TARGET ASPECT ESTIMATION FROM SINGLE AND MULTI-PASS SAR IMAGES

R. Meth & R. Chellappa

Center for Automation Research and
Department of Electrical Engineering
University of Maryland at College Park
College Park, MD 20742, USA

S. Kuttikkad

Etak, Inc.
1430 O'Brien Drive
Menlo Park, CA 94025, USA

ABSTRACT

A technique is presented for estimating the aspect of targets in SAR imagery for use in indexing, feature extraction and recognition. Aspect estimation is enhanced by combining multiple images of the same target. In order to properly combine the estimation of multiple passes, it is necessary to accurately register the images to a common coordinate frame. An algorithm for registering multiple high resolution SAR images, is presented. A global affine transformation derived from the sensor acquisition parameters is used to automatically register the images, followed by a refinement to correct for translational errors. The registered SAR images are used for improving the estimates of target orientation angles, detecting the presence of occlusion and indicating poor target segmentation.

1. INTRODUCTION

Estimation of target orientation is an important problem in automatic target recognition (ATR). The knowledge of a target's pose can enhance target indexing, feature extraction and recognition by reducing the size of the database that has to be searched for a possible match in model-based ATR. SAR imagery does not lend itself to classical image processing techniques, like edge detection, for pose estimation. The highly non-literal nature of the imagery, presence of speckle, and the lack of significant gradients at the edges of objects pose unique challenges in computing the orientation of objects in the image. Thus, it is difficult to obtain an accurate pose estimate even if a target is in the open and not camouflaged. Bhanu et al. [1] use principle component analysis to compute the major axis of a pattern of scattering centers. While this yields an estimate of the orientation of the scattering centers, there is no assurance that these centers are perfectly aligned with the object's true orientation. Another approach is to detect the target region using a constant false alarm rate (CFAR) detector [5] and to find the axis of least moment of inertia for this region [2]. Although this method gives a general indication of the orientation, the presence of shadows, which in general are not symmetric with respect to the target's alignment, contributes to the poor accuracy of this method. We present an approach that is not affected by the presence of shadows on the far range side of the target and is robust to inaccuracies in CFAR detection.

As in any estimation algorithm, there are error sources that effect the accuracy of the estimates. Some errors that affect aspect estimation are occlusion (eg. when a vehicle is situated near a tree) and inaccurate target segmentation. These errors can be detected and reduced by using multiple images of the same target. In order to properly combine the estimation of multiple wide area passes that include multiple targets, it is necessary to accurately register the images to a common coordinate frame. This ensures that the orientation estimates are all with respect to the same reference, as well as ascertaining that the correct targets are combined. Algorithms for registering and exploiting multiple pass data from satellite and shuttle-based radars have been developed by the remote sensing community (eg. see references in [4], ch. 9). However, techniques for registering satellite-based radar images cannot be directly applied to airborne radar images. The resolution of airborne SARs is at least an order of magnitude better than that of satellite-based systems. The higher resolution leads to more speckle as well as a need for more accurate pixel-level (as opposed to region) registration.

In this paper, we present target orientation estimation and SAR registration algorithms with results on real multi-pass airborne SAR imagery obtained from MIT Lincoln Laboratory. The use of multiple images of the same site at varying aspects improves the estimates of target orientation angles, and indicates the presence of occlusion and poor target segmentation.

2. DETERMINATION OF TARGET ASPECT

2.1. Algorithm

The basis for the following algorithm is that the leading edges of target signatures can be approximated by piecewise linear functions. By computing the slopes of these lines and projecting them to the ground plane, an estimate of the orientation can be obtained. The major axis can also be found by determining the longer axis of the target. In

Supported in part by the Defense Advanced Research Projects Agency under contract F49620-93-1-0576 (DARPA order no. A369) monitored by the Air Force Office of Scientific Research and the DARPA/AASERT contract DAAH-04-96-1-0219 monitored by the Army Research Office

practice, SAR imagery poses some challenges to direct implementation of this algorithm. Edges cannot be reliably extracted from SAR imagery. Target articulation also nullifies the piecewise linear assumption. Spurious pixels, due to multi-bounce and speckle, affect the piecewise linear approximation. In addition, the interval over which the linear function is fit significantly affects the orientation estimate.

