SIMILARITY RATIO BASED ALGORITHMS TO GENERATE SAR SUPERPIXELS

A THESIS SUBMITTED TO
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EMRE AKYILMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
GEODETIC AND GEOGRAPHICAL INFORMATION TECHNOLOGIES

APRIL 2017

Approval of the thesis:

SIMILARITY RATIO BASED ALGORITHMS TO GENERATE SAR SUPERPIXELS

submitted by **EMRE AKYILMAZ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Geodetic and Geographical Information Technologies Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver
Dean, Graduate School of **Natural and Applied Sciences** _____

Assoc. Prof. Dr. Uğur Murat Leloğlu
Head of Department, **Geodetic and Geographic Information Technologies, METU** _____

Assoc. Prof. Dr. Uğur Murat Leloğlu
Supervisor**, Geodetic and Geographic Information Technologies, METU** _____

Assoc. Prof. Dr. İlkay Ulusoy
Co-supervisor**, Electrical and Electronics Engineering Dept., METU** _____

Examining Committee Members:

Prof. Dr. Zuhal Akyürek
Civil Engineering Department, METU _____

Assoc. Prof. Dr. Uğur Murat Leloğlu
Geodetic and Geographical Information Technologies, METU _____

Assoc. Prof. Dr. Afşar Saranlı
Electrical and Electronics Engineering Department, METU _____

Assist. Prof. Dr. Emre Sümer
Computer Engineering Department, Baskent University _____

Assist. Prof. Dr. Fatih Nar
Faculty of Eng. and Architecture, Konya Food & Agriculture University _____

Date: 13 April 2017

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : EMRE AKYILMAZ

Signature:

ABSTRACT

SIMILARITY RATIO BASED ALGORITHMS TO GENERATE SAR SUPERPIXELS

AKYILMAZ, EMRE

Ph.D., Department of Geodetic and Geographic Information Technologies

Supervisor: Assoc. Prof. Dr. Uğur Murat Leloğlu

April 2017, 106 pages

Synthetic Aperture Radar (SAR) has the capability of working in all weather conditions during day and night that makes it attractive to be used for automatic target detection and recognition purposes. However, it has the problem of high amount of multiplicative speckle noise. Superpixel segmentation as a preprocessing step is an oversegmentation technique that groups similar neighboring pixels into regularly organized segments with approximately the same size. As boundaries of the objects are important elements to be traced, superpixels should adhere well to the edges. This can only be achieved by an algorithm robust to speckle noise. In this thesis, similarity ratio is first developed as a new metric that is robust to speckle noise. Secondly, Mahalanobis distance is used instead of Euclidian so that the superpixel can fit better to shapes in the real world. Thirdly, the constant determining the relative importance of radiometric and geometric terms is replaced with an adaptive function. The performance of combinations of similarity ratio with Euclidean distance (SREP), Mahalanobis distance (SRMP) and Mahalanobis distance with adaptive scheme (SRAMP) are evaluated by conducting experiments on real and synthetic images. The experimental results showed that similarity ratio

and adaptive Mahalanobis proximity (SRAMP) outperforms the other approaches in terms of uniformity, compactness and visual appearance.

Keywords: Image Processing, Synthetic Aperture Radar (SAR), Segmentation, Similarity Ratio, Superpixels, Remote Sensing

ÖZ

SAR SÜPERPİKSELLER ÜRETİMİ İÇİN BENZERLİK ORAN TABANLI ALGORİTMALAR

AKYILMAZ, EMRE
Doktora, Jeodezi ve Coğrafi Bilgi Teknolojileri Bölümü
Tez Yöneticisi: Doç. Dr. Uğur Murat Leloğlu

Nisan 2017, 106 sayfa

Sentetik Açıklı Radar(SAR) sistemleri gece, gündüz ve tüm hava şartlarında çalışabilen aktif bir radar tipidir. Bu nedenle, SAR görüntüleri otomatik hedef tespiti ve otomatik hedef teşhisi işlemlerinde oldukça fazla kullanılmaktadır. Bütün bunların yanı sıra SAR görüntülerinin kalitesi ve anlaşılabilirliği sahip oldukları benek gürültü nedeni ile oldukça düşük düzeydedir. Bir görüntünün anlamlı parçalara ayrıştırılmasını temel alan ayrıştırma işlemi görüntüde bulunan piksellerin komşularına bakılarak, benzer komşu piksellerin birleştirilmesine dayanıyorsa, bu işlem sırasında süperpiksel olarak ifade edilen bölütlenmiş parçalar ortaya çıkmaktadır. Nesne ve bölgelerin sınırları takip edilmesi gereken yapılar oldukları için oluşturulan alanların nesne ve bölgelerin sınırlarına düzgün oturmaları gerekmektedir. Bu nedenle, SAR görüntülerinde bulunan benek gürültüye karşı gürbüz bir algoritmaya ihtiyaç duyulmaktadır. Bu çalışmada, ilk olarak benzerlik oranı benek gürültüye karşı yeni gürbüz bir metrik olarak geliştirilmiştir ve sonrasında bu metrik üretilen süperpiksellerin görüntü üzerindeki şekillere daha iyi oturması için Öklid uzaklığı yerine, Mahalanobis uzaklığı ile birlikte kullanılmıştır. Daha sonra radyometrik ve geometrik terimleri arasındaki

önemi belirleyen sabit değer, adaptif bir fonkisyon ile değiştirilmiştir. Benzerlik oranı ile birlikte kullanılarak geliştirilen Öklid (SREP), Mahalanobis (SRMP) ve adaptif Mahalanobis (SRAMP) algoritmalarının performansları, gerçek ve sentetik görüntüler üzerinde deneyler yapılarak literatürde bulunan benzer algoritmalar ile karşılaştırılmışlardır. Deneysel sonuçlar, benzerlik oranı ve adaptif Mahalanobis yakınlık (SRAMP) yönteminin diğer yaklaşımlara göre tek biçimlilik, kompaktlık ve görsel görünüm açısında daha iyi performans sergilediğini ortaya koymuştur.

Anahtar kelimeler: Görüntü İşleme, Sentetik Açıklıklı Radar (SAR), Bölütleme, Benzerlik Oranı, Süperpiksel, Uzaktan Algılama

To my family

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor Assoc. Prof. Dr. Uğur Murat Leloğlu for his guidance, criticism, support, encouragement, insight and friendship throughout the years of my research.

I am grateful to all of the committee members for providing valuable criticism to improve my research.

I would like to thank my parents and my sister for their endless support and encouragement that I benefited greatly.

I also would like to thank to my beloved wife İpek for her continuous support that gave me courage and strong energy throughout the most difficult times in this thesis.

Finally, I would like to thank my son Umut who brings joy, happiness and hope to my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Segmentation is a widely-used technique that can be considered as one of the important tasks in automatic analysis of remote sensing images. If an algorithm partitions an image into non-overlapping homogeneous segments, many perceptually salient aspects of this complex imagery can be captured for further analysis. This type of segmentation groups pixels of an image into uniform regions that are called superpixels. In this introductory chapter, problems related with the superpixel algorithms are defined, then the major contributions of this thesis are stated and finally the organization of the thesis is presented.

1.1. Problem Description

Superpixelling has emerged as an oversegmentation approach over the last decade. This type of superpixel based segmentation transforms an image data from pixel based to segment-based data as opposed to nonsuperpixel-based segmentation. These segments are the elementary units of the superpixels rather than the pixels of the image. This process is fundamental for various applications of Synthetic Aperture Radar (SAR) image analysis since an effective segmentation can reduce image complexity and enhance interpretation and understanding of the segments by their defined boundaries.

There are many superpixel segmentation algorithms for various applications of target detection and recognition. The major issue associated with these approaches

is the lack of robustness. They are not robust enough to produce reliable segmentation of SAR images due to the existence of multiplicative speckle noise. For amplitude SAR images, information is available in the form of intensities. If these intensities are clearly identified, a reliable evaluation of their similarities would be obtained to be used to group them for successful segmentation of the image. The speckle noise appearing on SAR images is multiplicative in nature. In the presence of such noise, the intensities cannot be clearly defined. In this case the algorithm not only reduces an analyst ability to resolve fine detail on the image but makes automatic segmentation much difficult. Alternative to existing superpixel segmentation approaches, effective and robust superpixel generating approaches are proposed in this thesis.

1.2. Contributions of the Thesis

In this thesis, the first contribution is the development of a radiometric distance metric as similarity ratio term which is robust enough for multiplicative speckle noise in SAR images. Secondly, this term is combined with Mahalanobis distance instead of Euclidean term for Similarity Ratio and Mahalanobis Proximity (SRMP) algorithm. The use of Mahalanobis distance term results superpixels with elongated shapes to fit the complex structure of the real world. This type of formulation leads to better performance on SAR images. Finally, the balancing parameter that determines the relative importance of geometric and radiometric similarity distance terms whose best value should be chosen for the image is considered. Instead of a predefined constant, its value is adaptively determined for each superpixel pair as a function of the average values of the superpixels means. This both improved the performance of the algorithm and made parameter free.

To achieve these contributions, Similarity Ratio and Euclidean Proximity (SREP) and Similarity Ratio and Mahalanobis Proximity (SRMP) algorithms are formulated. In each formulation, similarity ratio and proximity terms are linearly combined and parameterized with a constant that controls the superpixel sizes and

their spatial extent. Since this parameter is kept constant for all clusters during the course of the algorithm, it might produce irregular shaped superpixels rather than compact ones within homogeneous regions. This is especially the case for heavily speckled SAR images. Therefore, the constant parameter needs to be determined during the course of the SRMP algorithm in an adaptive way. This is performed with a function developed and inserted into SRMP. In this way, the proposed similarity ratio based adaptive Mahalanobis proximity algorithm (SRAMP) adapts the intensity similarity and the spatial proximity for each pair of clusters. To explore the segmentation capabilities of the adaptive SRMAP as compared to SRMP, SREP and four other existing algorithms that are considered as state-of-the-arts methods in the literature, experiments are conducted by using speckled synthetic images and real SAR images.

1.3. Organization of the Thesis

After the introductory chapter, some basic SAR imaging essentials are presented in Chapter 2 for those researchers who have no background in radar imaging. The sole purpose of this chapter is to acquaint such readers with SAR imagery. Since SAR image tone varies between very dark to very bright depending upon the target scanned, the material presented in this chapter will be useful for interpreting the images and the results presented in this thesis.

An extensive literature review of the previous research efforts for nonsuperpixel-based and superpixel-based segmentation algorithms are presented in Chapter 3. As superpixel based segmentation algorithms are reviewed, a research outcome has clearly emerged: development of a robust superpixel segmentation algorithm to multiplicative speckle noise in SAR images that generates perceptually meaningful and representationally efficient superpixels.

Alternative to Euclidean distance, Mahalanobis distance is recommended and developed as another proximity metric in Chapter 4. For the two different spatial

distance metrics, similarity ratio based algorithms are formulated in Chapter 5. These are named as similarity ratio and Euclidean proximity (SREP) and similarity ratio and Mahalanobis proximity (SRMP) algorithms, respectively. The derivation of the similarity ratio is also presented in this chapter. As a final task in this chapter, adaptive Mahalanobis proximity (SRAMP) algorithm that adapts the similarity ratio and the spatial distance for each cluster pair is presented.

Segmentation performance of SRAMP as well as SREP and SRMP algorithms are evaluated and the results are presented in Chapter 6. For illustrative purposes, applications of superpixel-based clustering are given in Chapter 7. Finally, summary and conclusions with suggestions for further research are presented in Chapter 8.

CHAPTER 2

SAR IMAGING ESSENTIALS

In this chapter, some basic synthetic radar imaging essentials are presented. The purpose of this chapter is to acquaint the researchers who have no or little background in radar imaging with the basic concepts and definitions used in this thesis.

2.1. Historical Background

Airborne radar imaging has been available since 1960s. The lack of high spatial resolution in this type imaging prohibits obtaining coverage about spatial information and observations over a wide range of area.

Fine spatial resolution remote sensing of the Earth's surface was first accomplished from space by NASA in 1960s on military base using optical devices mounted on satellites. In 1973, NASA initiated a series of missions with LANDSAT to obtain fine resolution (in the standards of the day) optical images.

The mission of Seasat as the first civilian satellite launched by NASA in 1978 opened a new era for radar imaging. This satellite hosting a synthetic aperture radar (SAR) which has been designed for observing the oceans. The mission that lasted for three months concentrated on monitoring of oceanographic phenomena such as sea surface winds, temperature, wave heights, internal waves, sea surface features and ocean topography. Although lasted for a short period of time, the mission demonstrated the feasibility of using SAR to monitor timely oceanographic phenomena during day and night and in all weather conditions.

Since the launch of Seasat in 1978, many SAR satellites have been placed in orbit. Today, more than a dozen of satellites or satellite series such as ERS (Europe), JERS (Japan), TERRASAR (Germany), RADARSAT (Canada), ALOS (Japan), and ALMAZ (Russia) equipped with SAR sensors. It has been mentioned that while the systems has different configurations, the underlying operating principle for each system is similar (Mccandless and Jackson 2004)

2.2. Imaging Essentials

SAR differs from early imaging radars in its use of aperture synthesis to improve spatial resolution. Synthesized aperture is achieved by simulating the flight direction of the sensors as a large antenna.

High resolution capabilities coupled with operational advantage make SAR an ideal remote sensing instrument for many applications of earth observation. Such remote sensing had been recognized as an important tool in collecting data during day and night and in all weather conditions for certain programs initiated by the various agencies around the world. Some examples of such programs are ISTOP (Integrated Satellite Tracking of Pollution) program initiated by Canadian Centre for Mapping and Earth Observation agency of Canada, SEARCH (Study of Environmental Arctic Change) program initiated under the auspices of National Science Foundation of the U.S. and CORINE (Coordination of Information on the Environment) program initiated by European Environmental Agency of European Union. Among those CORINE program is worth to mention due to its recent use of SAR data. The data is collected via a SAR satellite launched in 2014 by the European Space Agency. The collected image data serve to map and monitor the land cover over the designated geographic areas of Europe. The satellite system made available to CORINE program is highly sophisticated in differentiating various land cover classes, surface deformation, disaster management and change detection. In a recent study (Baltzer et. al. 2015), it has been mentioned that 85 % accuracy is achieved in mapping five basic land cover classes over 75.000 km^2 of

Belgium, Netherlands and Germany. The segmentation of SAR images at preprocessing step can be important to differentiate various land cover classes. This segmented image can reduce image complexity in extracting those classes. Superpixel techniques can efficiently be used to oversegment the SAR images without loss of information on the region boundaries. To achieve this, algorithms must be robust to speckle noise which inherently exists in SAR images. The robust similarity ratio based Mahalanobis algorithm developed in this thesis can effectively be used for this purpose. The CORINE program as well as the other programs benefit to a great extent from the use of this robust technique in segmenting SAR images.

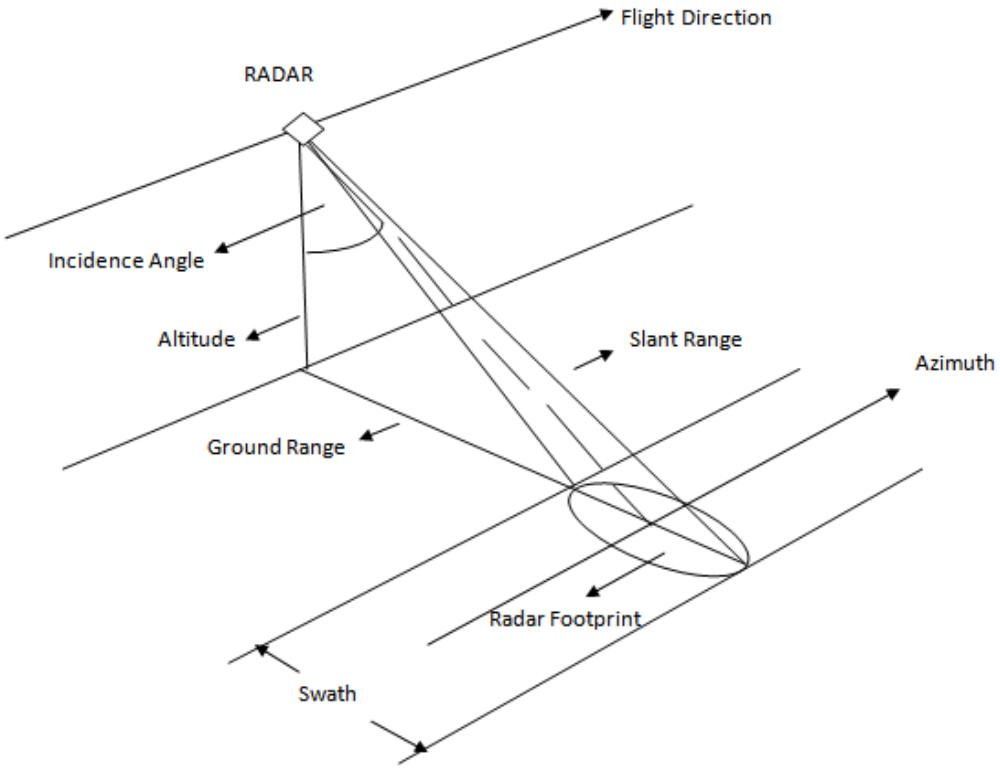


Figure 2.1 Geometry of SAR Imaging

The Figure 2.1 shows a simplified geometric configuration of a typical SAR system. Basically such a system consists of a pulsed microwave transmitter, an antenna used both for transmission and reception, a receiver, a data collection system and a signal processor for the construction of the image. The imaging SAR

system is an active radar system that is mounted on an airborne or spaceborne moving platform. It operates in a side-looking geometry shown in Figure 2.2. The system illuminates the terrain with emitted pulses from the transmitter and receives signal backscattered from the illuminated terrain by the receiver. This type of an active imaging technology makes this kind of sensor independent of day and night conditions (Lee and Pottier, 2009).

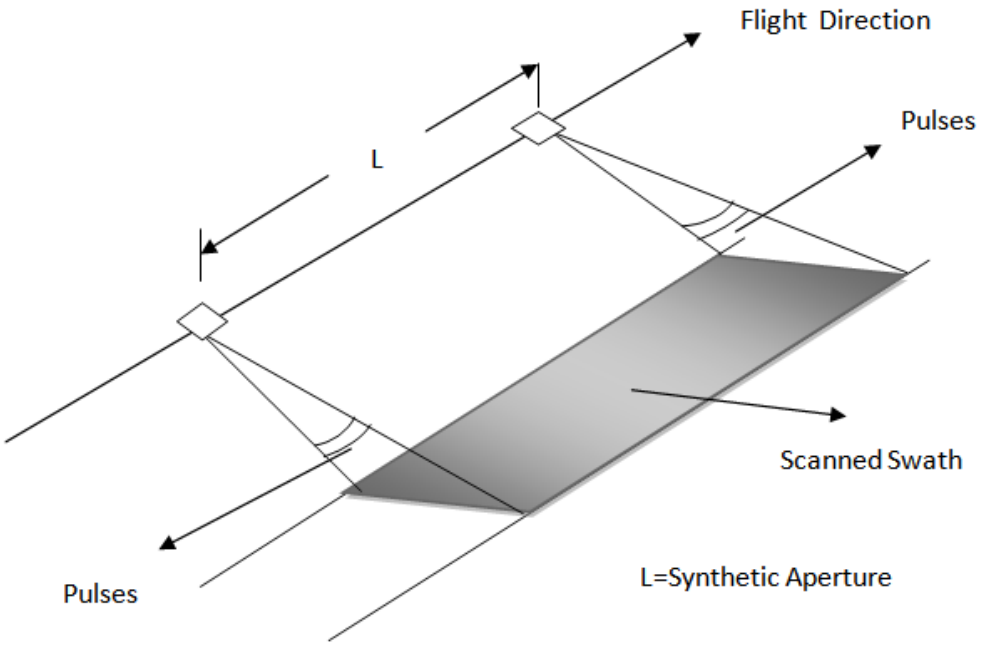


Figure 2.2 Synthetic Antenna along Flight Direction

Referring to Figure 2.2, as the SAR imaging system moves with a velocity V along a specified synthetic aperture length, it scans the area illuminated with radar pulses at a height H from the terrain. The antenna is aimed perpendicular to the flight direction referred to as azimuth and the antenna beam is directed slant-wise toward the ground with an angle of incidence. The radar line of sight is referred as slant range. The illuminated area in the terrain in the x and y directions is the antenna footprint which is scanned by the system. The area scanned by the antenna beam is the radar swath. The radar swath that covered by the reflected

pulses registered at all positions along the flight path are the essence for the construction of the image scene (Lee and Pottier, 2009).

To get some idea about the SAR imaging process, consider a moving platform travelling with a speed of in excess of 2000 km/hr. This system transmits more than a thousand of pulses per second with a speed of light, illuminate tens of millions of resolution cells (pixels) in the radar beam at each pulse time and thousands of processor operations per cell to construct an image.

The qualities of the constructed images are directly related with the ground range and spatial resolution. The spatial resolution in azimuth direction is directly related with the azimuth beam width of the antenna whereas spatial resolution in slant direction is directly related with the time resolution of different radar waves in the range direction. The quality of the constructed images are not much affected with the slant range but the projection of this onto the horizontal plane as ground range resolution is an important criterion (Curlander and McDonough, 1991).

The microwave frequency of the SAR provides the penetration of the beam pulses through the clouds and other weather conditions such as rain and snow and this makes it usable in all weather conditions for the global scale earth monitoring except the cases of high frequency ranges combined with extreme weather events like heavy rain. These radar systems provide perfect operation conditions for the observation of the physical properties of the objects and geographic formations.

Interpretation and understanding of a gray scale SAR image as shown in Figure 2.6 requires the analysis of the intensities of the pixels that image represents the type and the interpretation of reflection of the backscattered microwave from the area on the ground. Hence, the reflection types directly affect the visible shape of the target to be imaged. Different reflection types and their resultant effects are shown in Figure 2.3 to Figure 2.5.

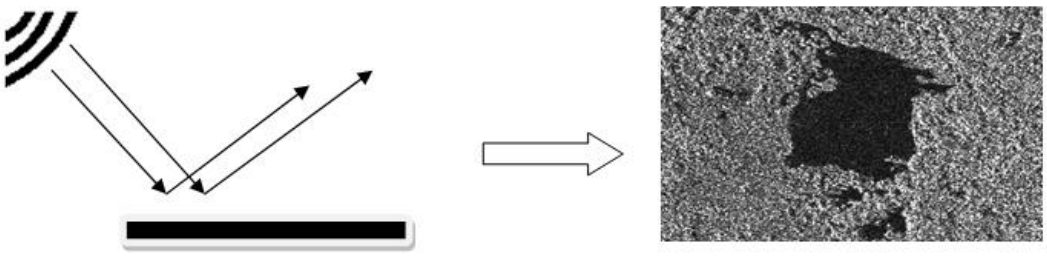


Figure 2.3 Specular Reflection

In case of specular reflection most of the microwave energy is reflected away so a very little or no energy is backscattered to the radar sensor as shown Figure 2.3. The resultant effect of such reflection is the appearance of the target area with dark image tone.

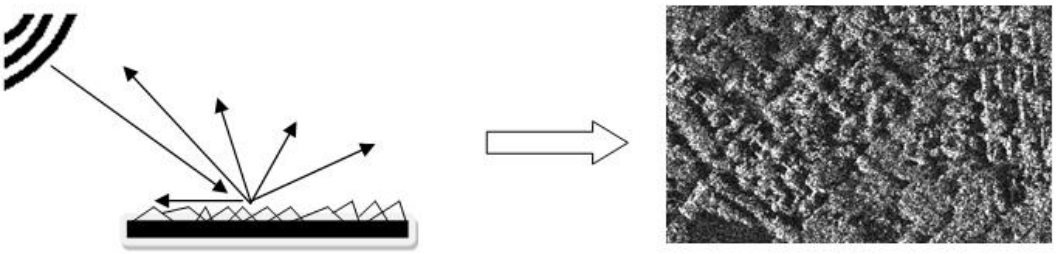


Figure 2.4 Diffused Reflection

Figure 2.4 is the case of reflecting radar pulses in all directions. Some of the radar energy is backscattered to the radar sensor. The resultant of such reflection is the appearance of the target area with medium gray image tone.

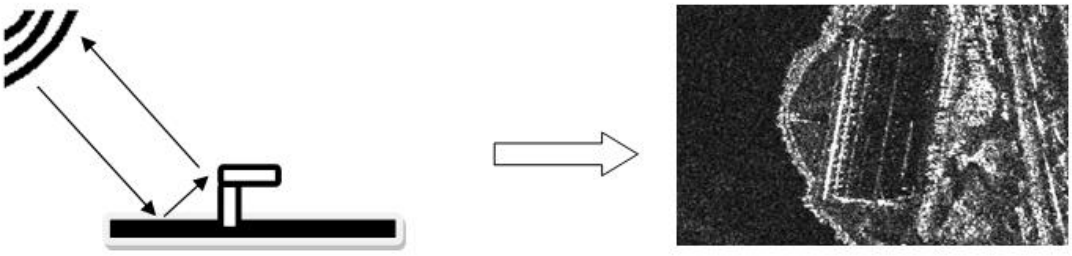


Figure 2.5 Corner Reflection

If two surfaces are at right or nearly right angle facing the radar beam, the beam bounces to both of the surfaces and nearly all of the radar energy is scattered back to the radar sensor as shown in Figure 2.5. The resultant of such reflection is the appearance of the reflected area with a bright tone.

