Wavelet and Fractal Analysis of Ground Vehicle Images

David Gorsich
U.S. Army National Automotive Center

Charles Tolle
Utah State University

Robert Karlsen and Grant Gerhart
U.S. Army Tank-Automotive Research, Development and Engineering Center

ABSTRACT

A large number of terrain images were taken at Aberdeen Proving Grounds, some containing ground vehicles. Is it possible to screen the images for possible targets in a short amount of time using the fractal dimension to detect texture variations? The fractal dimension is determined using the wavelet transform for these visual images. The vehicles are positioned within the grass and in different locations. Since it has been established that natural terrain exhibits a statistical 1/f self-similarity property and the psychophysical perception of roughness can be quantified by the same self-similarity, fractal dimension estimates should vary only at texture boundaries and breaks in the tree and grass patterns. Breaks in the patterns are found using contour plots of the dimension estimates and are considered as perceptual texture variations. Variation in the dimension estimate is considered more important than the accuracy of the actual dimension number. Accurate variation estimates are found even with low resolution images.

Keywords: Wavelet, Fractal Dimension, Image Segmentation

2. FRACTAL DIMENSION

There are two classes of self-similar fractals. One class of fractals is deterministically self-similar, meaning they are copies of themselves on different scales. Examples are the Cantor set and the Sierpinski triangle. The second class of fractals contains ones that are similar in a statistical sense.

Mandelbrot defined a fractal as a set whose Hausdorff-Besicovitch dimension strictly exceeds the topological dimension. This definition of dimension has no application to finite sets. There have been many other definitions applied to the dimension to incorporate the idea of self-similar sets but their relationships are not always clear. For example, the Minkowski dimension is an upper bound for the Hausdorff dimension and the Box dimension is an upper bound for the Minkowski dimension. But when finite sampling occurs, these inequalities may no longer hold.[2] To avoid these difficulties, we will turn to a statistical definition of a fractal defined below.[1]

We define a statistically self-similar fractal function, y(t) as:

$$y(t) \overset{d}{=} a^{-H} y(at) \text{ with a>0.} \tag{2.1}$$

The equality is defined in terms of second order finite-dimensional statistics. The function y(t) is a zero-mean Gaussian. It has been shown that these fractals are successful at modeling a number of natural phenomenon like texture variation in natural terrain.[1] These fractal textures and signals can be classified and segmented using both fractal dimension and lacunarity.[4,5] Functions of this form are considered 1/f

0-8194-2213-4/96/$6.00

processes if their power spectrum is proportional to $|w|^{-2H-1}$. When the Hurst coefficient, H, is 0.5 this function is ordinary Brownian motion. H is assumed to range from 0 to 1.
The fractal dimension, D, is given by H:

D=2-H for signals
D=3-H for images

We assume the discrete functions that follow are of the form (2.1) and have a dimension given by the Hurst coefficient.

Visual luminance images are not as easily represented by fractal functions as Synthetic Aperture Radar (SAR) imagery. It is easier to apply the fractal dimension parameter to SAR imagery, which can be made range independent, to segment the image. Lincoln Laboratory has shown the fractal dimension to be one of five optimal cue features in SAR imagery.[13] For visual images, the dimension values of the terrain will change based on viewing distance. The fractal dimension of these images also changes based on sky luminance, wind, and a number of other atmospheric parameters all of which have to be taken into account. This paper assumes all images were taken at about the same range, at the same viewing orientation, and with identical lighting and atmospheric conditions and instead focuses on determining variances in D using the wavelet transform. Provided a calibration image is provided, there does not appear to be any problems making these assumptions.

