AN UNSUPERVISED SONAR IMAGES SEGMENTATION APPROACH

Abdel-Ouahab Boudraa
IRENav, Ecole Navale & E^3I^2, (EA 3876) ENSIETA
Lanvéoc Poulmic, BP600, 29200 Brest−Armées, France

Jean-Christophe Cexus
IUT de Lannion
BP 30219, Rue Edouard Branly 22302 Lannion Cedex, France

Keywords: Image segmentation, Gabor filtering, Clustering, Feature selection, Texture, Sonar image.

Abstract: In this work an unsupervised Sonar (Sound navigation and ranging) images segmentation is proposed. Due to the textural nature of the Sonar images, a band-pass filtering that takes into account the local spatial frequency of these images is proposed. Sonar image is passed through a bank of Gabor filters and the filtered images that possess a significant component of the original image are selected. To calculate the radial frequencies, a new approach is proposed. The selected filtered images are then subjected to a non-linear transformation. An energy measure is defined on the transformed images in order to compute texture features. The texture energy features are used as input to a clustering algorithm. The segmentation scheme has been successfully tested on real high-resolution Sonar images, yielding very promising results.

1 INTRODUCTION

Synthetic aperture side scan Sonar imagery has long been a field of intense research interest for both military and civilian applications. An example of application is the sea mine detection and classification where traditionally a human operator would be required to carry out the analysis based on its expert knowledge. In high resolution Sonar imagery, three kinds of region can be visualized: echo, shadow, and sea-bottom reverberation. On images supplied by Sonar system, the echo features are generally less discriminant than the shadow shapes for the classification of objects lying on the sea-bed. For this reason, the detection of each object located on the sea bottom and its classification (as a wreck, a rock, a man-made object, etc.) is generally based on the extraction and the identification of its cast shadow (Collet et al., 1996). A number of unsupervised techniques have been proposed to segment Sonar images, including use of Markov Random Fields (MRF) to approximate the differing gray level regions and pixel correlations within the regions (Murino, 2001)-(Mignotte et al., 1999). The later technique provides an accurate segmentation, but is computationally intense. A fuzzy clustering method has recently been proposed to the segmentation problem using gray level information (Stitt et al., 2001). However, fuzzy clustering is very sensitive to speckle noise. Due the textural nature of the Sonar images and to the fact that they are strongly corrupted by speckle noise, a multi-channel filtering approach is suitable to take into account the local spatial frequency content of these images and to reduce the noisy components. In this work Gabor filters (Gabor, 1946) are used as band-pass filters (Daugman, 1985). Gabor kernels are commonly used for texture feature extraction. Their popularity is motivated by the mathematical and the biological properties of Gabor functions.

2 FILTER BANK DESIGN

Gabor filters provide simultaneous optimal resolution in both space and spatial-frequency domains (Daugman, 1985). In spatial domain, the complex impulse response of Gabor filters is given by

$$h(x,y) = \exp\left\{-\frac{1}{2}\left[\frac{(x-x_0)_r^2}{\sigma_x^2}+\frac{(y-y_0)_r^2}{\sigma_y^2}\right]\right\}\times \exp\left\{2\pi j(\mu_0 x+\nu_0 y)\right\} \quad (1)$$

Boudraa A. and Cexus J. (2007).
AN UNSUPERVISED SONAR IMAGES SEGMENTATION APPROACH.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IFP/IA*, pages 165-170

$h(x,y)$ is a complex sinusoid, known as the carrier, centered at the spatial frequency (μ_0, ν_0) in Cartesian coordinates and modulated by a 2D Gaussian-shaped function, known as the envelope. σ_x and σ_y are the space constants of the Gaussian envelope along the x and y axes, respectively. μ_0 and ν_0 are the central radial frequencies along the x and y directions respectively. These spatial frequencies can also be expressed in polar coordinates as (F_0, θ):

$$\mu_0 = F_0 \cos\theta \quad \text{and} \quad \nu_0 = F_0 \sin\theta \tag{2}$$

F_0 is the radial center frequency measured in cycles per pixel. The point (x_0, y_0) is he peak of the function $h(x,y)$, and r subscript sands for a rotation operation such that

$$\begin{bmatrix}(x-x_0)_r\\(y-y_0)_r\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix} \cdot \begin{bmatrix}x-x_0\\y-y_0\end{bmatrix} \tag{3}$$

Filter with an arbitrary orientation, θ, can be obtained via the rigid rotation of the $x-y$ coordinates system using relation (3). The angle θ is also the angle of rotation of the envelope. The 2-D Fourier transform, or spectral transfer function, of the real part of the Gabor function is given by (even-symmetric filters):

$$\begin{aligned} H(u,v) \quad = A\Bigg(& \exp\left\{-\tfrac{1}{2}\left[\frac{(u-u_0)_r^2}{\sigma_u^2} + \frac{v_r^2}{\sigma_v^2}\right]\right\} \\ & + \exp\left\{-\tfrac{1}{2}\left[\frac{(u+u_0)_r^2}{\sigma_u^2} + \frac{v_r^2}{\sigma_v^2}\right]\right\}\Bigg) \end{aligned} \tag{4}$$

where $A = 2\pi\sigma_x\sigma_y$, and (u,v) are the components of the frequency in the x and y directions respectively. $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$ are the extents of the Gaussian envelope in the spectral domain in the x and y directions respectively.

