Multi-Image Texton Selection for Sonar Image Seabed Co-Segmentation

J. Tory Cobb[a], Alina Zare[b]

[a] Naval Surface Warfare Center Panama City Division, Panama City, FL, USA;

[b] Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO, USA

ABSTRACT

In this paper we describe an unsupervised approach to seabed co-segmentation over the multiple sonar images collected in sonar surveys. We adapt a traditional single image segmentation texton-based approach to the sonar survey task by modifying the texture extraction filter bank to better model possible sonar image textures. Two different algorithms for building a universal texton library are presented that produce common pixel labels across multiple images. Following pixel labeling with the universal texton library, images are quantized into superpixels and co-segmented using a DP clustering algorithm. The segmentation results for both texton library selection criteria are contrasted and compared for a labeled set of SAS images with various discernable textures.

Keywords: Synthetic Aperture Sonar, Image Segmentation, K-Distribution, Parameter Estimation, Texton, Dirichlet Process

1. INTRODUCTION

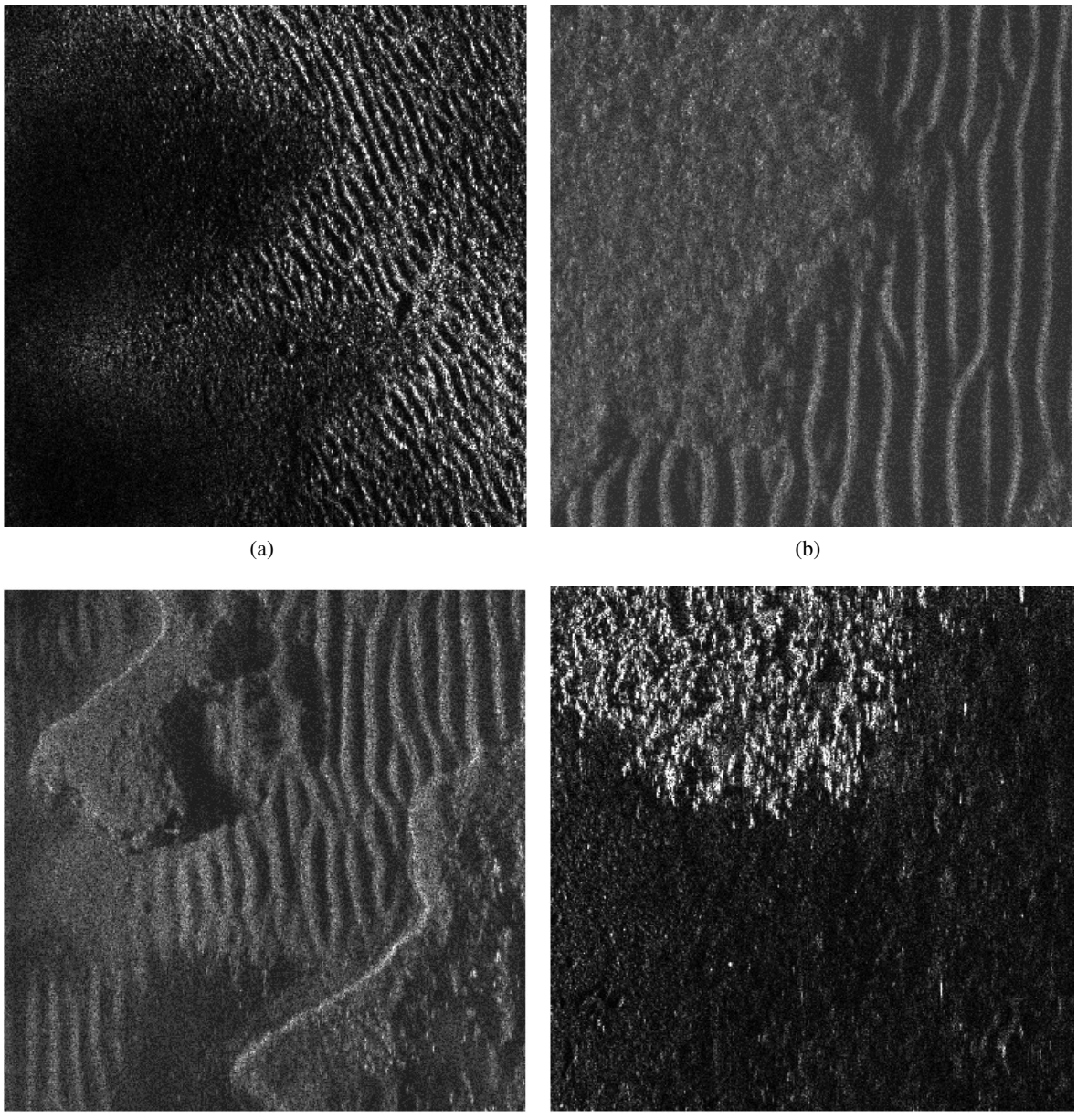
In the past ten years a sonar sensing approach known as stripmap synthetic aperture sonar (SAS) has matured to the point where autonomous underwater vehicles carrying these systems can produce highly-resolved seabed images.[1] These images clearly depict sand ripples, sea grass beds, and rock and coral beds in near photographic clarity. Additionally, the well-defined highlight and shadow boundaries of the patterns within these images yield rich features for distinguishing the general periodicity, shape, and height of seabed textures. Figure 1 shows a montage of various synthetic aperture sonar images of the seabed containing different seabed textures. Algorithms that segment and label these textured regions are useful for higher level automated tasks such as scene recognition, data mining from previously surveyed data, and seabed mapping.

