COVARIANCE BASED ANALYSIS OF RELEVANT SCATTERERS IN HIGH RESOLUTION SAR IMAGES

Houda Chaabouni-Chouayakh\(^1\) and Mihai Datcu\(^{1,2}\)

\(^1\)DLR, German Aerospace Center, Oberpfaffenhofen D-82234 Wessling - Germany
\(^2\)ENST, Ecole Nationale Supérieure des Télécommunications, 46 rue Barrault 75634 Paris Cedex 13 - France

ABSTRACT

The high diversity of man-made structures combined with the complexity of the scattering processes makes the analysis and information extraction from high resolution Synthetic Aperture Radar (SAR) images over urban areas non-trivial. In order to simplify interpretation and information extraction, the detection of the so-called Relevant Scatterers (RSs), is proposed in this paper. The advantage of such RSs, is that they have a stable description, allowing a better discrimination from the rest of the scene. This work addresses a RSs characterization problem for high resolution SAR Automatic Target Detection (ATR), based on the covariance analysis combined with the azimuth decomposition. Indeed, the covariance matrix and its spectrum of eigenvalues are of great interest in the analysis and modeling of experimental data. The detection/characterization is obtained by performing projections of the training data in the eigenspace generated by the covariance formalism. In this article, both a description of our work and a presentation of our preliminary performance results will be provided.

Key words: SAR, relevant scatterers, azimuth decomposition, covariance matrix, eigenspace.

1. INTRODUCTION

Synthetic Aperture Radar (SAR) is a coherent imaging mode in the microwave domain ([Schreider93, Elachi87]). Electromagnetic scattering in urban areas is characterized by a variety of single or multiple scattering mechanisms with a wide range of scattering signatures. Moreover, SAR images over urban areas are strongly affected by geometric distortion effects (as layover, shadowing) due to the combination of the SAR side-looking acquisition and the stepwise height variations within the scene. This makes the interpretation and information extraction from SAR images more complex.

In particular, with the increase of the SAR sensor resolution, the high resolution SAR images could include a large variety of real man-made objects (buildings, parking, ground vehicles,...). Also, the man-made objects have various designs and shapes that are too complicated and change too much to allow a realistic modeling in a fashion simple enough to lead to a practical technology. A full understanding of the behaviors of the different types of targets becomes thus, not easily reachable. A description of the backscattering behaviors of man-made targets, general scatterer classes and their significance, target modeling, is presented in more details in [Rihaczek96].

In the SAR Automatic Target Recognition (ATR), it is important to be able to reliably detect or classify a target in a manner which provides the largest possible robustness to target and clutter variability, with the highest possible discrimination capability.

To better detect/characterize the Relevant Scatterers (RSs) in urban areas, the azimuth sub-band decomposition was found in [Chaabouni06] to be a powerful tool since it exploits at most the azimuth spectrum, which is very rich in the case of the high resolution SAR images. In [Tupin04, Tison04], the azimuth sub-band decomposition was used to enhance areas of interest (stable or unstable man-made structures according to some automatic tools such as mutual information). the same approach was also used in [Schneider06] for the polarimetric and interferometric characterization of the point-like coherent scatterers in urban areas.

Among the ATR systems, the covariance based methods seems to be quite promising. In fact, [Kim01] addressed an adaptative target detection problem in radar imaging for which the covariance matrix of an unknown Gaussian clutter background is assumed to have a block diagonal structure. This block diagonal structure is the consequence of a target lying along a boundary between two statically independent clutter regions. Under this assumption, some detection strategies (generalized likelihood ratio and maximal invariant) were then investigated, and it was proved that each of these detectors could be of great help in image detection problems involving boundary and target interactions.
In [Pentland95], an unsupervised technique for visual target modeling, which is based on density estimation in high-dimensional spaces using an eigenspace decomposition. Such a decomposition was proved to be well-suited for the detection of facial features. Like the human faces, the SAR images over urban areas provide a high diversity of features. Thus, an eigenspace approach should be also adapted to the recognition of urban scenes in high resolution SAR images.