The image tone (digital number) in a SAR image may vary between very dark to very bright depending on the type of the target scanned. Flat surfaces such as paved surfaces as road and runways, calm water surfaces as rivers and oceans normally appear as dark areas due to specular reflectance as most of the emitted radar energy is reflected away from the radar sensor. Vegetated surfaces such as forestry areas which are moderately rough surfaces appear as moderately bright areas due to diffused reflectance as some of the radar energy is backscattered to the radar sensor. Note that surface roughness is a function of not only the surface geometry but also the wavelength. Built-up areas and man-made structures appear as bright patches due to corner reflectance where most of the radar energy is scattered back to the radar sensor.

Some SAR systems have the capability to send and receive radar energy with different polarizations. By emitting and receiving radar energy in combinations of polarizations setting, several images can be collected from the same series of pulses. Among those, HH (horizontal send-horizontal receive), HV (horizontal send-vertical receive) and VV (vertical send-vertical receive) are widely used in dual polarization systems. This polarimetric SAR imaging subject is beyond the scope of this thesis.

SAR images are formed by coherent interaction of the transmitted waves with the targets. The coherent summation of the waves reflected from many elementary scatters distributed randomly within each resolution cell(pixel) manifest itself as pixel-to-pixel variation in the form of granular noise pattern known as speckle noise (Lee and Pottier, 2009). That is why a radar image typically appears noisier than an optical image as shown Figure 2.6.

For proper interpretation and understanding of a SAR image, it is essential to suppress this speckle noise. Speckle reduction approaches can fall into two basic categories known as multilooking and filtering. Multilook processing can be accomplished during the formation of the images by averaging N number of single look images. The number of single look images is obtained by dividing the synthetic aperture length into N parts where each part regarded as a look. Each part is independently processed yielding single look images. Reduction of speckle noise is directly proportional to the number of looks. As the look number increases, the degree of reduction increases but at the expense of the image resolution (Lee and Pottier, 2009). On the other hand, filtering process can be performed after the images are formed by any available noise suppression algorithm such as Frost, Lee and Kuan filters which are commonly used in SAR images (Lee, 1983; Frost et. al., 1982; Kuan, et. al.,1985). The newer algorithms such as SRAD (Yu and Acton, 2002), FPD (Okman, et. al, 2012), SAR-BM3D (Parrilli, et. al 2012) are also effectively used for noise suppression on SAR images. These filters as well as the others have tendency to reduce the noise but at the same time degrade the resolution of the images to a lesser or a greater degree.

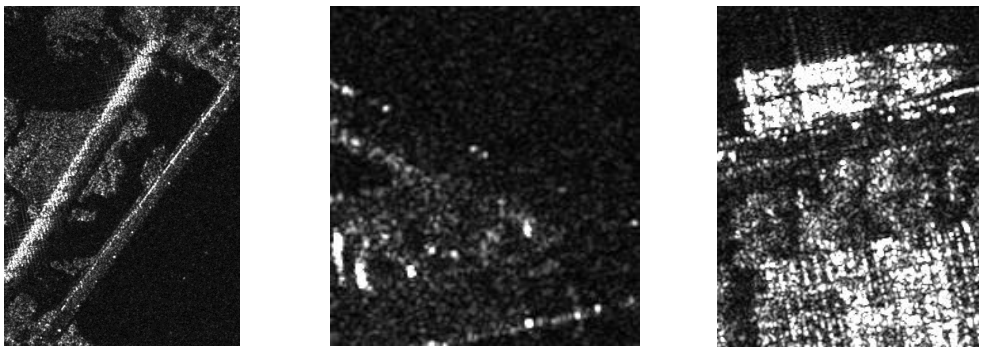


Figure 2.6 Example of High Resolution Spot TerraSAR-X Image Which is Taken from Vishakpatnam / India.

CHAPTER 3

BACKGROUND AND LITERATURE REVIEW

The core of the image segmentation process is to subdivide an image into meaningful homogenous regions that are made up from connected similar pixels according to their features such as intensities and spatial attributes. This type of segmentation is an important step that provides meaningful and useful information at an early stage of a Synthetic Aperture Radar (SAR) image analysis. In SAR images, the need for a reliable segmentation is crucial due the existence of multiplicative noise called speckle noise that highly degrades the image quality.

There are many ways that one can classify the segmentation techniques such as graph, nongraph, gradient, region, local and global based approaches some to mention. Since the work in this thesis is related with the superpixels, the segmentation algorithms can be broadly classified as superpixel and nonsuperpixel-based algorithms. In this chapter, following a brief review of the existing nonsuperpixel-based algorithms; the superpixel based algorithms are presented in detail to offset the stage for algorithms proposed in this thesis.

3.1. Nonsuperpixel-Based Algorithms

There are various approaches used for segmentation all of which can be considered as nonsuperpixel techniques. These can, further, be broadly classified as deterministic and stochastic models. In the deterministic approaches, the pixels intensities are used as discrete variables whereas in the stochastic models, they are used as random variables. In either case, the values of the entities are determined by the model parameters. Different variants of thresholding, graph-based

segmentation, region-based segmentation and k-means clustering are some widely used deterministic techniques in image analysis.

Thresholding followed by connected component analysis is the simplest image segmentation technique that can be used either in global or local scale. While the global approach is based on the histogram of the entire image, the local scale approach is based on local information such as co-occurrence matrix (Gonzales and Woods, 2002)

Graph-based methods partition an image into segments by minimizing an energy function. The approach is essentially a maximum flow problem where max-flow/min cut solution over the images represented as graphs are the partitioned segments (Greig et al., 1989; Ford and Fulkerson, 1956).

Region-based methods are another class of segmentation approaches all of which are local in nature. Among them, region growing is the best known. In this method, segmentation starts with a set of selected pixels as seed points and then the regions are grown from these seed points with a criterion on the correlation of pixels in that region (Gonzalez and Woods, 2002; Master and Franke, 1988).

The k-means clustering is the process of grouping pixels of an image into a small number of clusters by minimizing an objective function which is in the form of Euclidean L2 norm. The objective function represents the sum-of-the-squares of the distances of each pixel to its assigned center pixel. At each iteration of the minimization scheme of the algorithm, the center pixel of the cluster is taken as the mean of all the pixels belonging to that cluster (MacQueen, 1967).

Bayesian segmentation based on the Markov Random Field (MRF) is a stochastic technique that has been used extensively over the last couple of decades (Lankoande et. al., 2005; Besag, 1986; Li, 2009; Geman and Geman, 1984; Fujortoft et. al., 2003). This pixel labeling approach formulates a likelihood function about the given image data and a prior term based on spatial information of the pixels. In this type of formulation, information regarding the pixel

intensities as well as their locations are used to determine the label of the pixels. Specifically, the state of a given pixel is entirely determined by the states of its neighboring pixels. This local dependency is the basis of Markovian property.

The stochastic models are not well suited for the cases where the features of the image are not well represented with a single distribution. The inconsistency between the image data and the single probability distribution can be avoided by using a linear combination of the distributions known as finite mixture models. Gaussian mixture model as one of these models is probably the most widely used one (McLachlan and Peel, 2000; Blekas et. al, 2005; Nguyen and Wu, 2012). The image segmentation can be performed after the determination of the model parameters. Expectation-Maximization (EM) algorithm is a highly popular technique that is used for the estimation of mixture parameters (Dempster et. al., 1977; Celeux et. al., 2003; Carsen et. al., 2002). In addition to the estimation of mixture parameters, EM can also be used as a stochastic framework for segmenting the images.

3.2. Superpixel-Based Algorithms

In clustering, a homogenous region can be obtained by merging subhomogenous regions which is more efficient than dividing a large region into smaller homogenous regions. Hence, the clustering algorithms utilizing superpixel-based algorithms as a preprocessing tool divide the image into oversegmented homogenous areas with high boundary adherence. For this reason, superpixelling has emerged as an oversegmentation approach over the last decade. Essentially, this type of segmentation is based on partitioning an image into subregions that are made up from set of connected similar pixels. These connected similar pixels are the elementary units of compact and uniform superpixels rather than the pixels of the image (Bugden et. al., 2004). This superpixelling process is fundamental for various applications of SAR image analysis since an effective segmentation can reduce image complexity and enhance interpretation and understanding of the

segments as shown in Figure 3.1. In this figure, the left most image is the original SAR image taken by TerraSAR satellite, middle image shows the boundaries of the superpixels and in the right most image superpixel with intensity means are shown.

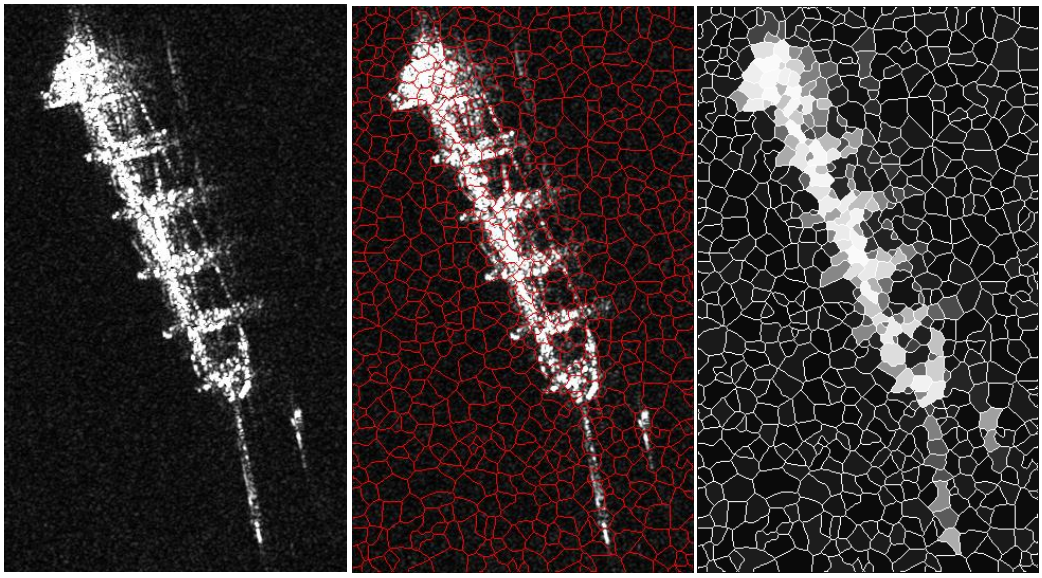


Figure 3.1 An Example of Superpixelling on SAR Images

Since superpixels are groups of connected pixels with similar features, they are perceptually meaningful and representationally efficient segments. They are perceptually meaningful because they are multiple segments of the image where each is composed of uniform pixel intensities. They are representationally efficient because they are multiple segments of the image where each can be visualized as separate entities by their defined boundaries. As boundaries of the objects are important elements to be traced, superpixels should adhere well to the boundaries of the segmented objects. If a proper adherence of the boundaries can be achieved, such a superpixel based segmented image can reduce the complexity in understanding and interpretation of the subsequent tasks. Due to the boundary adherence property of the superpixels such as shown in Figure 3.2, they are widely used in recent years. If the generated superpixels are clustered as a further image processing task, a final segmentation of the image can obtained.

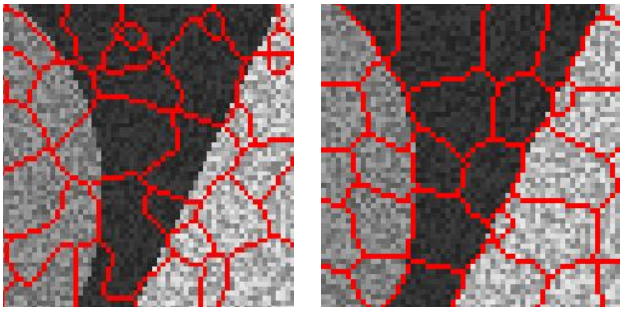


Figure 3.2 Boundary Adherence Example. Left image shows poor boundary adherence whereas right image shows better boundary adherence.

There are various algorithms to generate superpixels. The state-of-the-art performance of the algorithms can be addressed by the evaluation of their properties and the quality of the segmented images (Zhang, 2014). Algorithmic complexity and efficiency are the two major attributes to evaluate their properties. While the complexity is very much related with the simplicity in understanding and implementing the algorithm, the efficiency depends on the computational requirement. The quality of the segmented images can be explored by subjective or objective measures. The subjective measures are based on visual comparison of the partitioned images. The objective measures are related with certain metrics to determine the degree of boundary adherence and compactness of the superpixels generated. The two standard metrics which are widely used are the boundary recall and the undersegmentation error. Detailed descriptions of these two metrics are presented at the end of this chapter. As a summary, we can list the measures to evaluate the segmentation performance of the algorithms as follows:

- Boundary adherence and compactness,
- Computational complexity and efficiency,
- Visual appearance.

The aim of superpixeling is to group neighboring pixels into uniform regions by considering compactness. If compactness is not considered, highly irregular shaped and sized superpixels would be obtained.

Similarity and proximity has long been recognized as the two important factors for perceptual grouping and organization of the pixels of an image (Shi and Malik, 2000; Ren and Malik, 2003; Felzenshwalb and Huttenlocher, 2004; Wertheimer, 1938; Sarkar and Boyer, 1992). According to Wertheimer (1938), pixels with similar properties will be perceived as belonging to a group and pixels that are close to each other will appear also as a group. As a result of this type of grouping, an image can be segmented into set of partitions where each set consists of similar and proximate pixels. To clarify the way that these terms are used in this research, the similarity is meant to be the pixels with similar intensities (radiometric similarity) and the proximity meant to be the closeness of the pixels defined in terms of the geometric distance, that is, the spatial proximity between the two pixels of interest. Since similarity and dissimilarity can be regarded as complements of each other, then it is appropriate to use any of these two terms based on the context.

Among the early studies, it seems that the development of the normalized cut criterion by Shi and Malik (2000) is a significant advancement in generating perceptually grouped superpixels. This perceptual grouping algorithm is based on minimizing a "normalized cut" cost function to obtain optimal partitions of the image. The basis of this approach is graph partitioning method. In the graph partitioning methods, the image to be partitioned is represented as a weighted undirected graph consisting of pixels as nodes and boundaries as edges. Then the weight on each edge is formulated as a function of similarity between the nodes such as difference in intensities at those nodes. These edge weights are used to partition the image. In contrast to the use of the total edge weight connecting the two partitions in graph partitioning problem, the normalized cut algorithm computes the total cost defined as a function of edge connections over all the nodes in the graph. Optimal partitioning of the image is the one that minimizes this cut cost function. If the edge weight is defined as $w(i,j)$ between two pixels as nodes i and j, then the normalized cost function is obtained as the product of a similarity and spatial proximity terms that is formalized as,

- If $\|X(i) - X(j)\| < r$ pixels

$$w(i,j) = e^{-\|F(i)-F(j)\|^2} \times e^{-\|(X(i)-X(j)\|^2}$$

otherwise 0

where X and F designates the pixel locations and their features respectively.

The approach does not exhibit good boundary adherence and it is computationally slow in obtaining optimal partitions. However, it has been found that the results of the experiments on real and synthetic images are encouraging in extracting the objects from the image scenes. As a pioneering work in the area of superpixel segmentation, it does not include performance evaluation of the algorithm as compared with others.

Veksler et.al, (2010) formulated the superpixel generation problem in an energy minimization framework. An energy function is expressed as a weighted sum of data and smoothness terms to be minimized with the graph cut algorithm. Data term is a unary constraint that imposes the label to be constrained to the pixels of the image and smoothness term is a binary constraint that encourages smoothness among the labels of the neighboring pixels. This type of segmentation is essentially similar as the energy formulation in Markov Random Field for pixel labeling which consists of a likelihood function and smoothness of the pixels. The coefficient $w(i,j)$ in the smoothness term for the nodes i and j is related to the gradient magnitude between pixels i and i which is formulated as

$$w(i,j) = \frac{e^{-(I(i)-I(j))^2}}{dist(i,j)(2\sigma^2)} \tag{3.1}$$

The exponent term is the gradient magnitude between the pixels i and j where $I(i) - I(j)$ is the intensity difference that can be regarded as the similarity term and $dist(i,j)$ is the Euclidean distance that can be considered as the spatial proximity term for the neighboring pixels i and j. Explicit presentation of an energy function encourages regular superpixels obtained by means of placing

overlapping patches over the image and assigning each pixel to only one of those patches. In this work, two variants of the method are proposed as variable patch and constant intensity superpixel algorithms. The results of the experiments are compared in terms of boundary recall, undersegmentation error and computational efficiency.

Turbopixel approach developed by Levinshtein et al. (2009) is essentially a curve evolution technique to obtain superpixel boundaries. This curve evolution is a progressive process implemented by dilating a given number of seeds in an image using level-set based geometric flow. Dilation of those given image seeds is one of the essential steps in the algorithm. The level set of a smooth and continuous function is evolved by embedding as Euclidean distance between each pixel in one region to closest point on the boundary of another region. The algorithm generates visually pleasing compact superpixels that are conformal to lattice-like structure of the regions in an image. Uniformity of the boundaries of the superpixels is achieved by “proximity-based boundary velocity” and “image based boundary velocity" terms and the compactness of the superpixels is secured by placing seeds in a lattice formation where the distances between lattice neighbors are all approximately equal. Although the turbopixel framework allows superpixels to be constrained to have uniformity and compactness, this can be achieved at the expense of reduced undersegmentation and high computational cost. The superpixels generated by the algorithm exhibit low adherence to boundaries and the algorithm is slow in practice. Segmentation performance of the algorithm is evaluated in terms of boundary recall and undersegmentation error metrics with emphasis on the computational requirements.

Lattice approach proposed by Moore et al. (2008) is a method that produces superpixels conformal to a grid-like structure. Conformity of the superpixels to regular lattices is achieved by finding optimal paths that separates an image into vertical and horizontal regions. Regular lattice is guaranteed by finding optimal horizontal and vertical paths which cut the image. These optimal paths are

obtained from the given "boundary cost map" which is the major input to the algorithm. The superpixels are constructed incrementally by using these computed optimal paths. Edge weights in terms of features of those pairs of neighboring pixels are the basis of constructing optimal paths. The quality of the segmented image and the computational complexity of the algorithm are strongly affected by the pre-computed boundary cost maps. The algorithm is evaluated qualitatively and quantitatively. While qualitative evaluation is based on visual inspection, quantitative evaluation is based on an explained variation and accuracy as two metrics introduced by the authors. Based on these metrics the generated superpixels are conformal to a grid structure.

The approach proposed by Felzenshwalb and Huttenlocher (2004) is a graph-based method using a pairwise region comparison predicate. In graph partitioning methods, the image to be partitioned is represented as a weighted undirected graph consisting of pixels as nodes and boundaries as edges. Then, the weight $w(i,j)$ for each edge is based on the absolute difference between the pixels connected by that edge is formulated as

$$w(i,j) = \left| I(i) - I(j) \right| \tag{3.2}$$

where $I(i)$ and $I(j)$ are the intensities of pixels i and j. This difference function intensities is the dissimilarity term for relevant pixels. These edge weights are used to partition the image. A predicate is defined to explore the existence of a boundary between the two regions in an image. This predicate is built upon pairwise comparison of the dissimilarities between these elements along the boundary of the two regions and the elements within those regions. For this purpose, maximum edge weight in the minimum spanning tree of the region and minimum edge weight of the two regions are combined to be used for pairwise comparison for clustering. This clustering is performed in an agglomerative manner that produces regions with highly irregular shapes and sizes. However, the approach is capable of finding segments that capture many perceptually important

aspects of complex imagery. As normalized cut, this algorithm can also be regarded as another grouping algorithm for superpixel segmentation.

Mean shift (Comaniciu and Meer, 2002) is a nonparametric iterative mode seeking clustering approach. The essence of mode seeking technique is the convergence of all pixels to the same mode. In the mean shift algorithm, local maxima of a density function is aimed in a complex structured data space that delineates irregularly shaped clusters. As the pixels converge to the same mode, the superpixels are generated. In this approach, the number of clusters does not need to be known in advance. Since the structure of the generated superpixels may be arbitrary, the shapes of the superpixels are highly irregular with non-uniform sizes. The algorithm is evaluated qualitatively based on visual inspection.

Quick shift (Vedaldi and Soatto, 2008) is also a mode seeking clustering algorithm as mean shift. In this work, medoid shift, which is another mode seeking clustering algorithm in pattern recognition, is used instead of the mean shift. It has been shown that medoid shift is computationally faster than the mean shift. Unlike the mean shift, the quick shift searches for the mode as each pixel moves to the nearest neighbor for which there is an increment in the Parzen density. In this respect, the quick shift differs than the mean shift where the increment in density estimate is determined by the gradient to locate its mode. The approach produces superpixels with relatively good boundary adherence. The algorithm is evaluated by comparing its efficiency with the mean shift procedure and its variant, mean-medoid shift.

The simple linear iterative clustering (SLIC) algorithm proposed by Achanta et al. (2010) is based on searching for superpixels over a specified set of grids of sized S with the k-means algorithm. The basis of SLIC algorithm is a distance measure which is a linear combination of color and spatial proximities (actually dissimilarities) expressed in Euclidean norms formulated as

$$D_i = \sqrt{(I_k - I_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} + {}^{m}/_{S}\sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{3.3}$$

where the first is the color difference term, the second is the spatial distance term and m is the "compactness parameter". This distance measure imposes color dissimilarity as well as spatial distance in five dimensional space where two of them are Cartesian coordinates and three of them are CIELAB (lab) color space. In this sense, this algorithm, as normalized cut, can be regarded as another grouping algorithm. In this work, it has been demonstrated that it outperforms many of the above mentioned existing methods in terms of boundary adherence, segmentation speed, segmentation accuracy, control over superpixel compactness and control over the amount of superpixels.

The iterative edge refinement algorithm (Zhu et. al., 2015) is based on initializing the superpixels as regular grids first and then refining them iteratively by relabeling the pixels on the edges. In the method, the edge pixels on the grids are divided into stable and unstable pixels according to a criterion. The stable pixels are considered to be unchanged in the subsequent iterations and only the unstable pixels are relabeled. In the core of this algorithm, distance measure of SLIC algorithm with the default values of the parameters is used. In this sense, it seems that it is a variant of the SLIC algorithm.

The entropy rate superpixel segmentation (ERS) algorithm (Liu et. al. 2014) produces superpixels by optimizing an objective function which is linear sum of entropy rate of a random walk and a balancing term. Entropy rate defined over a graph is used as a criterion to obtain compact and homogeneous clusters. This rate is defined between the neighboring pixels of the image as nodes of the grid structure. The edge weights correspond to the similarities between the pair of pixels that can be represented as a similarity matrix. In case highly irregularly

shaped superpixels are generated and morphological algorithms are necessary to smoothen them.

The SEEDS algorithm developed by Van den Bergh et al., (2012) uses a similar optimization based on an energy function. This function consists of color likelihood term that favors colors uniformity and boundary shape term that encourages smooth boundaries. Color density distribution and the boundary term are approximated from the histogram of the superpixel labels. The optimization uses a hill-climbing approach in which the superpixels are refined as the objective function increases. This approach suffers from high shape irregularity and the complicated control of the number of superpixels.

In the last three algorithms, boundary recall and undersegmentation metrics are used to evaluate their segmentation performances. In Zhu et al. (2015), it has been found that iterative edge refinement and SLIC algorithms produce more compact and visually pleasing superpixels than those in Liu et al. (2014) and Van den Bergh et al.(2012). In addition, iterative edge refinement and SLIC algorithms exhibited similar segmentation performances.

All the above mentioned algorithms use, in one way or other, the similarities (or dissimilarities) between pairs of the pixels. The normalized-cut algorithm, the algorithm proposed by Felzenshwalb and Huttenlocher, the algorithm of lattice superpixels and the entropy rate algorithm use edge weights defined as the intensity differences of the pixels. Whereas, the algorithm proposed by Veksler et al., mean shift and quick shift algorithms employ gradient functions expressed as the intensity differences of the pixels. However, SLIC and turbopixel algorithms are based on the similarities defined in terms of the difference in the intensities of the pixels which are expressed in Euclidean norms. On the other hand, spatial proximities between the pairs of pixels are utilized in the edge weight functions of normalized-cut, in the gradient functions of that proposed by Veksler, et al. and in the distance measure of the SLIC algorithm. Since similarity and spatial proximity as two factors of perceptual grouping are used explicitly in the last three

algorithms, then the generated superpixels with these algorithms would be expected to be perceptually meaningful and representationally efficient.

From the above review of the existing superpixel based algorithms, it is evident that SLIC algorithm stands as the state-of-the-art superpixel segmentation algorithm in terms of boundary adherence and computational complexity.

3.3. Segmentation Performance Evaluation Measures

Boundary adherence and compactness, computational complexity and efficiency, and visual appearance are listed as the most commonly used segmentation performance measures in earlier in this chapter. Among those, the boundary adherence and compactness are related with how well the superpixels are adhered to the image boundaries in a regular and smooth manner; computational complexity and efficiency indicates how time and memory are used efficiently in understanding and implementing the algorithm.

The quantitative evaluation of the segmentation performance of superpixel generating algorithms depends on the boundary adherence and compactness of the generated superpixels. For this purpose, there are two widely used standard metrics: boundary recall and undersegmentation error. These metrics are extensively used to compare the segmentation performance of the existing superpixel algorithms. In order to use these two metrics, an additional human segmented image is required besides the superpixel segmented image. This manually segmented ground truth is essential for detecting the tightness of superpixels to image boundaries. If this ground truth is not available, which might be the case in SAR imagery, heavy reliance on visual appearance is obvious. For example, SAR image in Figure 2.6 is a heavily speckled image in which one can hardly visualize the anchored ship at the port. For this type of image, ground truth is extremely difficult to obtain in absence of which performance evaluation can solely be based on visual comparison.

