Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Contourlet Detection and Feature Extraction for Automatic Target Recognition

JoEllen Wilbur and Robert J. McDonald
Code T13: Computational Sciences
NSWC-PC
Panama City, FL, USA 32407
joellen.wilbur@navy.mil, robert.j.mcdonald@navy.mil

Jason Stack
321 OE: Ocean Engineering and Marine Systems
ONR
875 North Randolph St. Suite 1425
Arlington, VA, USA 22203-1995

***Abstract*—This research presents a contourlet based detection and feature extraction method for underwater targets. The method operates on Side Scan Sonar (SSS) images and is designed to automatically detect and generate target features for classification. Kernel based classifiers are used to determine the best boundary for separating targets and clutter. A statistically significant target data set is generated by embedding additional synthetic targets into SSS data collected during sea tests. Feature trade off studies show an improvement in classification results with the addition of directional based features.**

***Keywords*—contourlet, directional filterbank, kernel matching pursuit, relevance vector machine, support vector machince**

I. INTRODUCTION

Enhancement and representation of manmade submerged targets from Side Scan Sonar (SSS) for automated target recognition (ATR) is an active area of research. Processing of SSS images for ATR can involve using geometric features associated with pixel intensity. Conditions can arise on the seafloor which produces target returns with very little contrast over the background. In such cases the enhancement of SSS images requires preservation of object contour information for both the human operator and shape recognition. In SSS, image backgrounds often contain a broad pixel distribution with varying texture from seafloor effects, whereas manmade objects may display a low contrast or small highlight from the object, but with a prominent shadow. The shadow typically contains valuable shape and size information of the object. Together highlight and shadow contour provide distinctive features in identification of submerged targets. Manmade objects often have a high probability of exhibiting smoother contours along highlight and shadow boundaries. A crucial step in the analysis of acoustic returns from manmade objects for ATR is the accurate representation of these edges.

The curvelet introduced by Candes and Donohoe [1] offers a directional multiresolution transform for representing singularities along curves and edges. The curvelet is a fixed transform for coding of curved edges with significantly fewer coefficients than required by the wavelet. Recently, Do and Vetterli proposed a digital counterpart to the curvelet referred to as the contourlet transform [2,3].

In this paper we describe a method for automatic target recognition using contourlet based features of the target image. The contourlet is a multiresolution, multidirectional expansion using nonseparable filterbanks.

II. BACKGROUND

Contourlets effectively model the contours and ridges in an image. The lattice structure breaks the image into multi-resolutions, or multi-scales. Each scale is then grouped into contour segments using a set of skewing operations that form the mathematical equivalent of placing a directional filter bank on each output of the lattice. Grouping of the coefficients into contour segments gives rise to dominance of the contourlet coefficients along contours and ridges in the image. Where the wavelet distinguishes point singularities and effectively acts to separate point singularities at discrete levels, the contourlet will separate contour edges across scales [2,3].

The transform first performs a multiscale decomposition into octave bands by eliminating every second sample point across rows, then along columns of the image matrix yielding a coarser representation of the image. The eliminated points are replaced by prediction errors, or level differences from which the image can be reconstructed without loss.

For $x[\mathbf{n}]$ the original input image the Laplacian Pyramid outputs J bandpass (detail) images, $b_j[\mathbf{n}]\ \ j=1,2,...,J$, plus one lowpass (coarse) image, $a_j[\mathbf{n}]$.

Each bandpass image $b_j[\mathbf{n}]$, or prediction error, in $l^2(\mathbf{Z}^2)$, is further decomposed by an l_j-level directional filterbank (DFB) into 2^{l_j} bandpass directional images, $c_{j,k}^{(l_j)}[\mathbf{n}]$, where *j=scale, k=direction*.

For an l_j-level DFB there are 2^{l_j} directional subbands with $d_k^{(l_j)},\ 0 \le k < 2^{l_j}$, equivalent synthesis filters and downsampling matrices [2],

$$S_k^{(l_j)} = \begin{pmatrix} diag\left(2^{l_j-1},2\right)\ 0 \le k < 2^{l_j-1} \\ diag\left(2^{j},2^{l_j-1}\right)\ 2^{l_j-1} \le k < 2^{l_j} \end{pmatrix}, \qquad (1)$$

to form in $l^2(\mathbf{Z}^2)$ the directional basis

$$\{d_k^{(l_j)}[\mathbf{n}-S_k^{(l_j)}\mathbf{m}]\},\ 0 \le k < 2^{l_j},\ m \in \mathbf{Z}^2. \qquad (2)$$

The entire decomposition process is iterative and repeated until a base resolution, $a_j[\mathbf{n}]$, is reached. The image is described by the image coefficients [3]

$$\{a_j[\mathbf{n}], c_{j,k}^{(l_j)}[\mathbf{n}]\}. \qquad (3)$$

In this paper the contourlet decomposition is a *3*-level pyramid with directional subbands of order $\{l_1,l_2,l_3\}=\{2^3,2^3,2^1\}$ resulting in one coarse image and 18 directional images. Among several filter types the "9-7" filter produced the best results across SSS data sets and is used in the Laplacian pyramid. Furthermore, this satisfies the biorthogonality of the analysis and synthesis filters[3,4].