There are many methods to determine D, and many definitions of the dimension itself. We have reviewed the box dimension (Kolmogov), the Minkowski Cover, Hurst and Fourier Analysis, and the wavelet decomposition. We developed a fuzzy c-means and genetic algorithm approach to improve the box cover. The problems with these approaches are computation time, resolution and quantization effects. The genetic algorithm was the worst in terms of computation time, taking about a day to determine the dimension of a test 256x256 fractional Brownian texture. There are a number of papers comparing D estimation methods and their merits, although they all suffer terribly from resolution issues and computation time.[2,4,6,7,8] The wavelet method is appealing because of its speed and the additional information it can provide. Moreover, Wornell suggests that for functions of the form (2.1), the wavelet-based analysis is statistically optimal.[1]

3. WAVELET DETERMINATION OF D

The fractal dimension of signals and images are easily determined using the wavelet transform. This method of determining fractal dimension is fast and straightforward. Speed is very important since we are interested in sorting through large numbers of images. Various test signals and images were generated and used to test the ability of the wavelet transform to recapture fractal dimension variation (or the dimension itself, provided the resolution is reasonable). The synthesized signals and textures vary in fractal dimension and resolution. There were many calculations performed on these signals and only a few are displayed in the paper. Instead, the wavelet transform is applied to images of ground vehicles in natural terrain primarily composed of grass and trees.

The wavelet transform of a 1-dimensional (topological) signal is determined via[11]:

$$a_n^m = \sum_l h_0(l-2n)a_l^{m+1}$$
$$d_n^m = \sum_l h_1(l-2n)a_l^{m+1} \qquad (3.1)$$

where the h_0 are the low-pass filter coefficients, the h_1 are the high-pass filter coefficients and m is the resolution level which goes from 1 to M and corresponds to a signal length of 1 to 2^{m-1}.

For images the transform is defined by:

$$a_{i,j}^{m} = \sum_{l,k} h_0(l-2i)h_0(k-2j)a_{l,k}^{m+1}$$
$$dh_{i,j}^{m} = \sum_{l,k} h_1(l-2i)h_0(k-2j)a_{l,k}^{m+1}$$
$$dv_{i,j}^{m} = \sum_{l,k} h_0(l-2i)h_1(k-2j)a_{l,k}^{m+1}$$
$$dd_{i,j}^{m} = \sum_{l,k} h_1(l-2i)h_1(k-2j)a_{l,k}^{m+1} \qquad (3.2)$$

These are the average, horizontal detail, vertical detail, and diagonal detail coefficients respectively. We do not derive the filter coefficients. Instead we choose an existing wavelet which is orthogonal and has minimal support. Orthogonality allows us to use Wornell's decorrelation theorem.[1] Minimal support avoids boundary effects and experimentally was found to be more effective in the determination of D. Flandrin has shown that the number of vanishing moments, R, of the wavelet must satisfy $R \geq H$, where H is the Hurst coefficient . A larger number of vanishing moments provides a lower correlation between the wavelet coefficients.[3] We tested a number of wavelets including Coiflets, Symlets, and Biorthogonal wavelets with various supports, but based on the size of the support and initial results, we chose to use Daubechies' 4-tap filter coefficients.

The value of the dimension of a discrete signal is determined from the energy of the detail coefficients. For the fractals defined in (2.1), the majority of the information is contained in the detail coefficients instead of approximation coefficients.

In 1989, Mallat showed that given a random process as defined in (2.1), the fractal dimension can be calculated using the wavelet transform.[8] Moreover it is believed that the wavelet transform is the optimal method of analysis for processes of the form (2.1) because the wavelet transform breaks down the signal on a logarithmic scale and decorrelates coeffficients.[1] Mallat showed:

$$\sigma_m^2 = 2^{2H}\sigma_{m+1}^2,$$

where σ_m^2 is the energy of the detail coefficients. More recently, Wornell has shown that the fractal dimension can be derived using the variance of the detail coefficients instead of the energy of the detail coefficients.[1]

$$\text{Variance}(d_n^m) = \sigma^2 2^{-m(2H+1)} + \sigma_w^2,$$

where σ^2 is considered constant and σ_w^2 is the spectral density of white noise in the function. Broadband noise certainly alters the estimate of the fractal dimension, but we assume our images are noise free and set σ_w^2=0. Doing this allows us to avoid the Maximum Likelihood Approach outlined by Wornell.[1] Since the functions we will analyze have zero mean, we find the difference between Wornell's dimension and Mallat's dimension definition to be:

