

MODEL-BASED SUPERPIXEL SEGMENTATION OF SAR IMAGES

Koray Kayabol

Gebze Technical University
Electronics Engineering
Turkey

ABSTRACT

We propose a superpixel segmentation method for synthetic aperture radar (SAR) images. The method uses the SAR image amplitudes and pixels coordinates as features. The feature vectors are modeled statistically by taking into account the SAR image statistics. Nakagami and bivariate Gaussian distributions are used for amplitudes and position vectors, respectively. A finite mixture model (FMM) is proposed for pixel clustering. Learning and clustering steps are performed using posterior distributions. Based on the classification results obtained on real TerraSAR-X image, it is shown that the proposed method is capable of obtaining more accurate superpixels compared to state-of-the-art superpixel segmentation methods.

Index Terms— Superpixel segmentation, SAR image, finite mixture models

1. INTRODUCTION

Superpixels refer to irregular shaped large image regions that are obtained after the over-segmentation of an image. Superpixels are useful in reducing the complexity and the processing time of the images. In image understanding applications, superpixel segmentation is used as a pre-processing step to obtain a mid-level representation of the image. Superpixel segmentation methods are mostly proposed for natural optical images. In this study, we propose a superpixel segmentation method especially for SAR images. For this purpose, we use the theoretical statistics of the SAR images.

There are several superpixel segmentation methods proposed in the literature. One of the first proposed superpixel method [1] is a graph segmentation method that uses normalized cuts. In [2], another graph-based method is proposed. Beside the graph-based methods, clustering-based methods are recently proposed for superpixel segmentation [3], [4]. Quick-shift clustering algorithm is proposed in [3]. Quick-shift assigns each data point to its closest neighbor regarding to a kernel function. In [4], a k-means clustering algorithm called SLIC (simple linear iterative clustering) is proposed for

superpixel segmentation. SLIC algorithm uses a feature vector constituted by color values and the positions of the pixels.

All the aforementioned methods can be used for superpixel segmentation of the color images. Most of the methods use the Euclidean distance or Gaussian kernels to measure the similarity between the pixel intensities. The Gaussian distribution may be a good approximation to model the color image statistics but this is not the same case in SAR images. SAR images have some special statistical properties unlike the optical color images. For instance, theoretical intensity and amplitude statistics of a multi-look SAR image follow gamma and Nakagami distributions, respectively [5]. Since the theoretical statistical models for SAR images are obtained under the multiplicative noise assumption, they are consequently more convenient models to deal with speckle in SAR images.

In [6], a superpixel segmentation method is proposed for polarimetric SAR images by using a Wishart distance instead of Euclidean. In this study, we propose a SAR image specific superpixel clustering method. We interpret the clustering method in the light of gestalt-based perceptual grouping rules listed in [7]. We only use two rules called similarity and proximity. For similarity, we use Nakagami distribution as a statistical measure, i.e. that two pixels in the same cluster are assumed to be generated from the same Nakagami distribution. For proximity, we use the bivariate Gaussian distribution to model the spatial distances between the pixels.

Both similarity and proximity statistics are combined into a finite mixture model (FMM). FMMs have been already used in SAR image intensity and amplitude classification [8], [9], [10], [11]. Using these statistical measures within an FMM, the pixels are clustered around the superpixels' centroids.

Organization of the paper is as follows. Section 2 and 3 respectively present the proposed FMM-based superpixel model and related segmentation algorithm. The simulation results are reported in Section 4. Section 5 summarizes the conclusion and future work.

2. FINITE MIXTURE MODEL FOR SUPERPIXELS

We denote the pixel amplitudes by $a_n \in \mathbf{R}^+$ and the coordinates by $\mathbf{q}_n = [x_n, y_n]^T \in \mathbf{R}^2$ where $n = 1, \dots, N$ is the

This work is supported by Scientific and Technological Research Council of Turkey (TUBITAK) under Project No. 114E535.