Sonar images have many characteristics in common with natural scene optical images such as diffuse boundaries and large intraclass feature variation. Additionally, sonar image raster lines are generated by the envelope (amplitude or intensity) of sonar echoes, thus all images are grayscale. These complications alone make a typical binary segmentation task of finding a single object or texture within an image background difficult. Even more difficult is the task of finding similar seabed labels across multiple images. For example, common seabed categories such as a sand ripple field will not be defined by an affine transform between multiple images because the texture itself and not the ripple field boundary contains the important information.

In this paper we propose an unsupervised seabed co-segmentation algorithm for use in *survey-level* SAS image scene analysis tasks. In a SAS survey, hundreds or even thousands of stripmap images are collected then analyzed by a human operator. Viewing and labeling such a large amount of imagery is very time-consuming and an automated algorithm could save the expense and time of detailed human interaction. The overall goal of the task we present here is similar to the goal of co-segmentation[2] or shared scene segmentation[3] because we seek to label similar regions across all images in a single survey.

When segmenting single images, features that accurately separate different regions may not discriminate as well when applied to images with slight variations to the same textures. An example of this in sonar images is when encountering sand ripples at different orientations. A useful feature for survey-level segmentation would capture the salient features, e.g. size and period, but would be invariant to ripple orientation. Current single-image segmentation techniques using texton histograms as individual pixel features do not transfer easily to this task because they are tied to a common set of filter responses that are endemic to the textures within a particular image.[4,5] Here we propose two methods to create joint texton libraries to accurately categorize common textures across multiple SAS images.

Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII, edited by J. Thomas Broach, Jason C. Isaacs, Proc. of SPIE Vol. 8709, 87090H · © 2013 SPIE
CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2016427



(a) (b)

(c) (d)

Figure 1. Montage of SAS sample images with various textures. Some textures of interest are depicted clockwise from the upper left: (a) small sand ripples, (b) seagrass mixed with large sand ripples, (c) rock and hardpack sand, and (d) large sand ripples, seagrass, and hardpack sand.

In an automated sonar survey we also desire to accurately label common seabed regions without *a priori* knowledge of the number of categories. To circumvent this problem we propose the use of a Dirichlet process mixture model[6,7] to automatically compute the label cardinality of the segmentations. The Dirichlet process is becoming a very popular Bayesian clustering approach to natural-scene segmentation[3,8] due to its flexibility in modeling various feature distributions.

2. SONAR IMAGE SEABED FEATURES

Sonar image seabed texture descriptors have recently been characterized using bi-orthogonal[9] and Daubechies[10] wavelets, Haralick features,[11] and autocorrelation function parameters.[12] Other authors[13,14] have proposed single-point statistics such as the K-distribution shape parameter ν as a useful seabed texture descriptor when discriminating between vastly different seabed categories such as smooth sand and coral. Here we unite sonar image texture features and single-point image statistics within the framework of textons, a commonly used approach to segment similar regions in single images.[4]

2.1 Single-point Statistics

Borrowing from Oliver's work in synthetic aperture radar (SAR) image statistics,[15] single-point or single-pixel i.i.d. statistics of high-resolution side look sonar intensity images have been modeled as a product of a gamma random variable σ and an exponential random variable n[12,16] or

$$I(x,y) = n(x,y)\sigma(x,y). \tag{1}$$

In this compound model, the $\sigma(x,y)$ term models the underlying sonar cross-section (SCS) of the (x,y) resolution cell and $n(x,y)$ models the multiplicative noise-like qualities of speckle. In sonar image texture analysis, the SCS of the different seabed textures produces the backscatter received by the sonar, while speckle statistics are typically independent of any underlying scattering phenomena. Thus, in the work here we are interested in isolating the $\sigma(x,y)$ term to gain an understanding of the single-point statistics of the sonar cross section. In gross terms, the shape parameter of σ denoted ν_σ describes how heavy-tailed the distribution of image pixels is within a given sample size and is related to the image contrast. The larger the ν_σ value, the more uniform the texture appearance (e.g. smooth sand), and the smaller the value of ν_σ, the more bright-dark variation of a seabed texture (e.g. sand ripples).

A sample estimator for the shape parameter ν_σ can be easily derived from the second moment of a gamma pdf [15, Eq. 5.8]

$$E\left\{\sigma^2\right\} = \left(\frac{E\left\{\sigma\right\}}{\nu_\sigma}\right)^2 \frac{\Gamma\left(\nu_\sigma + 2\right)}{\Gamma\left(\nu_\sigma\right)}, \tag{2}$$

by substituting expectation operators with their respective sample estimates. This leads to the estimator