$$.5\left(-1+\frac{\log(2N-1/N-1)}{\log 2}\right),$$

which goes to zero for large N. N is the number of samples of discrete y(n). Since Mallat's approach is slightly faster, we chose to just determine the energy of the detail coefficients. This approach is tested on a number of artificially generated fractal signals and textures with zero mean and dimension D. The fractal functions that follow were generated using the midpoint method discussed by Russ and use the parameter H and a Gaussian random number generator.[2,7] There are a number of other methods to generate fractal functions, and we tested many of the ones given by Russ and found similar results.[2]

4. APPLICATIONS

We used the approach by Mallat to determine the fractal dimension of test functions with different dimensions and resolutions via the energy ratios of the wavelet coefficients. The new MATLAB Wavelet Toolbox scripts[10] were used to perform the decompositions defined by (3.1) and (3.2). The algorithm gave an accurate measure of D provided the signal had sufficient length. Below is an example of the typical trend of the fractal dimension estimate found by generating a signal with dimension 1.9 using the midpoint method with 4097 points. The wavelet transform using the Daubechies 4-tap filter coefficients was taken of a successively contracting window of the original signal. A new dimension estimate was computed after removing the last 32 points of the prior signal. Since the signal is statistically similar everywhere, we do not expect the dimension estimate to change when parts of the signal are cut off. Figure 1 is a typical plot of the dimension estimates, where the y-axis is the fractal dimension estimates and the x-axis is the amount of the signal that was removed.

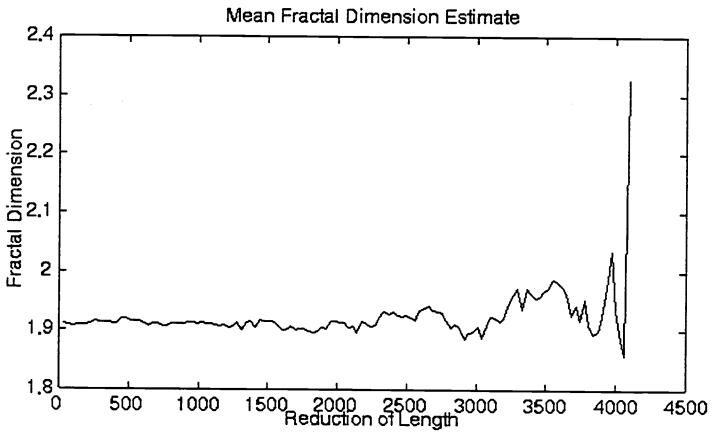

Figure 1: Fractal dimension estimate of D=1.9 fractal signal

Figure 1 was generated using the mean of the wavelet energy ratios of the first four wavelet levels. It was found that the above behavior was typical for all midpoint generated signals. More interesting were the dimension estimates of each ratio as the window was contracted. They varied as shown in Figure 2 below.

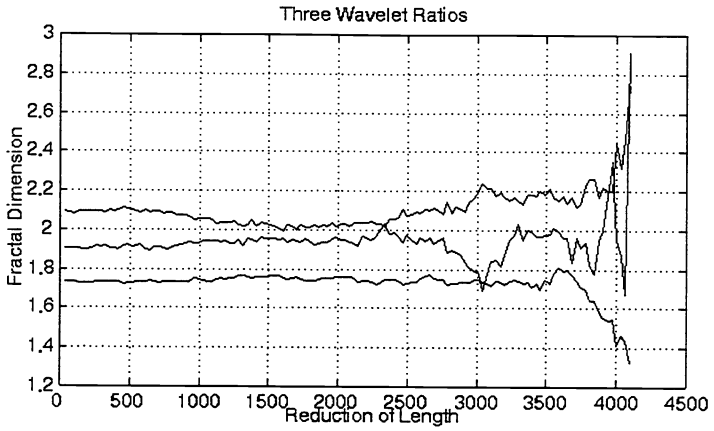

Figure 2: Three wavelet energy ratios estimating fractal dimension

The first ratio typically gave a higher estimate of the dimension, the second ratio a lower estimate, and the third ratio was usually the closest to the modeled dimension. The third ratio estimate is below the first and above the third in Figure 2. As can be seen, resolution of images less than 512x480 will be a problem. So we do not expect to have numerically accurate dimension estimates for all the ratios since our images are only 512x480, and we subdivide these images even smaller.