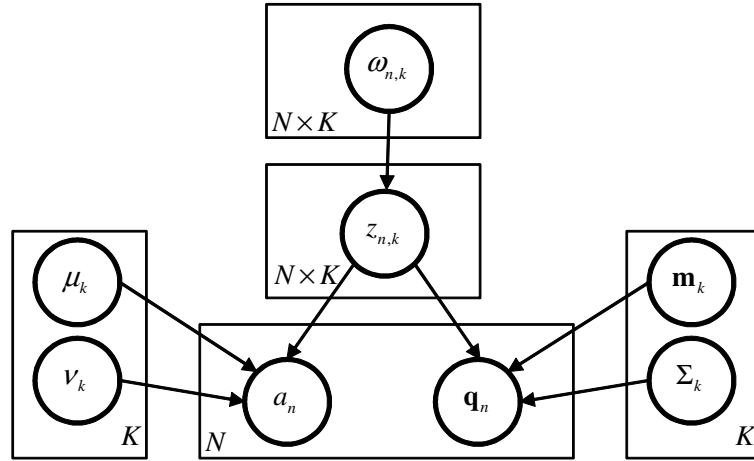


Fig. 1. Graphical representation of the proposed finite mixture model.

lexicographically ordered pixel index. The feature vector of the n th pixel is formed by $\mathbf{f}_n = [a_n, x_n, y_n]^T$. We assume that the amplitudes and the coordinates are statistically independent, i.e. $p(a_n, \mathbf{q}_n | \theta) = p(a_n | \theta)p(\mathbf{q}_n | \theta)$ where θ is the parameter set of the distributions.

We use Nakagami distribution to model the pixel amplitudes, that is a basic theoretical multi-look amplitude model for SAR images [5]. Nakagami distribution function is given by

$$p(a_n | \mu_k, \nu_k) = \frac{2}{\Gamma(\nu_k)} \left(\frac{\nu_k}{\mu_k} \right)^{\nu_k} a_n^{2\nu_k - 1} e^{-\nu_k \frac{a_n^2}{\mu_k}}. \quad (1)$$

where μ_k and ν_k are the parameters of the k th superpixel. We assume that the pixels are spatially distributed around the centroid of a superpixel according to the normal law as follows:

$$p(\mathbf{q}_n | \mathbf{m}_k, \Sigma_k) = \frac{1}{2\pi |\Sigma_k|^{\frac{1}{2}}} \exp \{ (\mathbf{q}_n - \mathbf{m}_k)^T \Sigma_k^{-1} (\mathbf{q}_n - \mathbf{m}_k) \} \quad (2)$$

where \mathbf{m}_k and Σ_k are the centroid and the covariance matrix of the k th superpixel, respectively. For the k th superpixel the parameter set is defined to be $\theta_k = \{\mu_k, \nu_k, \mathbf{m}_k, \Sigma_k\}$.

We aim to divide the image into mutually exclusive superpixels. Assuming that there are K number of superpixels in the image, we define a K -dimensional label vector $\mathbf{z}_n \in \{0, 1\}^K$ for each superpixel. The binary label vector \mathbf{z}_n has the property that $\sum_{k=1}^K z_{n,k} = 1$. We define that $\mathbf{z}_n \in \{[1, 0, \dots, 0], [0, 1, \dots, 0], \dots, [0, 0, \dots, 1]\}$. We also assume that \mathbf{f}_n 's are conditionally independent given the labels, \mathbf{z}_n 's. Since the natural prior for \mathbf{z}_n is a multinomial distribution, we may define the following prior for \mathbf{z}_n

$$p(\mathbf{z}_n | \omega_{1:K}) = \prod_{k=1}^K \omega_k^{z_{n,k}}$$

where $\omega_{1:K}$ are the parameters of the multinomial distribution.