In [Novak94], it was demonstrated that the eigenspace relative to the covariance matrix of the training images, provides a well-suited descriptive model of the scene. A principal component analysis is then performed on the generated eigenspace in order to identify the eigenimages that provide the best discrimination between the different classes. This approach was applied for radar target identification in a three-class-database formed by tanks, APCs and self-propelled guns. Such special targets are unfortunately not usual in urban areas. In fact, the mostly found classes in these areas, consist rather in high/small buildings, vegetation, roads, parking... Processing these classes is much more complex than tanks, APCs and guns, where the distribution variety of the targets is not too large.

In this paper, detection algorithms are developed for RSs in high resolution SAR images by combing the promising properties of the azimuth sub-band decomposition with the ones of the covariance matrix. These techniques have also the advantage that they analyze the complex image rather than only the intensity images. In fact, the intensity image produced by a radar, does not contain sufficient information for target identification or target detection under adverse circumstances.

A brief overview of the paper follows. Section 2 gives some basic principals about SAR data acquisition and image formation. Section 3 is dedicated to the description of the azimuth sub-band decomposition algorithm and its application for high resolution SAR image analysis. In section 4, the covariance formalism and the way we propose to use it are exposed. Finally, section 5 summarizes the important results of our study.

2. BASIC PRINCIPALS OF SAR

SAR synthesizes a long antenna by transmitting pulsed signals and coherently adding the successively reflected and received pulses to obtain high resolution in the flight (azimuth) direction. The resolution in range direction (orthogonal to the azimuth direction) is achieved by transmitting either very short or otherwise large bandwidth pulses, called chirp.

The fact that the antenna is moving in the azimuth direction during the illumination time results in a Doppler effect. This Doppler effect spreads the energy over the azimuth spectrum. A component of the frequency spectrum characterizes the amount of energy acquired in a given geometric configuration. When the beam axis is perpendicular to the flight direction, the average azimuth frequency called the Doppler centroid is null. But, if the antenna is tilted in the azimuth direction, the Doppler centroid differs from zero.

The basics of SAR theory are described in more details in [Schreider93, Elachi87, Jakowatz96].

3. AZIMUTH SUB-BAND DECOMPOSITION OF HIGH RESOLUTION SAR IMAGES

Several techniques of frequency analysis could be applied to a signal. For high resolution SAR images, the azimuth sub-band decomposition seems to be a promising tool to analyze the behavior of scatterers and to study some of their properties ([Chaabouni06, Tison04, Tupin04]). Indeed, unlike most of the natural images, SAR data are complex signals and their spectrum has a specific meaning. The azimuth direction is along the flight axis and each position corresponds to some frequency variations due to the Doppler effect. Each point in the scene, is illuminated many times by the radar beam. A selection of an azimuth sub-aperture corresponds thus, to a selection of a range of viewing angles or sensor positions.

Due to the particular fine backscattering phenomena in urban areas and the directivity property of the illuminated objects (depending on their orientations, the material of their surroundings surfaces,...), the signal of a sub-band aperture can be quite different from both, the full spectrum signal and the other generated sub-bands. For instance, rough surfaces are quasi-Lambertian and isotropic when the roughness is high according to the wavelength. Therefore, the same backscattering intensity should be observed in each sub-band. However, for some man-made objects in urban areas, such as a smooth wall or dihedral, the backscattered signal is highly dependant on the relative direction of the incidence wave and the object. In this case, the target could be faded or even disappear in some sub-bands, for which the object is badly orientated.

In our work, for sake of simplicity, we chose to undergo a division of the spectrum into two, but the cases of $n > 2$ could also be studied.

The 2 sub-aperture decomposition is made by:

- **Step 1**: Doppler centroid estimation and compensation of Doppler shift (in [Madsen89], three Doppler centroid estimators were proposed);
- **Step 2**: Unweighting in azimuth in order to obtain a uniform spectral density in the useful spectrum (in our work, we use a Hamming function for the unweighting step since a focused SAR image is usually weighted with a Hamming window);
- **Step 3**: Spectrum division into 2 sub-bands;
- **Step 4**: Centering the obtained sub-images;
Step 5: Zero-padding and Hamming weighting of each sub-band in order to suppress the sidelobes. This step is essential in urban areas due to the presence of many strong point-like scatterers.