If used alone, boundary recall or segmentation error do not yield sufficient information for the performance of the algorithms. For example, a segment with a high boundary recall exhibiting good boundary adherence might have low segmentation error. Thus, the boundary recall, if applicable, should be used together with undersegmentation error in order to evaluate how well the superpixel boundaries are adhered to image boundaries with a reasonable segmentation error.

When boundary recall and undersegmentation error used together with the other performance measures, the outperforming capabilities of the algorithms in generating the superpixels can be analyzed.

3.3.1. Boundary Recall Metric

To measure the extent of the adherence of the superpixel boundaries to image boundaries, the boundary recall is an appropriate standard metric to be used. This metric can be stated as the fraction of the ground truth edges falling within a prespecified, say *e*-pixels, strip along the superpixel boundary. Thus, it is the percentage of the ground truth border pixels within *e*-pixel vicinity of those superpixels. This measure as formulated by Neubert and Protzel (2013) as,

$$BR = \frac{TP}{(TP + FN)} \tag{3.4}$$

where

BR = Boundary recall,

TP = True Positives as the number of ground truth edge pixels which are falling within *e*-pixel strip along superpixel boundary and,

FN = False Negatives as the number of ground truth edge pixels which are away from *e* pixel strip along the superpixel boundary.

The pixels which are true positives and true negatives are illustrated in Figure 3.3.

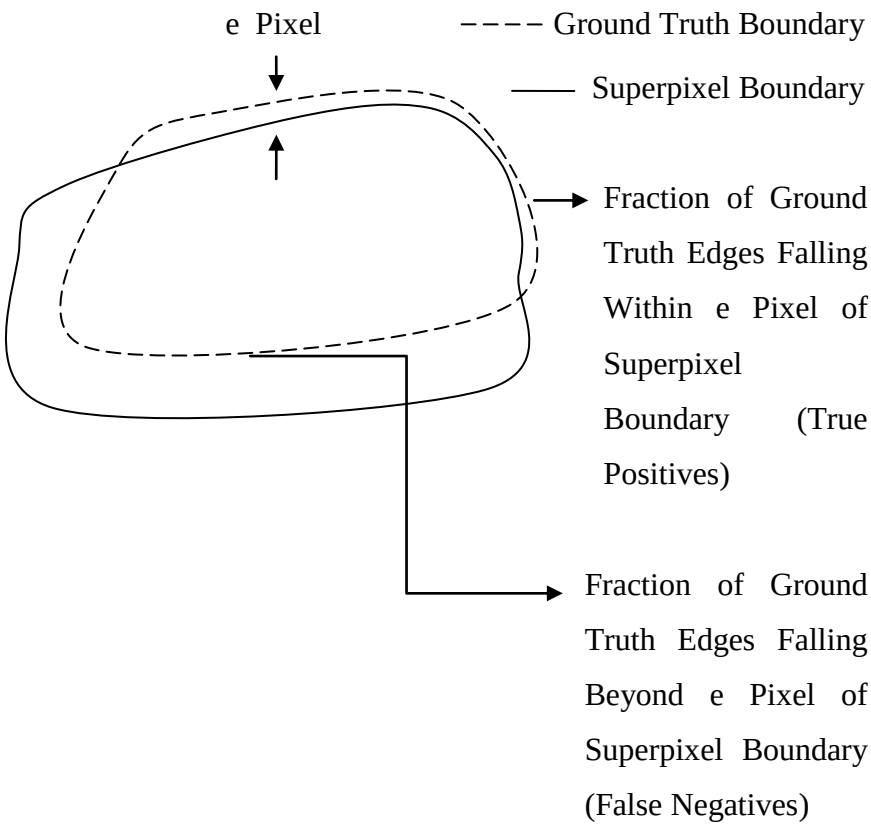


Figure 3.3 Illustration of Boundary Recall

3.3.2. Undersegmentation Error Metric

Undersegmentation error measures the extent of superpixel areas not overlapping the ground truth. Hence, it is an error that an algorithm makes in segmenting an image with respect to the ground truth of that image. Given ground truth segment areas as $G_i, \dots\dots, G_k$ and superpixel segment areas as $S_i, \dots\dots, S_m$, this measure can be formulated as in Levinshtein, et al. (2009):

$$UE = \sum \frac{\sum(S_j - G_i)}{G_i} \tag{3.5}$$

where $\forall S_j \cap G_i \neq \emptyset$ and UE is the undersegmentation error.

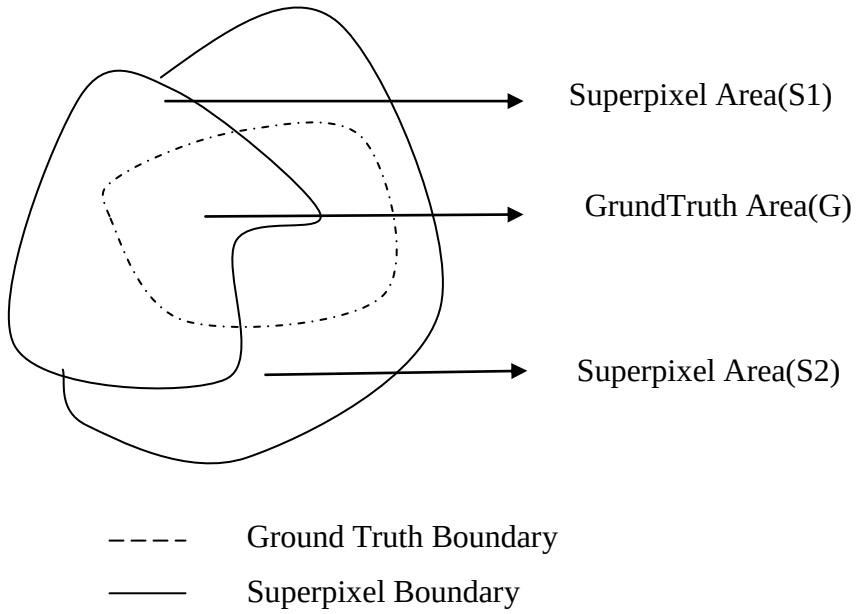


Figure 3.4 Illustration of Undersegmentation Error

For the example given in Figure 3.4 where there are two superpixels and one ground truth segment, undersegmentation error can be calculated as

$$UE = \frac{(S1 + S2 - G)}{G} \tag{3.6}$$

CHAPTER 4

MAHALANOBIS SPATIAL PROXIMITY METRIC

The way the similarity and the spatial proximity terms used within k-means clustering framework, SLIC algorithm generates perceptually meaningful and representationally efficient superpixels. To exploit the simplicity of k-means clustering, this approach is adapted in this work similar to SLIC algorithm.

The authors of SLIC investigated the segmentation performance of the algorithm by replacing the distance term with adaptively normalized and geodetic distance measures as two variants of the algorithm. It has been found that the distance measure as in SLIC algorithm outperforms the above proposed variants.

Due to its circular decision boundaries, the spatial Euclidean distance forces the superpixels to have a round shape. However, spatial Mahalanobis distance with elliptic decision boundaries adds more flexibility that lets the superpixels to fit to arbitrary shapes of the real world without undermining the regular distribution of the superpixels within the image.

In this chapter, the segmentation performance of the SLIC algorithm would further be investigated by replacing the spatial Euclidean distance term with Mahalanobis distance (MSLIC). After an overview of the SLIC algorithm is presented, the formulation of MSLIC is described. In order to show the segmentation performance measures, the experiments are conducted on Berkeley Data Set whose results are presented in subsequent chapters.

4.1. A Brief Overview of the SLIC Algorithm

Among superpixel generating existing algorithms in the literature that are discussed previously, SLIC stands out as the state-of-the-art superpixel generating algorithm. This algorithm is essentially an adaptation of k-means clustering algorithm. The basis of the objective function used in the k-means clustering is a linear combination of color dissimilarity and spatial distance terms. The color difference in the form of Euclidean distance in Lab color space is

$$D_C = \sqrt{(I_k - I_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \tag{4.1}$$

and the spatial Euclidean distance term is

$$D_S = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{4.2}$$

To combine the two Euclidean distances into a single measure, the spatial distance term is normalized with the cluster size S whereas a normalizing constant m ranging between 1 and 40 is used instead of normalizing the color difference term with its respective maximum value that differs significantly from cluster to cluster. The resulting expression is

$$D_i = D_C + {}^{m}/_{S}\, D_S \tag{4.3}$$

to be used over the search area.

In this formulation,

- i is the cluster center,
- x and y are the Cartesian coordinates,
- k is the pixel to be classified,
- l, a and b are the colors in CIELAB color space,

- m is the compactness parameter and
- S is the grid size.

The constant m regulates the relative importance between the color and the spatial distances. For large values of m, highly regular and compact superpixels would be obtained. On the other hand for small values of m, irregular shaped superpixels with high boundary adherence are obtained.

For SAR images, color difference term can be adapted as $D_c = \sqrt{(I_k - I_i)^2}$ resulting a distance measure as

$$D_i = \sqrt{(I_k - I_i)^2} + {}^{m}/_{S}\sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{4.4}$$

In this equation, the first term is the intensity difference as pixel dissimilarity between the pixel to be labeled and the centroid of the cluster and the second term is the spatial distance between those pixels based on the Euclidean metric. The equation enforces color similarity and spatial proximity such that the spatial extents of the clusters are approximately equal.

The above objective function should be minimized iteratively within the search areas of equally sized clusters to generate perceptually uniform superpixels. As the function is minimized, pixels based on the intensity similarity and the spatial proximity are clustered.

4.2. The k-means Clustering as Adapted in SLIC

The traditional k-means clustering is based on partitioning an image data into k number of specified clusters so that the within-cluster sum of the squares is minimized (MacQueen, 1967). For an image with n number of pixels, the algorithm seeks the centroids that minimize

$$\sum_{j=1}^{k} S_{\mathrm{j}} \tag{4.5}$$

where the within-cluster sum-of-squares for cluster j is

$$S_j = \sum_{i=1}^{n} (l_k - l_i)^{\,2} \tag{4.6}$$

in which l_k is the mean intensity of the cluster k and l_i is the intensity of the pixel i to be clustered. In this case, each labelled pixel belongs to only one cluster with the nearest mean.

The k-means clustering is probably the simplest and the widely used algorithm among those that are based on minimizing an objective function. Minimizing the sum of the distances to a nearest cluster centroid as an objective function is the essential step of the algorithm. Although there are no closed form solutions for the k-means problem to determine the local minima, an iterative minimization scheme is one popular heuristics to solve such problem. The main idea of this scheme is to classify the given data set through a predefined k number of clusters with one centroid for each cluster. In the iterative scheme, each point belonging to the given data set is associated to the nearest centroid. This type of procedure for finding the local minima turns the k-means clustering approach into an easy and well understood unsupervised classification algorithm. The basic steps of this heuristic approach consisting of an assignment and update steps to be followed iteratively can be summarized as follows:

- Assign l_i to cluster k if $(I_k - I_i)^2 \leq (I_k - I_j)^2$ for each pixel i of the image
- Update l_k as $\sum l_i$ / (number of pixels clustered as k).

The way the k-means clustering approach adapted in SLIC algorithm differs from traditionally used ones in two respects. First, it incorporates a distance measure that can be regarded as an objective function that is a linear combination of color

difference term and spatial distance measure, and second, the minimization is performed over a search area of size 2Sx2S that is proportional to the cluster size S. This proportionality between the search area and its cluster can be illustrated as follows:

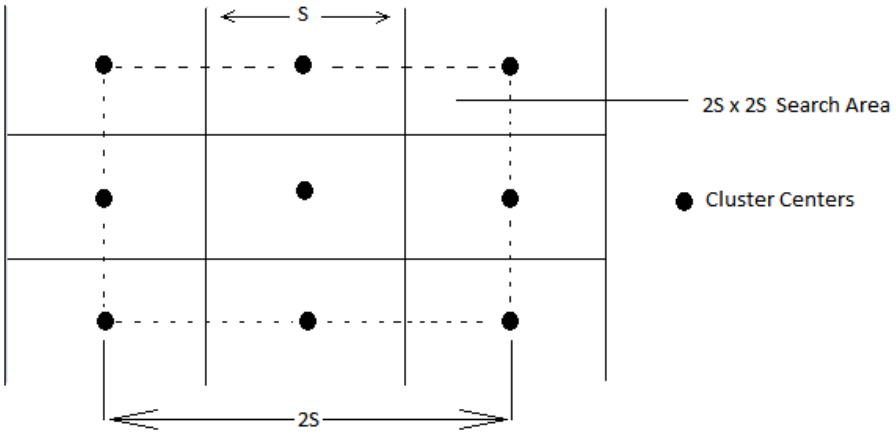


Figure 4.1 The Search Area and the Cluster Size S

Searching the data set within the limited areas rather than the whole data set appears an attractive property of the SLIC algorithm. In this manner, the simplicity of the k-means algorithm is further expanded and the grid structure is preserved.

In minimization sense, k-means clustering as adapted in SLIC is similar to the standard k-means clustering. The minimization of the objective function is over 2Sx2S sized overlapped search areas rather than the whole grid of the image data. Since each cluster has 2Sx2S sized search area, it has eight partially overlapping search areas each of which belong to the eight neighboring clusters of that given cluster.

At each step of the iterative minimization scheme, the pixels that have the minimum distance metrics to a given cluster center are assigned as that centroid. This can be summarized as,

- Assign l_i to cluster k if $(D_i) \leq (D_j)$ for each pixel i within the search area 2Sx2S.

In the way the distance measure is formulated as above, it can be considered as the k-means objective function. A similar minimization is performed until all the clusters with partially overlapping search areas are analyzed. This completes the first iteration of the algorithm. After the cluster centers are updated as $l_k = \sum l_i / k$ cluster size, the next iteration is continued. This minimization scheme given in SLIC is simplified as:

Table 4.1 Algorithm Framework of SLIC

Initialize cluster centers $G_i = [I_i, x_i, y_i]$ by sampling at regular grid size S.
Set all $I_i = \mu_i$ and label $l(i) = -1$ and label $d(i) = \infty$ for each pixel i
for $j = 1$ to number of iterations
for $i = 1$ to number of clusters,
compute D_i over 2Sx2S area around G_i
if $D_i < d(i)$, set $d(i) = D_i$ and $l(i) = k$, then
update cluster centers and superpixel means

4.3. Formulation of the MSLIC Approach

As an alternative to the distance metric in SLIC, in this thesis another distance measure, Mahalanobis is proposed. This distance metric is also widely used in cluster analysis and classification techniques. In contrast to Euclidean, the Mahalanobis distance metric automatically scales the coordinate axes. As the axes are scaled, distances are also scaled which does not require any normalization. The scaled distance measures can be expressed as

$$d_i = [z_x\ z_y]\sum^{-1}[z_x\ z_y]^T \tag{4.7}$$

where $z_x = x_k - x_i,\ z_y = y_k - y_i$ and $\sum^{-1}$ is the covariance matrix.

Since the Euclidean metric is normalized with S, the magnitudes of the distance range between zero and $\sqrt{2}$. To ensure the Mahalanobis distance values to be comparable with this range, these proximity values are need to be transformed to vary between zero and one. To ensure the proximity values to be less than the limiting values of the Euclidean distance metric, the magnitudes of the above Mahalanobis distance need to be transformed to vary between 0 and 1. This type of transformation can be done by first exponentiating Mahalanobis distances and then taking the inverses of the resulting expressions. Assuming equal color similarity terms at pixels i and j, the pixel k will be classified as pixel i if $d_i < d_j$where d_i and d_j are the distance measures for the pixels i and j, respectively. If both sides of the above inequality are exponentiated and then their inverses are taken, the resulting expression is obtained as

$$-e^{-d_i} < -e^{-d_j} \tag{4.8}$$

Thus, the proximity measure to be incorporated into distance measure, more formally, objective function would be $-e^{-d_i}$. The objective function developed to be used for MSLIC algorithm is:

$$D_i = \sqrt{(l_k - l_i)^2} + m\,d_i = |l_k - l_i| - m\,e^{-d_i} \tag{4.9}$$

CHAPTER 5

SIMILARITY RATIO BASED ALGORITMS FOR GENERATING SUPERPIXELS

The essence of the SLIC and MSLIC algorithms is an intensity difference, more formally intensity similarity, term combined with the Euclidean distance and Mahalanobis distance metrics as spatial distance (proximity) terms. The intensity similarity term, which is another Euclidean function in the radiometric space, has been demonstrated to be robust to additive noise but not to multiplicative speckle noise (Feng. et al, 2011). Hence, intensity difference metrics used in the algorithms are not robust enough for multiplicative noise inherent in SAR imagery. The sensitivity of these algorithms to speckle noise degrades the segmentation accuracy by hampering correct extraction of segments boundaries. Also this directly limits the segmented regions use in understanding and interpretation of SAR images. To avoid this drawback, a similarity ratio metric is proposed to be incorporated into the objective function. In this work, it has been demonstrated that the proposed approach is more robust to speckle noise in SAR images than those used in the SLIC and MSLIC algorithms.

In this chapter,

- the proposed similarity ratio is introduced,
- functional forms of the spatial proximity terms are defined,
- the way that these terms are formulated as an objective function is explained and
- the algorithmic framework within the minimization scheme of the objective function to generate superpixels is described.

In this thesis, the words cluster and grid, and the words measure and metric are used interchangeably.

5.1. Similarity Ratio

In spite of their merits, SLIC and MSLIC algorithms suffer from the multiplicative speckle noise inherit in SAR images. This is due to the sensitivity of the intensity similarity term in Euclidean norm to speckle noise. A similarity ratio metric as a more robust approach is proposed.

This metric has its basis on likelihood ratio. This ratio has been adopted for image segmentation by testing the hypothesis of the homogeneity of the segments within an image. In contrast the way the likelihood ratio adopted in this manner, the similarity ratio has no connection to any statistical testing and differs from those ratios used in image segmentation in two respects: its applicability in enhancing objective function and its suitability to measure the dissimilarity between clusters.

The likelihood ratio can be derived by assuming that two samples are drawn from a population. Suppose a sample X is $N(\mu_1 , \sigma_1^2)$, and another sample Y is $N(\mu_2 , \sigma_2^2)$. Let $(x_1, \ldots, x_m)$ be a sample space of X and $(y_1, \ldots, y_n)$ be a sample space of Y. If these two samples are drawn from population Z whose random space is $(x_1, \ldots, x_m, y_1, \ldots, y_n)$, then Z is $N(\mu , \sigma^2)$ as shown in Figure 5.1:

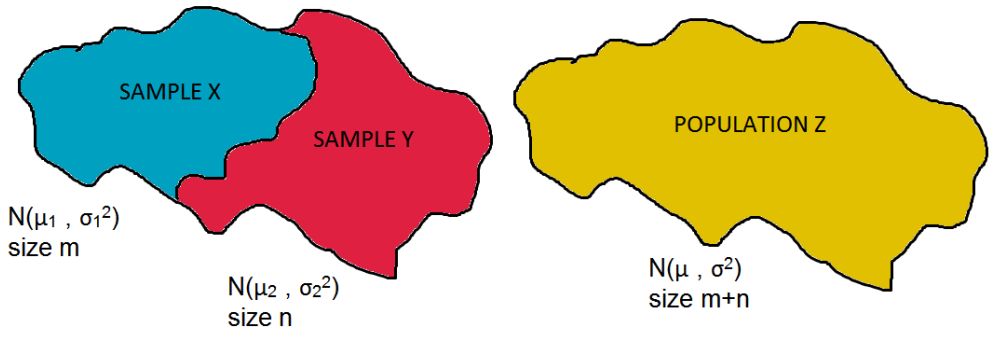


Figure 5.1 Illustration of Sample X, Y and Population Z

Moreover, if all random variables are assumed mutually independent then the likelihood functions for X, Y and Z would be as:

$$L_x = \prod_{i=1}^{m} \left(\frac{1}{\sqrt{2\pi{\sigma_1}^2}} * e^{-\frac{(x_i-\mu_1)^2}{2{\sigma_1}^2}}\right) = \left(\frac{1}{2\pi{\sigma_1}^2}\right)^{\frac{m}{2}} * e^{-\sum_{i=1}^{m}\frac{(x_i-\mu_1)^2}{2{\sigma_1}^2}} \quad (5.1)$$

$$L_y = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi{\sigma_2}^2}} * e^{-\frac{(y_i-\mu_2)^2}{2{\sigma_2}^2}}\right) = \left(\frac{1}{2\pi{\sigma_1}^2}\right)^{\frac{n}{2}} * e^{-\sum_{i=1}^{n}\frac{(y_i-\mu_2)^2}{2{\sigma_2}^2}} \quad (5.2)$$

$$L_z = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m+n}{2}} * e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu)^2+\sum_{j=1}^{m}(y_i-\mu)^2\right)} \quad (5.3)$$

The maximum likelihood estimates can be computed from the above likelihood functions as:

$$\mu_1 = \frac{1}{m}\sum_{i=1}^{m} x_i \ \ and\ {\sigma_1}^2 = \frac{\sum(x_i-\mu_1)^2}{m} \quad (5.4)$$

$$\mu_2 = \frac{1}{n}\sum_{i=1}^{n} y_i \ \ and\ {\sigma_2}^2 = \frac{\sum(y_i-\mu_2)^2}{n} \quad (5.5)$$

$$\mu = \frac{m*\mu_1 + n*\ \mu_2}{m+n} \quad and\ \ \sigma^2 = \frac{m*{\sigma_1}^2 + n*{\sigma_2}^2}{m+n} \quad (5.6)$$

The likelihood functions are simplified by using the maximum likelihood estimates as:

$$L_x = \left(\frac{1}{2\pi{\sigma_1}^2}\right)^{\frac{m}{2}} * e^{-\frac{m}{2}} \tag{5.7}$$

$$L_y = \left(\frac{1}{2\pi{\sigma_2}^2}\right)^{\frac{n}{2}} * e^{-\frac{n}{2}} \tag{5.8}$$

$$L_z = \left(\frac{1}{2\pi{\sigma_2}^2}\right)^{\frac{m+n}{2}} * e^{-\frac{m+n}{2}} \tag{5.9}$$

Since the two samples, X and Y, are drawn from the population Z, then their likelihood functions can be related as

$$\lambda = \frac{(L_x \times L_y)}{L_z} \tag{5.10}$$

where λ is the likelihood ratio which is the basic quantity used in hypothesis testing. If the likelihood functions above are placed into likelihood ratio formulation, the following is obtained:

$$\lambda = \frac{(\sigma)^{m+n}}{({\sigma_1}^n \times {\sigma_2}^m)} \tag{5.11}$$

Since superpixels are a group of similar neighboring pixels, it can be assumed that such segments are composed of fairly homogeneous areas. For a given SAR image, if the standard deviation to mean is constant (Lee and Pottier, 2009), then

$$\sigma_i = k\mu_i \tag{5.12}$$

where k is constant. Substituting this equation into the likelihood formulation, the constant k cancels out from this equation and the following similarity ratio is obtained.

$$R = \frac{(\mu)^{m+n}}{(\mu_1{}^m \times \mu_2{}^n)} \tag{5.13}$$

Since maximum likelihood estimation for the population is,

$$\mu = \frac{(m\mu_1 + n\mu_2)}{(n+m)} \tag{5.14}$$

Substituting this expression into Eq. 5.14, the final form of similarity ratio can be obtained as:

$$R = \frac{\left[\frac{(m\mu_1 + n\mu_2)}{(n+m)}\right]^{m+n}}{(\mu_1{}^m \times \mu_2{}^n)} \tag{5.15}$$

This ratio converges to unity from positive infinity as similarity between the means increases. Hence, it is a measure of the amount of deviation between the intensities of the superpixels, i.e., it is a measure of dissimilarity.

5.2. Proposed Similarity Ratio Based Algorithms

If the above similarity ratio for a pair of pixels is combined with the spatial distance terms denoted as P, the resulting distance measure would be

$$R(i,k) + \alpha P(i,k) \tag{5.16}$$

where

$$P(i,k) = \frac{\sqrt{(x_k - x_i)^2 + (y_k - x_i)^2}}{S} \tag{5.17}$$

for similarity ratio based Euclidean proximity algorithm (SREP), and

$$P(i,k) = e^{-d_i} \tag{5.18}$$

for similarity ratio based Mahalanobis proximity algorithm (SRMP) respectively.

In these expressions, the pixel i is the centroid of the cluster and pixel k is the one to be clustered. The parameter α is a regularizing parameter that would evidently be different for SREP and SRMP algorithms. These algorithm specific parameters should be determined by the user prior the implementation of the algorithm.

To exploit the simplicity of the k-means clustering algorithm, the above functions are used within reduced 2Sx2S sized search areas composed of S sized grids rather than over the whole grid of the data set. Within each search area all the pixels should be labeled at each iteration of the minimization scheme for k-means clustering to generate the superpixels. If regularizing parameters are properly defined, the boundary between the regions as well as the boundary between the grids are preserved during each iteration. The preservation of the boundaries around each grid is equivalent to the preservation of the structure of that grid. To illustrate the way the boundaries are preserved, let us consider the Figure 5.2. In this figure, assume that the pixel k will be classified either with cluster center i or j within the search area. Further, assume that the pixels k and i belong to the same region and pixels k and j belong to different regions separated by a region boundary. If pixel k is classified as pixel i, the region boundary is said to be preserved. Otherwise, the boundary between the grids of pixels i and j is said to be preserved so the structure of the grid of pixel i is preserved.

