate results, but when the number decreases, it also failed to retain boundaries of circles and thin bar. PHED obtains the finest results among the three methods. All the boundaries are well retained except the right part of the white building, as shown in Fig. 11(c). Hence it is clear that PHED can handle extreme cases with low desired numbers of superpixels with success.

However, the performances with different desired numbers is still a remaining question. Fig. 12 depicts the ASA and BR of three methods with the desired numbers of superpixels ranging from 100 to 1500. It is clear that higher accuracy is achieved when the number of superpixels

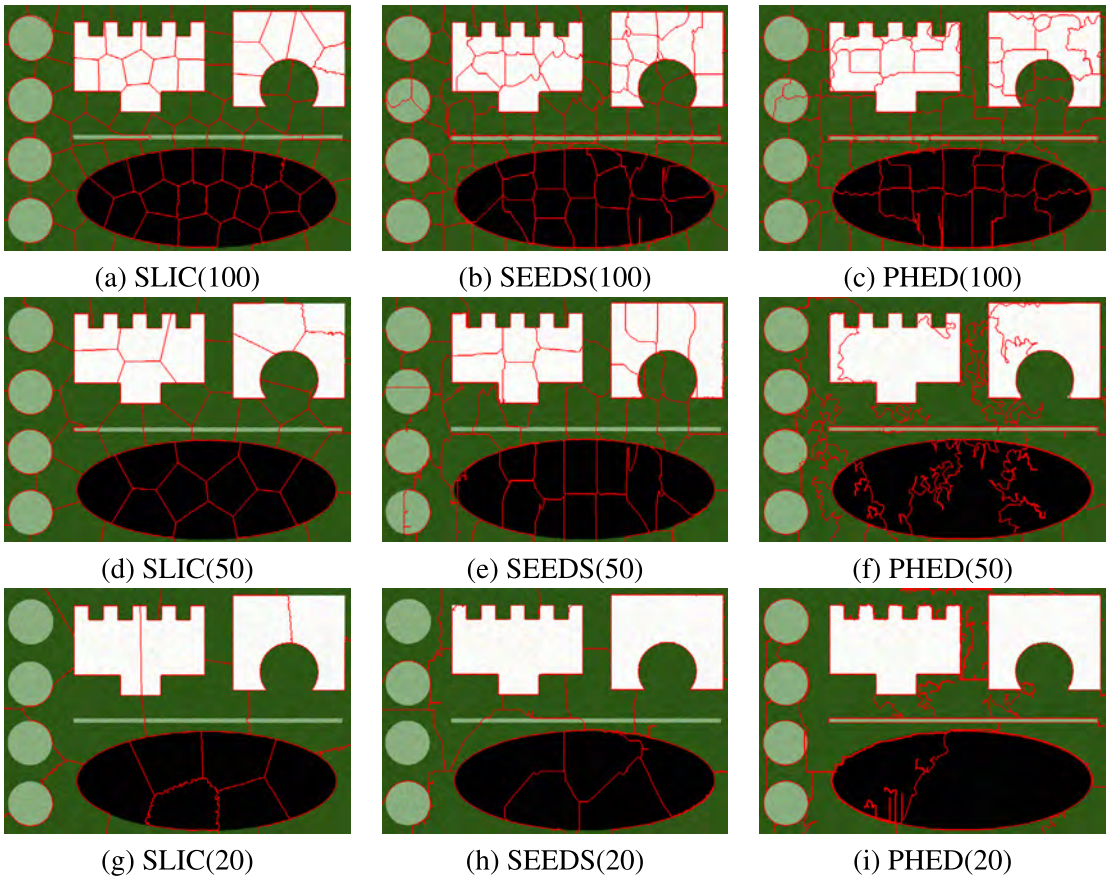

Fig. 11. Experimental results using the simulated data with low numbers of superpixels.