It is noted that, the azimuth resolution of the regenerated signals is degraded by a factor of 2 according to the original resolution.

A 2-azimuth sub band decomposition algorithm is described in the figure 1.

Figure 1. Steps of the 2 azimuth sub-band decomposition.

The figure 2 shows an example of a 2-azimuth decomposition of high resolution SAR images.

Figure 2. (a) original image, (b) sub-band left and (c) sub-band right, obtained after a 2-azimuth sub-band decomposition.

From the figure 2, many interesting effects can be observed:

1. Evidence of some details which were not in the original images: Since the full-band image corresponds to a complex average of the zero-padded images, there are configurations where the structures do not appear at all in the full resolution image, although they are clearly seen in a sub-band. For instance, most of the roofs of the buildings have different appearances depending on the sub-bands. They were already in the original image but their contributions are much more important in the two sub-bands (specially the left sub-band). This phenomena can be seen in the areas with red frames. Moreover, the backscattering could change from the left to the right sub-band in the case of the red frame in the top. Indeed, the backscatterers in this case have probably a more adapted orientation.

2. Loss or fading of some structures in the sub-band images: Some configurations (like the ones in the yellow frames) lose some particularities (geometry, aspect and shape) in the sub-bands in comparison to the original image (specially in the sub-band right). This is the case of the structures whose backscattering depends on the relative direction of the incidence wave and the object.

3. Low directivity of the corner reflectors: The corner reflectors (an example is shown in figure 2 as a blue frame) appear in all the sub-bands with a high intensity. In fact, their backscattering does not depend on the orientation or the position of the sensor.

4. COVARIANCE ANALYSIS

This section is dedicated to a SAR target classification based upon the eigen-image concept. Such a concept was demonstrated to be quite powerful for radar target identification in [Pentland95] and for automatic recognition of human faces in [Novak94].

This method is based on a covariance analysis formalism, from which an eigenspace is generated. After that, a principal components analysis is performed on the training images in order to determine those eigen-images that best account for the distribution of all the images within the space. The components of the projections in the generated eigenspace are then used to undergo the classification.

In our work, we will present a new version of this method which exploits the information given by the azimuth sub-band decomposition to improve the scatterers’ classification in high resolution SAR images. This section describes thus, the different steps of our new eigen-image classifier and presents some preliminary results performed on a five-class database.
4.1. General covariance algorithm description

Given a set of 2-D target images having \( n \) by \( n \) pixels, we can form a training set of vectors \( \{X_i\} \), where \( X_i \in \mathbb{C}^{n^2} \) by lexicographic ordering of the pixel elements of each target image. The database that we will use in our experiments is formed by Big Buildings (BB), Average Buildings (AB), Small Buildings (SB), Vegetation (V) and Water (W). The size of the whole database is 250 images (50 images of each class). The half of the target images of each class type are used for training the eigenimage classifier. The whole database will then be used to test the classifier.

The figure 3 presents a flowchart of the covariance algorithm.

![Flowchart of the covariance algorithm implementation.](image)

Let the training images be denoted by vectors \( X_1, X_2, ..., X_M \).

Then, we consider the averaged training images:

\[
\bar{X}_i = X_i - \bar{X} \quad ; \quad i = 1, 2, ..., M
\]

(1)

where \( \bar{X} \) is assumed to be the average image defined as:

\[
\bar{X} = \frac{1}{M} \sum_{i=1}^{M} X_i
\]

(2)

Let the training matrix \( \bar{X} \) be formed from the vectors \( \bar{X}_i \) as follows:

\[
\bar{X} = [\bar{X}_1, \bar{X}_2, ..., \bar{X}_M]
\]

(3)

The sample covariance matrix \( \Sigma \) is then computed as:

\[
\Sigma = \frac{1}{M} \bar{X} \bar{X}^T
\]

(4)

\( \bar{X}^T \) denotes the transpose conjugate operator.

