INTELLIGENT COMPRESSION OF SAR IMAGERY USING MULTIRESOLUTION MARKOV MODELS

Robert J. Bonneau[+] *and Glen P. Abousleman*[++]

[+]Radar Signal Processing Branch, Sensors Directorate, AFRL
Rome, New York, USA
Bonneaur@rl.af.mil

[++]Motorola, Integrated Information Systems Group
Scottsdale, Arizona, USA
p26994@email.mot.com

ABSTRACT

We present a system for coding synthetic aperture radar (SAR) imagery, whereby regions of interest and background information are coded independently of each other. A multiresolution constant-false-alarm-rate (CFAR) detection scheme is utilized to discriminate between target regions and natural clutter. Based upon the detected target regions, we apply less compression to targets, and more compression to background data. This methodology preserves relevant features of targets for further analysis, and preserves the background only to the extent of providing contextual information. The resulting system dramatically reduces the bandwidth/storage requirements of the digital SAR imagery, while preserving the target-specific utility of the imagery.

1. INTRODUCTION

Representation of digital synthetic aperture radar (SAR) imagery requires a vast amount of raw data. Increased sensor resolution and image delivery rates only increases the burden for storage and transmission of the image data. For example, 2048 X 2048 32-bit SAR imagery requires approximately 134 Mbits of digital information. Clearly, the data associated with modern high-resolution SAR scanners is far beyond the capability of current and proposed wireless data links.

To address the problem of increased data rates associated with all forms of digital imagery, mathematical lossy compression algorithms can be utilized to reduce the overall number of bits required to represent the digital imagery, while adhering to desired subjective and/or quantitative image fidelity criteria. In other words, the overall number of bits is reduced at the expense of discarded image information. Examples of lossy compression algorithms include JPEG/JPEG-2000, MPEG variants, H.263, etc.

Lossy compression algorithms typically introduce distortion is the coded image uniformly. That is, as the specified bit rate is reduced, the quality throughout the image degrades by the same amount. These visual artifacts usually manifest themselves as blurriness, fuzziness, blockiness, etc., depending upon the compression algorithm that is utilized. Lossy coders can also be designed to introduce distortion in a non-uniform fashion. For example, regions of the image that have the greatest energy can be coded with more bits than those regions with less energy. Alternatively, regions with more edges can be assigned more bits, while regions with few edges can be assigned fewer bits. This type of region classification can be based upon a variety of metrics such as those mentioned above, or others such as fractal dimension, average gray level, statistical variance, etc.

Although the afore-mentioned region-adaptive compression methods can yield significantly improved compression performance for a wide range of imagery, they lack sufficient intelligence to discern those regions where fidelity truly needs to be maintained, from regions where fidelity is less important. In the present work, we have developed an intelligent compression system whereby regions of interest (ROI) and background information are coded independently of each other. We apply less compression (more bits) to regions of interest (targets), and more compression (fewer bits) to background data. This methodology preserves relevant features of targets for further analysis, and preserves the background only to the extent of providing contextual information. The resulting system dramatically reduces the bandwidth/storage requirements of the digital imagery, while preserving the target-specific utility of the imagery.

In the current study, we will investigate the use of multiscale Markov models to distinguish between man-made objects and natural clutter in synthetic aperture radar imagery, when used in conjunction with a state-of-the-art image-coding scheme. Specifically, these models are used to define a multiresolution discriminant as the likelihood ratio for distinguishing between the two pixel types.

This model-based approach can be used in two general ways to detect targets of interest. First, the Markov detection scheme can be used to separate wavelet coefficients corresponding to targets, from those corresponding to clutter. The target coefficients can then be quantized with high resolution, and the background coefficients can be quantized with low resolution. This approach can greatly increase the overall SAR compression efficiency if the aggregate target area is much smaller than the remaining clutter area.