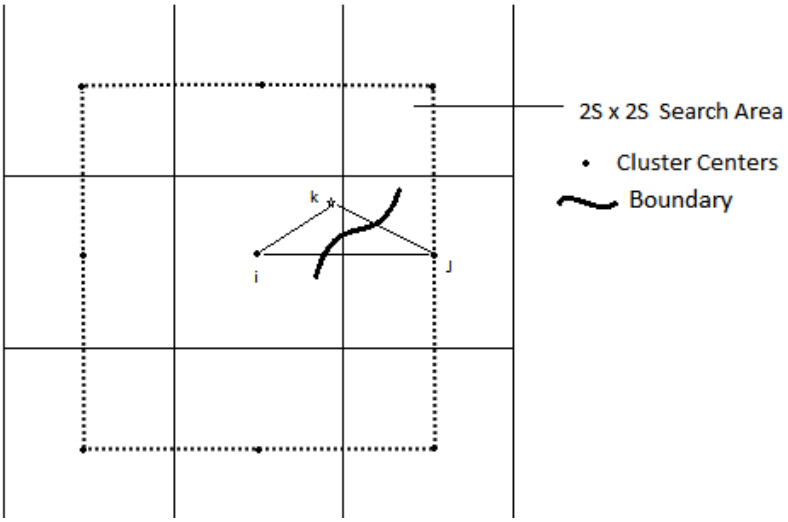


Figure 5.2 Boundary versus Cluster Centers

This structural relation between pixels i, j and k would be frequently referred throughout this thesis.

5.3. Algorithmic Framework for Generating Superpixels

The similarity ratio based algorithms proposed in the previous section can be used within 2Sx2S sized search areas composed of overlapping S-sized grids by utilizing k-means clustering approach.

This is the essence of the algorithms for the generation of the superpixels. At each iteration of the algorithm, the center pixels of the search areas are kept fixed but updated at the end of that iteration. All other pixels within the search areas are investigated for labeling. For the similarity ratio to be adaptive to this type of minimization scheme, the center pixel of the grid is assumed as one segment consisting of a single pixel at the initial iteration and a group of pixels including the pixel to be labeled as another segment. This level of grouping makes the similarity ratio as robust to speckle noise since the extracted similarities would be more reliable than those obtained using pair of single pixels. For computational purposes, 3x3 window as a group of pixels seems adequate that can be illustrated as follows:

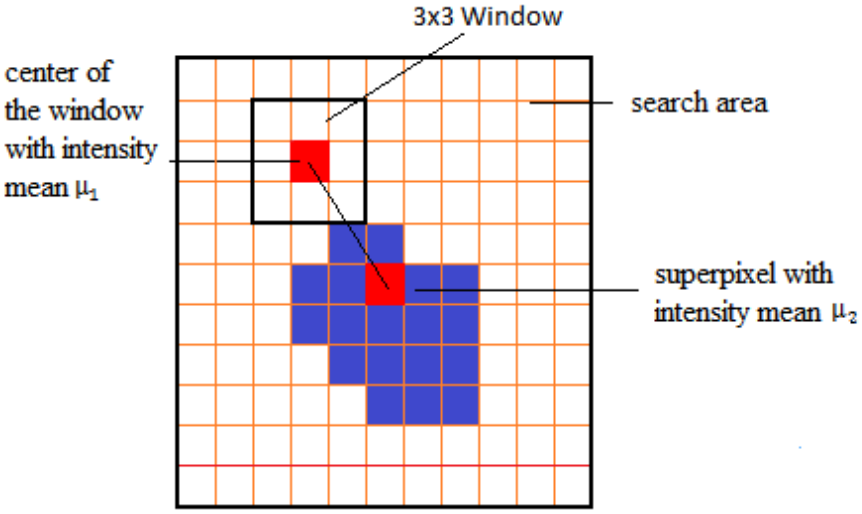


Figure 5.3 Window and 2Sx2S Search Area

After the center pixel of the grid is updated in the subsequent iterations, the center pixel of the grid is no longer assumed as one segment but its actual size for n is used.

The grids sized S are located within the search areas of size $2S$. The similarity ratio R_i is computed for the center pixel of the cluster i and the center pixel of the window k within the search area. The ratio computed in this manner is linearly combined with spatial proximity term. This computation is continued for the next window by sliding it a pixel to the right. Once, all distance measures are computed, the superpixels can be generated for the overlapping grids. For an image of size M, the number of clusters to be considered would be $K = \frac{M}{S^2}$. The complete algorithm is presented Table 5.1.

Table 5.1 Algorithm Framework

Initialize cluster centers $G_i = [x_i, y_i]$ by sampling at regular grid size S.
Set all $\mu_i = image(x_i, y_i)$
Set label image $l(k) = -1$ and distance image $d(k) = \infty$ for all pixels
for $j = 1$ to number of iterations
for $i = 1$ to number of cluster centers
for each pixel k in the 2Sx2S search area
compute μ_k as the mean of 3X3 window
compute R, P and $D_k = R + \alpha P$
if $D_k < d(k)$ then set $d(k) = D_k$ and set $l(k) = i$
update cluster centers and superpixel means

In this algorithm and in most of the existing approaches, the grid size S is a user-specified parameter that would be constant throughout the course of the algorithms. Hence, the proper choice of this parameter has great importance. As mentioned in Chapter 2, the spatial resolution varying within the scanned area affects the quality of the results of segmentation process. Since imaging SAR system operates in a side-looking geometry, the actual resolution is not constant. If slant-range images are used instead of ground-range images, the performance can be different in near range and far range. One such example is that the cluster with a specified size might allow the superpixels to be adhere well to boundaries of the homogenous regions in the near range whereas the boundary adherence highly hampered in areas that are in the far range. This seems one of the limitations of the superpixel generating algorithms in remote sensing.

The similarity ratios are in the form of exponentiated pixel intensities with the cluster sizes as exponents. The pixel intensities as well the cluster sizes are relatively large integer numbers. If not impossible, the calculation of the similarity ratios using these large integer numbers is extremely tedious and difficult task.

For example, pixel intensity of 150 raised to a power of 125 cannot be calculated in reasonable time. For this reason, it is more convenient to calculate the similarity ratios on logarithmic scale.

Similarity ratios for various pair of pixel intensities are calculated and tabulated in Table 5.2. For clarity of illustration, the ratios higher than one are shown in red color whereas the ratios lower than one are shown in blue color. The entries in this table correspond to the ratios between i or j center pixels of the clusters with size of $n = 1$ and the k center pixels of 3x3 windows with size of $m = 9$. The intensity values at eight bit format vary between 20 and 240 which is a typical range of values that can be experienced with a typical SAR image.

Table 5.2 Similarity Ratios for Various Pixels ***k*** to be Clustered with Pixels ***i*** or ***j***

i -j / k	20	40	60	80	100	120	140	160	180	200	220	240
20	0	0.260	0.725	1.237	1.755	2.263	2.754	3.229	3.681	4.116	4.534	4.935
40	0.180	0	0.082	0.260	0.481	0.725	0.979	1.237	1.497	1.755	2.011	2.263
60	0.409	0.066	0	0.040	0.135	0.260	0.404	0.561	0.725	0.893	1.065	1.237
80	0.607	0.180	0.035	0	0.024	0.082	0.164	0.260	0.367	0.481	0.601	0.725
100	0.776	0.298	0.103	0.021	0	0.016	0.056	0.113	0.182	0.260	0.345	0.435
120	0.922	0.409	0.180	0.066	0.014	0	0.011	0.040	0.082	0.135	0.194	0.260
140	1.050	0.512	0.260	0.122	0.047	0.010	0	0.008	0.030	0.063	0.104	0.151
160	1.164	0.607	0.335	0.180	0.088	0.035	0.008	0	0.006	0.024	0.050	0.082
180	1.266	0.694	0.409	0.239	0.133	0.066	0.027	0.006	0	0.005	0.020	0.040
200	1.360	0.777	0.478	0.298	0.180	0.103	0.052	0.021	0.005	0	0.004	0.016
220	1.445	0.851	0.544	0.354	0.228	0.141	0.082	0.042	0.017	0.004	0	0.004
240	1.524	0.922	0.607	0.409	0.274	0.180	0.113	0.066	0.035	0.014	0.03	0

If contrast is regarded as an attribute that describes the difference between the intensity levels of the pixels, then we can define a contrasting pixel pair as a pair of low and high intensities. In case the deviation between low and high intensities of a pair is substantial, the pair can be regarded as a high contrasting pair otherwise a low contrasting pair. From this definition, it is clear that low and high intensity pixels of the pair belong dark and light regions in an image, respectively. For example, the pixels with 20 and 160 intensities as a high contrasting pair have

a ratio of 3.23 which indicates that the former pixel might belong to a dark region and the latter pixel might belong to a bright region. However the pixels having 140 and 160 intensities which is a low contrasting pair have a ratio of 0.01 which indicates that they both might belong to a relatively homogeneous bright region.

If only similarity ratio is used as a metric to generate superpixels, highly irregularly shaped superpixels are produced which exhibit a high boundary recall that results good boundary adherence. Although this is desirable over heterogeneous areas, over homogeneous areas preservation of the squared structure of the grids is essential for convenient clustering. This can be clearly illustrated with the aid of the figure given in Figure 5.2.

If the pixel intensities for the points i, j and k are 20, 140 and 160 respectively, then pair $i-k$ is a high contrasting and $j-k$ is a low contrasting pixel pair as defined above with similarity ratios 1.16 and 0.01 respectively. If only similarity ratio is used, it is most likely that the pixels j and k having intensities 140 and 160 respectively are to be clustered since $R_i > R_j$. In this manner, the boundary between the pixels of the high contrasting pair is preserved. If the intensity of the pixel i would be 120 instead of 20, the pair $i-k$ would be a low contrasting pair with ratio 0.03. Again, if only similarity ratio is used, clustering of the pixels j and k would be expected. In this case, the boundary is again preserved. Since these low contrasting pairs are within relatively homogeneous area, the preservation of the squared structure of the grid is more essential then preservation of the boundary for convenient clustering of the superpixels to obtain a proper segmentation to be used for further analysis. This can be achieved by augmenting the similarity ratio with a spatial distance measure so that the pixels are clustered depending upon their proximities to the cluster centers within homogeneous regions. Referring to the above example, if the ratio of the pixel pair $j-k$ is augmented with a proximity measure, say 0.05, by virtue of $R_i < R_j + 0.05$, it would be expected that the pixel k is clustered with the pixel i. In this case, the grid boundary shown as a solid line in the given in Figure 5.2 is preserved.

5.4. Similarity Ratio Based Adaptive Mahalanobis Proximity (SRAMP) Algorithm

In SRMP approach, the parameter α is constant during the course of the algorithm. If the image is homogeneous in certain regions and heterogeneous in others, the parameter alpha being constant seems improper to produce compact superpixels in homogeneous regions and well boundary adhered superpixels in the heterogeneous regions. Since this value of alpha can be different for various parts of the image, a global parameter cannot be representative for all those parts. This is the main weakness of SRMP algorithm especially for the case of heavily speckled SAR images. For this purpose, a different value for each neighboring superpixel pair should be determined to obtain its optimum value from mean intensity values of the superpixels.

If parameter alpha is properly determined then the number of pixels misclassified would be minimum. The pixels are misclassified if they are clustered with other neighboring pixels rather than the one to be clustered. Evidently, the number of pixels misclassified which are spread over a region of two clusters would be different for various alpha values. The two clusters can be differentiated with the intensity levels of their cluster centers. Therefore, the alpha values can be expressed in terms of the intensity differences of those two clusters. If these alpha values correspond to the minimum misclassified pixels, then the parameter alpha is said to be properly determined for the two neighboring clusters.

To establish the relation between the two variables, namely parameter alpha and the intensity difference, an experiment is designed. For this design, a VV single polarized TerraSAR-X image and its manually derived ground truth shown in Figure 5.4 are divided into S sized grids.

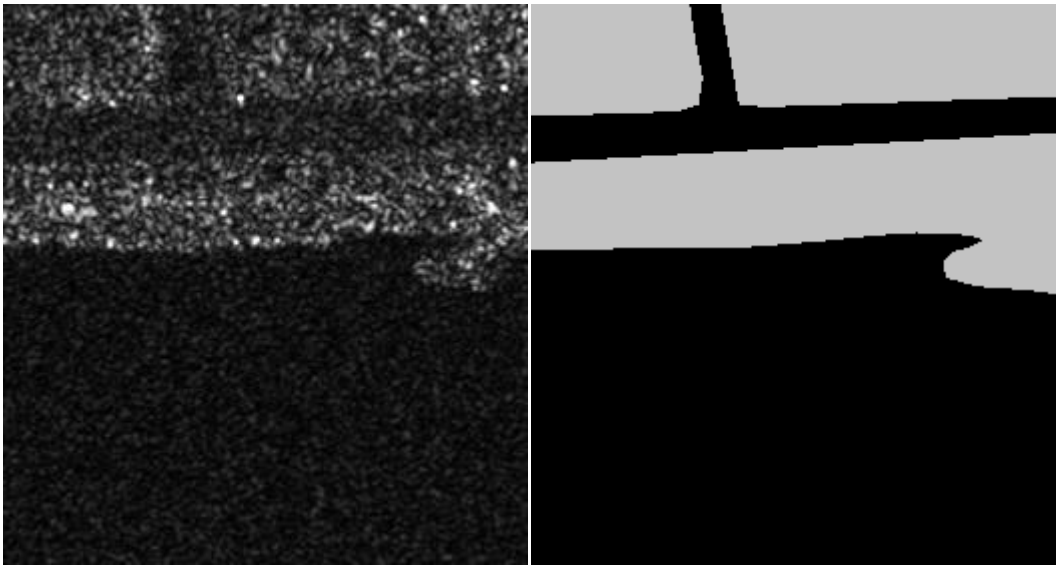


Figure 5.4 The SAR Image and Its Groundtruth

The grid pairs that have a boundary in the ground truth are cut out as 2SxS or Sx2S sized patches, so that the small images composed of two superpixels with a known boundary between them are obtained. For each patch, SRMP algorithm is applied for alpha values ranging between zero and one with increments of 0.01 and the misclassified pixels are counted. The best alpha values are plotted against intensity differences as shown in Figure 5.5. The minimum number of misclassified pixels for each patch and its intensity difference is plotted as shown below:

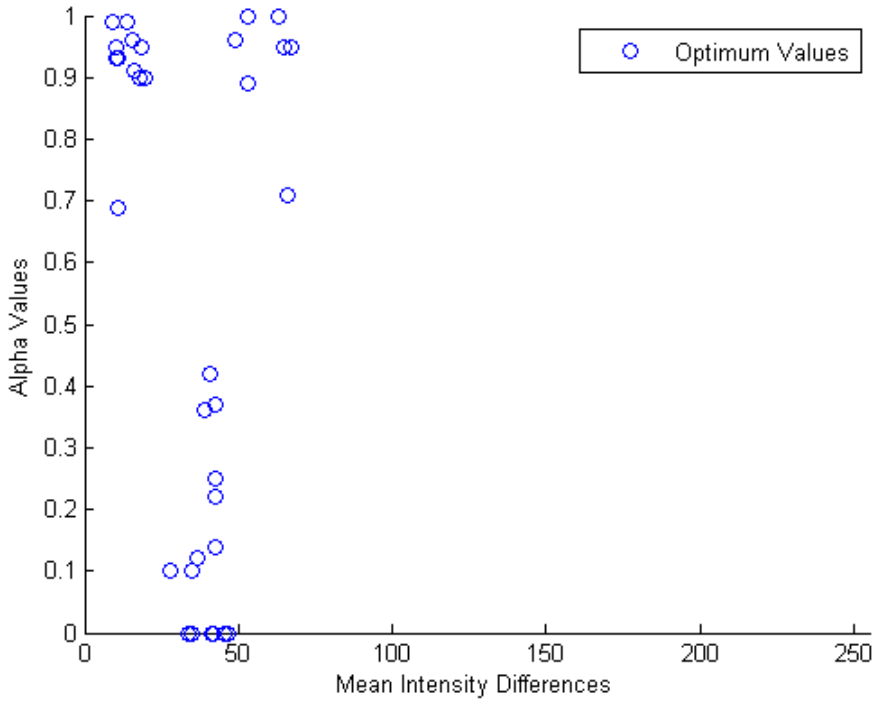


Figure 5.5 Optimum Alpha Values for Various Intensity Differences

This pattern can be expected, because, when the difference is very low, the algorithm should rely only on spatial proximity to have reasonable shape instead of following the borders. However, if the intensity difference is too large, the spatial term need to be increased as well to preserve the boundary of the shape.

If a downward sloping function for those points on the left and an upward sloping function for those points on the right parts of the above figure are utilized, a reasonable smooth pattern for the variables would be obtained. Since sigmoid function is an S curve with finite limits varying from zero and one, it can properly be used as a downward or an upward sloping curve as the signs are reversed. The sum of two sigmoid functions with reverse signs shown below can be appropriately represents the alpha values as distributed in Figure 5.5.

$$\frac{1}{1+e^{abs(\Delta I)}} \quad , \quad \frac{1}{1+e^{-abs(\Delta I)}} \tag{5.19}$$

To define a smooth function that follows this pattern, we have to specify the values for low and high intensity differences that may change from image to image. By analysis of the simulation results, it is found that that the values smaller than $\mu - \sigma$ are low intensity values whereas the values larger than $\mu + \sigma$ are high intensity values where μ is the mean intensity and σ is the standard deviation of the image. Therefore, the range is one standard deviation of the mean.

Since the above sigmoid functions produce smoothly shaped transition between finite limits varying from zero and one, the sum of those two functions with opposite signs can appropriately represent the pattern. If one standard deviation range is inserted into the sum of the two sigmoids, the resulting adaptive function is:

$$alp\square a(i) = \frac{1}{1+e^{0.5\times(abs(\Delta I)-(\mu-\sigma))}} + \frac{1}{1+e^{-0.5\times(abs(\Delta I)-(\mu+\sigma))}} \tag{5.20}$$

where ΔI the intensity difference between cluster i and its neighboring cluster j, μ is the mean and σ is the standard deviation of the image. The graph of the function is given in Figure 5.6:

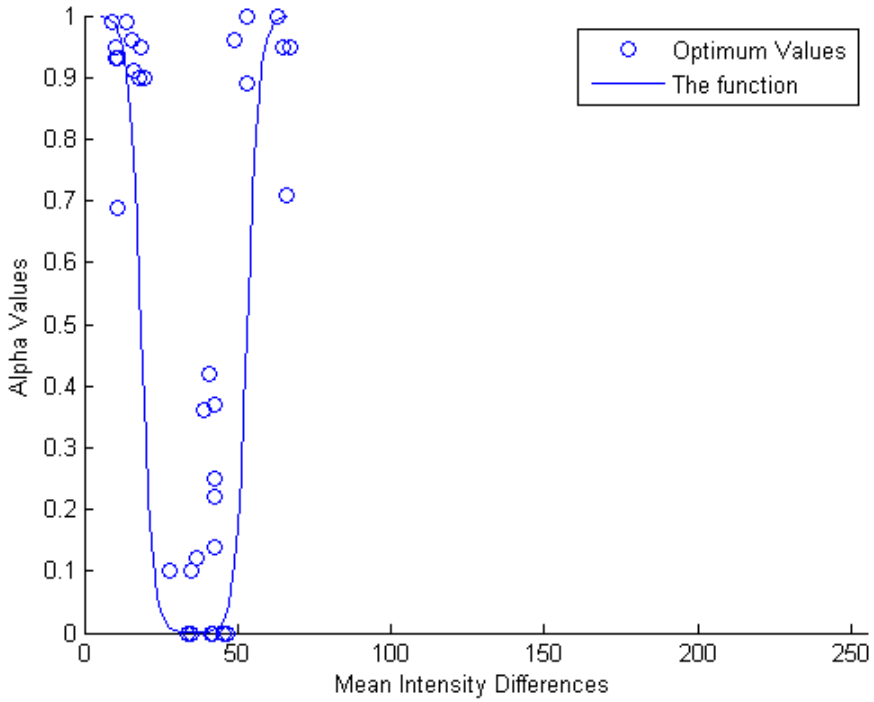


Figure 5.6 Optimum Alpha Values and the Estimated Function

If the distribution is multimodal, it fails to determine the proper value of alpha parameter. Since in the multimodal probability density function the difference between the mean and variance of each class is different than that of the mean and variance of the image. This causes variations in the estimated low and high contrasting pairs. This can be a limitation of the SRAMP algorithm.

CHAPTER 6

EXPERIMENTAL RESULTS

In this chapter, the performing capabilities of the recommended approaches are investigated by conducting experiments on real images and synthetic images. The experiments are performed on MATLAB 2012a. The MATLAB codes for SLIC and morphological cleanups are adapted from Kovesi (2013) to be applicable to SAR images. For the boundary recall evaluation, the *e*-pixel neighborhood is prespecified as 1 pixel which is generally used in comparing of all state-of-the-art algorithms.

6.1. Comparison of Mahalanobis and Euclidean Proximities

Due to the limited capability of Euclidean proximity in capturing the proper classification of the pixels near or along the borders of the two pixels, Mahalanobis proximity is recommended as an alternative metric and inserted into SLIC algorithm. To evaluate the segmentation performance of Mahalanobis-SLIC (MSLIC) algorithm, it is compared with SLIC, Turbopixels, ERS and SEEDS algorithms whose codes are publicly available. For this purpose, experiments are conducted on twelve distinct images from Berkeley Data Set. Various grid sizes are chosen in all the calculations to obtain different number of superpixels. The parameter m is taken as equal to 10, the one used as default in Achanta et al (2010).

Boundary recall and undersegmentation error are computed and plotted in Figure 6.1 and Figure 6.2 respectively.

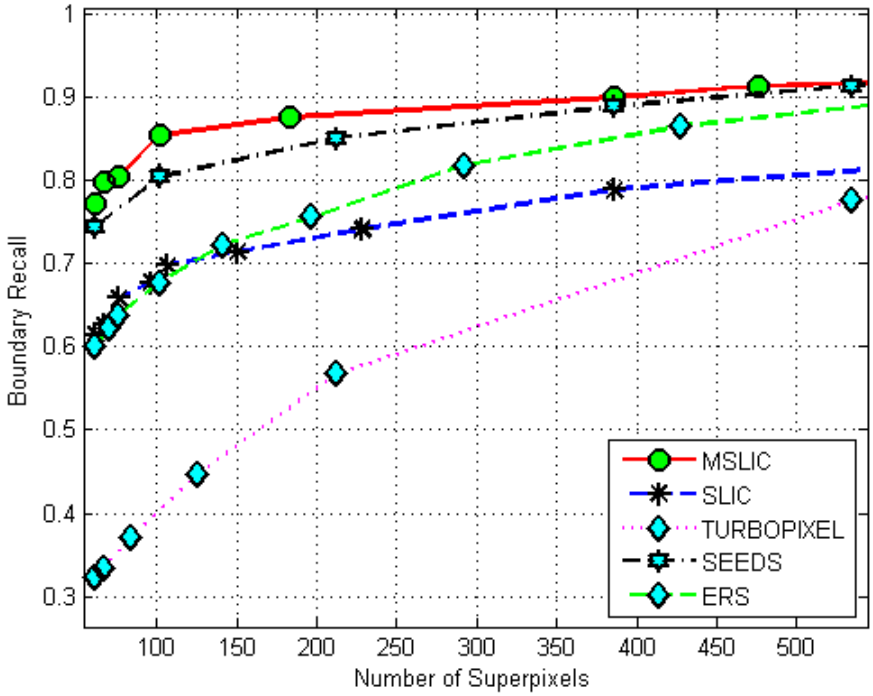


Figure 6.1 Boundary Recall Results for Berkeley Data Set

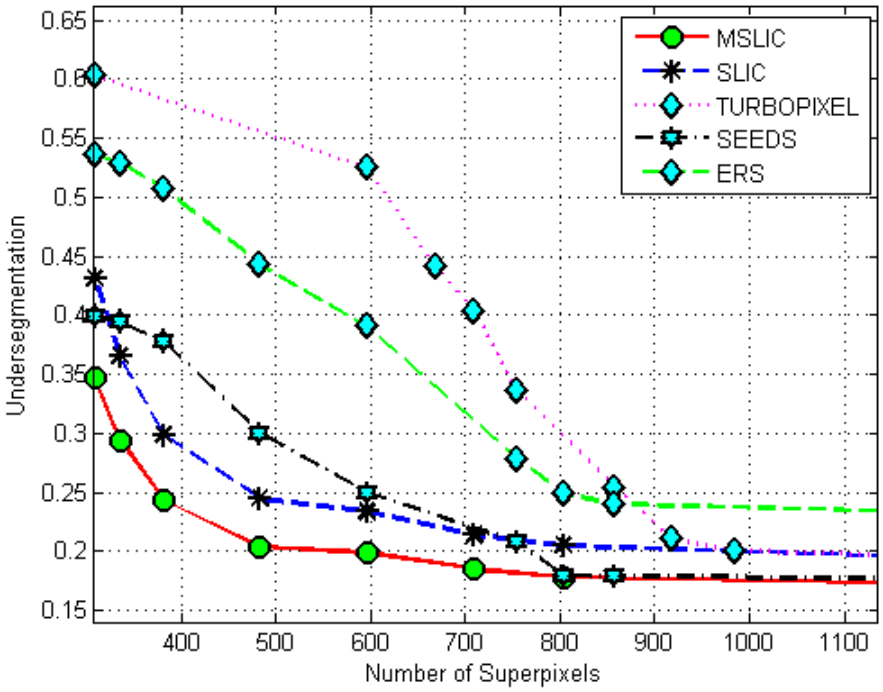


Figure 6.2 Undersegmentation Error Results for Berkeley Data Set

The purpose of the experiments on optical images such as Berkeley Data Set is to illustrate the way the Mahalanobis distance proximity term is capable to capture those pixels which are sensitive to the liming value of the Euclidian metric using the groundtruth data of the images.