After these definitions, we may show that the marginalization of the joint density

$$p(\mathbf{f}_n, \mathbf{z}_n | \theta_{1:K}, \omega_{1:K}) = p(\mathbf{f}_n | \mathbf{z}_n, \theta_{1:K}) p(\mathbf{z}_n | \omega_{1:K})$$

yields a finite mixture density as follows:

$$p(\mathbf{f}_n | \theta_{1:K}, \omega_{1:K}) = \sum_{\mathbf{z}_n} \prod_{k=1}^K [p(f_n | \theta_k) \omega_k]^{z_{n,k}} \quad (3)$$

where ω_k corresponds to mixture proportion. Graphical representation of the proposed FMM is given in Fig. 1.

3. INFERENCE

We need to perform a posterior inference for the proposed probabilistic model. In the model, there are two key variables to be inferred, namely the superpixel labels $\mathbf{z}_{1:N}$, and the parameters of the Nakagami and Gaussian distributions $\theta_{1:K}$. We assume the mixture proportions to be fixed $\omega_k = 1/K$. We resort to iterated conditional mode (ICM) algorithm because the inference from the joint posterior $p(\mathbf{z}_{1:N}, \theta_{1:K}, \omega_{1:K} | \mathbf{f}_{1:N})$ is not tractable. In ICM algorithm, conditional densities of variables are maximized iteratively. The joint posterior of the variables is factorized as follows

$$p(\mathbf{z}_{1:N}, \theta_{1:K}, \omega_{1:K} | \mathbf{f}_{1:N}) \propto p(\mathbf{f}_{1:N} | \mathbf{z}_{1:N}, \theta_{1:K}) p(\mathbf{z}_{1:N} | \omega_{1:K}) \quad (4)$$

where the likelihood term (the first term on the righthand side) can be factorized as follows:

$$p(\mathbf{f}_{1:N} | \mathbf{z}_{1:N}, \theta_{1:K}) = \prod_{n=1}^N \prod_{k=1}^K (p(a_n | \theta_k) p(\mathbf{q}_n | \theta_k))^{z_{n,k}}$$

The joint prior distribution of the labels $p(\mathbf{z}_{1:N} | \omega_{1:K})$ in

(4) is given by

$$p(\mathbf{z}_{1:N}|\omega_{1:K}) = \prod_{n=1}^N \prod_{k=1}^K \omega_k^{z_{n,k}} \quad (5)$$

In order to perform a posterior inference, we use the block ICM algorithm. Unlike the ICM algorithm, block ICM updates the same kind of variables together at a time. In this way, we aim to obtain a faster algorithm than conventional ICM [12]. We update the variables along the iterations in the following order:

$$\begin{aligned} \mathbf{z}_n^t &\leftarrow \max_{\mathbf{z}_n} p(\mathbf{f}_n|\mathbf{z}_n, \theta_{1:K}^{t-1}) p(\mathbf{z}_n|\omega_{1:K}^{t-1}) \\ \theta_k^t &\leftarrow \max_{\theta_k} p(\mathbf{f}_{1:N}|\mathbf{z}_{1:N}^t, \theta_k) \end{aligned}$$

where $n = 1, \dots, N$, $k = 1, \dots, K$ and t is the pseudo time index.

4. EXPERIMENTAL RESULTS

We present a comparison of the proposed FMM-based SAR superpixel and two state-of-the-art superpixel methods. These algorithms are SLIC [4] and Quickshift [3]. We use VLFeat open source library [13] for implementation of SLIC and Quickshift algorithms.

4.1. Practical Issues

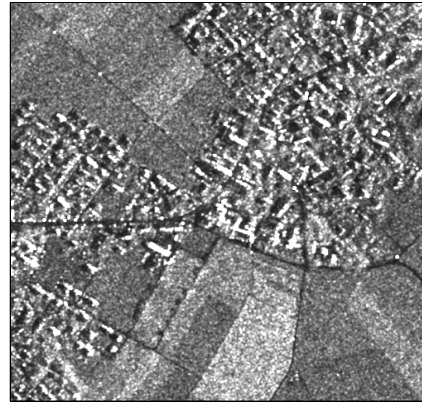
In this section, we give the details of the practical implementation of the proposed algorithm. We initialize the superpixels by dividing the image into $S \times S$ ($S < N$) sub-image blocks.