The second scheme involves detection following quantization. Here, the autoregressive coefficients, and hence the likelihood ratio will be altered, depending upon the rate and type of coding. As we will show, multiscale discrimination still significantly im-

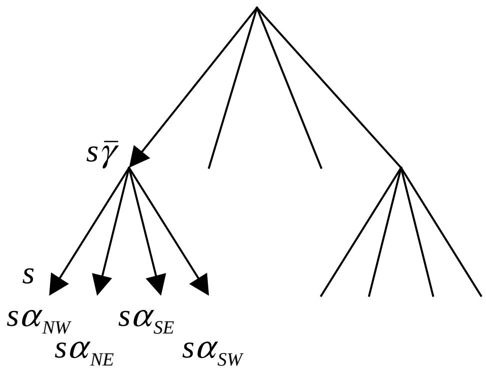

Figure 1: Markov tree decomposition.

proves the detection performance as compared to single-scale detection schemes.

This paper is organized as follows. Section 2 provides a description of the multiresolution detection algorithm. Section 3 outlines the image coding system that uses pre-detection to preserve relevant target areas. Detection results are presented in Section 4, and a conclusion is given in Section 5.

2. MULTIRESOLUTION TARGET DETECTION

Recent studies have shown the optimality of the wavelet transform domain for target detection [1]. Accordingly, we must first represent our signal and noise processes with appropriate data models. This type of model is known as the wavelet Markov random field (MRF) model.

Referring to the quad tree shown in Figure 1, we can represent a Markov random field as follows [1, 2]. We define a given node in the quad tree structure as s, its children nodes as $s\alpha_{NW}$, $s\alpha_{NW}$, $s\alpha_{NW}$, and $s\alpha_{NW}$, and its parent node as $s\gamma$, where $\bar{\gamma}$ shifts the wavelet coefficients from parent, $s\gamma$, to child, s. A K^{th}-order model defined on the multiresolution structure is defined in either one or two dimensions with $t \in \{1, 2, \ldots, K(T+1)\}$.

Now, defining a MRF on a 2^N x 2^N lattice, a state at the m^{th} level represents the values of the MRF at $16(2^{N-m} - 1)$ points. This set of points is denoted as Γ_s, and is the union of four mutually exclusive subsets. In general, we can divide Γ_s into four sets of $4(2^{N-m(s)} - 1)$ points each, in a similar fashion. We denote these subsets as $\Gamma_{s,i}$, $i \in$ {NW, NE, SE, and SW}.

The multiresolution detection scheme presented here is similar in spirit to the development in [3]. The model identification consists of three general steps. First, our multiscale stochastic models are restricted to be autoregressive in scale. We then choose the appropriate regression coefficients based on a simple optimization criterion. Finally, we characterize the statistical distribution of the model driving noise.

Once we define the Markov structure from the wavelet transform, we take the individual coefficient elements and represent them using an autoregressive set of equations. A target is represented by the polynomial coefficients, A(s), added to a Gaussian noise component, w(s), represented by the B(s) coefficients:

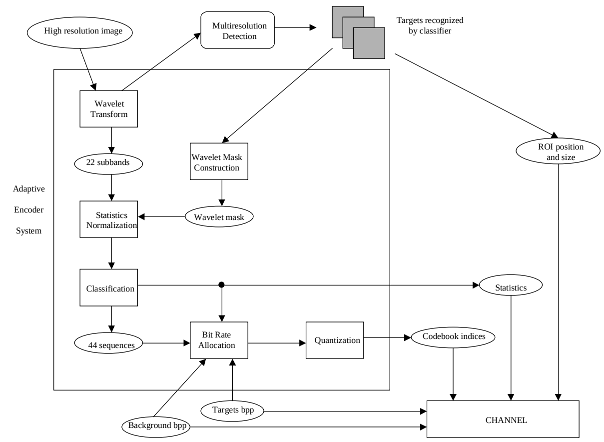

Figure 2: Intelligent compression system block diagram.

$$x(s) = A(s)x(s\bar{\gamma}) + B(s)w(s) \,. \tag{1}$$

In the image context, we can represent the elements of the image Markov structure in terms of the recursive scale structure shown in Figure 1. Given a multiresolution image sequence, $I_0, I_1, \ldots I_L$, we characterize the joint statistical distribution of pixel values in the sequence. The SAR pixel value residing at node *s* (i.e., *I(s)*) is related to its ancestors by a linear autoregression in scale, which is given by:

$$I(s) = a_{1,m(s)} I(s\bar{\gamma}) + \cdots + a_{R,m(s)} I(s\bar{\gamma}^R) + w(s) \,, \tag{2}$$