As can be seen in Figure 6.1 and Figure 6.2, the proposed MSLIC algorithm has higher boundary recall and lower undersegmentation error than SLIC and the other three approaches. This implies that Mahalanobis distance proximity is capable of capturing those pixels where Euclidean distance proximity fails. Thus, it can be said that MSLIC has better segmentation performance than SLIC and the other three approaches.

6.2. Visual Comparison of Mahalanobis and Euclidean Proximities

Mahalanobis distance proximity is better suited to real world shapes than the Euclidean distance proximity. In the Figure 6.3, an example of MSLIC and SLIC algorithm results are shown.

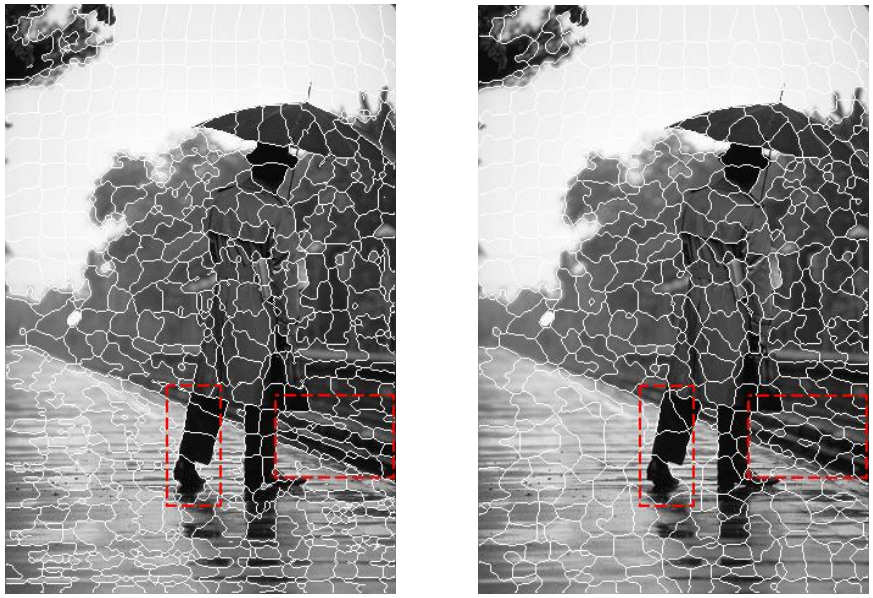


Figure 6.3 Visual Comparison of Mahalanobis and Euclidean Proximities, first shown MSLIC and then SLIC. (The image-250087- from Berkley Data Set)

A closer look at the areas that are enclosed by red rectangles show that the MSLIC algorithm fits better to all kinds of shapes from rectangle to elongated ones as shown in Figure 6.4.

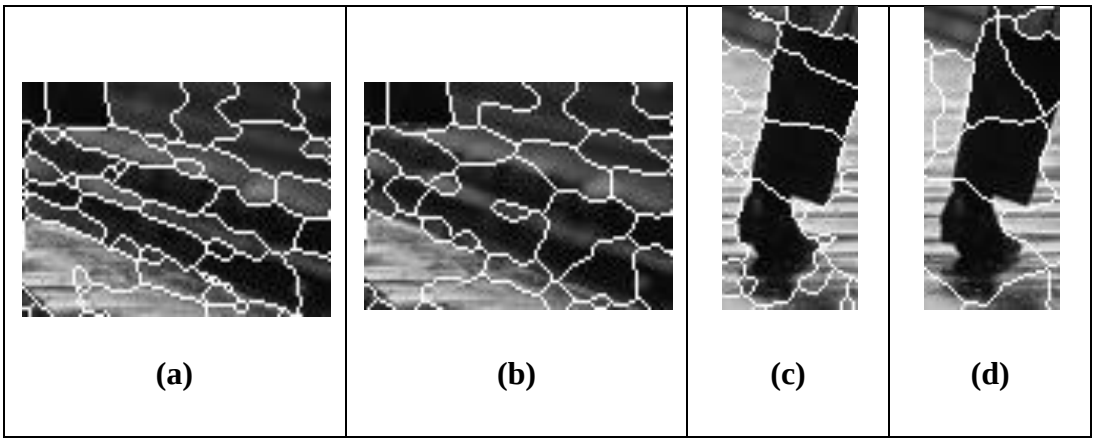


Figure 6.4 Selected areas, (a) MSLIC, (b) SLIC, (c) MSLIC, (d) SLIC

6.3. Performance Evaluation of SREP and SRMP on Synthetic Images

For the two different spatial proximities, namely Euclidean and Mahalanobis similarity ratio based algorithms have been formulated as SREP and SRMP. To evaluate the segmentation performance of these algorithms, experiments are conducted on two distinct synthetic images. The reason for the choice of synthetic images is to use their ground truth data. The synthetic images are shown in Figure 6.5.

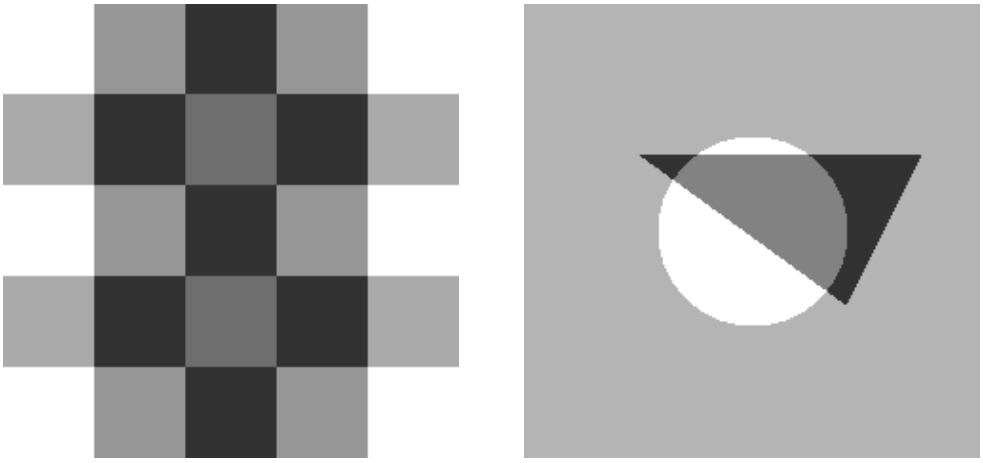


Figure 6.5 Synthetic Images

In both of the proposed approaches, the parameter α regulates the degree of compactness and the boundary adherence. When the parameter is large, spatial proximity is more important than intensity similarity ratio resulting more compact superpixels. In this case, the grid structure of the superpixels are preserved. When the parameter is small, intensity similarity ratio is more important than spatial proximity resulting superpixels adhered well to the boundaries. In this case, the boundary is preserved. In the calculations, α is determined as equal to 0.5 which yields the most visually pleasing superpixels. The purpose of the experiments on the synthetic images is to show the sensitivity of all the approaches, hence their robustness, to the speckle noise level.

To evaluate boundary adherence, the boundary recall is computed on the synthetic image that are speckled at 0.05 noise level .As in the previous experimentation, same grid sizes are used. Computed results are plotted in Figure 6.6. In this figure, it can be seen that the boundary recall of the proposed SRMP approach is higher compared to others indicating that the proposed approach is less sensitive to noise than the other algorithms. Thus, it can be said that SRMP algorithm is capable of obtaining better boundary adherence than as compared to the other algorithms in noisy images.

To evaluate the compactness, the undersegmentation error is also calculated and plotted for the same noisy synthetic images as shown in Figure 6.7. In this figure, the undersegmentation error of SRMP approach is lower than the other algorithms implying that it is capable of generating more compact superpixels than the others.

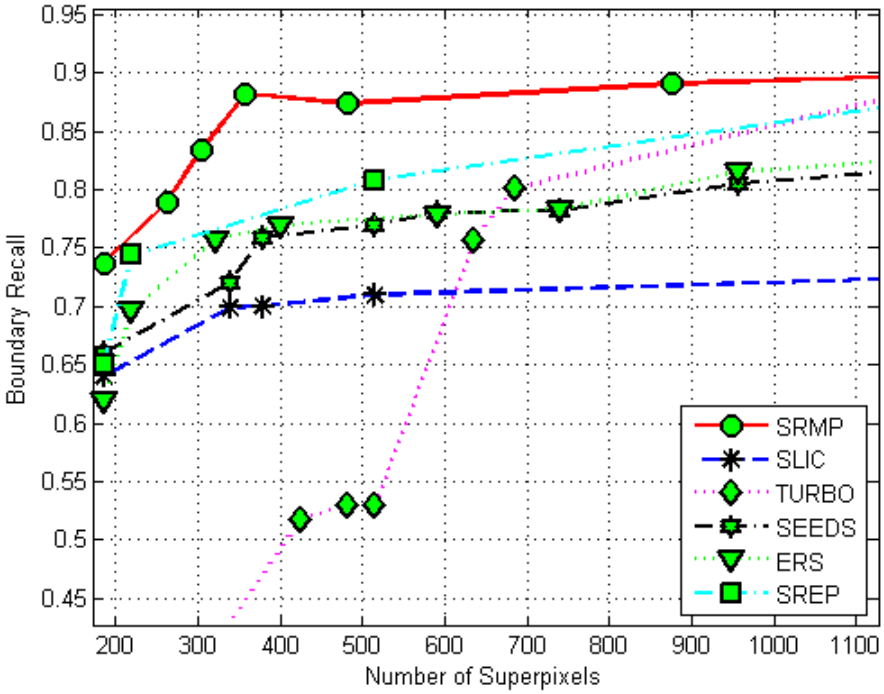


Figure 6.6 Boundary Recall Results for Synthetic Images

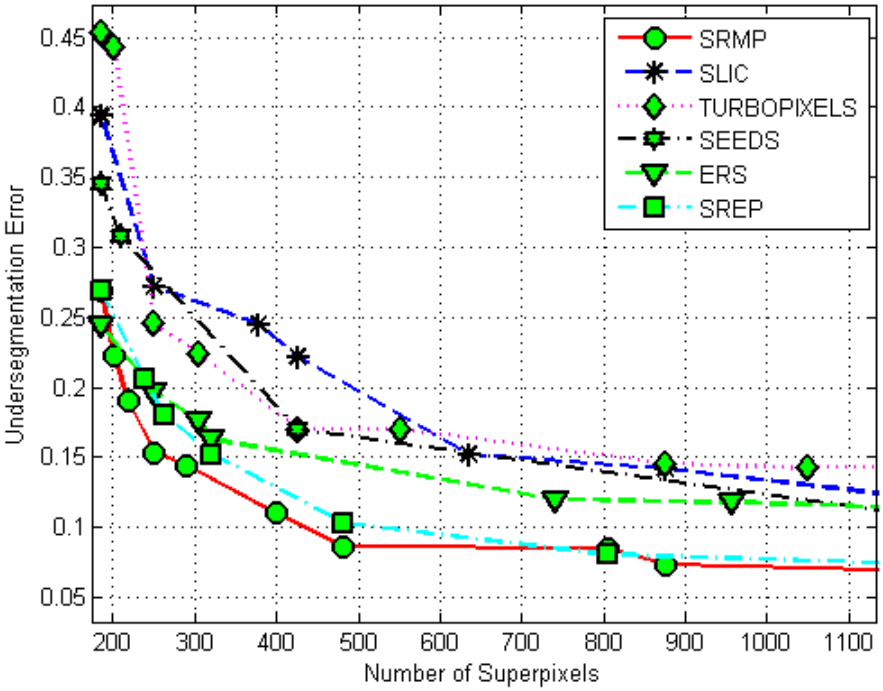


Figure 6.7 Undersegmentation Error Results for Synthetic Images

6.4. Comparison of SRMP and SRAMP on Synthetic Images

SRAMP algorithm is an adaptive version of SRMP approach where similarity ratio and Mahalanobis proximity are combined with an adaptively determined parameter. To evaluate the segmentation performance of SRMAP and SRMP algorithms, experiments are conducted for the same two synthetic images that are again speckled at 0.05 noise level by using the same grid sizes as in the previous experimentations. The boundary recall and undersegmentation error are computed and plotted in Figure 6.8 and Figure 6.9 respectively.

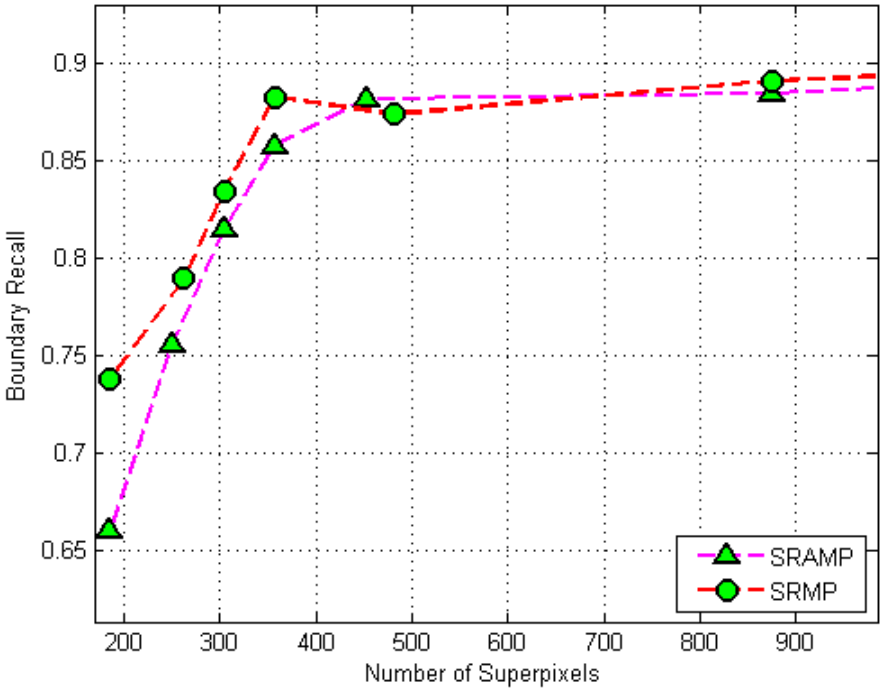


Figure 6.8 Boundary Recall Results for Synthetic Images

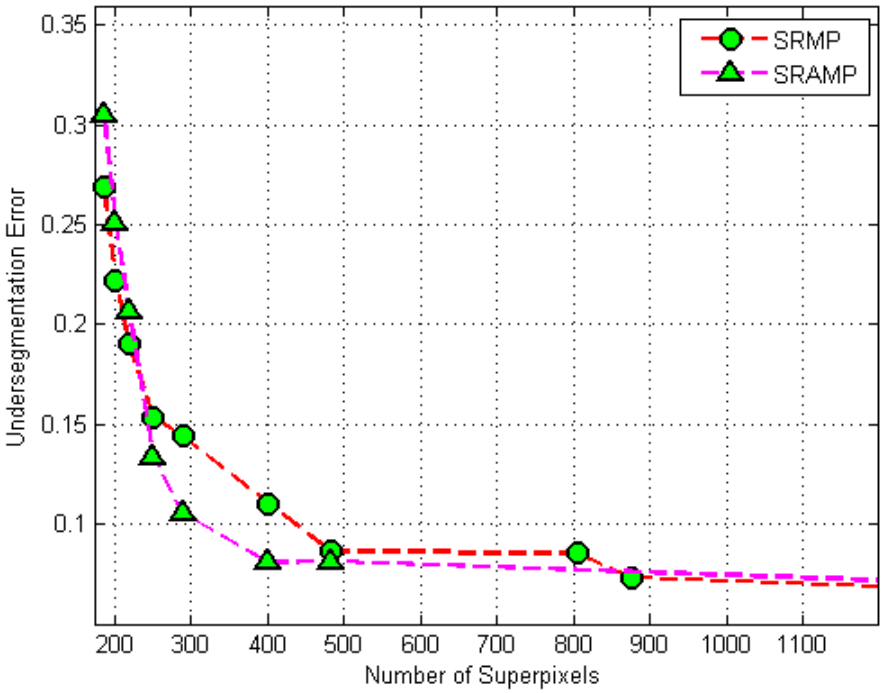


Figure 6.9 Undersegmentation Error Results for Synthetic Images

From these figures, it is evident that SRAMP has comparable boundary recall but lower segmentation error than SRMP implying it has high segmentation performance.

6.5. Performance Evaluation of SREP and SRMP on Real SAR Images

The experiments conducted on the speckled synthetic images indicated that SREP and SRMP approaches have higher segmentation performance in terms of boundary recall and undersegmentation error than those state-of-the-art methods. To evaluate the segmentation capability of these two approaches for SAR images, experiments are conducted on six distinct real SAR images shown in the Appendix A. For this purpose, ground truth of each SAR image is prepared manually and presented in Appendix B. In the computations, 0.5 is used for the balancing parameter α and same grid sizes are used as in the previous experimentations. The results for boundary recall and undersegmentation error are presented in Figure 6.10 and Figure 6.11 respectively.

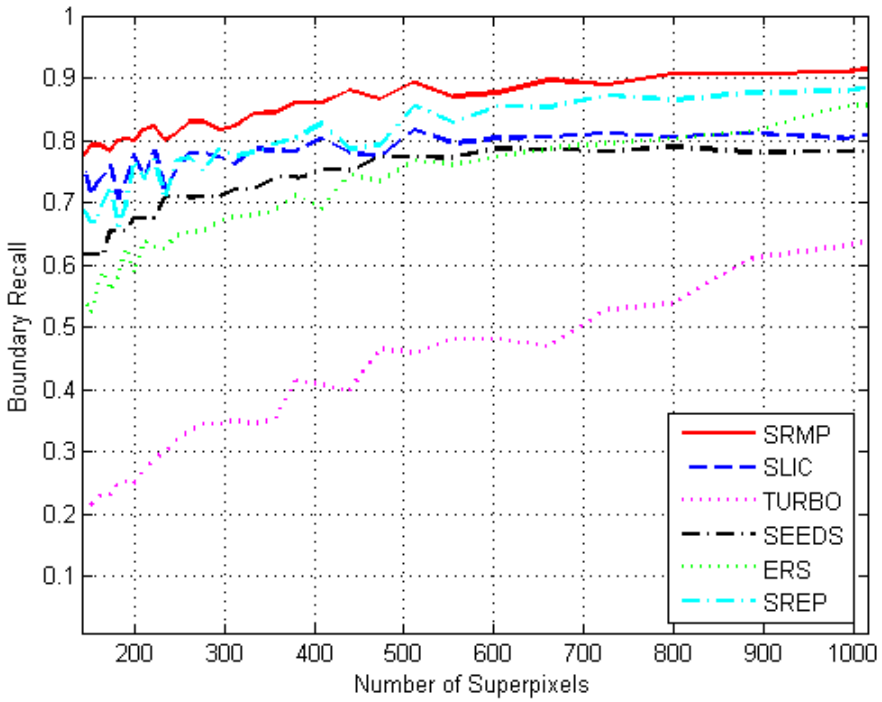


Figure 6.10 Boundary Recall Results for Real SAR Images

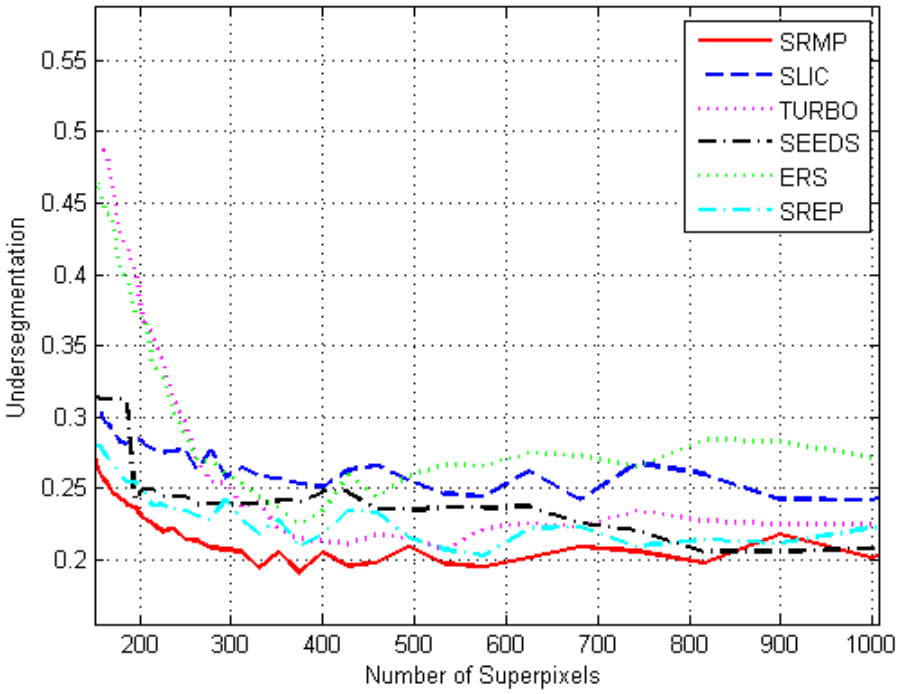


Figure 6.11 Undersegmentation Error Results for Real Images

As expected, SRMP has the highest boundary recall and lowest undersegmentation error than the other algorithms and SREP stands as the second best in terms of these metrics.

6.6. Comparison of SRMP and SRAMP on Real SAR Images

The segmentation capability of SRMP and SRAMP algorithms are compared by conducting experiments on the same real SAR images used in the previous section. The balancing parameter and the grid sizes are kept same as those in the previous calculations. The computed boundary recall and undersegmentation error results are plotted in Figure 6.12 and Figure 6.13 respectively.

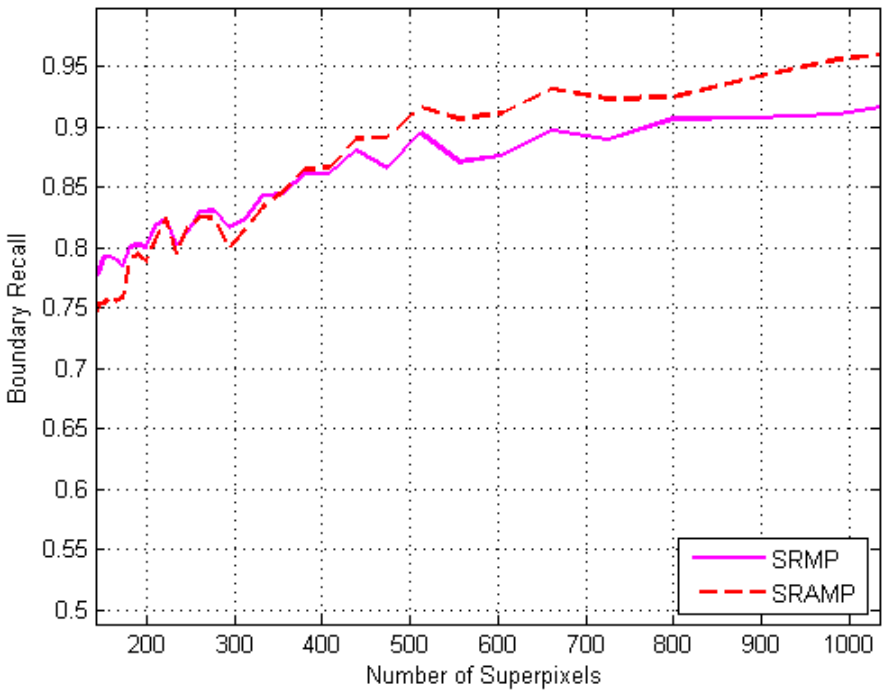


Figure 6.12 Boundary Recall Results for Real SAR Images

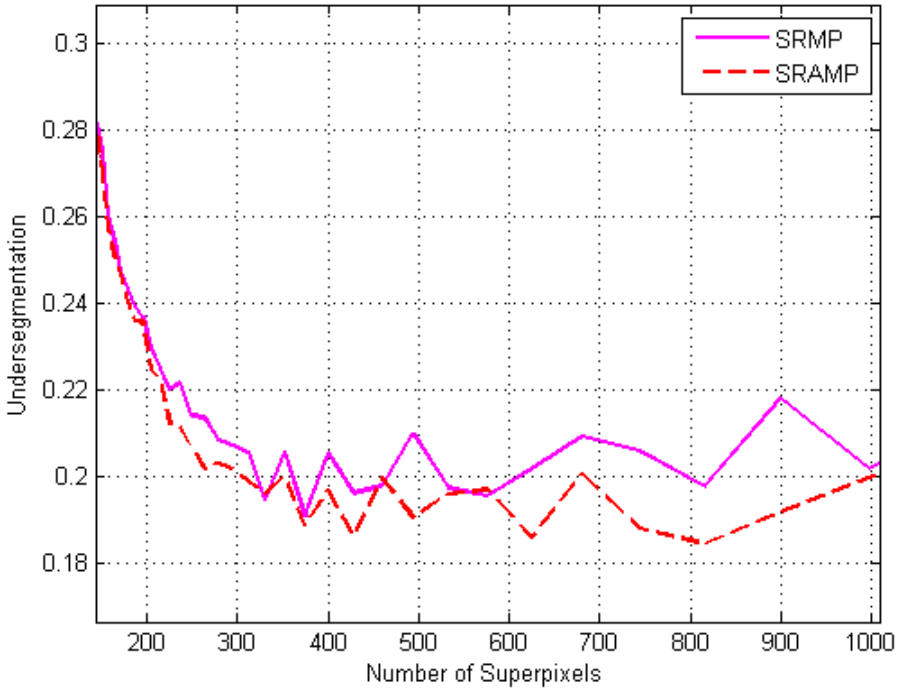


Figure 6.13 Undersegmentation Error Results for Real Images

From these figures, it can be concluded that the boundary recall of SRAMP is comparable but its undersegmentation error is lower than that of SRMP. This result is similar to that one related with the speckled synthetic images.