Since we did not expect the images of the trees and grass to be homogeneous, we were not as concerned about the actual dimension number as we were about the variance of the dimension estimate over parts of the image. A mixed fractal was generated using H=0.1 for the first part of the signal at H=0.6 for the rest of the signal. The entire signal is shown on the left in Figure 3, consisting of 4097 points, and the dimension estimates are shown on the right. The 128 dimension estimates were determined at three ratio levels with a contracting window as above. Once the window reached half the size of the signal, the dimension estimates jumped up, as hoped. With mixed fractals, the three ratios departed from each other more noticeably. Notice the entire signal has zero mean, but the two halves do not, causing a problem in the estimator. In Figure 3, the first ratio is the highest estimate on the left, and the third ratio is the lowest.

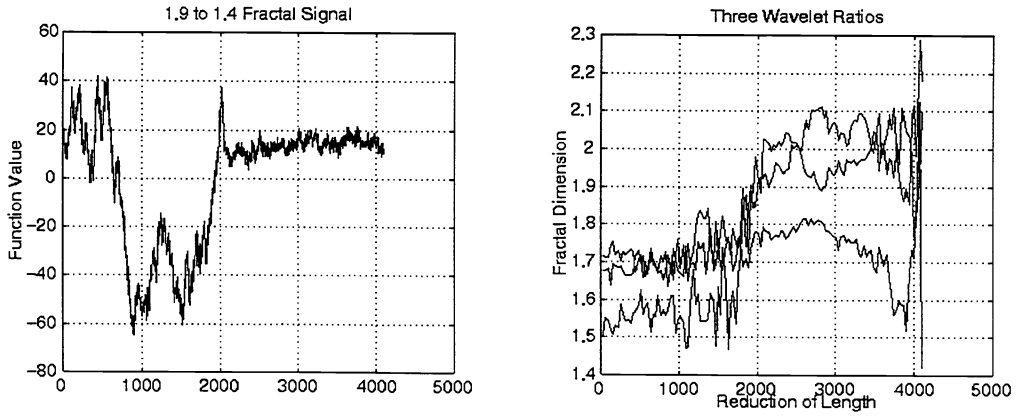

Figure 3

4.2 Analysis of Tank Images

The results from our one-dimension cases were easily expanded to two dimensions. Similar test images were created using the midpoint method and tested with similar results. The most significant concern about the algorithm was its performance on low resolution tank images. A number of tank images were taken with a resolution of 512x480 pixels at Aberdeen Proving Ground. The pixels range from 0 to 255

grayscale and vary linearly with luminance. The images were originally in PCX format and were converted to TIF and then read into a MATLAB matrix where all the computation were done.

Because of the lack of resolution, and the reality that the images are certainly not homogenous, we do not expect a single number, a regional fractal dimension, to determine the location of a tank. Even with the addition of lacunarity, the ability of this cue feature to extract the non-natural object is doubted. Instead, since even at varying resolution levels, the trends in the fractal dimension ratio estimates usually stay the same, more information is contained in the variance of the fractal dimension based on position than on the actual number. The images are divided into sixteen sub-images, each 128x120 pixels. Each image is decomposed several times and three dimension estimates are determined using the energy of the detail wavelet coefficients at four levels. The diagonal, as well as horizontal and vertical details are used. This gives nine cues for each of the 16 sub-images. Three tank images and their corresponding 16 sub-image windows are displayed in Figures 4 through 6.