Target Segmentation
In order to detect candidate target pixels, an order statistic CFAR detector is applied to all pixels in the image [3]. A low false alarm rate is desirable to reduce the number of false alarms (non-target pixels) while a high false alarm rate is desirable to reduce misses (of target pixels). Therefore, a two-pass CFAR is employed. In the first pass, a low false alarm rate (10^{-6}) is used on all pixels in the image. This detects mainly target pixels with a minimum number of false alarms. In the second pass, only those pixels within a given neighborhood around pixels that were detected in the first pass are considered. A high false alarm rate (10^{-1}) is used to ensure that all target pixels (in the neighborhood of target regions) are detected. Non-target region false alarms are reduced by the first pass while a maximum number of target pixels are retained by the second pass. A filter is then passed over the image which removes all pixels that do not have at least N detected pixels in an nxn neighborhood around that pixel. This removes spurious pixels that do not belong to any target region. Isolated pixels (regions) that are close to the target region may pass through this filtering process and are eliminated based on their size.

Leading Edge Divided Into "Edge" and "Center" Regions
A target may be approximated as a rectangle in the ground plane (which is skewed when projected onto the slant plane). Depending on its orientation, one or two sides will be facing the radar sensor. The most accurate estimate will come from the longer side. Thus, it is desirable to retain as many pixels that belong to that side and reject all pixels that are not in that side. While the target pixel closest to the sensor is a good candidate for separating the two sides, resulting estimates are poor for a number of reasons. This pixel is not always unique; this pixel is not always indicative of the separation point between sides, especially near orthogonal orientations; and the ends of the target region may have large jumps due to poor segmentation, target hide, and/or incorrect rectangular approximation. Therefore, each target is initially viewed as concatenation of three regions: edge regions containing a small fraction of the target at the low and high cross range ends, and the remaining center region.

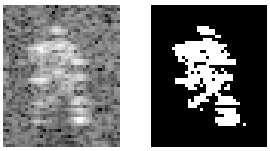
Test for Orthogonal Orientation
An approximate leading contour of the target is obtained by retaining the target pixel that is closest to the sensor along each range line. Before any attempt is made to divide this contour into line segments, the end regions are examined to determine whether a near orthogonal orientation is present. Within each edge region, the difference between the furthest range pixel on the far side and the nearest range pixel on the near side is found. This is compared to the maximum difference in the leading edge range locations for the center region. If both end region differences are greater than the center region difference, this indicates a near orthogonal orientation. The rationale for this is that orthogonal orientations have flat contour slopes in the center and steep slopes at both edges, whereas oblique orientations have steeper slopes in the center than *both* of the ends. The approximations of the end slopes use end regions and maximum differences to exaggerate the presence of a steep edge since they may not otherwise be detected for the reasons mentioned above. A linear fit is then performed on the center region. If the resulting slope indicates that the orientation is not orthogonal, a new estimation is performed using a smaller edge region. The major axis is found by comparing the target span in the range and cross range dimensions.

Extract Largest Edge for Orientation Estimate
The maximum difference in range values within each of the three regions in computed. If either of the end region range differences are greater than the center region difference, this indicates an oblique orientation. If one of the end regions has a large slope, that edge region is removed from consideration since that large jump is an artifact that would detract from the estimation. The pixel closest to the sensor is then used to divide the facing contour into two segments. (In case of multiple maxima, the lowest and highest cross range locations of these maxima are used for the segment endpoints and the middle region is ignored.) A linear fit is then performed on the segment that contains the most pixels. (If both segments contain the same number of pixels, a linear fit is performed on both segments and the line which has the lowest mean absolute error from the data points is used.) The major axis is chosen to be the axis of the larger segment.

The major axis determination used above has an ambiguity of 180°. No attempt was made to resolve which end of the target was the head or tail. In addition, at some poses, a 90° ambiguity may be present since both axes may appear to be approximately the same size. (This can happen due to shadowing from the target or other sources, as well as having a target that is approximately square.) This can many time be resolved if additional images of the target are present from different viewpoints.

2.2. Robustness

This algorithm contains robustness to some of the problems posed above. While articulation of component on a given target may effect the signature, such as a turret or gun facing the radar, the corresponding contour ascends and descends. The same is true for spurious pixels that are included in the contour in that the contour increases as it comes to that pixel and decreases as it leaves that pixel. The longer the line segment is, the less effect this spurious pixel has on the orientation estimation. This is demonstrated in Figure 1 where an orientation of 172.5° was computed for a ground truth of 176.0° when the contour significantly deviates from the piecewise linear ideal. A median line fit is sometimes used to remove the effects of outliers, but experimentation has shown least squares fits



(a) (b)

Figure 1: Example of segmented target that significantly deviates from the piecewise linear ideal: (a)SAR image, (b)segmented target pixels

to perform well.

2.3. Experimental Results

The above algorithm was applied to targets in the Lincoln Lab ADTS public release target array. A histogram of the absolute error of the aspect estimates is shown in Figure 2. The targets that contain large errors are due to the presence of occlusion (e.g. tree shadows) and the resulting inaccurate segmentation.