6.7. Comparison of Filtered SLIC and Similarity Ratio Based Algorithms

The algorithms SREP, SRMP and SRAMP employ similarity ratios that utilize 8-neighbours of the pixel to be labeled as the mean of the superpixels. This neighborhood grouping which is 3x3 window slided over each pixel within each search area sized 2S is the basis of the algorithmic computations. Although this type of local operation is similar to average filtering, the algorithms have no relation with that filtering. However, it raises the question of producing superpixels sensitive to speckle noise if an image is filtered first before the application of the SLIC, ERS, SEEDS and TURBOPIXELS algorithm. In this

section, the extent of boundary adherence of such superpixels are compared to those produced with the proposed algorithms. For this purpose, the two synthetic images are speckled at 0.05 noise level. As in the previous experimentations, same grid sizes are used.

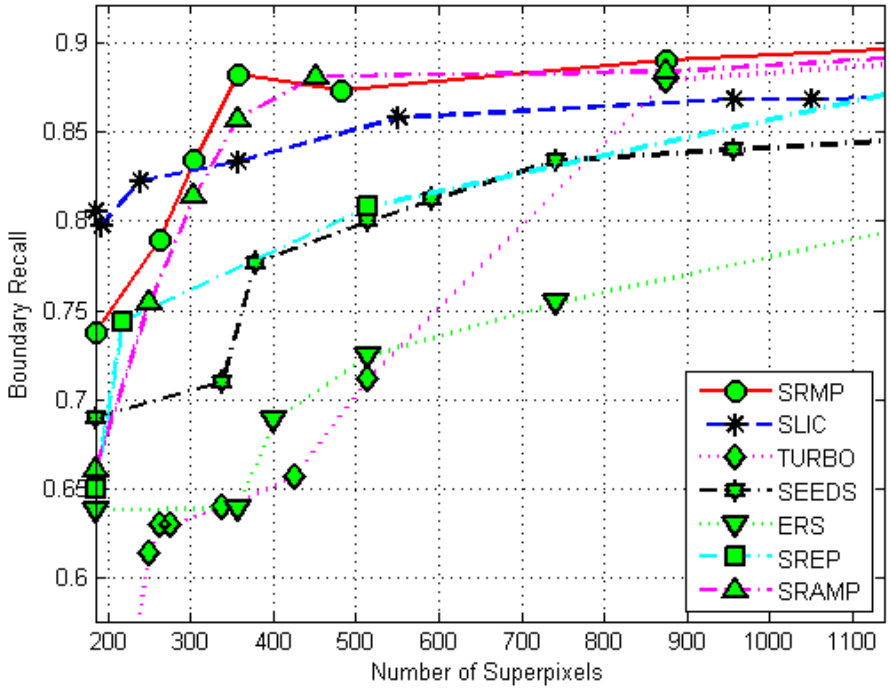


Figure 6.14 Boundary Recall Results for Similarity Ratio Based Algorithms and Filtered Superpixel Algorithms

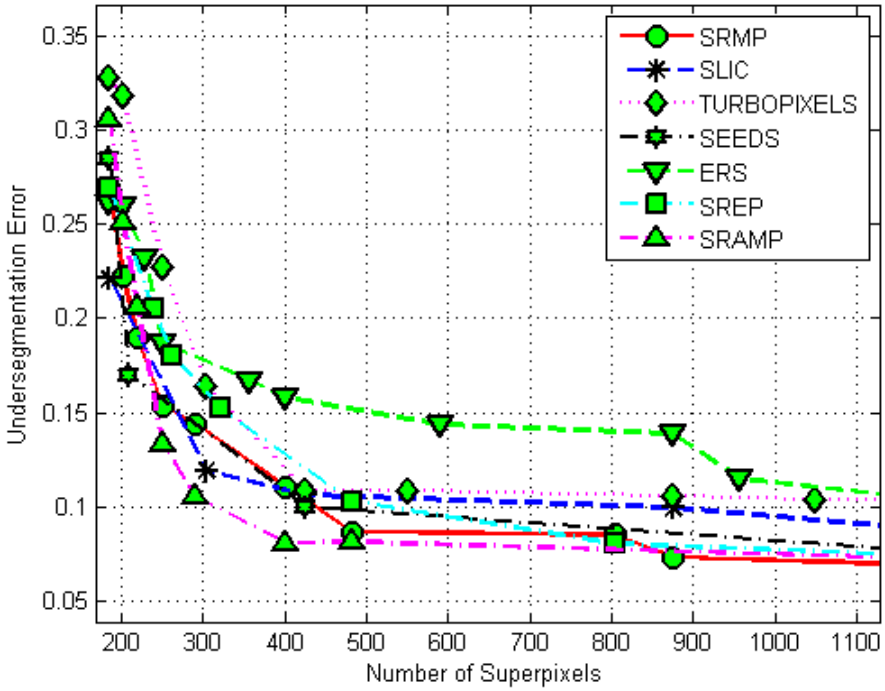


Figure 6.15 Undersegmentation Error Results for Similarity Ratio Based Algorithms and Filtered Superpixel Algorithms

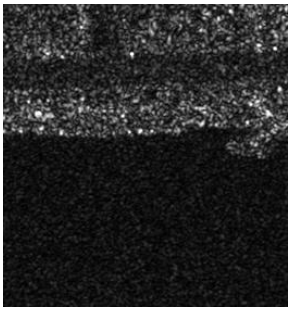
It is known that filtering techniques reduce the noise but at the same time smooth away edges to a greater or a lesser degree. From Figure 6.14 and Figure 6.15, it is evident that the proposed algorithms produce superpixels with higher boundary recall and lower undersegmentation error than those produced by filtered superpixel algorithms. This result is expected due to the propagation of error resulted at the filtering state to the subsequent tasks of the algorithms used. Hence, a more reliable superpixel segmentation can be obtained by utilizing a single robust approach rather than employing a filter as a preprocessing tool for any superpixel algorithm. In this respect, it can be said that the local operation used for the computation of the similarity ratio has no resemblance to average filtering.

6.8. Visual Comparison of SAR Superpixel Segmented Images

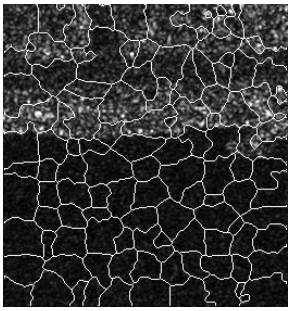
For visual comparison of the superpixels that are produced by the algorithms used in the previous sections, experiments are conducted on six real SAR images. The grid sizes chosen as 20 in all computations. The results are presented in from Figure 6.16 to Figure 6.21. For this figure, each page is allotted to one original image and seven other images showing superpixels produced by the various algorithms.

As remotely sensed radar images, any SAR image contains different landcover classes such as urban areas, man-made structures, vegetated areas and others. These landcover classes can be visualized by analyzing the adherence of the superpixels to the boundaries of those regions. If explicit information is required about these areas, the superpixels should be clustered using them as elementary units rather than the pixels. This superpixel-based clustering is discussed later in Chapter 7.

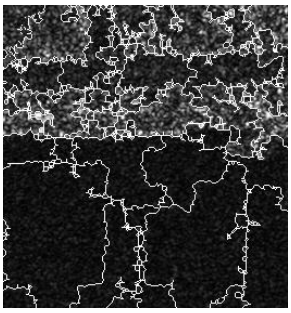
The aim of visual comparison of the images specific for this research can be twofold. First, the degree of robustness of the approach can be assessed by comparing the segmented image with the original one. If randomly spread highly irregular shaped superpixels are generated, then it can be said that the algorithm is not robust enough to speckle noise in the image. Second, boundary adherence and compactness can be determined by comparing the extent of the deviation of the superpixel boundaries from the edges of the objects.



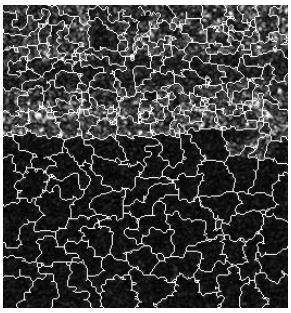
Original Image



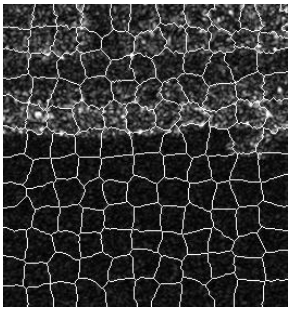
SLIC



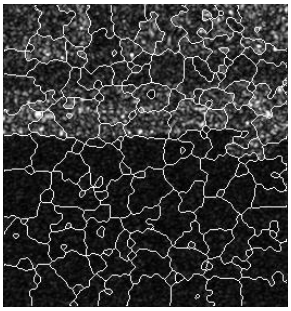
SEEDS



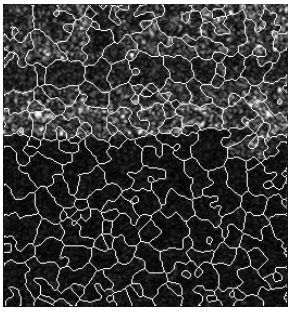
ERS



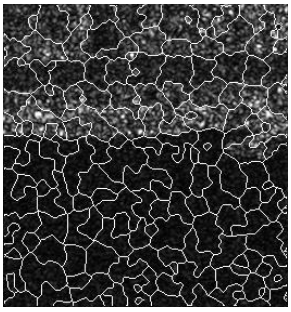
TURBOPIXELS



SREP

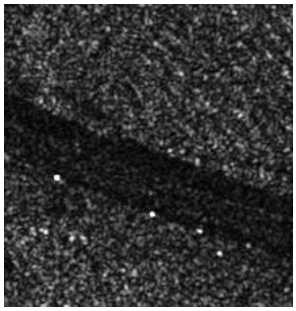


SRMP

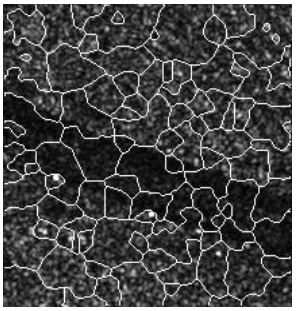


SRAMP

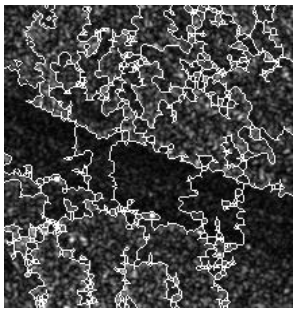
Figure 6.16 SAR Image 1 and Its Superpixel Segmentations



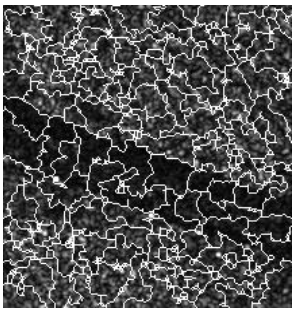
Original Image



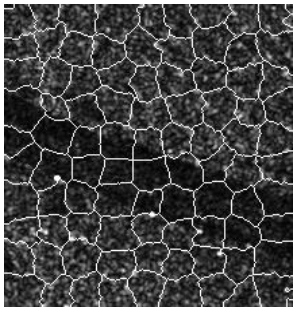
SLIC



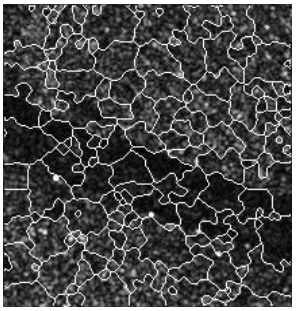
SEEDS



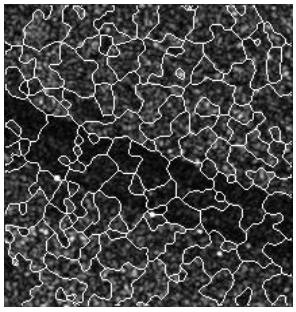
ERS



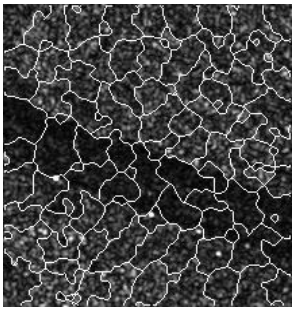
TURBOPIXELS



SREP

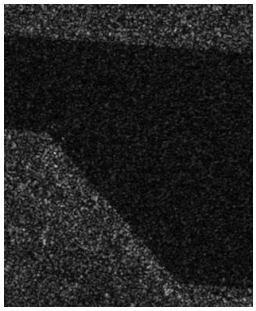


SRMP

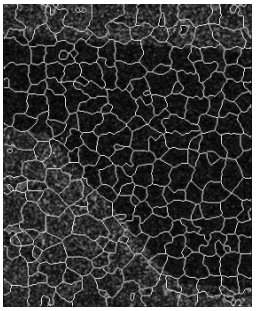


SRAMP

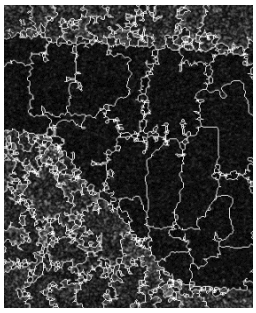
Figure 6.17 SAR Image 2 and Its Superpixel Segmentations



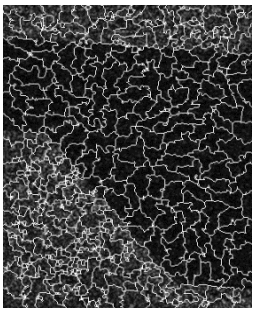
Original Image



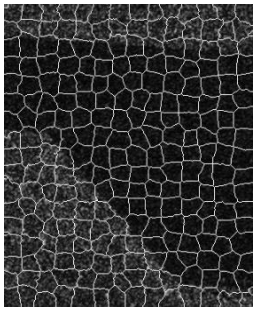
SLIC



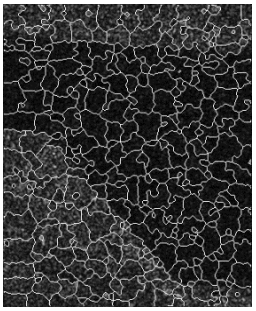
SEEDS



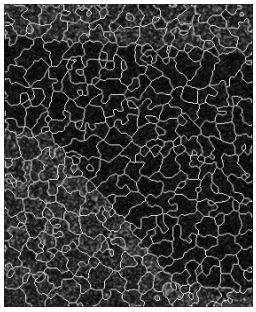
ERS



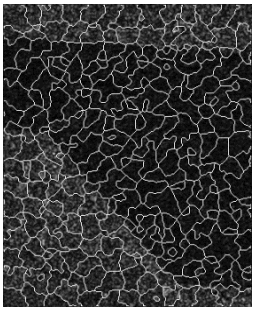
TURBOPIXELS



SREP

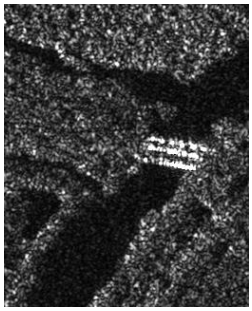


SRMP

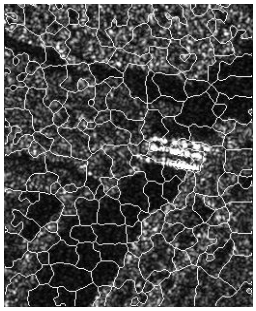


SRAMP

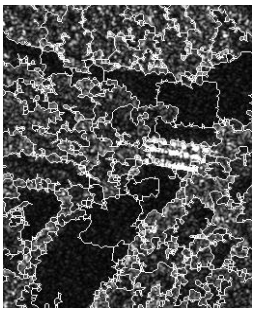
Figure 6.18 SAR Image 3 and Its Superpixel Segmentations



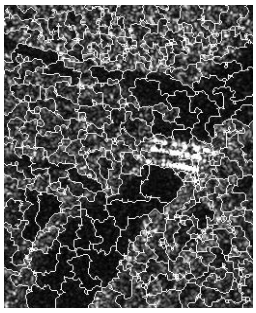
Original Image



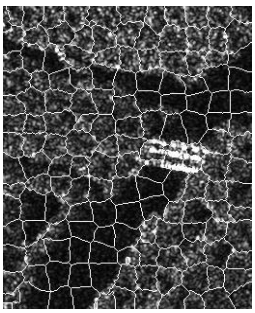
SLIC



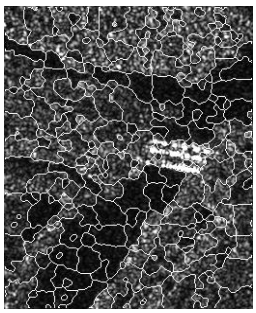
SEEDS



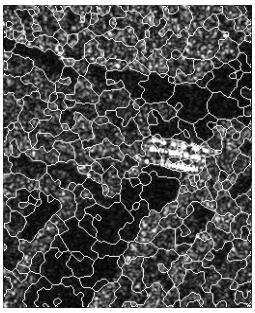
ERS



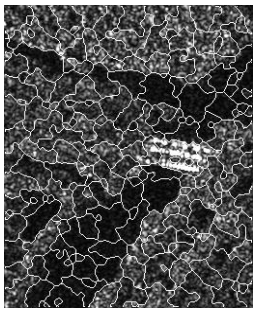
TURBOPIXELS



SREP

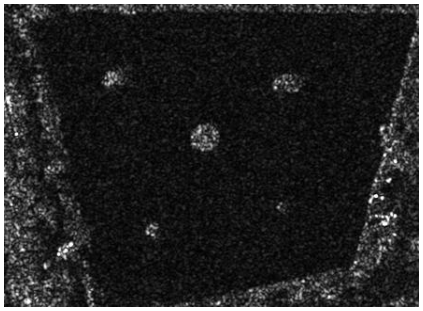


SRMP

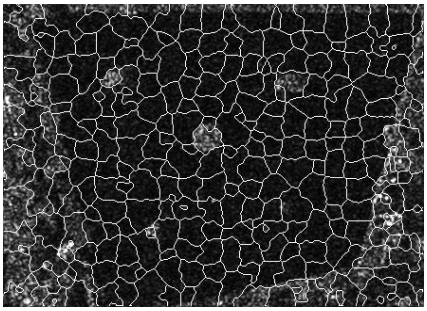


SRAMP

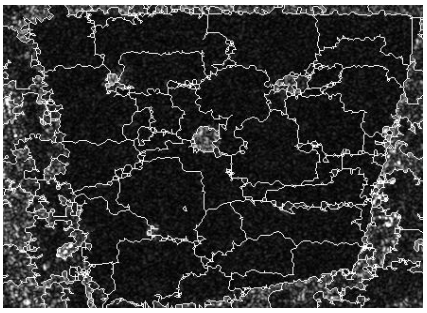
Figure 6.19 SAR Image 4 and Its Superpixel Segmentations



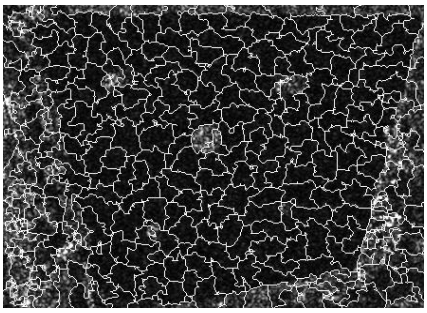
Original Image



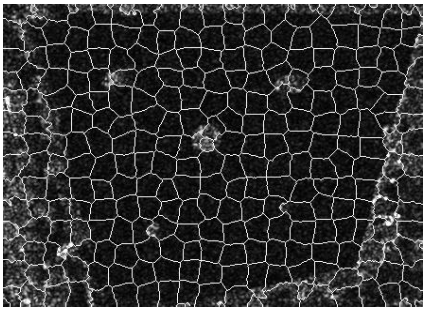
SLIC



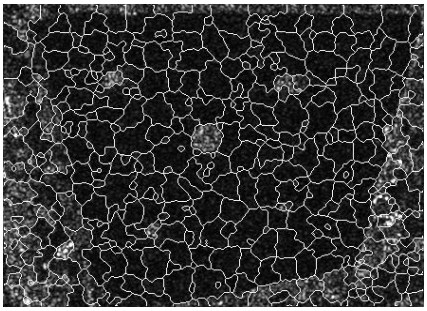
SEEDS



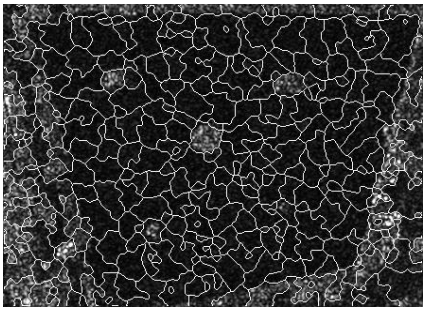
ERS



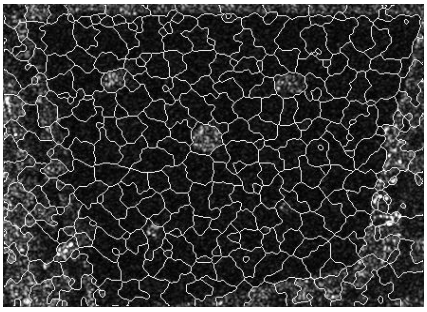
TURBOPIXELS



SREP

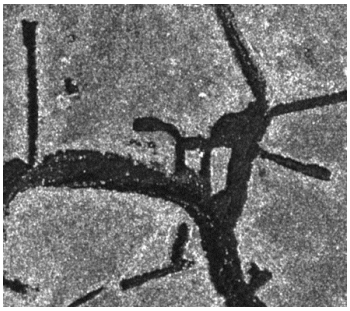


SRMP

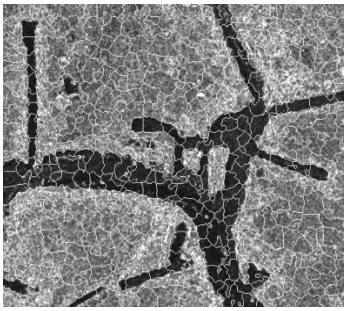


SRAMP

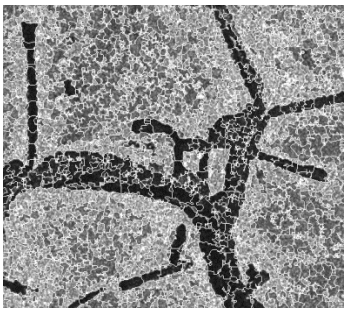
Figure 6.20 SAR Image 5 and Its Superpixel Segmentations



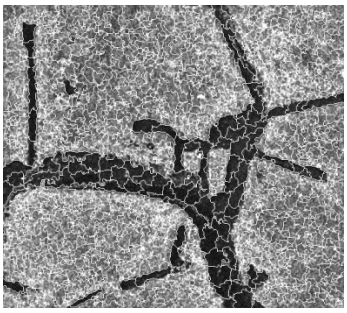
Original Image



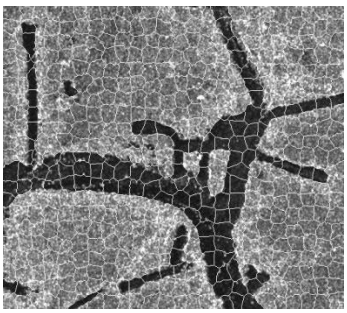
SLIC



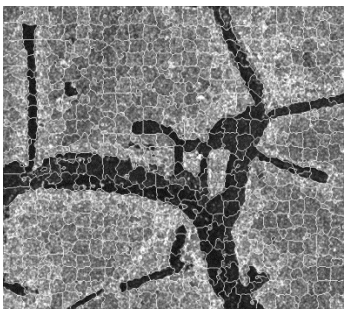
SEEDS



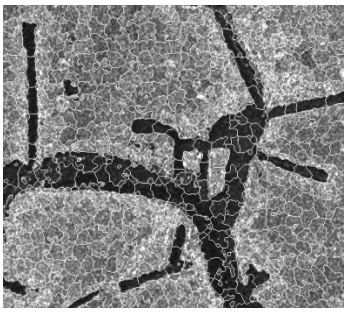
ERS



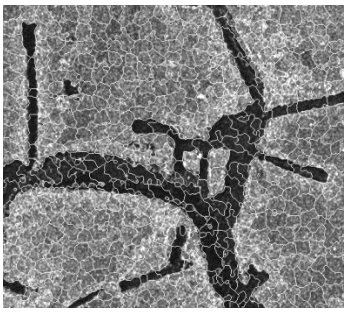
TURBOPIXELS



SREP



SRMP



SRAMP

Figure 6.21 SAR Image 6 and Its Superpixel Segmentations

As it can be seen from the results, the superpixels adhere well to the boundaries of strong edges such as coastal line. The strong edges of man-made structures such as roads and bridges are also well preserved with the boundary adherence of the superpixels. In homogenous areas like sea, algorithms tend to produce more similar superpixels whereas in vegetated areas they produce more irregularly shaped superpixels. When the segmented images are compared to the original ones, it can be observed that SREP, SRMP and SRAMP approaches produce more robust images than those generated by the others. This is expected because of the similarity ratio employed by SREP, SRMP and SRAMP algorithms. When segmented images are compared among themselves, the superpixels produced by SRMP and SRAMP approaches look more uniform and compact than the SREP due to the well adherence of the superpixel boundaries to the image edges of objects. In addition, the superpixels produced by these approaches are more visually pleasing than the others. Although, the superpixels generated by adaptive SRAMP approach are comparable to those of SRMP with respect to uniformity and compactness, they are relatively more visually pleasing than those generated by SRMP.