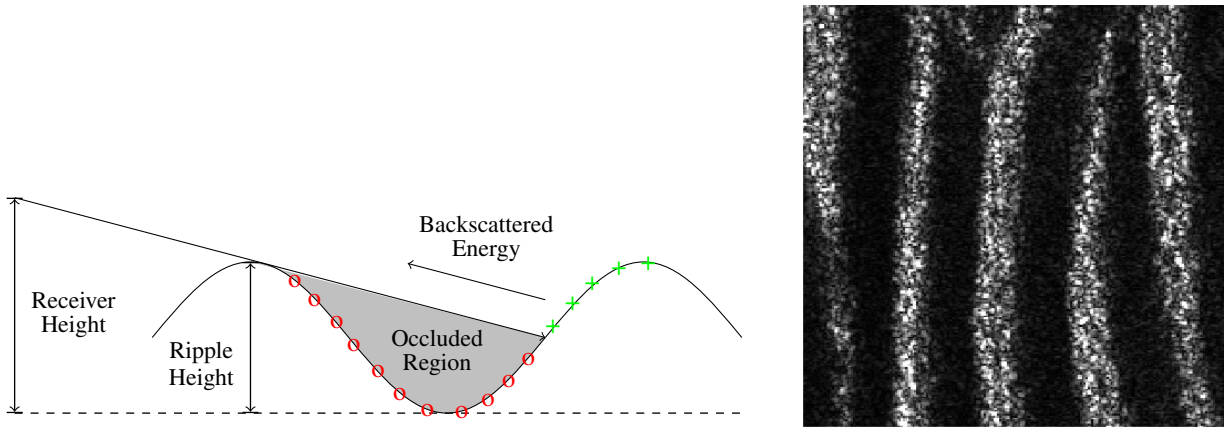
$$\hat{\nu}_\sigma = \frac{(\hat{\mu}_\sigma)^2}{\hat{var}(\sigma)}, \tag{3}$$

where $\hat{\mu}_\sigma$ and $\hat{var}(\sigma)$ are the sample mean and sample variance operators respectively. The estimate $\hat{\nu}_\sigma$ will be used in Section 3 along with $\hat{\mu}_\sigma$ as features for the segmentation algorithm.

2.2 Sonar Image Texture Filterbank

Malik, Belongie, Leung, and Shi introduced a set of texture features based on clustered responses from a filterbank of various orientations, spatial frequencies, and scales generated from Gaussian derivatives.[4] While this filterbank represents many edge orientations and sizes in typical electro-optic image settings, we limited the orientations and spatial frequencies to better suit common sonar image object representations.

In side-look sonar imagery a single horizontal raster scan is created by an echo from the seafloor, also called backscatter, received broadside from the sonar. Stacking the raster scans vertically forms a matrix of echoes which is interpreted as an image. Bright spots on the seafloor signify high-energy backscatter and dark spots signify low-energy backscatter or the absence of sound. In typical sonar textures of interest such as coral beds or sand ripples, the bright-dark variation is a result of the alternating high-energy backscatter from seabed facets perpendicular to the impinging sound followed by seabed facets occluded by the preceding bottom topography. The texture produced from a sand ripple bathymetric profile is depicted in Figure 2 (a) and (b). In this example, bright-dark variations occur regularly from left to right across the sonar



(a) Ripple Sensing Geometry (b) Ripple Sonar Image

Figure 2. A simple model of sand ripple bathymetry (a) and a corresponding sonar image of sand ripple (b). In the model, the occluded resolution cells (o) return little energy to the receiver, while the resolution cells on the visible face of the sand ripple (+) return a lot of energy. This manifests as quasi-periodic bright-dark variations in the sonar image (b).

image. The sensing geometry also limits the type of bright-dark variations that may occur, specifically an alternating bright-dark pattern from top to bottom across the sonar image. Backscattered energy will not produce an alternating pattern when "looking down the trough" because the angles of incidence of the transmitted acoustic waves are the same in the sensed region even though the bathymetric profile may vary. The knowledge of the back-scattering physics can thus be used to constrain filterbank libraries, i.e. no filters oriented around 0^o.

To capture the seabed textures common in sonar images we constructed a filterbank similar to texton-based segmentation,[4] but constrained the set of possible orientations to better conform to side-look sensing geometry. The filterbank consists of oriented odd-symmetric and even-symmetric filter pairs and radially-symmetric filters. Depicted in Figure 3 (a), the oriented filter pairs are rotated through 5 angles, $\theta = \{45^o, 60^o, 90^o, 120^o, 135^o\}$ at 4 scales for a total of 40 filters. Depicted in 3 (b), the radially-symmetric filters comprise Gaussian filters at 4 scales and Laplacian of Gaussian filters at 8 scales for a total of 12 filters. Incorporating our knowledge of side-look sonar image formation, we limited the oriented filters' rotation angles to represent only those seabed textures created by occlusion from bathymetric variation. By restricting the set of orientations yet increasing the number of scales we aim to produce better fidelity in the discriminating seabed texture features.

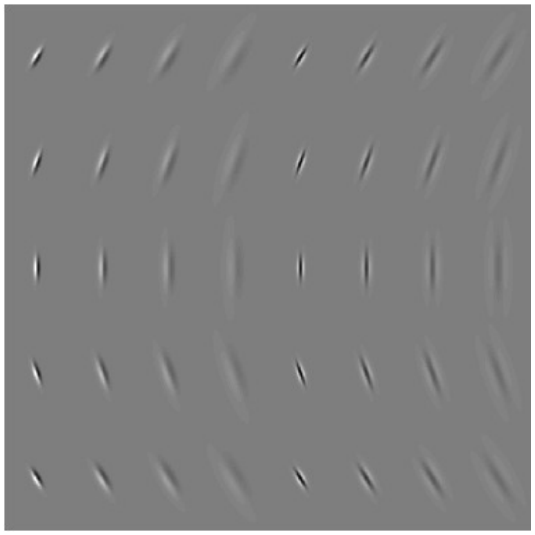
3. SEGMENTATION APPROACH

Our co-segmentation scheme borrows from texton library generation for image segmentation[4] and contour detection,[5,17] superpixel image representation,[18,19] and image segmentation via Dirichlet process clustering.[3] In our sample case we have 40 1000×1000 sonar images produced from a high-resolution, high-frequency side-look SAS system. The images contain a variety of seabeds including: seagrass, coral, scattered rocks, bright sand, mud, and sand ripple fields with two distinct periods.