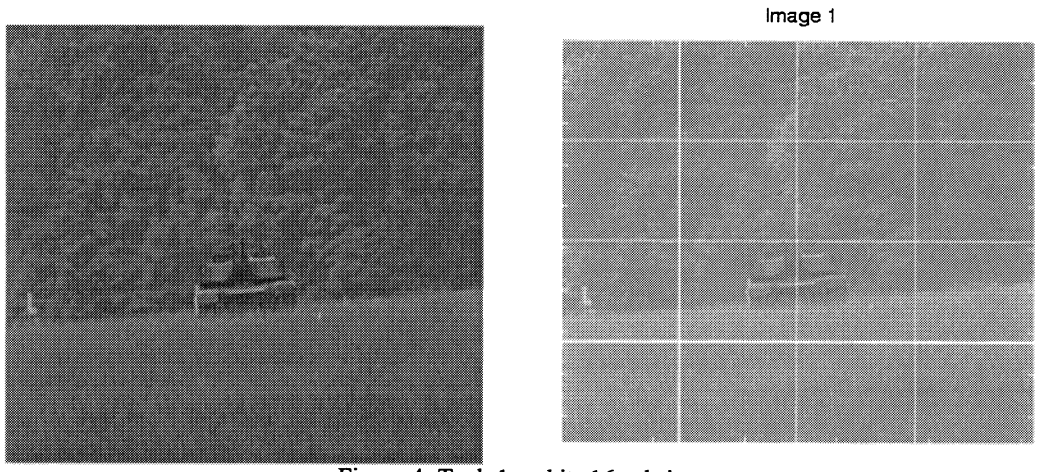

Figure 4: Tank 1 and its 16 sub-images

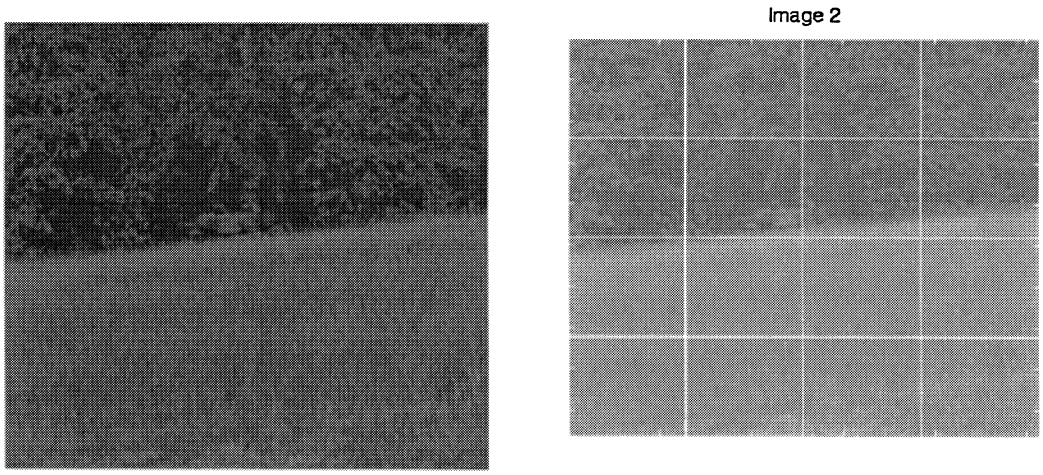

Figure 5: Tank 2 and its 16 sub-images

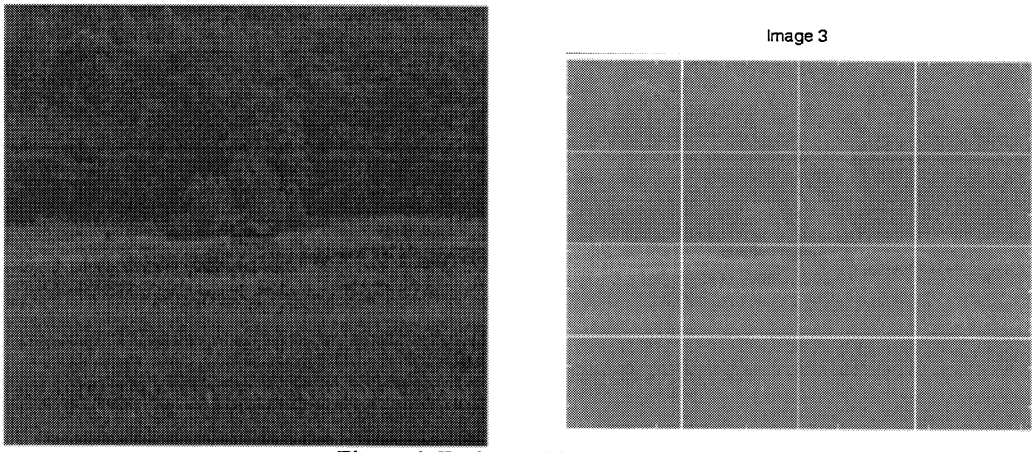

Figure 6: Tank 3 and its 16 sub-images

The resulting dimension estimates were plotted as surfaces consisting of 16 points, each point corresponding to the center of a sub-image. These surfaces were projected as interpolated contours on the actual images as seen in Figure 7. Figure 7 shows the first tank sub-images with the contour plot of the third ratio of the diagonal detail coefficients on the left and the third ratio of the horizontal details on the right. The actual tank gets cut in half by our windowing method, but