where *R* is the order of the regression, $a_{1,m(s)}, a_{2,m(s)}, \ldots, a_{R,m(s)}$ are the scalar regression coefficients, and w(s) is the residual error in the prediction of *I(s)*. We can characterize a given signal or texture in the image by first solving for the autoregressive coefficients, and then using these coefficients to predict the target signal in a given input signal, x(s). The residual between the target signal and the input signal is then w(s). This model assumes that the target signal is uncorrelated with the input signal. If $\mathbf{a}_k = \begin{bmatrix} a_{1,k} & a_{2,k} & \cdots & a_{R,k} \end{bmatrix}^T$, the regression coefficients can be found via the following relation:

$$\mathbf{a}_k = \arg \min_{\mathbf{a}_k \in \mathbf{R}^R} \left\{ \sum_{\{s; m(s)=k\}} [I(s) - a_{1,k} I(s\bar{\gamma}) - \ldots - a_{R,k} I(s\bar{\gamma}^R)]^2 \right\} . \tag{3}$$

The Markovian structure of the models leads to simple, explicit likelihood expressions. Let H_0 and H_1 denote the hypothesis that the ROI represents natural clutter and a man-made object, respectively. Let M_0 and M_1 denote the coarsest and finest scales, respectively, for which we have observations (i.e., $M_0 = M - L + 2$, and $M_1 = M$). We define Y to be a vector containing all the observations, y(s). Additionally, we let $a_{k,H0}$ and $a_{k,H1}$ denote the k^{th}-scale regression coefficients for the natural-clutter model and the man-made model, respectively. We let $w_{Hi}(s)$ denote the residual in the autoregressive prediction of the pixel value, I(s), using the model underlying H_i. $w_{Hi}(s)$ is given by

$$w_{Hi}(s) = y(s) - a^T_{m(s),Hi} x(s\bar{\gamma}) \tag{4}$$

The multiresolution discriminant can be expressed as

$$\log(P_{Y|H_1}(Y \mid H_1)) - \log(P_{Y|H_0}(Y \mid H_0)) \ , \tag{5}$$

where

$$\log[P_{Y|H_i}(Y \mid H_i)] = \sum_{k=M_0}^{M_1} \sum_{\{s;m(s)=k\}} \log[P_{w(s)|H_i}(w_{k,H_i}(s) \mid H_i)] \tag{6}$$

and i = 0, 1. Additional details related to the detection scheme can be found in [4].

3. SYSTEM DESCRIPTION

The overall SAR compression system configuration is illustrated in Figure 2. The SAR image is first decomposed into 22 subbands using a 2-D discrete wavelet transform (DWT). That is, a standard 4-level dyadic decomposition is performed, with one additional level of decomposition being performed on the highest-frequency components following the first decomposition level. Regions of interest are determined by the multiresolution detection scheme, which classifies image pixels as target or clutter. For each subband, the DWT coefficients corresponding to the same class (either target or clutter) are grouped into sequences (subband-class sequences) to be encoded using entropy-constrained trellis-coded quantization (ECTCQ). The wavelet mask construction consists of mapping the ROI from the original image to each subband. Figure 3 shows a sample ROI wavelet mask. Referring to Figure 3, the white areas correspond to targets to be coded at high resolution, while the dark portion corresponds to clutter, which is coded at low resolution. All obtained subband-class sequences are normalized by subtracting their mean, and dividing by their respective standard deviation. Since the average number of bits per pixel (bpp) for the targets and the background is independent, and can be specified a priori, the statistics of each subband-class sequence is used separately by a rate allocation procedure.

The probability distribution of each sequence to be encoded is modeled by the so-called *Generalized Gaussian Distribution* (GGD), whose probability density function is given by

$$f_X(x) = \left[\frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right] \exp\{-[\eta(\alpha,\sigma) \mid x \mid]^{\alpha}\} \ , \tag{7}$$

where

$$\eta(\alpha,\sigma) \equiv \sigma^{-1}\left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{1/2} . \tag{8}$$

The shape parameter, α, describes the exponential rate of decay, σ is the standard deviation of the associated random variable, and Γ(·) is the gamma function. Distributions corresponding to α = 1.0 and α = 2.0 are Laplacian and Gaussian, respectively. It can be shown that there is one-to-one mapping between α and the Kurtosis of a given sequence. Thus, the Kurtosis can be used to determine the appropriate α of a particular sequence [5].