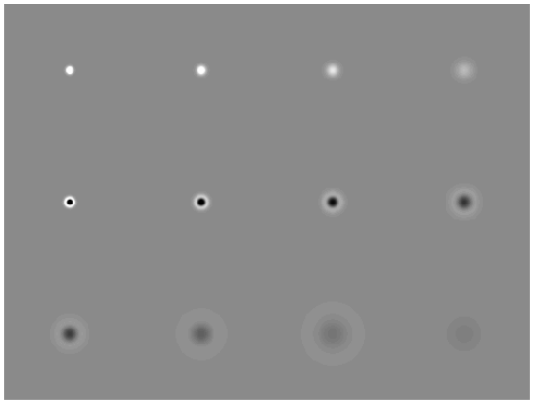
Image co-segmentation is carried out in six steps. First, each image is preprocessed by median filtering and normalized to a common mean. Next the filterbank coefficients and textons for each image are calcuated and stored. The larger set of single-image textons is then concentrated into a smaller set of multi-image or universal textons by additional clustering. The multi-image texton set is then applied across all images to build texton histograms centered on each image pixel. These histograms are then clustered into superpixels for a final dimensionality reduction step prior to segmentation. Finally, the image superpixels are segmented using a Dirichlet process clustering algorithm. A block diagram of the approach is depicted in Figure 4.

3.1 Image Preprocessing

Sonar images are first median-filtered with a 3×3 median filter to reduce multiplicative speckle and isolate the underlying SCS of seabed textures. The single-point shape parameter estimate $\hat{\nu}_\sigma$ is then calculated pixelwise in a moving window



(a) Oriented Filters



(b) Radially Symmetric Filters

Figure 3. Plot of the filterbank used to extract sonar image texture features. Oriented odd-symmetric and even symmetric filter pairs (a) are depicted at 4 scales and rotated through 5 angles $\theta = \{45^o, 60^o, 90^o, 120^o, 135^o\}$ for a total of 40 filters. Radially-symmetric filters (b) comprise Gaussian filters at 4 scales and Laplacian of Gaussian filters at 8 scales for a total of 12 filters.

across the entire image and stored for later use. Following estimation of $\hat{\nu}_\sigma$, the logarithm is taken of each de-speckled amplitude image and the mean removed. The log of a synthetic aperture radar or sonar image is a popular format for extracting texture information[20,21] because it transforms the pixel data into a more amenable form for calculating Gaussian statistics such as correlation, variance, and mean.

3.2 Single Image Texton Selection

Following preprocessing, each image is filtered with the filterbank described in Section 2.2. The filtered image coefficients are concatenated with the estimated shape parameter $\hat{\nu}_\sigma$ calculated in 16×16 and 32×32 windows, and pixel mean calculated in 32×32 and 64×64 windows resulting in a vector of length 56 centered at each pixel. These feature vectors are then clustered with the k-means algorithm using $k = 45$. The reader should note that this modified texton representation also includes information from single-point statistics (mean and $\hat{\nu}_\sigma$) and is not the same representation as described in traditional texton segmentation approaches.[4] Nonetheless, this step adequately quantizes salient texture information.