Note that the sample covariance matrix has dimension \( n^2 \times n^2 \) but only the rank \( M \), thus there will be only \( M \) meaningful eigenvectors corresponding to the \( M \) positive eigenvalues of the sample covariance matrix (the remaining eigenvalues are all zeros). The principal components analysis requires that we solve the eigenvalues equations:

\[
\Sigma Z_i = \lambda_i Z_i \quad ; \quad i = 1, 2, ..., M
\]

(5)

where \( \lambda_1, \lambda_2, ..., \lambda_M \) are supposed to be the \( M \) positive eigenvalues of \( \Sigma \).

The subset of the eigenvectors corresponding to the largest eigenvalues, provides the best representation of the training images. We refer to this subset of eigenvectors as eigen-images.

To avoid the long computations due to the high dimensionality of the covariance matrix \( \Sigma \), we will use a lower dimensional (in our case \( M \times M \)) matrix to calculate the eigenvalues and their corresponding eigenvectors. The approach can be described as follows:

1. Firstly, the eigenvalues equations (6) of the \( M \) by \( M \) matrix \( (\bar{X}^* \bar{X}) \) are solved.

\[
(\bar{X}^* \bar{X}) V_i = \lambda_i V_i \quad ; \quad i = 1, 2, ..., M
\]

(6)

Indeed, the eigenvalues computed from the previous equations are identical to the eigenvalues of the covariance matrix \( \Sigma \).

2. The desired eigenvectors or eigenimages are then given by:

\[
Z_i = \bar{X} V_i \quad ; \quad i = 1, 2, ..., M
\]

(7)

Then, the classifier is implemented as follows:

- A test image \( X_t \) is projected into the eigenspace spanned by the \( M \) eigenimages. The feature projection vector is given by the following equation:

\[
w_t = \begin{bmatrix} w_{1} \\ w_{2} \\ \vdots \\ w_{M} \end{bmatrix}
\]

(8)

where:

\[
w_i = Z_i^* (X_t - \bar{X}) \quad ; \quad i = 1, 2, ..., M
\]

- For our five-class problem, the training images of each class are projected as in equation (8) and then averaged to calculate average feature vectors \( \bar{w}_{BB}, \bar{w}_{AB}, \bar{w}_{SB}, \bar{w}_{V} \) and \( \bar{w}_{W} \).

- Finally, to perform the classification, euclidian distances from an unknown test target to the five average training feature vectors are calculated by:

\[
d_k = \| w_t - \bar{w}_k \| ^2 \quad ; \quad k = 1, 2,.., 5
\]

(9)

where 1, 2, ..., 5 correspond respectively to BB, AB, SB, V and W.
4.2. Covariance with azimuth sub-band decomposition algorithm description

From section 3, we concluded that the azimuth sub-band decomposition of high resolution SAR images over urban areas, gives more information and finer description of the man-made structures than when applying only on the original images. Thus, using the output of the azimuth decomposition as input to the covariance analysis algorithm should yield superior classification performance and better discrimination between the different classes in urban areas.

In this paper, we propose to combine the attractive properties of these two techniques (azimuth sub-band decomposition and covariance formalism) to build a new version of the previous covariance algorithm, more adapted to man-made structures classification.

The flowchart of the covariance with azimuth decomposition algorithm implementation is described in figure 4.

\[
\begin{align*}
\text{Im} & \rightarrow \text{2-Azimuth sub-band Decomposition} \\
& \rightarrow \text{Matrix to vector Converter} \\
& \rightarrow X_i^k \rightarrow \text{Covariance Algorithm} \\
& \rightarrow \text{Output} \\
\end{align*}
\]

Figure 4. Flowchart of the covariance with azimuth decomposition algorithm implementation.

Instead of working directly on the full-spectrum training images \(X_i (i = 1, 2, ..., M)\), we propose to use \(X_i^{Az} (i = 1, 2, ..., M)\) defined as follows:

\[
X_i^{Az} = \begin{bmatrix} X_i^1 \\ X_i^2 \end{bmatrix} ; \quad i = 1, 2, ..., M \tag{10}
\]

where \(X_i^k (k = 1, 2)\) denote the two sub-bands obtained after a 2-azimuth sub-band decomposition of the training image \(X_i\).