ECTCQ codebooks were designed for generalized Gaussian distributions with α values of 0.5, 0.75, 1.0, 1.5, and 2.0 [5]. Training sequences consisted of 100,000 samples derived from

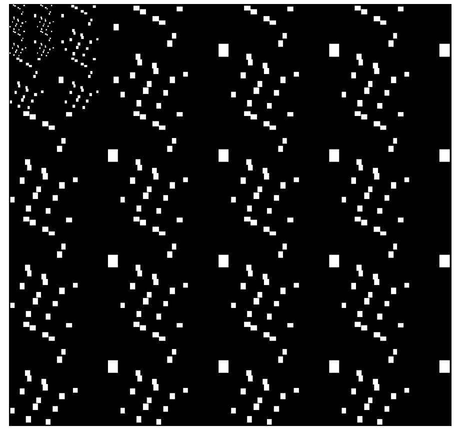

Figure 3: Example ROI wavelet mask.

generalized Gaussian pseudo-random number generators, each tuned to the appropriate α value.

The obtained wavelet mask is used to classify the wavelet coefficients in each subband into either a target class or a background class. Within each subband, coefficients corresponding to the same class are grouped into a single one-dimensional sequence. Since we assume a 22-subband decomposition, we would then obtain 44 different sequences (two per subband). The overall bit rates in bits per pixel for the targets and for the background are either provided by the user, or are determined by some automated means. Any rate allocation procedure can be used to allocate rate among the wavelet coefficients for both targets and background. In the present work, the following rate allocation algorithm is used [6], where the coefficient sequences are coded at an average rate of R_1 bpp for the background, and R_2 bpp for the targets. The overall MSE is represented by:

$$E_s = \sum_{i=1}^{K} \alpha_i \sigma_i^2 E_i(r_i) \ , \tag{9}$$

where σ_i^2 is the variance of sequence i, $E_i(r_i)$ denotes the rate-distortion performance of the quantizer at r_i bits/sample, K is the number of data sequences, and α_i is a weighting coefficient to account for the variability in sequence length. For a 22-band decomposition, $K = 22$.

The rate allocation vector, $B = (r_1, r_2, \ldots, r_K)$, is chosen such that E_s is minimized, subject to an average rate constraint:

$$\sum_{i=1}^{K} \alpha_i r_i \le R_1 (or\, R_2) \ \text{bits/coefficient.} \tag{10}$$

The solution, $B^*(r_1^*, r_2^*, \ldots, r_K^*)$, to the unconstrained problem,

$$\min_{B} \left\{ \sum_{i=1}^{K} (\alpha_i \sigma_i^2 E_i(r_i) + \lambda \alpha_i r_i) \right\} , \tag{11}$$

minimizes E_s, subject to $\sum_{i=1}^{K} \alpha_i r_i \le \sum_{i=1}^{K} \alpha_i r_i^*$. Thus, to find a solution to the constrained problem of equations (9) and (10), it suffices to find λ such that the solution to equation (11) yields

$\sum_{i=1}^{K}\alpha_i r_i^* \le R_1 (or\, R_2)$. We can perform a bisection search for the λ corresponding to the required B.

For a given λ, the solution to the unconstrained problem is obtained by minimizing each term of the sum in (11) separately. If S_j is the set of allowable rates for the j^{th} quantizer, and r_i^* is the i^{th} component of the solution vector, B^*, then r_i^* solves

$$\min_{r_i \in S_j} \left\{\alpha_i \sigma_i^2 E_i(r_i) + \lambda \alpha_i r_i\right\}. \quad (12)$$

4. RESULTS

The detection performance of the proposed system was evaluated by processing the synthesized 256 X 256 8-bit SAR image shown in Figure 4. Figure 5 shows the constructed target mask by using the multiresolution detection scheme. As indicated previously, light areas corresponding to target regions are coded at high resolution, and dark areas corresponding to natural clutter are coded at low resolution.