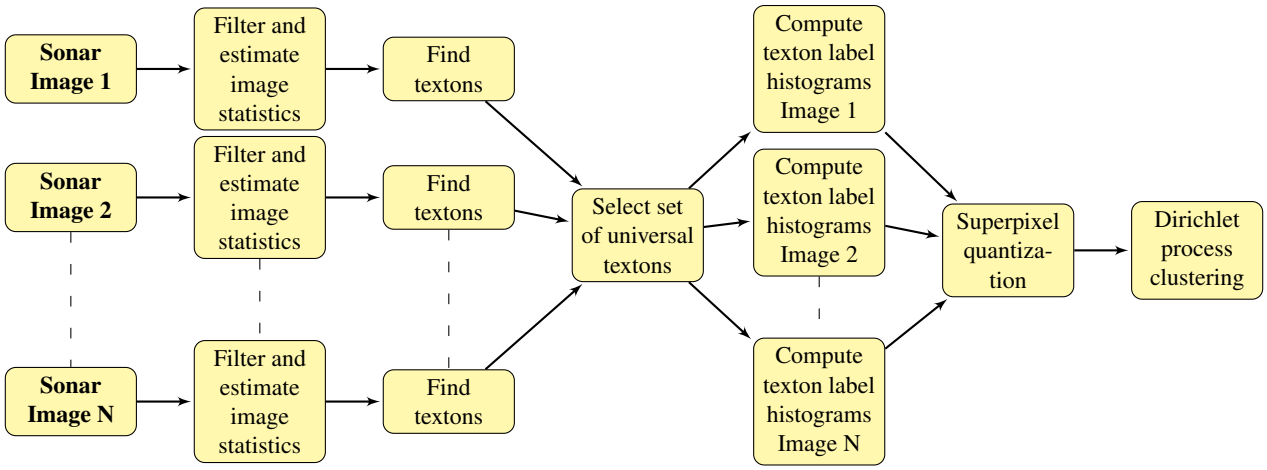


Figure 4. Block diagram of the co-segmentation approach. First, filter coefficients and image statistics are found for each sonar image. Next, a set of image-specific textons is then found for each image using the k-means algorithm. A set of universal textons is found from the larger set of image-specific textons using either the minimum distortion or maximum scatter criteria. The universal textons are then used to generate texton label histograms centered at each pixel. These histograms are quantized to into superpixels as a data reduction and spatial feature clustering step. Finally, the superpixels are clustered using a Dirichlet process mixture model. The label assignments generated by the DP clustering are the seabed segmentation labels.

3.3 Multi-Image Texton Selection

After each set of single-image textons is found, the next task is to seek a universal set of textons that can represent common seabed regions that appear in multiple images. Previous work by Martin[5] addressed this idea of *universal textons* but only offered the solution of clustering all the image filter coefficients at one time, which is computationally burdensome in the case of the amount of data collected from a typical sonar survey. If this approach was applied to our example here, we would have to cluster 40 million vectors of length 56 for 40 1000×1000 sonar images, not an impossible task. However, in a real-life scenario a typical sonar survey produces several hundred 4000×4000 images, making consolidated clustering of all the images a much more difficult task. So to demonstrate a realistic approach to sonar survey image co-segmentation here we test two techniques for combining the single-image textons into a universal set of textons using the 40 images in our test set.

In the first approach we choose the universal texton set based on a *minimum distortion* criteria. Here we cluster the 40 sets of 45 textons using the k-means algorithm with $k = 45$ once again. This technique will of course quantize the textons into a smaller set of universal textons with the aim of finding a multi-image representative texton that is closest to neighboring textons in Euclidean distance. This approach will combine clusters of single-image textons into representative samples but will not guarantee any global properties of the universal textons with respect to the distance between cluster centroids.

In the second approach we choose the universal texton set based on a *maximum scatter* criteria. In this approach we seek a universal set of textons that covers that greatest volume in feature space or produces maximum separation between the individual textons. Here we initialize the set of universal textons by randomly choosing a texton from the set of 1800 single-image textons. The next texton t_i added to the universal set $\mathbf{T}_U$ is the one that produces maximum scatter or

$$\arg\max_{t_i \in \mathbf{T}} \sum_{t_k \in \mathbf{T}_U} (t_i - t_k)^2 . \tag{4}$$

We continue to add textons one at a time that satisfy (4) until the set of universal textons contains 45 vectors. The algorithm continues by choosing a texton at random from the universal set found by the greedy search and replaces it with a texton that satisifies (4). This random selection and replacement continues until the set of textons does not change for several select/replace cycles. In practice we find that the universal texton set becomes stable after about 200 iterations of random selection and replacement.

3.4 Superpixel Quantization

After finding a set of universal textons we assign the pixels in the original images the label of the universal texton closest to its feature vector of filterbank coefficients and statistical parameters. A 60×60 window of labels centered at each pixel is then binned into a histogram of length 45. These histograms are then finally quantized into superpixels before segmentation. Superpixel quantization serves as a final data reduction step prior to segmentation. Our superpixel quantization approach uses the intuitive idea of a superpixel from Ren[18] and Mori[19] but implements the superpixel clustering using a k-means algorithm similar to Jain's texture segmentation approach using Gabor filters[22] where pixel spatial location plays a role in each clustered feature vector. The superpixel representation is obtained by breaking each 1000×1000 histogram image into 4 blocks of 250×250. The smaller 250×250 blocks of histograms are then concatenated with a scaled (x, y) location so that contiguous spatial regions are favored in the superpixel centroid assignments. 100 superpixels are created for each 250×250 block using a k-means algorithm with $k = 100$, thus quantizing each sonar image into 1600 superpixels. The entire sonar image set of 64,000 superpixels (40 images $\times$ 1600 superpixels/image) is then segmented in the final stage of the algorithm. The concatenated (x, y) location is removed from each superpixel prior to clustering.

3.5 Segmentation via Dirichlet Process Clustering

A Dirichlet process (DP) clustering algorithm segments the sonar image superpixels into labeled regions. A DP mixture model[7] has two key properties that make it applicable in this co-segmentation task, 1) no prior specification of label cardinality is required and 2) an appropriate probability model, i.e. multinomial in our case, can be chosen to best fit the feature vectors of interest. In a DP mixture model,[23] the data vector or superpixel $\mathbf{s}_i$, $i \in \{1, 2 \ldots, N\}$, is assumed to be drawn from a mixture of component distributions with mixing parameter π_k, $k \in \{1, 2 \ldots, \infty\}$. The prior mixing parameter variable π_k is drawn from a Dirichlet process with concentration parameter α and base distribution G_o. To cluster the superpixels, we infer the labels c_i given the superpixels $\mathbf{s}_i, i \in \{1, 2 \ldots, N\}$ using a Gibbs sampling algorithm.[24]