We follow the same strategy used in [4], i.e. when maximizing the superpixel labels, we only take into account the neighboring pixels but not the entire superpixel. The neighborhood is defined to be a $2S \times 2S$ region around the superpixel center.

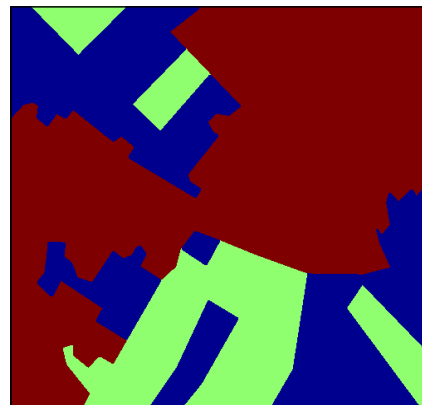
Since the methods are based on clustering, they do not take into consideration the spatial dependency of the pixel. Therefore, a superpixel may include more than one connected component. At the end of the each algorithm, we remove the small connected components by merging them with the adjacent big segment. For this purpose, we use the code provided in [14].

4.2. SAR Image

We have tested the algorithms on the test image TSX (see Fig. 2(a)): 500×500 pixels, HH polarized, TerraSAR-X SpotLight (8.2 m ground resolution) which was acquired over Rosenheim in Germany on January 27, 2008 (©Infoterra). The ground-truth map is generated manually as seen in Fig. 2(b) by assuming three kind of land cover classes. One of them is urban zone. The remaining two regions are two different land zones which are seen dark and bright in Fig. 2(a).



(a) TSX



(b) Ground-truth

Fig. 2. SAR image and its ground-truth.

4.3. Comparisons

Superpixel segmentation results of three algorithms are given in Fig. 3. The numbers of superpixels obtained by FMM, SLIC and QS are 100, 102 and 102, respectively. As seen from the segmentation maps in Fig. 3, the superpixels found by proposed FMM-based method fits the real SAR image boundaries better than those found by SLIC and Quick-shift.

We use under-segmentation error (UE), boundary recall (BR) and boundary precision (BP) for the numerical evaluation of the performances of the algorithms. Although BR is the mostly used measure for superpixel evaluation, high BR values do not always correctly measure the goodness of fit of the edges. If the superpixel boundaries are unnecessarily overfitted to the images details, BR may indicate a good edge segmentation result despite a bad edge representation. Therefore, we use BP along with UE and BR. The UE used in this paper is defined by