The detection results of the proposed system are shown in Figure 6. From the figure, we see that when the proposed multiresolution detection scheme is applied following decompression, the detection performance begins to decline at approximately 0.25 bits/pixel, and falls to zero at approximately 0.03 bits/pixel. In sharp contrast, the detection performance remains at 100% at the illustrated compression rates when the multiresolution scheme is applied prior to compression. In the latter case, since all detected targets are well preserved due to the high-resolution coding, additional automatic or human analysis is facilitated, following decoding of the compressed bit stream.

5. CONCLUSIONS

We have presented an intelligent system for compression of SAR imagery. The proposed system uses a multiresolution detection scheme to identify target regions within the image. The determined target regions are then coded at high resolution, while the remaining background region is coded at low resolution. Since the target regions are well preserved, further target analysis is facilitated following image decompression. As our detection performance results illustrate, extreme compression can be realized, while preserving the target-specific utility of the SAR imagery.

6. REFERENCES

[1] C. H. Fosgate, H. Krim, W. W. Irving, W. Karl, and A. S. Willsky, "Multiscale segmentation and anomaly enhancement of SAR imagery," *IEEE Trans. Image Proc.*, vol. 6, no. 1, Jan. 1997.

[2] M. Luettgen, W. Karl, A. S. Willsky, and R. Tenny, "Multiscale representations of Markov random fields," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, pp. 3377-3396, Dec. 1993.

[3] W. Irving, A. S. Willsky, and L. M. Novak, "A multiresolution approach to discriminating targets from clutter in SAR imagery," in *Proc. of the SPIE*, vol. 2487, pp. 272-299, 1995.

[4] R. J. Bonneau, "A constant probability of detection model for image quantization," in *Proc. of IEEE AIRP*, Oct. 2000.

[5] G. P. Abousleman, "Adaptive coding of hyperspectral imagery," in *Proc. of IEEE ICASSP*, Mar. 1999.

[6] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. 36, pp. 1445-1453. Sept., 1988.

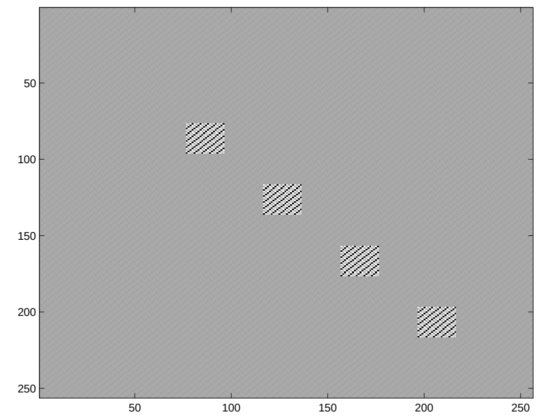

Figure 4: 8-bit synthesized SAR image.

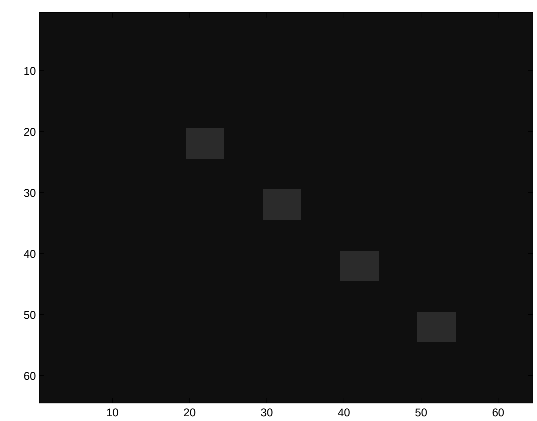

Figure 5: Target mask.

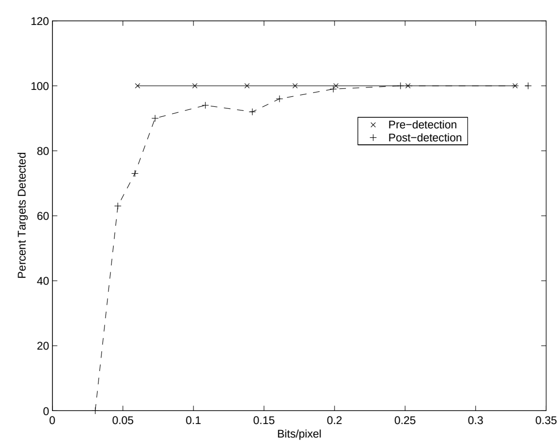

Figure 6: Detection performance of proposed intelligent compression system.