For the image data we chose a multinomial probability model with parameter vector $\mathbf{p}$, denoted $Mult(\mathbf{p})$ for the mixture components because it is best suited for clustering histograms. Additionally we choose the Dirichlet distribution, denoted $Dir(\boldsymbol{\gamma}_o)$, for the base distribution G_o with hyperparameter $\boldsymbol{\gamma}_o$ because it is the conjugate prior of the multinomial distribution and simplifies some of the sampling calculations. A summary of the DP mixture model construction follows

$$\begin{aligned}
\mathbf{s}_i \mid \mathbf{p}_k, c_i = k &\sim Mult(\mathbf{p}_k) && (5)\\
\mathbf{p}_k \mid \boldsymbol{\gamma}_k, c_i = k &\sim Dir(\boldsymbol{\gamma}_k) && (6)\\
c_i &\sim \sum_{k=1}^{\infty} \pi_k \delta_k && (7)\\
\pi_k &\sim DP(G_o, \alpha) && (8)\\
G_o &\sim Dir(\boldsymbol{\gamma}_o). && (9)
\end{aligned}$$

Clustering or inference of the data labels c_i is carried out via a series of Markov chain (MC) sampling steps using the model described above. The MC sampling algorithm implements a form of DP inference known as stick-breaking.[24] The inference or clustering algorithm is summarized below:

1. Initialize all mixture component parameters, hyperparameters, mixing parameters.
2. Assign label $c_i = k$ to $\mathbf{s}_i$ based on probability $\propto \pi_k L(\mathbf{s}_i \mid \mathbf{p}_k)$ where
$$L(\mathbf{s}_i \mid \mathbf{p}_k) = \prod_j p_k(j)^{s_i(j)}. \quad (10)$$
3. Update $\boldsymbol{\gamma}_k \mid c_i = k$ where
$$\boldsymbol{\gamma}_k \mid c_i = \boldsymbol{\gamma}_o + \sum_{c_i = k} \mathbf{s}_i. \quad (11)$$
4. Draw
$$\mathbf{p}_k \sim Dir(\boldsymbol{\gamma}_k). \quad (12)$$

5. Draw
$$v_k \sim \beta(1 + n_{i=k}, \alpha + n_{i>k}), \tag{13}$$
where $\beta(a, b)$ denotes a beta random variable with parameters a and b and the notation $n_{i=k}$ means the number of data points assigned to cluster k.

6. Update
$$\pi_k = v_k \prod_{i=1}^{\infty}(1 - v_i). \tag{14}$$

7. Repeat Steps 2-6 until convergence.

8. The label c_i denotes cluster assignment of superpixel $\mathbf{s}_i$.

Note: in practical implementation of the inference algorithm, the number of components k is limited to some large integer $K >> k$ thus the product in (14) has a finite number of elements.

4. RESULTS

The co-segmentation scheme described in Section 3 was applied to 40 sonar images which contained several discernable textures: shadow, dark sand, bright sand, seagrass, large sand ripple, rock, and small sand ripple. The textured regions were hand labeled by a sonar image analyst. The DP clustering algorithm parameters were set to $\alpha = 0.012$ and $\boldsymbol{\gamma}_o = \mathbf{1}$. Figure 5 depicts 14 images (first and fourth column) alongside their co-segmentations using the minimum distortion criteria (second and fifth columns) and maximum scatter criteria (third and sixth columns) for the universal texton library selection. The co-segmentation results are also listed in Tables 1 and 2. Since the DP clustering algorithm does not generate a predetermined number of clusters an additional label category of "Other" is appended to the end of the tabulated results that contains any pixel labels that do not match the other 7 label assignments.

The co-segmentation results produced by the minimum distortion selection criteria yields a slightly higher co-segmentation accuracy across all label categories. Close examination of the images reveals that the labeled regions contain erroneous labels well within a homogeneous region's boundary. The minimum distortion criteria does discriminate large ripple, seagrass, and shadow regions fairly well. The minimum distortion criteria does not discriminate well between rock and small ripple. In some cases the maximum scatter criteria produces more visually homogeneous segmentation labels but erroneously groups large portions of some label groups together. Notably, the bright sand labels are combined with seagrass labels and rock labels are combined with small sand ripple labels. These two co-segmentation schemes appear to have tradeoffs between higher overall accuracy (minimum distortion) and homogeneous region labels (maximum distortion) thus the choice of the scheme may depend on a follow-on application using the labeled data.

Table 1. Minimum Distortion Texton Selection Co-Segmentation Results (in Pct.)

	Segmentation Label							
True Label	SH	DS	BS	SG	LR	RK	SR	Other
Shadow (SH)	48.7	4.1	0.2	3.9	6.9	10.0	1.1	25.1
Dark Sand (DS)	21.6	25.6	5.9	15.0	0.7	25.4	3.4	2.4
Bright Sand (BS)	4.1	10.7	22.6	43.7	1.5	14.0	2.4	1.0
Seagrass (SG)	2.6	2.1	0.1	55.1	6.8	22.5	8.6	2.2
Large Rip. (LR)	0.1	0.2	0.0	7.0	84.8	3.1	3.8	1.0
Rock (RK)	0.1	0.3	0.0	8.9	2.7	5.2	82.5	0.3
Small Rip. (SR)	0.0	0.0	0.0	14.6	0.0	1.6	83.8	0.0