III. Contourlet Detection and Feature Extraction

A. Functional Block Diagram

The contourlet detector algorithm is designed to rapidly decompose an image into a multi-resolution, multidirectional series of image shells and combine features extracted from those image shells. A feature vector is created for each block in which a detection is called. The feature vector is then sent to a set of kernel based classifiers.

Figure 1 provides a functional block diagram illustrating the approach. The processor first performs pixel squaring and target enhancement with a chirplet filter over the entire image. Pixel squaring improves computational speed of the contourlet processor.

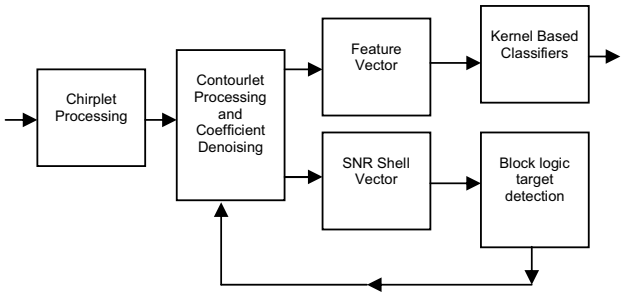


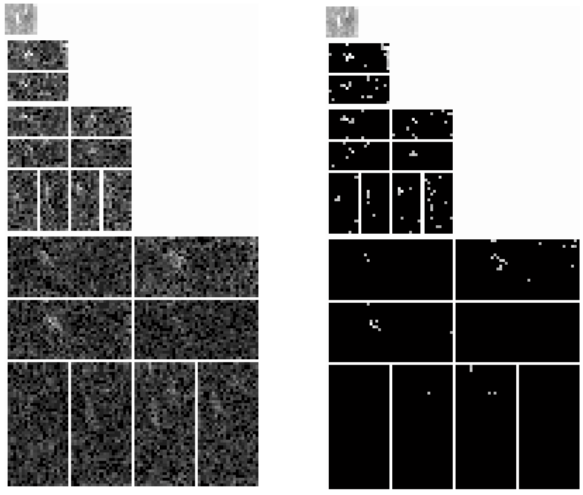
Figure 1. Block diagram for contourlet detection and feature extraction.

Overlapping blocks of the chirplet processed image are applied to a detection and feature extraction algorithm. The contourlet based detector/feature extractor performs block processing on the image and assigns a detection weight, or score, to each detected block. Contourlet shell decomposition is performed on each image block. Contourlet denoising is applied to the shell coefficents in the directional subbands given by the $c_{j,k}^{(l_j)}[\mathbf{n}]$ in equation (3). The denoising filter uses a hard limiter with a noise threshold of

$$\eta_{thr} = l_j\left(\frac{1}{length(c_n)}\sum_n (c_n - \bar{c}_n)^2\right)^{\frac{1}{2}}, \qquad (4)$$

where c_n is the vector of directional coefficients of length $length(c_n) = J2^{l_j}$.

Figure 2 gives the *3-level* directional contourlet decomposition for a target buried in sand before and after coefficient denoising. There are J bandpass images, where each image is further decomposed into 2^{lj} directional subbands plus one lowpass image. For the *3-level=$\{2^3,2^3,2^1\}$* decomposition pictured below this corresponds to 18 directional subbands plus 1 coarse, or lowpass, image.



(a) (b)

Figure 2. (a) Curvelet Shells (b) Curvelet Shells after application of denoising algorithm

B. *Target Detection*

The detection decision is calculated from the combined signal-to-background ratios within each contourlet shell taken from the denoised contourlet coefficients for that particular shell. The score is calculated from a combination of the mean SNR for the denoised directional shell coefficients and the SNR of the reconstructed filtered image snippet of the processed image block after denoising.