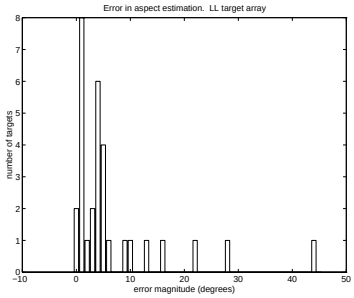


Figure 2: Histogram of the error magnitudes for orientation estimation

3. TARGET ASPECT FROM MULTI-PASS IMAGERY

Aspect estimation may be enhanced by combining estimates from multiple images of the same target. In order to properly combine the estimation of multiple passes, it is necessary to accurately register the images to a common coordinate frame. This transformation is derived below and applied to real multi-pass SAR imagery to improve upon the aspect estimation of the previous section.

3.1. Registration

Since scatterers at a constant range from the radar are mapped into the same point, SAR image formation is a many-to-one projection of 3-D space onto a 2-D plane. While it is possible to derive the complete transformation of any 3-D point into its corresponding 2-D point in a SAR image it is not possible to register two arbitrary SAR images, without making some simplifying approximations. The first approximation is the "flat earth" assumption, i.e. the imaged terrain is assumed to be flat. Some parts of the image may have significant slope, resulting in foreshortening and possible layover (buildings, trees, mountainsides), but these elements are intractable from a single image. By making the flat earth assumption, it is possible to register most of the image; regions of misregistration will then correspond to structures with significant height and to regions that are occluded by other object at some aspects (eg. targets that are shadowed by trees). The second approximation is that the depression angle θ is constant across the image swath. This enables a single transformation to be used for all the points in the image irrespective of their range locations. Once again, assuming a large range-to-swath-width ratio justifies this approximation. Third, the arc of the circle of constant range that extends from a point in the slant range dimension to its ground range location is approximated to be linear, since the subtended angle at the sensor is small.

Two SAR acquisition geometries may differ in resolution, depression angle, sensor heading, and reference ground location. In the image domain, these correspond to differences in scaling, projection to ground, rotation, and translation, respectively. A point in one SAR image can be transformed to its corresponding point in another SAR image through the following steps: a rescaling to normalize for the range and cross range resolutions of the first image; a projection to the ground plane using a linear approximation (equivalent to dividing the range coordinate by the cosine of the depression angle); a rotation and translation in the ground plane; a projection to the slant plane of the second image; and a rescaling to normalize for the resolution of the second image.

Thus, a 2-D point $p^{(1)}(=[x^{(1)}\ r^{(1)}]^T)$ in the first image can be transformed into the corresponding point $p^{(2)}(=[x^{(2)}\ r^{(2)}]^T)$, in the second image via the affine transformation

$$p^{(2)} = Ap^{(1)} + b \tag{1}$$

where

$$A = \begin{bmatrix} \frac{1}{\delta x^{(2)}} & 0 \\ 0 & \frac{1}{\delta r^{(2)}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \cos\theta^{(2)} \end{bmatrix} \tag{2}$$
$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\cos\theta^{(1)}} \end{bmatrix} \begin{bmatrix} \delta x^{(1)} & 0 \\ 0 & \delta r^{(1)} \end{bmatrix}$$

and $b = [t_x\ t_r]^T$ is the translation (cross-range and range) vector required to align the images. Here, $\delta x^{(i)}, \delta r^{(i)}$ are the respective pixel resolutions in the cross-range and range dimensions, $\theta^{(i)}$ are the depression angles, and $\phi = \phi^{(2)} - \phi^{(1)}$ is the difference in sensor headings between the two images. For convenience, the affine transformation of (1) can be written in the general matrix formulation

$$\begin{bmatrix} x^{(2)} \\ r^{(2)} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_r \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ r^{(1)} \\ 1 \end{bmatrix} \tag{3}$$

All the parameters in (2) are typically available, to some degree of accuracy, for airborne SAR data. The only unknowns in (3) are t_x and t_r. In order to compute the translation, it is necessary to know the actual ground location of at least one point in each image. Sometimes, Global Positioning System (GPS) information is available as a reference for the scene location. In case this information is

unavailable, a point correspondence may have to be found manually or by an automatic scheme that matches point features across images.