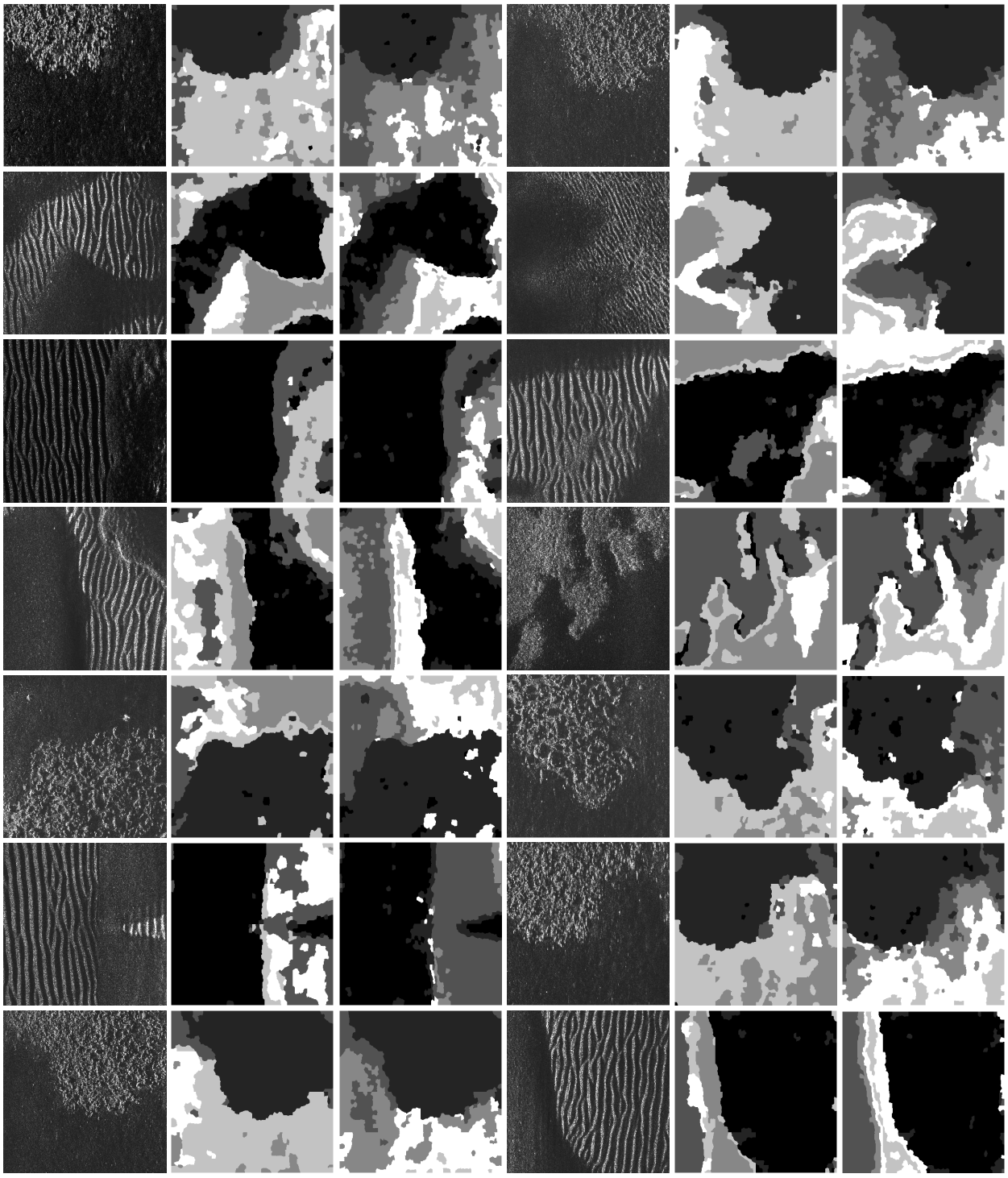


Figure 5. A sample of co-segmentation results from the set of 40 labeled sonar images. The first and fourth column from the left are the original images. The segmentation results using the minimum distortion criteria for universal texton selection are displayed in the second and fifth column from the left. The segmentation results using the maximum scatter criteria for universal texton selection are displayed in the third and sixth column from the left.

Table 2. Maximum Scatter Texton Selection Co-Segmentation Results (in Pct.)

	Segmentation Label							
True Label	SH	DS	BS	SG	LR	RK	SR	Other
Shadow (SH)	41.2	3.4	3.8	17.0	5.5	9.1	3.5	16.5
Dark Sand (DS)	10.9	25.0	33.3	5.0	0.5	18.8	4.9	1.6
Bright Sand (BS)	1.7	8.9	72.8	5.0	0.9	3.1	7.4	0.2
Seagrass (SG)	1.5	5.8	46.4	11.7	3.3	3.7	27.0	0.6
Large Rip. (LR)	0.1	0.3	3.2	3.3	78.4	0.1	14.6	0.0
Rock (RK)	0.1	0.5	6.3	3.7	2.6	0.2	86.6	0.0
Small Rip. (SR)	0.0	0.0	8.0	0.4	0.1	0.0	91.5	0.0

5. CONCLUSIONS

We have presented an unsupervised approach to co-segment multiple sonar images collected from seabed surveys into commonly labeled regions. We constrained the filterbank-based approach of textons to the limitations of side-look sonar sensing geometry and offered two different approaches for building universal texton libraries. The two approaches to universal texton selection, the minimum distortion criteria and maximum scatter criteria offered tradeoffs in accuracy versus label homogeneity. A Dirichlet process clustering approach allowed us to accurately model feature distributions and eased the requirement of specifying the number of label clusters *a priori*.