$$UE = \sum_k^K \sum_l^L \frac{|S_l \cap G_k^C|}{|S_l|} \quad (6)$$

where S_l is a superpixel and G_k^C is the complement of the

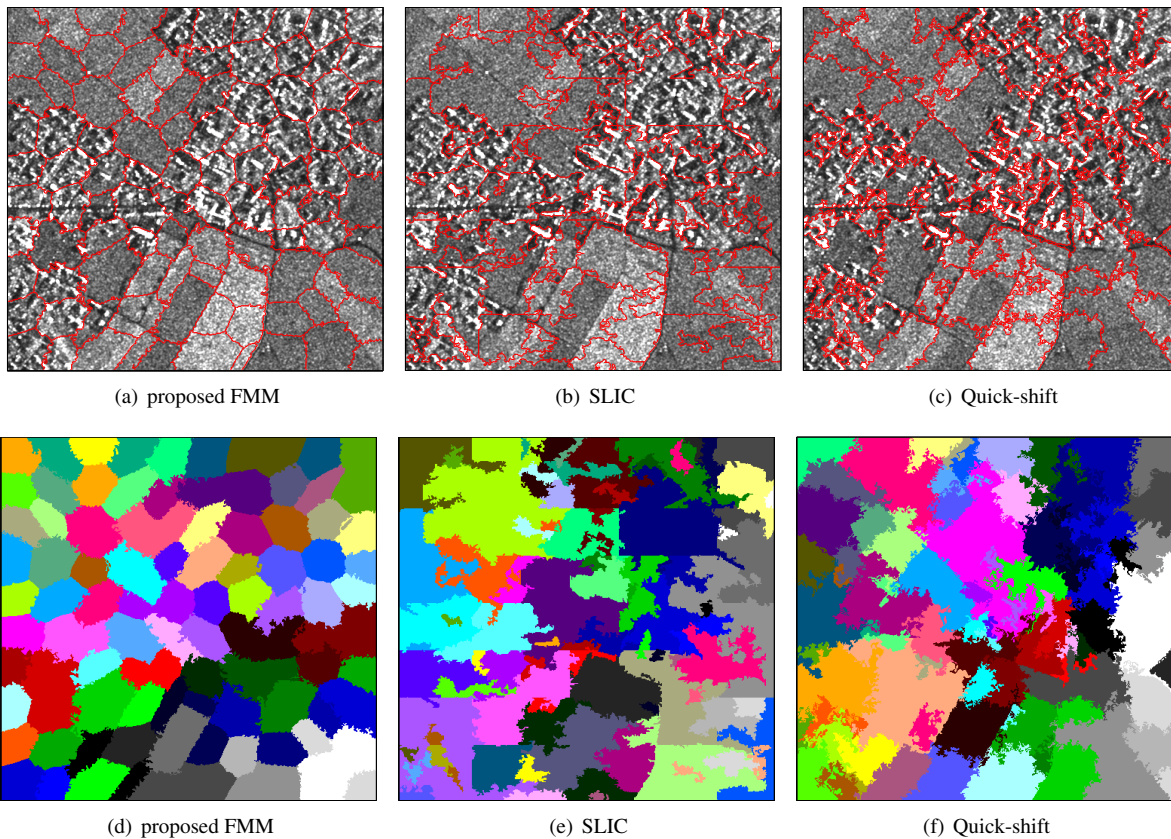


Fig. 3. Superpixel segmentation results obtained by proposed FMM, SLIC and Quick-Shift methods.

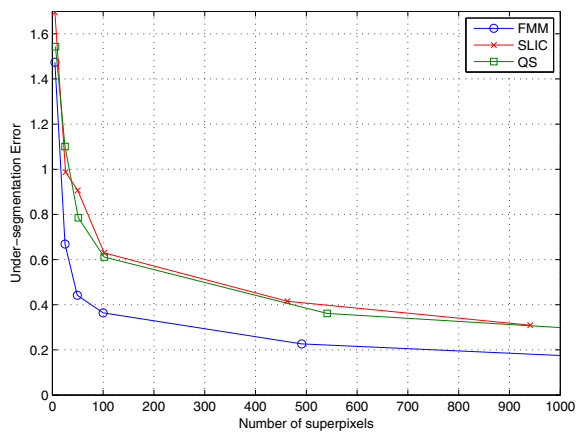


Fig. 4. Under-segmentation error as a function of number of superpixels.

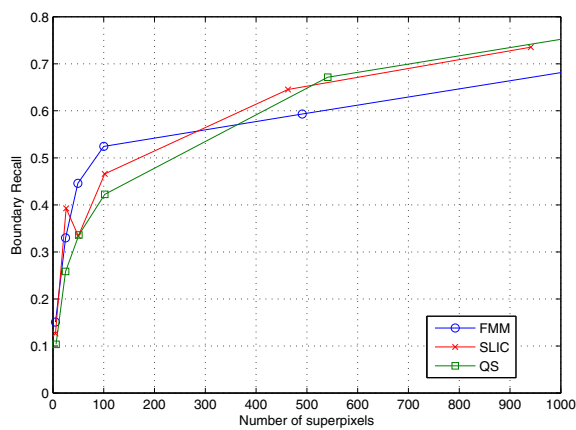


Fig. 5. Boundary recall as a function of number of superpixels.