3 SELECTION OF PARAMETERS

By passing the original image through a Gabor filter, we obtain all those components in the image that have their energies concentrated near the spatial frequency point $(\pm u_0)$ within a frequency bandwidth of B_r octaves and orientation bandwidth of B_θ degrees. A Gabor filter bank is usually designed to cover all available the frequency spectrum (Jain and Farrokhnia, 1991),(Manjunath and Ma, 1996),(Guo et al., 2000). In general the Gabor filter set is constructed such that the half-peak magnitude ($\eta = 0.5$) of the filter in the frequency spectrum touch each other. In this work, we define the η-peak magnitude of the filter to compute B_r as follows:

$$B_r = 2\log_2\left(\frac{\mu_0 + \sqrt{(-2\ln\eta)}\sigma_u}{\mu_0 - \sqrt{(-2\ln\eta)}\sigma_u}\right) \tag{5}$$

where $\eta \in [0,1]$ is the filter magnitude where neighboring filters along the μ axis intersect. To calculate radial frequencies, a new approach, different than that proposed by (Jain and Farrokhnia, 1991), is presented. Let f_l and f_h be the lower and the higher frequency respectively ($u_0(0) = u_0 \in [f_l, f_u]$). We suppose that the radial frequencies follows a logarithmic scale and the frequencies of the ith and (i-1)th filters are such that:

$$\frac{u_0(i)}{u_0(i-1)} = 2^{B_r} \text{ and } B_r = \frac{\ln(\frac{f_l}{f_u})}{n\ln 2} \tag{6}$$

In the frequency domain the spread $\sigma_u(i)$ and $\sigma_v(i)$ of the ith filter are given by:

$$\begin{aligned} \sigma_u(i) = 2^{i\times B_r} \text{ and } \sigma_v(i) = \frac{K_b}{K_a}.\sigma_u(i) \\ K_b = \tan\left(\frac{\pi}{2K_\theta}\right) \text{ and } K_a = \frac{2^{B_r-1}}{2^{B_r+1}} \end{aligned} \tag{7}$$

where $\sigma_v(0) = \sigma_0$ and K_θ is the number of orientations.

4 MULTI-CHANNEL FILTERING

For an original image $I(x,y)$, the output of the Gabor filter, $I_h(x,y)$ is given by

$$I_h(x,y) = I(x,y) * h(x,y) \tag{8}$$

where $*$ denotes the convolution product. The image $I(x,y)$ is filtered using a bank of $L = n \times K_\theta$ filters where n is the number of scales used. The n value can be given by the number of radial frequencies μ_0 used or the width of the image, N. Thus, if N is a power of 2, the frequencies selected are: $1\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, 8\sqrt{2}, \ldots, (N/4)\sqrt{2}$ cycles per width.

The filters orientation ($\theta \in [0,\pi]$) is given by

$$\theta_m = m\frac{\pi}{K_\theta} \quad \text{with} \quad m \in \{0,1,\ldots,K_\theta - 1\} \tag{9}$$

The angular bandwidth, B_θ, is given by

$$B_\theta = \frac{\pi}{K_\theta} \tag{10}$$

Before generating the feature images, each filtered image is transformed in a new image, $I_{NL}(x,y)$, using a non-linear function:

$$I_{NL}(x,y) = \psi(\alpha \mid I_h(x,y) \mid) \tag{11}$$

where

$$\psi(t) = \tanh(\alpha t) = \frac{1-\exp(-2\alpha t)}{1+\exp(-2\alpha t)} \tag{12}$$

where α is a constant (Jain and Farrokhnia, 1991) ($\alpha = 0.25$ is this study). A texture measure is defined over a small Gaussian window, with a standard deviation $\sigma = \frac{N}{2F_0}$ and of size $M \times M$, around each transformed pixel in the selected filtered images. M is inversely related to u_0, and is the smallest odd integer larger than or equal to 5σ, where $\sigma = 0.25N/u_0$ (Jain and Farrokhnia, 1991). More formally, the feature image $e_k(x,y)$, corresponding to the kth filtered image $r_k(x,y)$, is given by

$$e_k(x,y) = \frac{1}{M^2} \sum_{(a,b)\in W_{xy}} | \psi((r_k(a,b)) | \qquad (13)$$

where W_{xy} is a window of size M^2 centered at pixel (x,y).