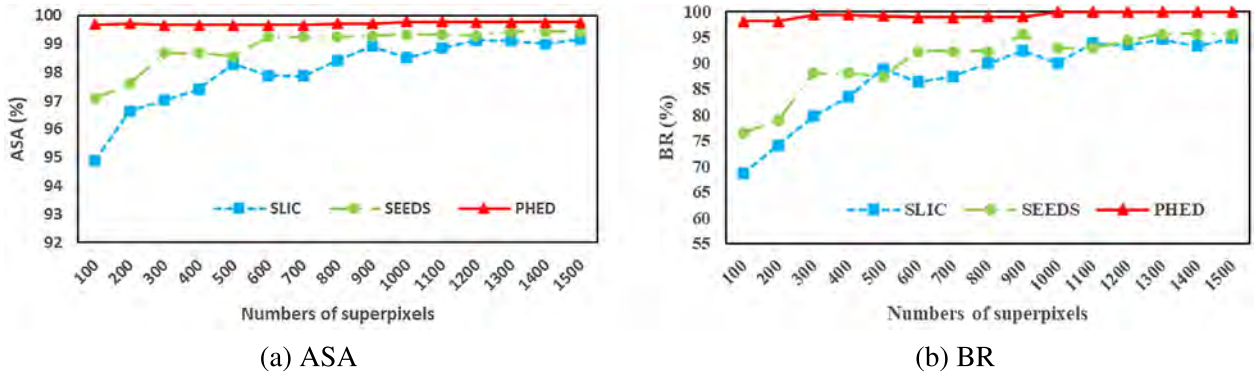

Fig. 12. The ASA and BR of the three methods with respect to the number of superpixels.

increases for both SLIC and SEEDS. Nevertheless, the accuracy of PHED remains high regardless of the number of superpixels. For traditional PolSAR superpixel generation methods, the accuracy is severely influenced by the desired numbers of superpixels. PHED outperforms the other two methods especially for the low number of superpixels. It is evident that PHED is a better method that ensures accuracy, robustness, and adaptability. Results of BR ($\epsilon = 1$) as shown in Fig. 12(b) reconfirms this conclusion.

Fig. 13 demonstrates the influence of boundary factor and minimum confidence on PHED. It is clear that the impact of boundary factor is low for both the accuracy and boundary recall. As for the minimum confidence, results of Fig. 13(c) indicates that ASA is also stable when changing the minimum confidence. However, a relatively low BR is achieved when the minimum confidence is 0.3. The decrease of BR reveals less effective of PHED when we set high minimum confidence. Nevertheless, it still achieves a BR higher than 97.5%, which is a fine result. Based on our experimental results, we conclude that PHED is a robust superpixel generation method.

4. Conclusion

Superpixel has been applied in many fields in PolSAR data processing, and the accurate extraction of superpixel has drawn lots of attention of researchers. The inner speckles severely influence the superpixel generation, which makes it still a challenging task. In this paper, we raised a PolSAR Hierarchical Energy Driven method inspired by the Superpixels Extracted via Energy-Driven Sampling, which fully utilizes all the information included in the PolSAR data. This method applies histogram intersections during the block-level update and introduces a novel Wishart energy as the main metric for the pixel-level