the ratio estimates still vary significantly for the target windows. For the image on the right, notice the variance of the dimension estimates at the grass-tree boundary and the symmetry.

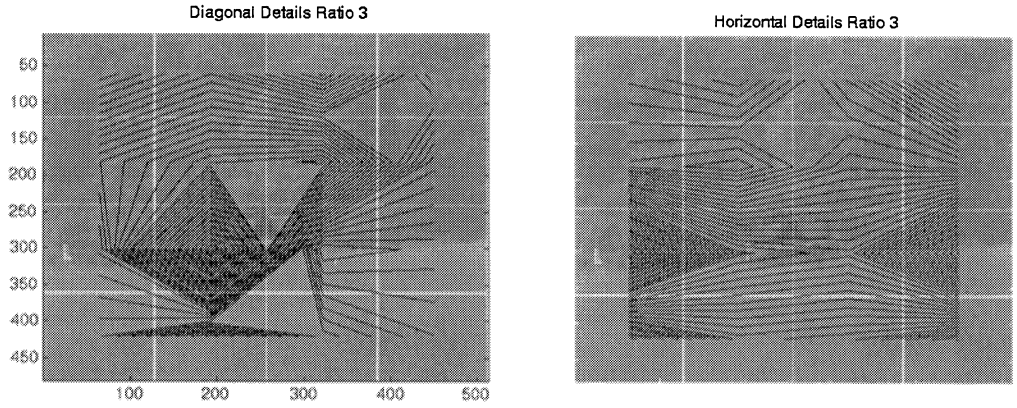

Figure 7: Contour plots over the 16 sub-images of tank 1.

Figure 8 shows the first three ratio contour plots. The corners of each contour plot represent the center of the subimage, as seen when comparing the diagonal ratio 3 below to the contours in Figure 7. Notice that not all ratios provide meaningful information. For the images analyzed, ratio 3 seemed to give the most reliable information.

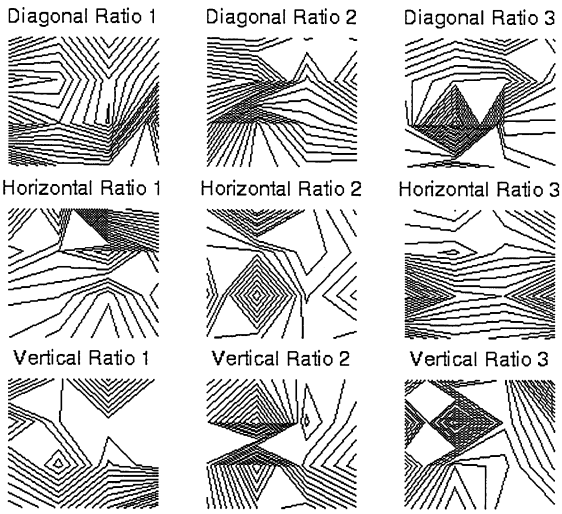

Figure 8: All contour plots of the 16 sub-images of Tank 1.