The composite of the mean directional SNR and the reconstructed denoised snippet image SNR form the score for each block within the image. The detection threshold, $d_{thr}(i)$, adjusts to the mean score and deviation over processed image blocks for that image as follows

$$d_{thr}(i) = \langle score_{block} \rangle + \delta \times std\{score_{block}\} \qquad (5)$$

$$score_{block} = \frac{1}{J2^{l_j - 1}} \sum_j \sum_k \frac{\hat{c}_{j,k}{}^{l_j}[n_0] \hat{c}_{j,k}{}^{l_j^*}[n_0]}{(\sigma_{j,k}^{l_j})^2} + SNR[\hat{i}_{block}]$$

where $\hat{c}[n_0]$ is the peak coefficient in the j,kth directional shell post processing and $\hat{i}_{block}$ is the filtered, reconstructed image block from the denoised contourlet coefficients. The deviation parameter δ is dependent on the operational requirements. A feature vector is generated for all detections above the threshold and fed to a feature based classifier. Figure 3 illustrates the threshold process on SSS data collected during sea trials. The threshold adjusts to the image as defined in (5) and is shown in Figure 3b by the horizontal line across detection blocks for the image in 3a. The detection mapping in 3c gives the corresponding feature boxes for each target location. Only those processed blocks above the detection threshold (which for the image in Figure 3 is .39) form detection blocks.

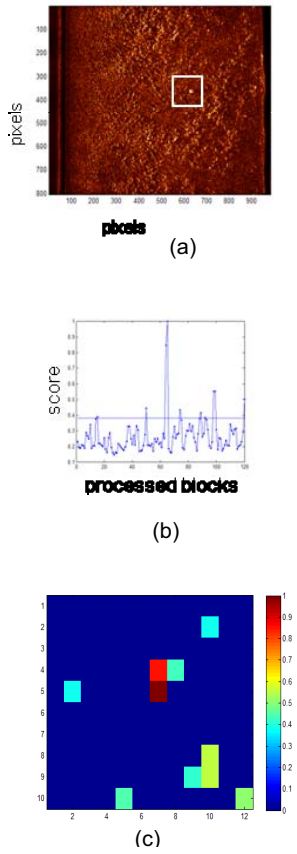


Figure 3. (a) Sonar image showing detection block (b) adapting threshold, adjusts to image mean score and deviation (c) detection map.

C. *Feature Extraction for Classification*

A feature vector is calculated for each detection block. The detection blocks have been re-centered in the image about the peak energy of each target. The detection block is 176x176 pixels for the case study presented here. More generally, the location and size of the detection block is defined by operational requirements and choice of contourlet filter lattice parameters. The denoised contourlet coefficients are calculated for each detection block in the image and the feature vectors are calculated from these contourlet coefficients.

The features are provided in Table I below. The energy, entropy, and mean of the denoised coefficients are calculated because these features have been shown to yield high classification rates when derived from contourlet coefficients [5]. The additional features from the lowpass shell are added to discriminate peaked sand ridges from targets.

Table I

Features 1-7	**Features 8-25**
SCORE: Out of detector **SNR**: Avg of SNRs in each shell **CDET**: SNR from denoised, filtered reconstructed image box **LPE**: Energy of coeff peaks in spatial lowpass shell **LPMU**: Mean of coeff peaks in spatial lowpass shell **LPNPKS**: Number of coeffs > threshold =mean+3*sigma in lowpass shell **LPP2PW**: Clustering of coeffs peaks above threshhold in lowpass shell	**BPE**: Energy of coeff peaks in each directional shell (1 by 18 vector) **Features 26-43** **BPMU**: Mean of coeff peaks in each directional shell (1 by 18 vector) **Features 44-62** **LPETP**: Entropy in spatial lowpass shell **BPETP**: Entropy in directional shells (1 by 18 vector)

Statistical testing of the contourlet features requires an extensive set of targets as well as non-targets in the sonar imagery. One means of accomplishing this is by inserting synthetic targets into actual SSS images to create an object set large enough for statistical testing of the features. Using the methodology described in [6,7], simulated target returns were generated for the particular SSS used during the data collection process. The simulated targets were then scaled and embedded into the real SSS images, thus producing an image set comprised of both real and synthetic targets. The purpose for adding synthetic targets is to generate enough targets for statistical testing of the features.

Figure 4 below is a typical example of an image with synthetic targets and man-made clutter (lobster traps) used for testing the various contourlet features.

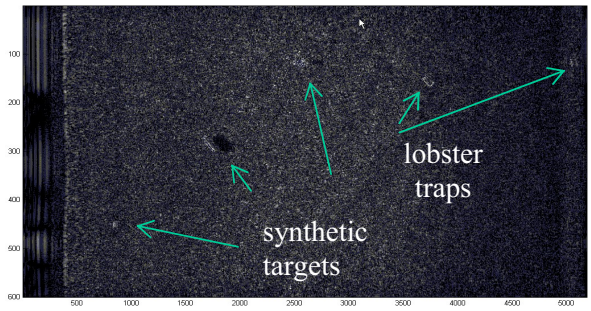


Figure 4. Sonar image with synthetic targets inserted into a background containing man-made clutter.