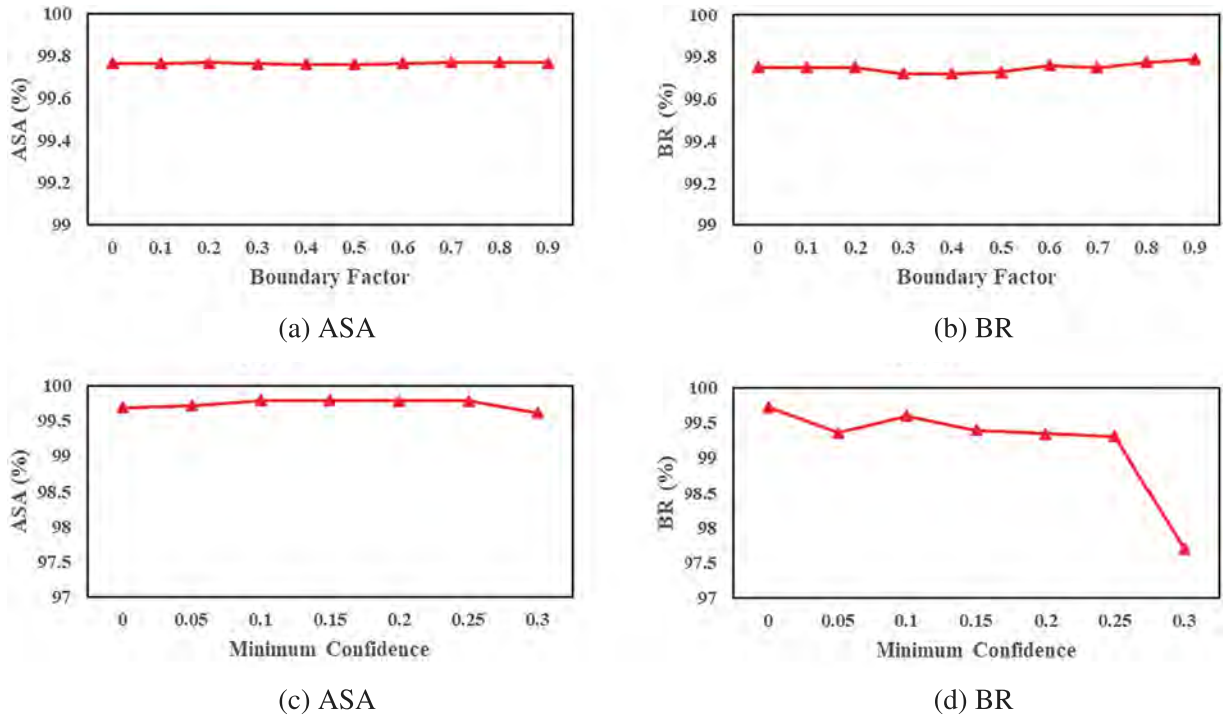

Fig. 13. Comparisons of PHED with different boundary factor and minimum confidence. (a) The ASA of PHED with respect to boundary factor; (b) The BR of PHED with respect to boundary factor; (c) The ASA of PHED with respect to minimum confidence; (d) The BR of PHED with respect to minimum confidence.

update. The optimization is achieved by a hill-climbing procedure, which maximizes the energy. The potential local optimization is also effectively avoided by the hierarchical structure.

The Wishart energy characterizes the similarity of ground objects and able to differentiate categories with success. Further experiments are carried out with simulated data, ESAR, and AIRSAR data. The comparison confirms the effectiveness of PHED when compared to SLIC, LSC, TP, and SEEDS. The robustness of PHED is also confirmed by the success of handling an extremely low number of superpixels.

The sensitivity and uncertainty are evaluated and PHED exhibits less sensitivity to noise and shape of the ground objects compared to SLIC, LSC, TP, and SEEDS. It is clear that the impact of boundary factor is low for both the accuracy and boundary recall. The algorithm offers a flexible framework that is easy to enforce constraints such as compactness and smoothness. The polarimetric term can be replaced with another distance based energy or matrix similarity.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China under Grant No. 41471355 and 41771467. In addition, Shuai Yang is supported by the China Scholarship Council Grant No. 201706410044 for his visit to the University of North Texas.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Sustrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34 (11), 2274–2282.

An, W.T., Zhang, W.J., Yang, J., Hong, W., Cao, F., 2009. On the similarity parameter between two targets for the case of multi-look polarimetric SAR. Chin. J. Electron. 18 (3), 545–550.

Anfinsen, S.N., Jenssen, R., Eltoft, T., 2007. Spectral clustering of polarimetric SAR data with the wishart-derived distance measures. In: Proc. POLinSAR, pp. 1–9.

Feng, J.L., Cao, Z.J., Pi, Y.M., 2014. Polarimetric contextual classification of PolSAR images using sparse representation and superpixels. Remote Sens. 6 (8), 7158–7181.

Foucher, S., Lopez-Martinez, C., 2014. Analysis, evaluation, and comparison of polarimetric SAR speckle filtering techniques. IEEE Trans. Geosci. Remote Sens. 23 (4), 1751–1764.

Goodman, J.W., 1976. Some fundamental properties of speckle. J. Opt. Soc. Amer. 66 (11), 1145–1150.

Hou, B., Kou, H.D., Jiao, L.C., 2016. Classification of polarimetric SAR images using multilayer autoencoders and superpixels. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (7), 3072–3081.