There are several avenues for future work using this approach. Geospatial information can be added to bias clustering via the Spatial Distance Dependent Chinese Restaurant Process.[8] Since each pixel in the map is tagged with latitude and longitude information, pixels grouped in the same image and across images that are repeat passes of the same seabed could be clustered together. A hierarchical DP clustering approach[25] would probably need to be adopted to handle cross-image co-segmentation of labeled regions in this case because geospatially grouped regions of similar labels would need to be grouped globally to identify a survey-level set of seabed labels. Further improvement of universal texton selection could also be possible by examining the sensitivity of results to the various choices for k values in the k-means clustering steps or by using other clustering techniques that better approximate the distribution of texton values rather than restrictive unit-variance assumptions of the presented approaches. Finally, it is clear that some label groups are difficult to distinguish whatever the clustering or texton selection critieria, e.g. rock and small ripple, thus future work should focus on identifying useful features for segmentation of these texture groups.

ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research Code 321OE.

REFERENCES

1. D. Brown, D. Cook, and J. Fernandez, "Results from a small synthetic aperture sonar," in *Proc. IEEE OCEANS*, **1**, pp. 1–6, (Boston, MA), Sept. 2006.
2. C. Rother, T. Minka, A. Blake, and V. Kolmogorov, "Co-segmentation of image pairs by histogram matching - Incorporating a global constraint into MRFs," in *Computer Vision and Pattern Recognition (CVPR 2006)*, **1**, pp. 993–1000, 2006.
3. E. Sudderth and M. Jordan, "Shared segmentation of natural scenes using dependent Pitman-Yor processes," in *Neural Information Processing Systems (NIPS) 2008*, **22**, 2008.
4. J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *Int'l J. Computer Vision* **43**, pp. 7–27, June 2001.

5. D. Martin, C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, pp. 530–548, May 2004.
6. R. Neal, "Bayesian mixture modeling," in *Proceedings of the Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis*, **11**, pp. 197–211, 1992.
7. M. Escobar and M. West, "Bayesian density estimation and inference using mixtures," *Journal of the American Statistical Association* **90**, pp. 577–588, 1995.
8. S. Ghosh, A. Ungureanu, E. Sudderth, and D. Blei, "Spatial distance dependent Chinese Restaurant Process for image segmentation," in *NIPS*, pp. 1476–1484, 2011.
9. D. Williams, "Unsupervised seabed segmentation of synthetic aperture sonar imagery via wavelet features and spectral clustering," in *Proceedings of the 16th IEEE International Conference on Image Processing (ICIP)*, pp. 557–560, Nov. 2009.
10. T. Celik and T. Tjahjadi, "A novel method for sidescan sonar image segmentation," *IEEE J. Ocean. Eng.* **36**, pp. 186–194, Apr. 2011.
11. M. Lianantonakis and Y. Petillot, "Sidescan sonar segmentation using texture descriptors and active contours," *IEEE J. Ocean. Eng.* **32**, pp. 744–752, July 2007.
12. J. T. Cobb, K. C. Slatton, and G. J. Dobeck, "A parametric model for characterizing seabed textures in synthetic aperture sonar images," *IEEE J. Ocean. Eng.* **35**, pp. 250–266, Apr. 2010.
13. J. Dunlop, "Statistical modeling of sidescan sonar images," in *Proc. MTS/IEEE Oceans Conf. and Exhibition*, **1**, pp. 33–38, (Halifax, Nova Scotia), Oct. 1997.
14. J. T. Cobb and K. Slatton, "Dynamic tree segmentation of sonar imagery," in *Proc. SPIE Defense and Security Symposium*, **6553**, 2007.
15. C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, SciTech, Raleigh, NC, 2004.
16. D. Abraham and A. Lyons, "Simulation of non-Rayleigh reverberation and clutter," *IEEE J. Ocean. Eng.* **29**(2), pp. 347–362, 2004.
17. P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**, pp. 898–916, May 2011.
18. X. Ren and J. Malik, "Learning a classification model for segmentation," in *Ninth IEEE Int'l Conference on Computer Vision (ICCV '03)*, **1**, pp. 10–17, Oct. 2003.
19. G. Mori, "Guiding model search using segmentation," in *Tenth IEEE Int'l Conference on Computer Vision (ICCV '05)*, **2**, pp. 1417–1423, Oct. 2005.
20. R. Frankot and R. Chellappa, "Lognormal random-field models and their applications to radar image synthesis," *IEEE Transactions on Geoscience and Remote Sensing* **GE-25**, pp. 195–207, Mar. 1987.
21. P. Lombardo and C. Oliver, "Estimating the correlation properties of K-distributed SAR clutter," *IEE Proc. Radar, Sonar Navig.* **142**, pp. 167–178, Aug. 1995.
22. A. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition* **24**(12), pp. 1167–1186, 1991.
23. R. Neal, "Markov chain sampling methods for Dirichlet process mixture models," *Journal of Computational and Graphical Statistics* **9**, pp. 249–265, June 2000.
24. H. Ishwaran and M. Zarepour, "Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models," *Biometrika* **87**(2), pp. 371–390, 2000.
25. Y. Teh, M. Jordan, M. Beal, and D. Blei, "Hiearchical Dirichlet processes," *Journal of the American Statistical Association* **101**, pp. 1566–1581, Dec. 2006.