CHAPTER 7

APPLICATIONS OF SUPERPIXEL-BASED CLUSTERING ON SAR IMAGES

The purpose of this chapter is to illustrate the way the preprocessed superpixels can be used in subsequent tasks of an image analysis. The task considered here is clustering of the SAR superpixels into larger segments. This superpixel-based clustering reduces the search space to be explored for high-level abstraction of the objects based on superpixels. One important benefit in the reduction of search space is the decrease in computational complexity of the task which makes it tractable thus providing timely allowances to explore many important aspects of complex imagery. On the other hand, as image complexity is reduced, the defined boundaries of clustered superpixels delineate the borders of the regions of interest which can be used to extract them. Such clustering can find wide applications in the area of remote sensing images to identify different land cover according to certain homogeneity-based criteria. Based on such criteria, the number of clustered superpixels is expected to be proportional to the land cover classes. In case the land cover over an image consists of limited number of classes such as sea and land, the number of clusters would be low as compared to that obtained from an image with large number of land cover classes. If specific land cover classes are aimed to be extracted from the image or to control the extent of clustering, the cluster number can be specified before the task. For example, if linear structures such as canals are of interest, then two or three can be used as the cluster number.

The k-means and the superpixel density-based spatial (SPDBSCAN) (Kovesi, 2013) approaches are used in this application to cluster the generated superpixels. The former approach needs the specification of the cluster numbers while the latter do not need such requirement. This is the reason behind the choice and the comparison of these two approaches.

The traditional k-means clustering have been briefly described in Chapter 4. There are various methods to determine the value of k which are all discussed in Pham et al. (2004). In this chapter, k is set as equal to number of land cover classes through visual inspection.

SPDBSCAN is essentially a merging approach using density-based spatial clustering (DBSCAN) algorithm (Ester, M., 1996) which is efficient on large data bases and capable of capturing clusters with arbitrary shapes. In this approach, adjacent superpixels are merged to form clusters of superpixels using the distance of the means of those superpixels as the distance measure.

Among the SAR images utilized in Chapter 6, three of them are chosen whose superpixels are to be clustered. The original images as well the superpixels of these images generated with SREP, SRMP and SRAMP algorithms in Chapter 6 are once more shown here in Figure 7.1. To facilitate the comparison of these images, the superpixel-based clustered segments are all shown together in Figure 7.2. The value of k is taken as two for the k-means clustering.

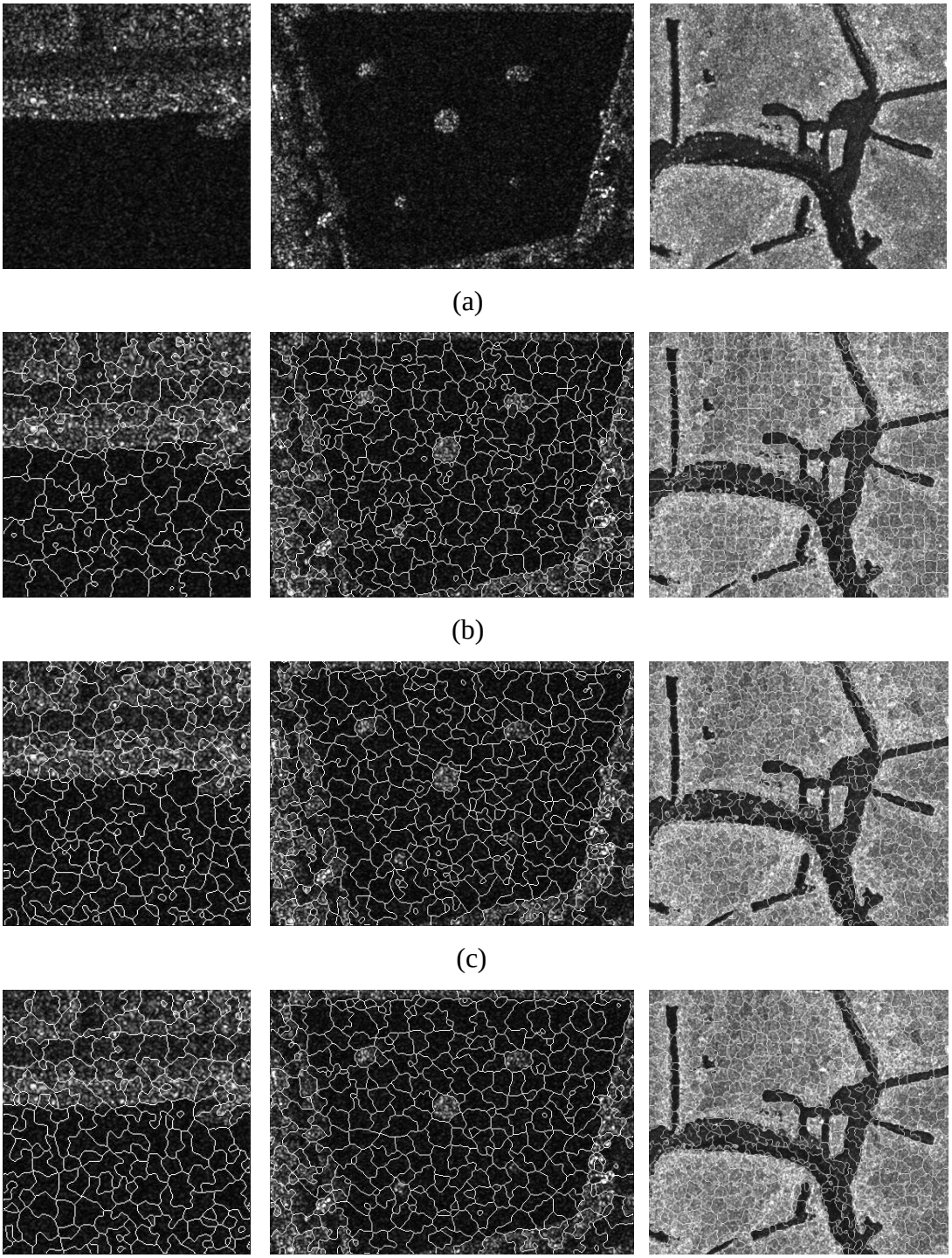
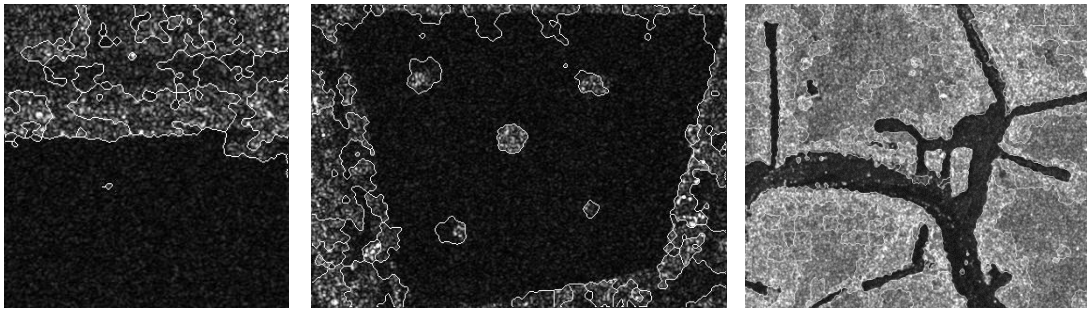
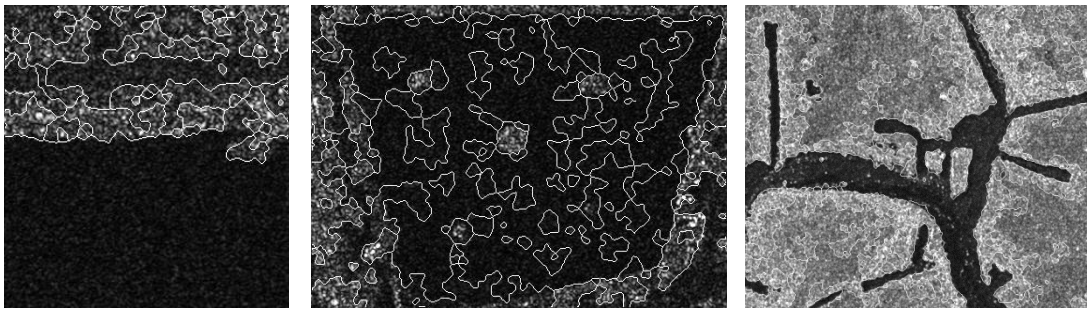


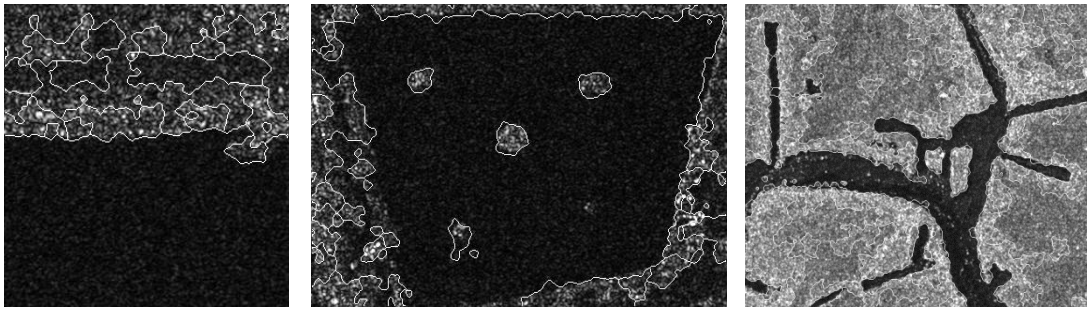
Figure 7.1 Images Used for Applications. (a) Original images (b-d) Superpixels generated by SREP, SRMP and SRAMP respectively.



(a)

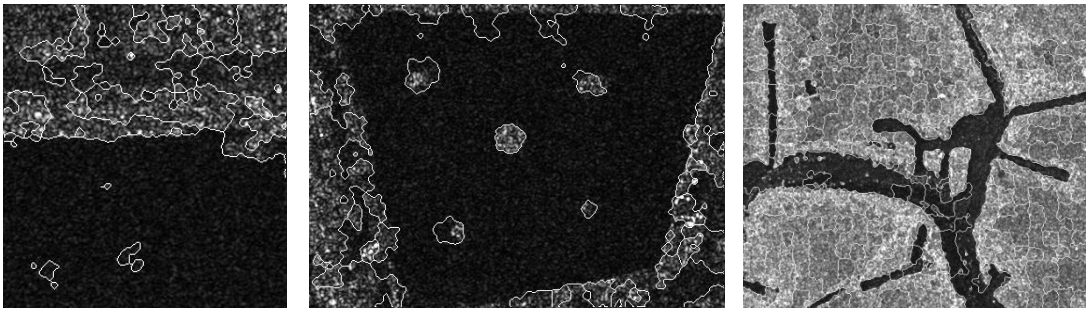


(b)

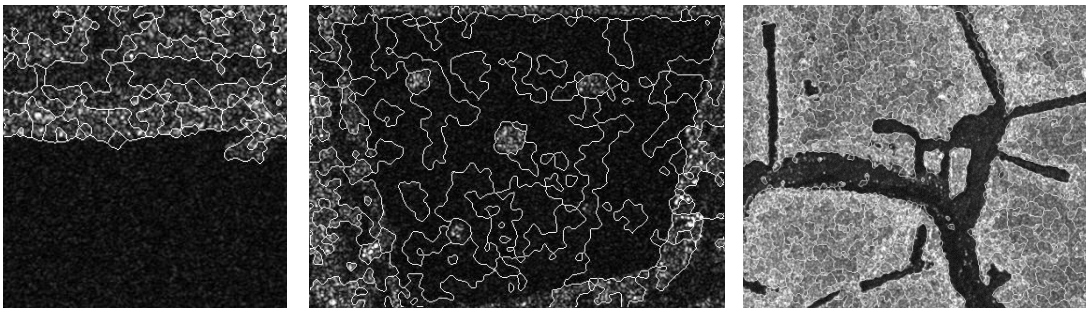


(c)

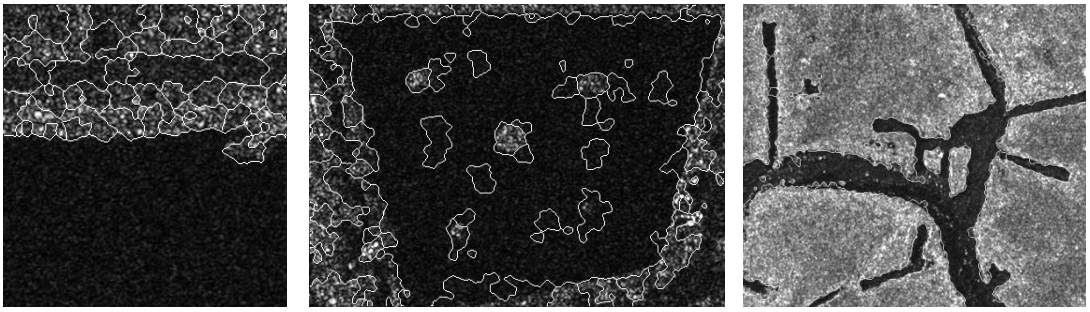
Figure 7.2 Clustered Results with k-means of the Superpixels (a-c) generated by SREP, SRMP and SRAMP.



(a)



(b)



(c)

Figure 7.3 Clustered Results with SPDBSCAN of the Superpixels (a-c) generated by SREP, SRMP and SRAMP.

As can be seen from these images, the edges are properly delineated between the land cover classes. If a number for land cover classes is specified, the edges for those classes are defined and the segments are extracted accordingly. As in the k-means clustering, the images are segmented into two regions defined by the value of k. In the absence of any value for k, the superpixels are expected to be clustered according to the means of the adjacent superpixels. In any case, superpixel-based clustering is efficient in extracting the land cover classes by their defined boundaries. Therefore, superpixel segmentation as a pre-processing tool can be successfully used in the subsequent tasks of any SAR image analysis such as automatic target detection and recognition algorithms.

CHAPTER 8

SUMMARY AND CONCLUSIONS

In this thesis, similarity ratio based algorithms have been described as new superpixel generating algorithms for SAR images. The core of the proposed algorithms is a linear combination of similarity ratio and spatial proximity terms parameterized with a constant. This parameter regulates uniformity and compactness of the generated superpixels. Standard k-means clustering is a minimization scheme adapted to be used locally over reduced search area rather than the whole image. Among those proposed algorithms, it has been found that similarity ratio and adaptive Mahalanobis proximity (SRAMP) approach has relatively high segmentation performance.

As the existing superpixel generating algorithms are reviewed, intensity similarity and spatial proximity are found to be the two important factors for perceptual grouping. As a result of this review, it has been concluded that the algorithms which group the pixels using those two factors produce perceptually meaningful and representationally efficient superpixels.

If Euclidean distance metric is used for spatial proximity, the approach is regarded as similarity ratio-based Euclidean proximity algorithm designated as SREP. If Mahalanobis as another spatial proximity metric is used instead of Euclidean metric, the approach is referred as similarity ratio based Mahalanobis proximity algorithm designated as SRMP.

Among those grouping algorithms, SLIC is the simplest and the most efficient one. Its simplicity is due to the use of k-means clustering and its efficiency is due to the linear combination of intensity similarity and Euclidean proximity terms parameterized with a compactness parameter. This parameter regulates the degree of the boundary adherence and the extent of the compactness of the superpixels. To exploit the simplicity of k-means clustering, this approach is adapted in this thesis as in a manner similar to SLIC algorithm.

In homogeneous areas, Euclidean proximity term negatively affects grouping of those pixels near or along the borders of the two superpixels. To avoid any possible misclassification of those pixels during the course of the algorithm, Mahalanobis distance is recommended and is used instead of Euclidean distance. The formulated Mahalanobis-SLIC (MSLIC) and SLIC are then tested by using the Berkley Data Set. It has been found that MSLIC produces superpixels with higher boundary recall and lower segmentation error than those of SLIC.

Although Euclidean proximity is replaced with Mahalanobis proximity in the formulation of MSLIC, the intensity similarity term is still the essence of SLIC and MSLIC algorithms. The intensity distance term formulated as the sum-of-square differences of intensities has been demonstrated to be robust to additive noise but not to multiplicative speckle noise inherent in SAR imagery. To avoid this drawback of the intensity distance term, similarity ratio is developed. Combining this ratio with Euclidean and Mahalanobis distances separately, new superpixel generating algorithms are formulated as similarity ratio based Euclidean proximity (SREP) and similarity ratio based Mahalanobis proximity (SRMP). Although these approaches can effectively be used for single band single polarized images, one evident limitation is their applicability within multi polarized setting.

To evaluate the segmentation performance of the above newly formulated methods as well as three additional popular algorithms, experiments are conducted on two

speckled synthetic and six real SAR images. These three are TURBOPIXELS, ERS and SEEDS algorithms whose codes are publicly available. The reason to include these three algorithms into comparison is due to their computational requirements that are all at a competitive level with SLIC. These experiments demonstrated that SRMP algorithm is more robust to speckle noise and capable of generating well boundary adhered and compact superpixels as compared to other four algorithms.

In SRMP approach, similarity ratio and Mahalanobis proximity metrics are parameterized with a constant parameter that is kept constant for all clusters during the course of the algorithm. Even the constant parameter is properly determined, it might produce irregular shaped rather than compact superpixels within homogeneous regions. This is especially the case for heavily speckled SAR images. Therefore, the constant parameter needs to be determined during the course of the SRMP algorithm in an adaptive way. This is performed with a function developed and inserted into SRMP. The compactness parameter is determined as a function of the difference of the mean values of two neighboring superpixels. In this way, the proposed similarity ratio based adaptive Mahalanobis proximity algorithm (SRAMP) adapts the intensity similarity and the spatial proximity within each cluster. To explore the segmentation capabilities of this adaptive SRMAP and SRMP algorithms, experiments are conducted by using the same two speckled synthetic and six real SAR images. It has been found that SRAMP and SRMP are comparable in terms of boundary recall but SRAMP has lower segmentation error than SRMP implying that it has high segmentation performance.

For visual comparison of the superpixels produced by SLIC, TURBOPIXELS, ERS, SEEDS, SREP, SRMP and SRAMP approaches, experiments are conducted on six different real SAR images. The comparison of the results indicates that the superpixels generated by SRMP and SRAMP approaches are more visually acceptable than those produced by the other algorithms. This result is expected based on the findings with the synthetic images. If only those produced by SRMP

and SRAMP are compared, it can be said that SRAMP generates more visually pleasing superpixels than that are produced by SRMP. This segmentation performance of SRAMP approach is due to the use of an adaptive parameter that regulates the compactness in homogeneous regions and boundary adherence in heterogeneous regions. Evidently, this outperforming capability of SRAMP is achieved at the expense of computational cost. This cost is relatively higher than that of SLIC, SREP and SRMP algorithms.

Despite the efforts devoted towards the development of a robust superpixel generating algorithm for SAR images in this dissertation, there are several areas that still need to be researched. One area might be related with the size of grids to be used as search areas. The user specified grid sizes determine the number of superpixels (K). For large values of K, the boundaries of the grid structure are properly preserved but not the region boundaries; on the other hand, for small values of K, the region boundaries are properly preserved but not the boundaries of the grid structure. Thus, searching for means to optimize this parameter automatically is one possible future research direction. The other area might be related with searching the ways to decrease the computational requirement. Although the proposed algorithm is simple to understand and implement, its computational cost can be considered as a drawback when compared to SLIC as the current state-of-the-art algorithm. Another possible area of future work is to investigate the use of nonlinear combination of similarity ratio and spatial proximity to explore the outperforming capabilities of such combination.

REFERENCES

[1] Achanta, Radhakrishna, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. "SLIC Superpixels." Tech. rep., EPFL, 2010.

[2] Achanta, Radhakrishna, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Süsstrunk Sabine. "SLIC Superpixel Compared To State-of-the-Art Superpixel." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012): 2274-2281.

[3] Akyilmaz, Emre, and Ugur Murat Leloglu. "Segmentation of SAR images using similarity ratios for generating and clustering superpixels." *Electronics Letters* 52 (2016): 654-656.

[4] Balzter, Heiko, Beth Cole, Christian Thiel, and Christiane Schmullius. "Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests." *Remote Sensing* 7 (2015): 14876-14898.

[5] Besag, Julian. "On the statistical anaysis of dirty pictures." *Journal of the Royal Statistical Society. Series B* 48 (1986): 259-302.

[6] Bishop, C. M. *Neural Networks for Pattern Recognition.* Oxford, U.K: Oxford Univ.Press., 1995.

[7] Blekas, K., A. Likas, N. P. Galatsanos, and I. E. Lagaris. "A spatially constrained mixture model for image segmentation." *IEEE Trans. Neural Netw.* 16 (2005): 494-498.

[8] Bugden, J. L., J. Andrey, and P. J. Howarth. "A SAR process model for land-cover mapping." *Canadian Journal Of Remote Sensing* 30 (2004): 195-204.

[9] Carson, C., S. Belongie, H. Greenspan, and Jitendra Malik. "Blobworld: Image segmentatio using expectation-maximization and its applications to image querying." *IEEE Pattern Anal. Mach. Intell.* 24 (2002): 1026-1038.

[10] Celeux, G., F. Forbes, and N. Peyrard. "EM procedures using mean field like aproximations for Markov model-based image segmentation." *Pattern Recognition* 36 (2003): 131-144.

[11] Comaniciu, Dorin, and Peter Meer. "Mean Shift : A Robust Approach Toward Feature Space Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002): 603-619.

[12] Curlander, C. John, and N. Robert McDonough. *Synthetic Aperture Radar: Systems and Signal Processing.* John Wiley and Sons Inc., 1992.

[13] Dempster, A. P., N. M. Laird, and D. B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society. Series B* 39 (1977): 1-21.

[14] Dutta, A., and K. K. Sarma. "SAR image segmentation using wavelets and Gaussian mixture model." *IEEE Int. Conf. Signal Process. Integ. Netw*. 2014. 766-770.

[15] Ester, M., H. P. Kriegel, J. Sander, and X. Xu. "A density-based algorithm for discovering clusters in latge spatial databases with noise." *Proc. of Second International. Conf. on Knowledge Discovery and Data Mining.* 1996. 226-231.

[16] Felzenszwalb, Pedro F., and Daniel P. Huttenlocher. "Efficient Graph-Based Image Segmentation." *Inernational Journal of Computer Vision* 59 (2004): 167-181.

[17] Feng, Hongxiao, Biao Hou, and Maoguo Gong. "SAR Image Despeckling Based on Local Homogeneous-Region Segmentation by Using Pixel-Relativity Measurement." *IEEE Transactions on Geoscience and Remote Sensing* 49 (2011): 2724-2737.

[18] Fjortoft, R., Y. Delignon, Wojciech Pieczynski, Marc Sigelle, and Florence Tupin. "Unsupervised classification of radar images using hidden Markov chains and hidden random fields." *IEEE Trans. Geosci. and Remote Sens.* 41 (2003): 675-686.

[19] Ford, L. R., and D. R. Fulkerson. "Maximal flow through a network." *Canadian Journal of Mathematics* 8 (1956): 399-404.

[20] Frost, V. S., J. A. Stiles, K. S. Shanmugam, J. C. Holtzman, and S. A. Smith. "An adaptive filter for smoothing noisy radar images." *Proceedings of the IEEE* 69 (1981): 133-135.

[21] Gauthier, Marie France, Laurie Weir, Zigiang Ou, Matt Arkett, and Roger De Abreu. "Integrated satellite tracking of pollution: A new operational program." *Geoscience and Remote Sensing Symposium.* 2007.

[22] Geman, Stuart, and Donald Geman. "Stoachastic relaxation: Gibbs distributions and the Bayesian restoration of images." *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (1984): 721-741.

[23] Gonzalez, Rafael C., and Richard E. Woods. *Digital Image Processing.* Upper Saddle River, NJ: USA: Prentice Hall, 2002.

[24] Greig, D. M., B. T. Porteous, and A. H. Seheult. "Exact maximum a posteriori estimation for binary images." *Journal of the Royal Statistical Society. Series B* 51 (1989): 271-279.

[25] Hammersley, J. M., and P. Clifford. "Markov Fields on Finite Graphs and Lattices." 1971.

[26] Kalti, Karim, and M. A. Mahjoub. "Image segmentation by Gaussian mixture models and modified FCM algorithm." *Int. Arab Jour. Inf. Tech.(IAJIT)* 11 (2014): 11-18.

[27] Kovesi, Peter. "Image Segmentation using SLIC SuperPixels and DBSCAN Clustering." 2013.

[28] Kuan, D. T., A. A. Sawchuk, T. C. Strand, and P. Chavel. "Adaptive restoration of images with speckle." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 35 (1987): 373-383.

[29] Lankoande, Ousseini, Majeed M. Hayat, and Balu Santhanam. "Segmentation of SAR Images Based on Markov Random Field Model." *2005 IEEE International Conference on Systems, Man and Cybernetics,*. 2005. 2956-2961.

[30] Lee, Jong-Sen. "A simple speckle smoothing algorithm for synthetic aperture radar images." *IEEE Transactions on Systems, Man and Cybernetics* SMC-13 (1983): 85-89.

[31] Lee, Jong-Sen, and Eric Pottier. *Polarimetric Radar Imaging From Basic to Applications.* Flori, d: CRC Press Taylor and Francis Group, 2009.

[32] Levinshtein, Alex, Adrian Stere, Kiriakos N. Kutulakos, David J. Fleet, and Sven J. Dickinson. "TurboPixels : Fast Superpixels Using Geometric Flows." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2009): 2290-2297.