The test image \(X_t\) will also be replaced by:

\[
X_t^{Az} = \begin{bmatrix} X_t^1 \\ X_t^2 \end{bmatrix} \tag{11}
\]

where \(X_t^k (k = 1, 2)\) are the two sub-bands obtained after a 2-azimuth sub-band decomposition of the test image \(X_t\).

It is noted that:

\[
X_i^{Az} \in C^{2 \times n^2} ; \quad i = 1, 2, ..., M
\]

and

\[
X_t^{Az} \in C^{2 \times n^2}
\]

5. RESULTS AND DISCUSSION

The performance of the classification were computed as a function of the size \(n\) of the used images. Such a parameter is very important in the case of high resolution SAR images. In fact, taking into account the surrounding area (high \(n\)) and its different structures could improve the classification for some classes and worsen the classification for others.

To evaluate the classification performance, we computed the Percentage of the Good Classification (\(PGC_k\)) defined as:

\[
PGC_k = \frac{RC_k}{C_k} \times 100 ; \quad k = 1, 2, ..., 5 \tag{12}
\]

where:

- \(RC_k\): number of the well recognized images of the class \(k\).
- \(C_k\): total number of the test images of the class \(k\).
- 1, 2,..., 5 correspond respectively to BB, AB, SB, V and W.

In our experimentation, we chose to vary image size \(n\) from 20 to 60 in order to have significant results. In fact, each image (either for training or for testing) should have enough pixels (not less than \(20 \times 20\)) to perform a significant decomposition of the spectrum in the azimuth direction. Also, we fixed the maximum image size to \(60 \times 60\) to avoid confusions between classes since over this threshold, it would be difficult to find a homogenous image describing just one class.

The figures 5, 6, 7, 8 and 9 show respectively the percentage of good classification of the BBs, ABs, SBs, V and W in function of the size \(n\) of the images.

![Figure 5. Big building classification as a function of the size of the images using the covariance formalism with and without the azimuth sub-band decomposition.](image-url)
Figure 6. Average building classification as a function of the size of the images using the covariance formalism with and without the azimuth sub-band decomposition.

Figure 7. Small building classification as a function of the size of the images using the covariance formalism with and without the azimuth sub-band decomposition.

Figure 8. Vegetation classification as a function of the size of the images using the covariance formalism with and without the azimuth sub-band decomposition.

Figure 9. Water classification as a function of the size of the images using the covariance formalism with and without the azimuth sub-band decomposition.

The following observations could be made from the figures 5, 6, 7, 8 and 9:

- The covariance with azimuth sub-band decomposition algorithm outperforms the general covariance algorithm for all the classes for almost all the image sizes. In particular, the azimuth decomposition improves advantageously the quality of the covariance-based classification of the vegetation where the amelioration reaches sometimes about 40%. It is also worth to note that the lowest level of good classification of the vegetation (PGC4 = 44%) obtained when using the azimuth decomposition, is close to the best level of good classification achieved without azimuth decomposition (only PGC4 = 48%).

- By using only the covariance formalism, the algorithm is not able to recognize the average buildings (less than 30% of good classification in most of the cases). But, with the azimuth sub-band decomposition, the recognition becomes much better (more than 40% of well-classified average buildings when the image size n > 28).

- The azimuth sub-band decomposition has no effective amelioration on water classification. In fact, the class of water is quite stable and its contents do not have special behaviors regarding the sensor positions (the azimuth decomposition does not provide more information and finer characterization of the targets in this case). Nevertheless, combing it with the covariance formalism seems to provide more flexibility to find a common image size that gives a good discrimination between the different classes.

- A good classification of both the average and the small buildings requires a large size of the images (the high PGCs are reachable more easily for the high n than for the low n). The surrounding area in this case seems to react as a relevant characteristic
to the corresponding class. Indeed, unlike the optical images, the response of the buildings in SAR images is more complex to identify/recognize since it depends highly on the orientations, and the materials of the surrounding area and the objects that may exist on the roofs (good/bad reflectors). Also, a building includes in general, different small sub-classes (vegetation, cars, roads, lights,...) which have several backscattering behaviors, and thus requires a sufficiently large number of pixels to be well described.