D. Results

The features are tested using the following kernel machines: the kernel matching pursuit KMP described [7,8,9], the support vector machine algorithm (SVM) [10,11] and the relevance vector machine described in [9,12].

The receiver operating curve (ROC) was calculated and the area under the curve (AUC) was used as the performance metric. (An AUC=1.0 indicates perfect classification with no false alarms.) Blind testing over independent trials yielded performance values in terms of average AUC (area under the ROC curve). Figure 5 gives the average AUC from each of the feature based classifiers where average AUC corresponds to the areas under the ROC curves averaged over 10 independent trials.

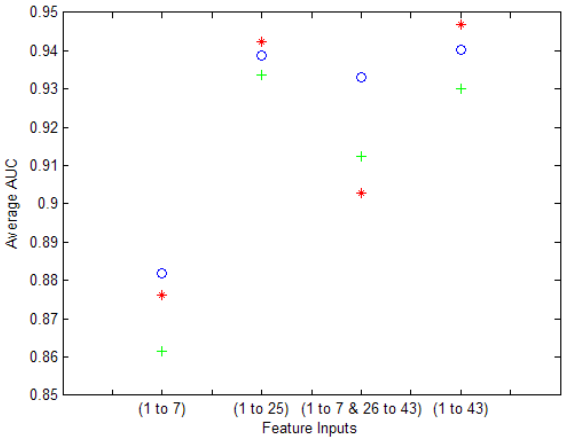


Figure 5. Average AUC output relative to feature input vector where O=KMP, *=SVM, + =RVM

In all cases each classifier eliminated the entropy in the directional subbands during the learning process. Addition of the energy in the directional shells offered worthwhile improvement in performance. Additional features beyond directional energy had little, if any, gain in average AUC.

IV. Conclusions

In this paper we showed a contourlet based method for detecting and identifying underwater targets in sonar images. The contourlet coefficients are used to generate a compact set of features that are fed to a set of feature based classifiers. The classifiers used for evaluating the contourlet features were the KMP, RVM, and the SVM. All classifiers were in agreement as to the level of feature input relative to performance.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research, Code 321 Maritime Sensing.

REFERENCES

[1] E.J. Candès and D.L. Donoho, "Curvelets – A Surprisingly Effective Nonadaptive Representation For Objects with Edges," *Saint-Malo Proceedings*, pp. 1-10, Vanderbilt University Press, Nashville, TN. 2000

[2] T.T. Nguyen and S. Oraintara, "Multiresolution Direction Filterbanks: Theory, Design, and Applications," *IEEE Transactions on Signal Processing*, Vol. 53, No. 10, Oct 2005

[3] M.N. Do and M. Vetterli, "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation," *IEEE Transactions on Image Processing,* vo.l 14, no. 12, pp. 2091-2106, 2006

[4] Do, M.N., and Vetterli, M., "Framing Pyramids," *IEEE Trans. on Signal Processing*, Vol. 51, No. 9, Sept. 2003.

[5] Semler, L. and Detorri, L., "Curvelet Based Texture Classification of Tissues in Computed Tomography," *Proceedings IEEE on Computer Tomography.*

[6] G. Sammelman, J. Christoff, J. Lathrop, "Synthetic Images of Proud Targets," *Proc. IEEE MTS OCEANS 2004,* pp. 266-271, Brest France, September 1994.

[7] T. Cobb and J. Stack, "In Situ Adative Feature Extraction for Underwater Target Classification," *Applied Image Pattern Recognition Workshop,* 36th IEEE, pp. 42 – 47, Oct. 10-12, 2007.

[8] J.R. Stack, G. Dobeck, X. Liao, and L. Carin, "Kernel matching pursuits with arbitrary loss functions," IEEE Transactions on Neural Networks, vol. 20, no. 3, pp. 395-405, Mar. 2009.

[9] X. Liao, H. Li, and B. Krishnapuram, "An M-ary KMP Classifier for Multi-Aspect Target Classificatio," *ICASSP-2004*, vol II, no 61, 2000

[10] T. Joachims, "Optimizing Search Engines using Clickthrough Data," Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, Edmonton, Alberta, Canada

[11] pp. 133-142, 2000. T. Joachims, "Making Large-Scale SVM Learning Practical," Ch. 11., Advances in Kernel Methods – Support Vector Learning, MIT Press, Cambridge, USA 1998.

[12] M. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine," *J. Machine Learning Research,* 1, 2001.