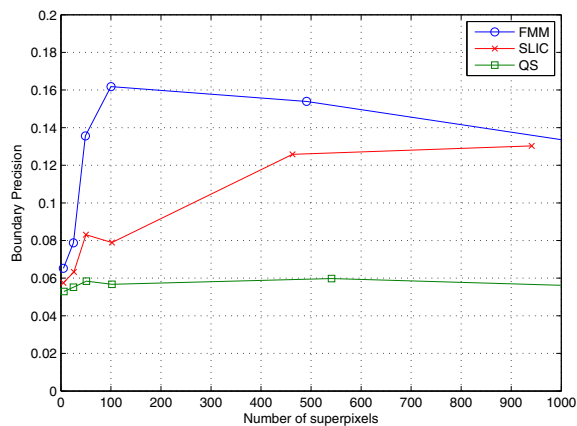


Fig. 6. Boundary precision as a function of number of superpixels.

ground-truth segment of k th class. BR is the fraction of the ground-truth edges that hit the superpixel boundaries within at least two pixels. BP is the fraction of the superpixel boundaries hit by the ground-truth edges. Fig. 4 shows the UE values obtained by three algorithms for different number of superpixels. As seen from the UE and BP plots in Fig. 4 and 6, the proposed FMM performs significantly better than the other two methods. According to BR plots in Fig. 5, the performance of the proposed algorithm is better when the number of superpixels is less than 300.

5. CONCLUSION

In this study, we have proposed a FMM-based superpixel segmentation method that takes into account the SAR image statistics. The proposed method yields significantly better superpixel segmentation for SAR images compared to state-of-the-art superpixel methods. By using proposed method, without any preprocessing as noise removing, we are able to obtain superpixel segmentation of SAR images. The proposed method provides a mid-level representation of the images that may be used for higher-level understanding of SAR images. We use theoretical Nakagami distribution to model the SAR image statistics but other empirical distributions may be used as well.

Acknowledgment

The author would like to thank A. Vedaldi, B. Fulkerson, and P. D. Kovesi for providing their code online.

REFERENCES

- [1] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 888–905, 2000.

- [2] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, 2004.
- [3] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in *European Conf. Computer Vision, ECCV*, 2008, vol. LNCS 5305, pp. 705–718.
- [4] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 34, no. 11, pp. 2274–2281, 2012.
- [5] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Norwood, 1998.
- [6] F. Qin, J. Guo, and F. Lang, "Superpixel segmentation for polarimetric SAR imagery using local iterative clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 13–17, 2015.
- [7] M. Wertheimer, *A Sourcebook of Gestalt Psychology*, chapter Laws of organization in perceptual forms (partial translation), pp. 71–88, Harcourt, Brace and Company, 1938.
- [8] J. M. Nicolas and F. Tupin, "Gamma mixture modeled with "second kind statistics": application to SAR image processing," in *Int. Geosci. Rem. Sens. Symp. IGARSS'02*, 2002, pp. 2489–2491.
- [9] G. Moser, J. Zerubia, and S.B. Serpico, "Dictionary-based stochastic expectation-maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 1, pp. 188–199, 2006.
- [10] K. Kayabol and J. Zerubia, "Unsupervised amplitude and texture classification of SAR images with multinomial latent model," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 561–572, 2013.
- [11] K. Kayabol, "Histogram-based contextual classification of SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 33–37, 2015.
- [12] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. R. Stat. Soc. B*, vol. 36, no. 2, pp. 192–236, 1974.
- [13] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," <http://www.vlfeat.org/>, 2008.
- [14] P. D. Kovesi, "MATLAB and Octave functions for computer vision and image processing," Centre for Exploration Targeting, School of Earth and Environment, The University of Western Australia, Available from: <<http://www.csse.uwa.edu.au/~pk/research/matlabfns/>>.