Lee, J.S., Ainsworth, T.L., Wang, Y.T., 2014. Generalized polarimetric model-based decompositions using incoherent scattering models. IEEE Trans. Geosci. Remote Sens. 52 (5), 2474–2491.

Lee, J.S., Pottier, E., 2009. Polarimetric Radar Imaging: From Basics to Applications, vol. 142. CRC Press, Hoboken.

Lee, J.S., Schuler, D.L., Lang, R.H., Ranson, K.J., 1994. K-distribution for multi-look processed polarimetric SAR imagery. In: IGARSS '94 - 1994 International Geoscience and Remote Sensing Symposium Volumes 1-4, vol. 4, pp. 2179–2181.

Levinshtein, A., Stere, A., Kutulakos, K.N., Fleet, D.J., Dickinson, S.J., Siddiqi, K., 2009. Turbopixels: Fast superpixels using geometric flows. IEEE Trans. Pattern Anal. Mach. Intell. 31 (12), 2290–2297.

Li, Z.Q., Chen, J.S., 2015. Superpixel segmentation using linear spectral clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1356–1363.

Liu, B., Hu, H., Wang, H., Wang, K., Liu, X., Yu, W., 2013. Superpixel-based classification with an adaptive number of classes for polarimetric SAR images. IEEE Trans. Geosci. Remote Sens. 51 (2), 907–924.

Pan, X., Zhou, Y.F., Chen, Z.G., Zhang, C.M., 2019. Superpixels by bilateral geodesic distance. IEEE Trans. Multimed. 1.

Qi, Z.X., Yeh, A.G.-O., Li, X., Lin, Z., 2012. A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data. Remote Sens. Environ. 118, 21–39.

Qin, F.C., Guo, J.M., Lang, F.K., 2015. Superpixel segmentation for polarimetric SAR imagery using local iterative clustering. IEEE Geosci. Remote Sens. Lett. 12 (1), 13–17.

Russell, S.J., Peter, N., 2016. Artificial Intelligence: A Modern Approach. Pearson Education Limited, Malaysia, pp. 111–114.

Shi, J.B., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. (8), 888–905.

Stutz, D., Hermans, A., Leibe, B., 2018. Superpixels: An evaluation of the state-of-the-art. Comput. Vis. Image Underst. 166, 1–27.

Van den Bergh, M., Boix, X., Roig, G., Van Gool, L., 2015. SEEDS: Superpixels extracted via energy-driven sampling. Int. J. Comput. Vis. 111 (3), 298–314.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13 (6), 583–598.

Xiang, D.L., Ban, Y.F., Wang, W., Su, Y., 2017. Adaptive superpixel generation for polarimetric SAR images with local iterative clustering and SIRV model. IEEE Trans. Geosci. Remote Sens. 55 (6), 3115–3131.

Xiang, D.L., Tang, T., Zhao, L.J., Su, Y., 2013. Superpixel generating algorithm based on pixel intensity and location similarity for SAR image classification. IEEE Geosci. Remote Sens. Lett. 10 (6), 1414–1418.

Xu, Q., Chen, Q.H., Yang, S., Liu, X.G., 2016. Superpixel-based classification using k distribution and spatial context for polarimetric SAR images. Remote Sens. 8 (8), 619.

Yang, S., Chen, Q.H., Yuan, X., Liu, X.G., 2016. Adaptive coherency matrix estimation for polarimetric SAR imagery based on local heterogeneity coefficients. IEEE Trans. Geosci. Remote Sens. 54 (11), 6732–6745.

Yang, S., Zhang, Q.Q., Yuan, X., Chen, Q.H., Liu, X.G., 2017. K-distribution for multi-look processed polarimetric SAR imagery. In: 2017 IEEE International Geoscience and Remote Sensing Symposium. IGARSS, pp. 3700–3703.

Zhou, Y.F., Pan, X., Wang, W.P., Yin, Y.L., Zhang, C.M., 2016. Superpixels by bilateral geodesic distance. IEEE Trans. Circuits Syst. Video Technol. 27, 2281–2293.