- In the case of vegetation, the size of the used images seems also to be determining for the classification performance. In fact, when combing the azimuth sub-band decomposition with the covariance formalism, more than $40 \times 40$ pixels are needed to provide a well-suited description of the vegetation ($PGC_4 > 70\%$). Indeed, the fact that the vegetation could sometimes be considered as a sub-class for the buildings, results in a kind of confusion between classes for the low image sizes.

It seems that our newly developed algorithm is more adapted to the classification of vegetation and water (an average of 70% of good classification is reached for many image sizes $n$). However, for the buildings (big, average, small), the average of the good classification is only about 60%. It is clear that the use of azimuth sub-band decomposition in the covariance formalism has improved advantageously the classification results but they are still not enough good and need some improvements.

To find an optimal size of the images, we propose to find the image sizes $n_{opt}$ which obey to the following constraints system:

$$\begin{align*}
PGC_k(n) &\geq 60\% \quad \text{for} \quad k = 1, 2, 3 \\
PGC_k(n) &\geq 70\% \quad \text{for} \quad k = 4, 5
\end{align*}$$

(13)

In fact, 60% corresponds to the averaged good classification of the big buildings ($k = 1$), average buildings ($k = 2$) and small buildings ($k = 3$). 70% is the averaged good classification for both vegetation ($k = 4$) and water ($k = 5$).

The optimal image size (solution for the constraints system (13)) is:

$$n_{opt} = 58$$

In the following experimentations, the classifier ($n = 58$) that yielded the best results will be analyzed in-depth. The analysis is presented in terms of classification diagram given by figure 10.

From figure 10, several observations could be made:

- The best-recognized class (vegetation, $PGC_4 = 80\%$) is well discriminated from the big buildings. In Fact, the big buildings are mainly formed by strong scatterers that appear as very bright points in the SAR images, which is not the case for vegetation. It could happen also that a big building sample image contains also vegetation in the surrounding area, but since the size of the window is relatively high ($n = 58$), the amount of vegetation could not be above a certain threshold (otherwise the sample image would belong to the vegetation class rather than the big building class).

- The vegetation is mistaken for both the average buildings and small buildings (more for small buildings). Indeed, when dealing with a sufficiently large image size ($n = 58$), these two classes include in general the vegetation among their surrounding area. The vegetation/small buildings confusion could also result from the fact that the roads and cars are common sub-classes for the two classes.

- The water is rarely confused with all kinds of buildings (only with small buildings with relatively low misclassification percentages). However, 22% of the used water sample were recognized as vegetation when the image size is equal to 58. This is probably due to a confusion between the cars that may exist in the vegetation samples and the ships that may exist in the water samples.

- The buildings (big, small, average) are mistaken between each other. This confusion is probably due to the fact that these classes include almost the same structures (houses, vegetation or gardens, roads, cars, lights,...) with different occurrences. It seems that the covariance formalism is not able to discriminate between so close classes. More specific features are needed in this case.

The figure 11 shows the first well-recognized images for each class.
6. CONCLUSIONS

In this article, a preliminary classification of high resolution SAR images has been performed on a five-class database (big buildings, average buildings, small buildings, vegetation and water).

The proposed method aims at exploiting the rich information provided by the azimuth sub-band decomposition and to combine it with the promising properties of the covariance analysis, in order to get both a superior classification performance and a better discrimination between the different classes.

To evaluate the quality of the classification, a study on the optimal image size was carried out. It was demonstrated that the azimuth decomposition provides more flexibility in the choice of the optimal size of the images. Besides, the combination between the covariance formalism and the azimuth sub-band decomposition was shown to be worthwhile mainly for big buildings, vegetation and water classification.

ACKNOWLEDGMENTS

This work was performed in the frame of the CNES/DLR/ENST center of competence on information extraction and image understanding for earth observation. The authors would like to acknowledge Gottfried Schwarz for many valuable contributions.

REFERENCES