Figure 9 shows the diagonal ratio 3 dimension estimates for the second tank sub-images. The tank in this image is covered by grass and hidden by the texture boundary, but the variance in the dimension is still substantial at the target location.

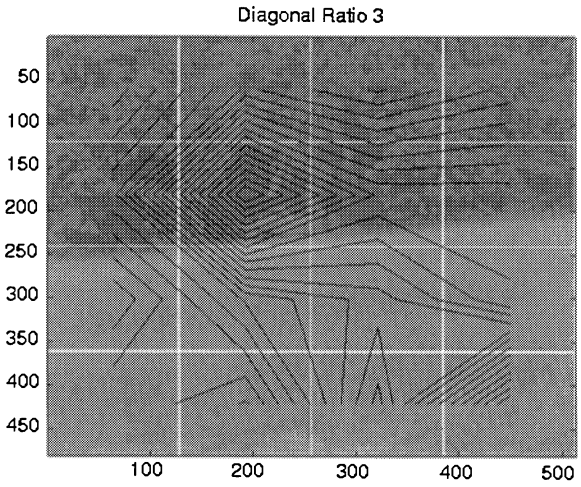

Figure 9: Contour plot over 16 sub-images of Tank 2.

Figure 10 shows all dimension estimates for the second tank image. Again, the third ratio estimates seem to give more information, and the texture boundary is clearly seen in the dimension variation.

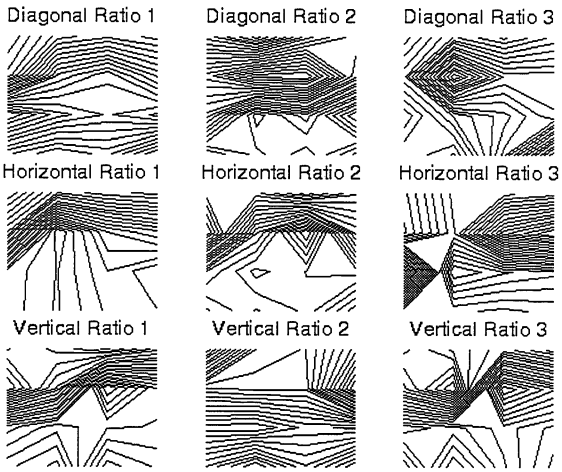

Figure 10: All contour plots of 16 sub-images of Tank 2.

The final tank image is seen below in Figure 11 with the third diagonal dimension. There is a larger amount of haze in this image in addition to three texture variations and a smaller target, but the target subimage still has a larger dimension variance. Figure 12 goes on to show the large number of dimension variances on the target sub-image windows.

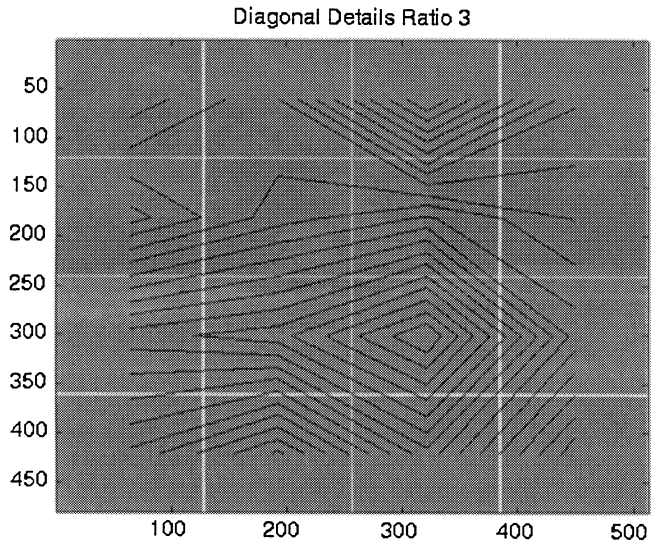

Figure 11: Contour plot of 16 sub-images over Tank 3.