[33] Li, Junhua, and Wenjun Chen. "Clustering synthetic aperture radar (SAR) imagery using an automatic approach." *Canadian Journal Of Remote Sensing* 33 (2007): 303-311.

[34] Li, Stan Z. *Markov random fields and modelling in Image Analysis.* Lond, o: UK: Springer, 2009.

[35] Liu, Ming Yu, O. Tuzel, Srikumar Ramalingam, and R. Chellappa. "Entropy-rate clustering: cluster analysis via maximizinga submodular function subject to a matroid constraint." *IEEE Trans.Pattern Anal. and Mach.Intell* 36 (2013): 99-112.

[36] Macqueen, J. "Some Methods For Classification and Analysis of Multivariate Observations." *Proceedings of 5th Berkeley Symposium on Mathematiacal Statistics and Probability.* 1967. 281-297.

[37] Matas, J., O. Chum, M. Urban, and T. Pajdla. "Robust wide baseline stereo from maximally stable extremal regions." *British Machine Vision Conference.* 2002. 384-396.

[38] McAndrew, Alasdair. *An introduction to digital image processing with Matlab.* Boston, Mass., USA: Thomson/Course Tech, 2004.

[39] Mccandless, Samuel W. Walt, and Christopher R. Jackson. "Principles of Synthetic Aperture Radar." Chap. Chapter 1 in *National Oceanic and Atmospheric Administration*, 1-23. 2004.

[40] McLachlan, G., and D. Peel. *Finite Mixture Models.* New, York: USA: Wiley, 2000.

[41] Mester, Rudolf, and Uwe Franke. "Statistical Model based image segmentation using region growing contour relaxation, and

classification." *Proc. SPIE Cambridge Symposium on Visual Communication and Image Processing.* SPIE, 1988. 624-626.

[42] Moore, Alastair P., Simon J. D. Prince, Jonathan Warrell, Umar Mohammed, and Graham Jones. "Superpixel Lattices." *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008.* 2008.

[43] Murtagh, Fionn, Adrian E. Raftery, and Jean Luc Starck. "Bayesian inference for multiband image segmentation via model-based cluster trees." *Image and Vision Computing* 23 (2005): 587-596.

[44] Neubert, P., and P. Protzel. "Superpixel Benchmark and Comparıson." Tech. rep., Chemnitz University of Technology, Dept. of Elec. Eng. and Inf. Tech, 2013.

[45] Nguyen, T. M., and Q. M. J. Wu. "Gaussian mixture model based spatial neighborhood relationships for pixel labeling problem." *IEEE Trans. Syst., Man, Cybern.* 42 (2012): 193-202.

[46] Nikhil, R. Pal, and K. Pal Sankar. "A review on image segmentation techniques." *Pattern Recognition* 26 (1993): 1277-1294.

[47] Nikou, Christophoros, N. P. Galatsanos, and A. Likas. "A class-adaptive spatially variant mixture model for image segmentation." *IEEE Trans. Image Process.* 16 (2007): 1121-1130.

[48] Okman, O. E., F. Nar, C. Demirkesen, and M. Çetin. "Feature Preserving SAR Despeckling and Its Parallel Implementation with Application to Railway Detection." *EUSAR. 9th European Conference.* 2012.

[49] Osher, Stanley, and J. A. Sethian. "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations." *Journal Of Computational Physics* 79 (1988): 12-49.

[50] Parrilli, Sara, Mariana Poderico, Cesario Vincenzo Angelino, and Luisa Verdoliva. "A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage." *IEEE Transactions on Geoscience and Remote Sensing* 50 (2012): 606-616.

[51] Pham, D. T., S. S. Dimov, and C. D. Nguyen. "Selection of K in K-means clustering." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science.* 2004.

[52] Prince, Simon J. D. *Computer vision: models, learning and inference.* Cambridge University Press, UK, 2012.

[53] Pyun, K., J. Lim, C. S. Wun, and R. M. Gray. "Image segmentation using hidden markov Gauss mixture models." *IEEE Trans. Image Process.* 16 (2007): 1902-1910.

[54] Ren, Xiaofeng, and Jitendra Malik. "Learning a Classification Model for Segmentation." *Proceedings. Ninth IEEE International Conference.* 2003. 10-17.

[55] Richard, John A., and Xuiping Jia. *Remote Sensing Digital Image Analysis.* Verlag, Berlin: Springer, 2006.

[56] Sanjay, G. S., and T. J. Hebert. "Bayesian pixel classification using spatially variant fixture model for image segmentation." *IEEE Trans. Image Process* 7 (1988): 1014-1028.

[57] Sarkar, S., and K. L. Boyer. "Computing perceptual organization using voting methods and graphical enumeration." *Inter. Conf. Pattern Recognition.* 1992.

[58] Shi, Jianbo, and Jitendra Malik. "Normalized Cuts and Image Segmentation." *EEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000): 888-905.

[59] Silverman, Judith F., and D. B. Cooper. "Bayesian Clustering for Unsupervised Estimation of Surface and Texture Models." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10 (1988): 482-495.

[60] Titterington, D. M., A. F. M. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions.* Haboken, N: Wiley, 1985.

[61] Van den Bergh, Michael, Xavier Boix, Gemma Roig, Benjamin de Capitani, and Luc Van Gool. "SEEDS : Superpixels Extracted via

Energy-Driven Sampling." *12th European Conference on Computer Vision.* 2012. 13-26.

[62] Vedaldi, Andrea, and Stefano Soatto. "Quick Shift and Kernel Methods for Mode Seeking." *Proc. European Conf. Computer Vision.* 2008. 705-718.

[63] Veksler, Olga, Yuri Boykov, and Paria Mehrani. "Superpixels and Supervoxels in an Energy Optimization Framework." *11th European Conference on Computer Vision.* 2010. 211-224.

[64] Webb, A. R. *Statistical Pattern Recognition.* West, Sussex: John Wiley and Sons, Ltd., 2002.

[65] Wertheimer, Max. "Laws of Organization in Perceptual Forms." *A Sourcebook of Gestalt Psychology*, 1938: 71-88.

[66] Yu, Peter, A. K. Qin, and David A. Clausi. "Unsupervised Polarimetric SAR Image Segmentation and Classification Using Region Growing With Edge Penalty." *IEEE Trans. Geosci. Remote Sensing* 50 (2012): 1302-1317.

[67] Yu, Yongjian, and Scott T. Acton. "Speckle Reducing Anisotropic Diffusion." *IEEE Transactions on Image Processing* 11 (2002): 1260-1270.

[68] Zaart, Ali El, Djemel Ziou, Shengrui Wang, and Qingshan Jiang. "Segmentation of SAR images." *Pattern Recognition* 35 (2002): 713-724.

[69] Zhang, H., Q. M. J. Wu, T. M. Nguyen, and X. Sun. "Synthetic Aperture Radar image segmentation by modified Student's t-mixture model." *IEEE Trans. Geosci. Remote Sensing* 52 (2014): 4391-4403.

[70] Zhang, Y., M. Brady, and S. Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation– maximization algorithm." *IEEE Trans. Med. Imag.* 20 (2001): 45-57.

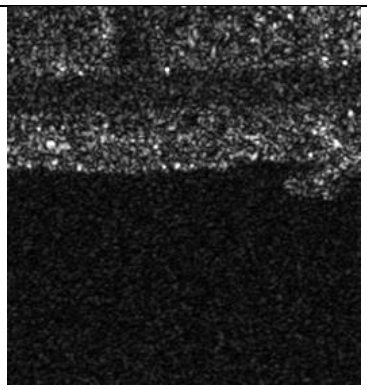
[71] Zhu, Song, Danhua Cao, Shixiong Jiang, Yang Wu, and Pan Hu. "Fast superpixel segmentation by iterative edge refinement." *IEEE Electronics Letters* 51 (2015): 230-232.

APPENDICES

APPENDIX A

DATA USED IN EXPERIMENTS

In this appendix, the detailed information about real SAR images which are used in this thesis are given. Since it is not always easy to interpret SAR images, the images that are taken from Google Earth™ are also provided.



TerraSAR-X

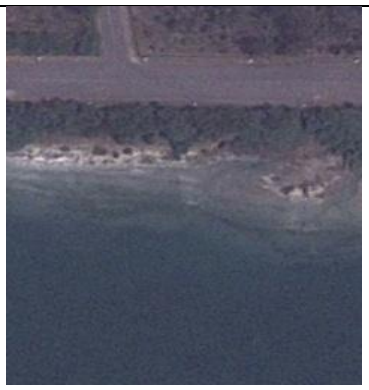


Image 2016 CNES/Astrium

File Name:
TSX1_SAR__GEC_SE___HS_S_SRA_20081012T123311_20081012T123312
Location: Vishakpatnam / India
Image Look Direction: Right
Image Polarization: VV
Coordinates:

- Upper Left: (8528, 7106) Lower Right: (8790, 7386)

Content:

- Man made structure: Road
- Sea, costline and vegetation

Figure A.1 Original Image with Its Optical Correspondence

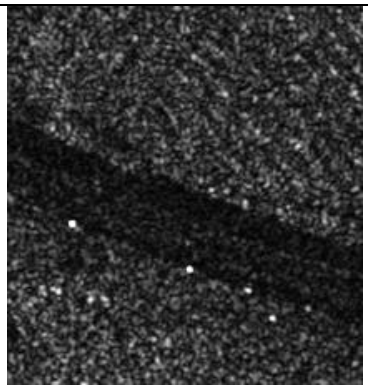
<table>
<tr>
<td> 
TerraSAR-X</td>
<td> 
Image 2016 CNES/Astrium</td>
</tr>
<tr>
<td colspan="2">File Name:
TSX1_SAR__GEC_SE___HS_S_SRA_20081012T123311_20081012T123312
Location: Vishakpatnam / India
Image Look Direction: Right
Image Polarization: VV
Coordinates:
• Upper Left: (8077, 6285) Lower Right: (8291, 6513)
Content:
• Man made structure: Road
• Arid land</td>
</tr>
</table>

Figure A.1 (continued)

TerraSAR-X

Image 2016 CNES/Astrium

File Name:

TSX1_SAR__GEC_SE___HS_S_SRA_20081012T123311_20081012T123312

Location: Vishakpatnam / India

Image Look Direction: Right

Image Polarization: VV

Coordinates:

- Upper Left: (3222, 5144) Lower Right: (3535, 5530)

Content:

- Man made structure: Airfield
- Arid land

Figure A.1 (continued)

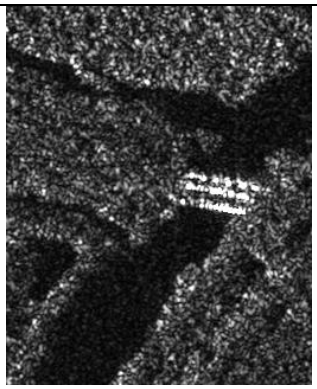
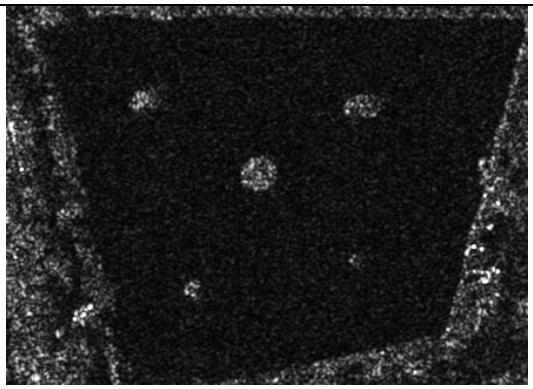
<table>
<tr><td> 
TerraSAR-X</td><td> 
Image 2016 CNES/Astrium</td></tr>
<tr><td colspan="2">File Name:
TSX1_SAR__GEC_SE___HS_S_SRA_20081012T123311_20081012T123312
Location: Vishakpatnam / India
Image Look Direction: Right
Image Polarization: VV
Coordinates:
• Upper Left: (5537, 8005) Lower Right: (5795, 8325)
Content:
• Man made structures Road, Bridge
• Vegetation and River</td></tr>
</table>

Figure A.1 (continued)



TerraSAR-X

Image 2016 CNES/Astrium

File Name:

TSX1_SAR__GEC_SE___HS_S_SRA_20081012T123311_20081012T123312

Location: Vishakpatnam / India

Image Look Direction: Right

Image Polarization: VV

Coordinates:

- Upper Left: (10474, 6407) Lower Right: (5795, 8325)

Content:

- Arid Land, Mud / Swamp

Figure A.1 (continued)

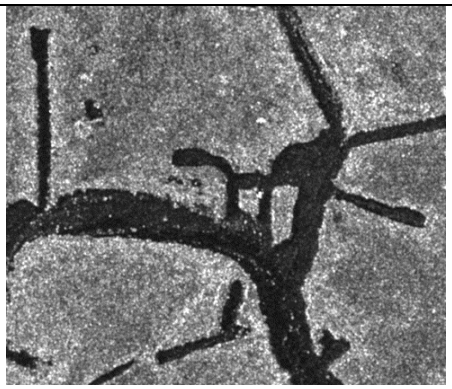
<table>
<tr><td> 
TerraSAR-X</td><td> 
Image Landsat / Copernicus</td></tr>
<tr><td colspan="2">File Name:
TSX1_SAR__GEC_RE___SM_S_SRA_20080201T235718_20080201T235726
Location: Mississippi / USA
Image Look Direction: Right
Image Polarization: HH
Coordinates:
• Upper Left: (8113, 3217) Lower Right: (8674, 3714)
Content:
• River, Vegetation and Marsh</td></tr>
</table>

Figure A.1 (continued)

APPENDIX B

SAR IMAGE GROUNDTRUTHS

The groundtruths of the SAR images that are used in this thesis are shown below. These groundtruths are manually preapared using optical correspondence of the SAR images.

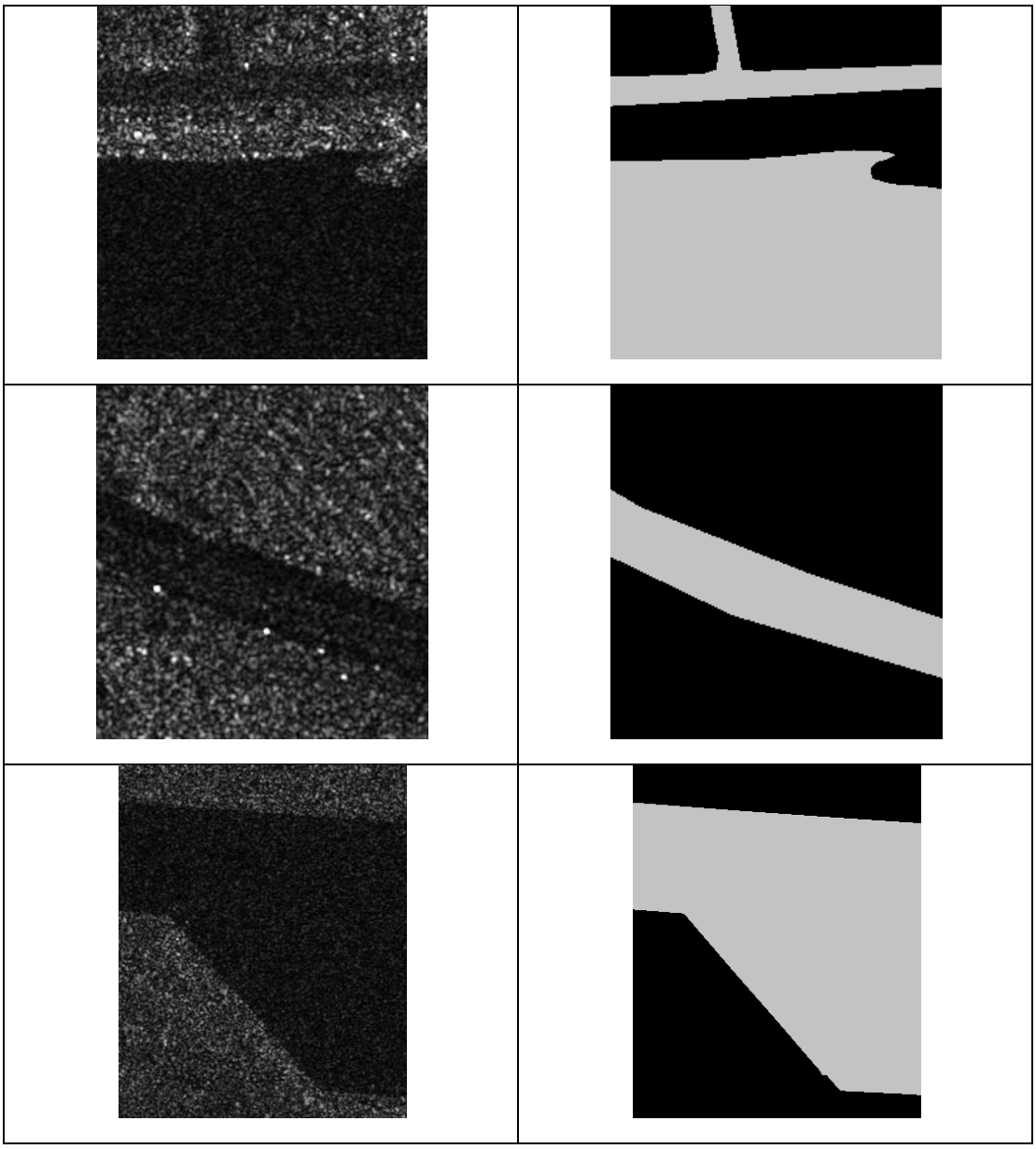


Figure B.1 Original Images with Their Groundtruths

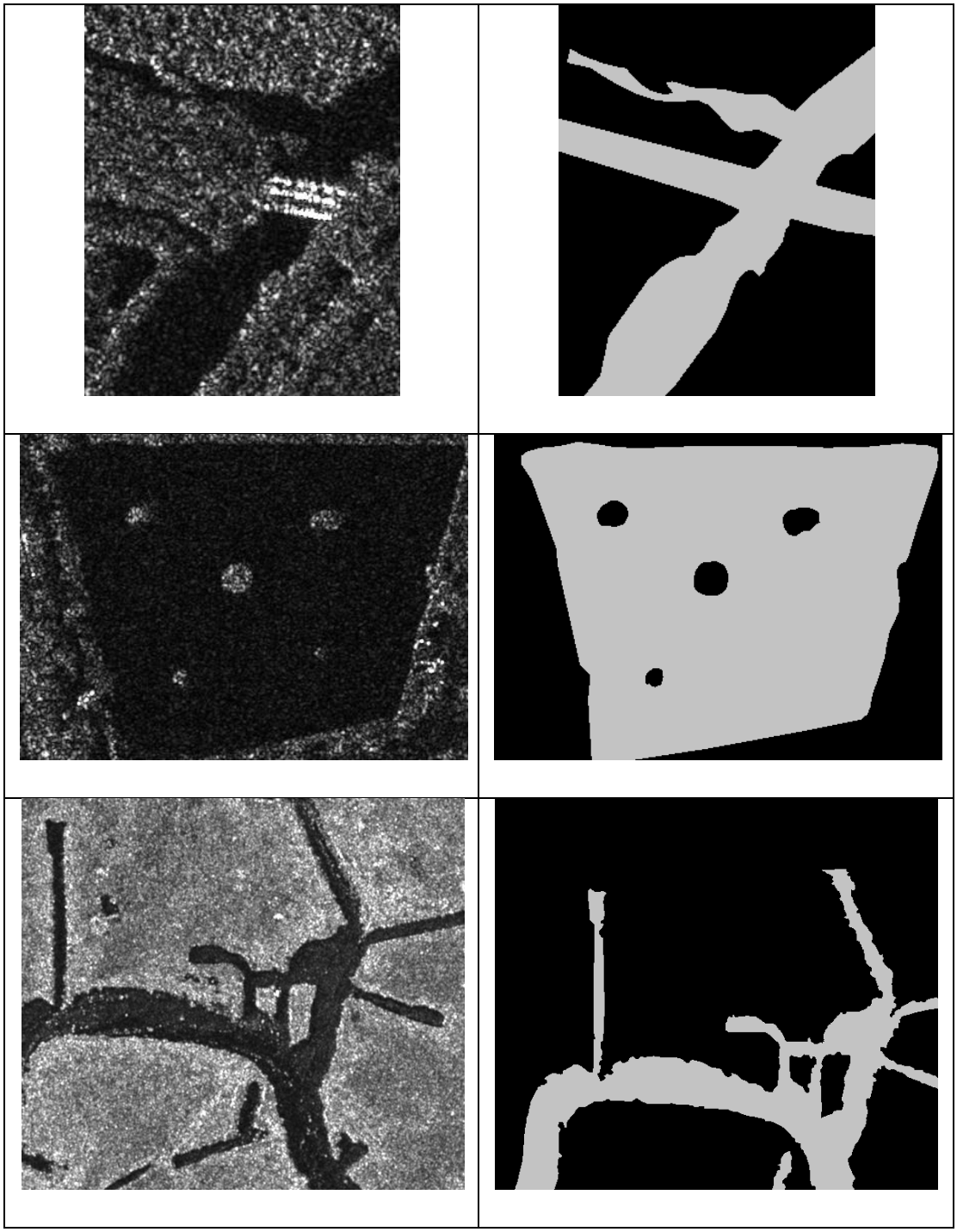


Figure B.1 (continued)

APPENDIX C

ALGORITHM RUN TIMES

The run times for the similarity ratio based and SLIC algorithms for the generation of the various number of superpixels can be visualized from the following chart.

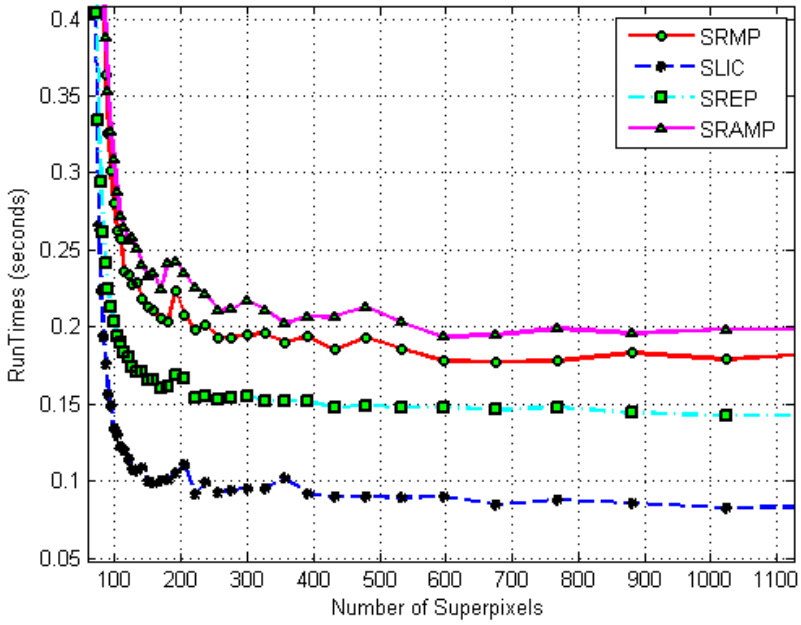


Figure C.1 Algorithm Run Times

The average of all those run times for each algorithm are shown in the table below.

Table C.1 Averaged Algorithm Run Times

	Average	**Min**	**Max**
SRMP	0.2552	0.1654	0.9930
SLIC	0.1229	0.0802	0.4498
SREP	0.1847	0.1376	0.5190
SRAMP	0.2762	0.1794	0.9937

The run times are calculated on Intel(R) Core(TM) i7-4770 CPU @ 3.40 GHz processor with 8 GB DDR3 RAM.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Akyılmaz, Emre
Nationality: Turkish (TC)
Data and Place of Birth: 12 September 1985, Ankara
Marital Status: Married
Phone: +90 533 416 2015
E-mail: emreakyilmaz@gmail.com

EDUCATION

Fall 2009 – Spring 2017	Ph.D in Geodetic and Geographic Information Systems, Middle East Technical University
Fall 2007 – Spring 2009	M.Sc. in Software Engineering Middle East Technical University
Fall 2003 – Spring 2007	B.Sc in Computer Engineering Bilkent University

WORK EXPERIENCE

December 2015 – (present)	**Senior Software Engineer:** TAI (Turkish Aerospace Industries, Inc.)
October 2011 – December 2015	**Senior Software Engineer:** SDT (Space & Defense Technologies) Inc.
March 2008 – April 2011	**Software Engineer:** SDT (Space & Defense Technologies) Inc.
September 2007 – February 2008	**Software Test Engineer:** Aydın Yazılım – AYESAS, Ankara

PUBLICATIONS

1. "Automatic and Semi-automatic Extraction of Curvilinear Features from SAR Images", Emre Akyılmaz, O. Erman Okman, Fatih Nar, and Müjdat Çetin, SPIE Security + Defence Symposium, SAR Image Analysis, Modeling, and Techniques Conference, Edinburgh, United Kingdom, September 2012.

2. “Interactive Ship Segmentation in SAR Images”, Emre Akyılmaz, Can Demirkesen, O. Erman Okman, Fatih Nar and Müjdat Çetin, IEEE Conference on Signal Processing and Communications Applications, Cyprus, April 2013.
3. “Information-theoretic noisy band detection in hyperspectral imagery ”, Mustafa Ergül, Fatih Nar, Emre Akyılmaz, Nigar Şen and Müjdat Çetin, IEEE Conference on Signal Processing and Communications Applications, Malatya, April 2015.
4. "Segmentation of SAR images using similarity ratios for generating and clustering superpixels", Electronics Letter, IET, April 2016
5. "Similarity Ratio Based Adaptive Mahalanobis Distance Algorithm to Generate SAR Superpixels", Canadian Journal of Remote Sensing, (in revision)