The accuracies of the resolution and depression angle parameters are usually high, because they are part of the SAR system design. Aircraft heading information is also expected to be reasonably accurate, but GPS location may result in significant errors (the Precise Positioning System mode of GPS has an accuracy on the order of 20 meters). We have found this to be the case with the Lincoln Lab ADTS dataset. High-frequency errors due to aircraft motion are expected to be local within the image, with adequate compensation for gross errors during image formation. It may be argued that the errors due to inaccuracies in the angular parameters have the least effect at the center of rotation (increasing with distance), whereas location errors affect all pixels in the image equally.

3.2. Automatic Refinement of Registration

In order to compensate for errors in GPS-derived location, we extract a number of point features from each image and refine the translation parameters from them. The features chosen should lie in (or near) the ground plane, so that there are no layover effects that would affect different views differently. They should also be easy to detect and should persist across images. We have chosen the centroids of clusters of bright pixels as our point features. These bright returns result from metallic objects and other specular reflectors in the scene which may be embedded in non-homogeneous background clutter. In the images we experimented with, they consist of stationary vehicles and other strong reflectors, like dihedrals and trihedrals. The Order Statistic Constant False Alarm Rate (OS CFAR) [5] technique is used to detect bright pixels in spatially varying clutter. Terrain backscatter is modeled as complex Gaussian, resulting in a Rayleigh magnitude distribution. After initial registration, distances between each feature point in one image and all feature points in the other image are computed. A search is then performed to find the maximum number of one-to-one matches that result in the same approximate translation.

3.3. Experimental Results

The above registration transformation was applied to images from the Lincoln Lab ADTS public release target array data. The aspect estimation algorithm was then used on the registered target chips. The ADTS wide area images were cropped which cause a couple of the targets to be clipped at the edge. The resulting aspects are shown in Table 1. Some targets had a 90° ambiguity in which the major axis could not be confidently identified, which was resolved using information from another pass. The standard deviation of the estimated values indicates the confidence in the accuracy of the average estimate. The moderately large standard deviations for targets 2 and 11 indicate the low confidence in those estimates. If the two closest estimates are retained, a more accurate orientation is obtained. The very large standard deviation for targets 5, 8 and 9 indicate target occlusion in some of the images, and therefore, unreliable estimates. Visual inspection of the images corroborated these assertions (5 was covered by trees in some aspects, and 8 was cut off at the end of the image in one of the passes). A closer inspection of target 9 revealed that the orientation estimation was good, but that the major axis was identified incorrectly since the target was clipped in one image.

Target/	Orientation				Std
Truth	pass 1	pass 2	pass 3	Avg	Dev
1/83.0	86.9	79.3*a	81.6	82.6	3.9
2/318.5	319.7	334.1x	314.0	322.6	10.4
3/356.0	352.0*	355.6	352.5	353.4	2.0
4/322.0	323.4	323.5	325.2*	324.0	1.0
5/67.0	80.5x	57.8*x	95.3x	77.9	18.9b
6/275.0	272.9	269.6*	275.3	272.6	2.9
7/85.0	80.1	79.4	75.1	78.2	2.7
8/136.0	131.0	92.2x	N/A	111.6	27.4b
9/89.0	358.4x	N/A	92.6	45.5	66.6b
10/319.0	320.1	323.7	318.3*	320.7	2.7
11/138.0	139.4	132.4	115.6	129.1	12.2

(*) Contains 90 degree ambiguity
(x) Target occluded, cut off, or poorly segmented
(N/A) Target not detected in that image
(a) 90 degree ambiguity removed
(b) Indication of target occlusion

Table 1: Target orientation estimates

4. CONCLUSION

A technique for estimating target orientation angles in SAR imagery has been presented. We have also presented an algorithm for registering high resolution airborne SAR images through an affine transformation. An algorithm for automatically refining inaccuracies in translation was described. The registration algorithm was applied to the Lincoln Lab ADTS dataset. Estimates of target orientation angles were obtained and refined using multi-pass data. The use of multi-pass registered images enhanced target aspect estimation and provided indications of inaccurate estimates due to target occlusion.

5. REFERENCES

[1] B. Bahnu, G. Jones, J. Ahn, M. Li, and J. Yi, "Recognition of Articulated Objects in SAR Images," *Proceedings ARPA Image Understanding Workshop*, pp. 1237–1250, February 1996.

[2] B. K. P. Horn, *Robot Vision*, The MIT Press nad McGraw-Hill Book Company, 1986.

[3] S. Kuttikkad and R. Chellappa, "Non-Gaussian CFAR Techniques for Target Detection in High Resolution SAR Images," *IEEE International Conference on Image Processing*, pp. 910–914, Nov 1995.

[4] F. W. Leberl, *Radargrammetric Image Processing*, Norwood, MA: Artech House, Inc., 1990.

[5] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-19, pp. 608–621, July 1983.