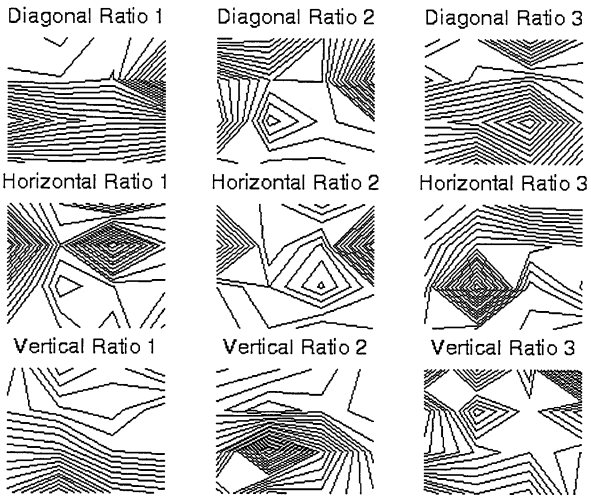

Figure 12: All contour plots of 16 sub-images of Tank 3

5. CONCLUSIONS

The wavelet transform was performed on several images and the trends in the energy ratios of the detail coefficients were compared. The horizontal and vertical details were examined along with the diagonal details. Even with the low resolution of the images, significant variations in the dimension estimates were found in the sub-images containing the M1A1 tank. The nine contour plots of each image provided information about the tank location most of the time. This is seen as very positive since the resolution of the images was so low. This approach to examining images can be used to screen out no-tank terrain images. Further discussion of the resolution effects and the windowing method are needed, along with an automated method of weighting the contour plots to give a positive or negative tank-presence result (neural net.) The actual performance of this approach based on false positives and false negatives needs to be studied as well. Initial comparisons of our results with that of others shows similar accuracy with much higher calculation speeds. Since the fractal dimension can quantify the perception of roughness of fractal textures, and the Human Visual System uses Gabor wavelets, a comparison of the Gabor wavelet is needed as another next step.

6. REFERENCES

[1] G. Wornell, *Signal Processing with Fractals: A Wavelet-Based Approach*, Prentice Hall, 1996.

[2] J. Russ, *Fractal Surfaces*, Plenum Press, 1994.

[3] P. Abry, P. Goncalves and P. Flandrin, "Wavelets, spectrum analysis and 1/f processes," *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, pp 15-30, Springer-Verlag, 1995.

[4] Q. Huang, J Lorch, R. Dubes, "Can The Fractal Dimension of Images be Measured?," Pattern Recognition, pp 339-349, 1994.

[5] J. Keller and S. Chen, "Texture Description and Segmentation through Fractal Geometry," Computer Vision, Graphics, and Image Processing, pp 150-166, 1989.

[6] X. Jin, S. Ong, Jayasooriah, "A practical method for estimating fractal dimension," Pattern Recognition Letters, pp 457-464, 1995.

[7] H. Hastings and G. Sugihara, *Fractals: A User's Guide for the Natural Sciences*, Oxford University Press, 1993

[8] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Transactions on Pattern Analysis and Machine Intellegence, July 1996.

[9] P. Beaver, S. Quirk, J. Sattler, "Object identification in grayscale imagery using fractal dimension," *Fractal Reviews in the Natural Sciences*, M. Novak, pp 63-73, Chapman & Hall, 1995.

[10] M. Misiti, Y. Misiti, G. Oppenheim and J. Poggi, *MATLAB Wavelet Toolbox*, The Math Works, 1996.

[11] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[12] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992

[13] D. Kreithen, S. Halversen, and G. Owirka, "Discriminating Targets from Clutter," MIT Lincoln Laboratory Journal, pp 101-116, Vol. 6, No. 1, Spring, 1993.