5 SPACE REDUCTION

The values in the L feature images corresponding to a given pixel form an L-dimensional feature vector representing the pixel. Features are normalized to zero mean and unit standard deviation. Some filtered images may show similar response to different textures because the textures may share the same spatial frequency properties. Hence, the L filtered images are not all of practical interest and thus space reduction is necessary to discard irrelevant image features. Principal Components Analysis (PCA) is commonly used reduction technique (Jolliffe, 1986). Given a set of data, PCA finds the linear lower-dimensional representation of the data such that the variance of the reconstructed data is preserved. Intuitively, PCA finds a low-dimensional hyperplane such that, when we project the data onto the hyperplane, the variance of the data is changed as little as possible (maximum of data variance). In general there is no standard rule for deciding how many principal components should be used to represent the data adequately, but a useful heuristic is to choose a fraction (0.8 in this study) of the inertia I_q to be retained by computing:

$$I_q = \frac{\sum_{j=1}^{q\leq p} \lambda_j}{\sum_{i=1}^{p} \lambda_i} \qquad (14)$$

λ_j denotes eigenvalues and p denotes the number of eigenvalues.

6 CLUSTERING

K-means is well known method for clustering data (Jain and Dubes, 1988). However, like all partitional algorithms, the k-means requires the number of clusters before starting the clustering process. Due to the low contrast and the speckle of Sonar images, the estimation of the number of clusters is very difficult (Figs. 2(a),9(a),10(a)). To avoid this problem the k-means is started with an overestimated number of clusters, *KC*, and combined with the dendrogram (Jain and Dubes, 1988).

7 RESULTS

Experiments have been conducted on real Sonar images (Figs. 2(a),9(a),10(a)). Sonar images are provided by a side-scan Sonar with frequency around 500 kHz. The size of these images is 256×256 pixels corresponding to a sea floor surface of 25 by 25 m. Thus, for $N = 256$ a filter bank can be created with a total of $n = 7$ central frequencies or scales associated to a given orientation. The number of orientations K_θ is set to 5. Finally, we start with $L = 35$ Gabor filters. For each pixel (x,y) is associated a feature vector of 35 features. Using the dendrogram the number of clusters is reduced to 2. Figure 2(a) displays a man made object (Trolley) lying on the sea bed. Gaobr filter with tuned radial frequency and orientations is applied to "Trolley" image (Fig. 2(a)). Figure 3 shows 20 filtered images (among 35) corresponding to five orientations: 0°, 36°, 72°, 108°, 144°. The filtered images clearly show that filter responses in object-echo regions are different from those in the non object-echo regions. Feature images are obtained by transforming the filtered images using a non-linearly relation (Eq. 11) followed by a Gaussian filtering (Eq. 13). Result of 15 feature images is shown in figure 4. Not that the feature values corresponding to object (Trolley) are consistently high. A PCA analysis performed for space reduction is shown in figure 5. This figure shows that many feature images are irrelevant and this is confirmed by the calculated eigenvalues 7. Indeed, the plot of figure 7 shows that most the information of the original image ("Trolley") is concentrated on few eigenvalues and thus only a reduced number of feature images is of practical interest. From the reduced number of feature images, feature vectors are formed and clustered using the k-means with *KC* set to 10. Result of clustering is shown in figure 6. Figure 8 shows the dendrogram obtained using the *KC* cluster centers generated by the k-means. The extracted shadow of the trolley ex-

hibits, as expect, a regular shape (Fig. 2(b)). Figure 9(a) displays a real Sonar image of sandy floor with the cast shadow of a manufactured object (cylinder). As in Figure 2, the shadow region est well segmented (Fig. 9(b)). Figure 10(a) displays a real Sonar image involving an object and rock shadows. The segmentation result is shown in figure 10(b). The shadows of the rock and the manufactured object are well detected. However, a post processing such as connected components analysis is necessary to the two shadows. The obtained results (Figs. 2(b),9(b),10(b)) are in good agreement with the ground truth provided by an expert. Note that figures 2(c), 9(c) and 10(c) display the class or cluster corresponding to set formed by echo and sea-bottom reverberation. The accuracy in extracting and preserving the borders of the cast shadows is very appealing in the prospect of as further classification step. Furthermore the proposed scheme exhibits good robustness against speckle noise.

8 CONCLUSIONS

In this paper an unsupervised segmentation method to distinguish, from Sonar images, man-made and natural objects lying on the sea-bed is presented. The obtained results show the interest of multi-channel filtering, based on the Gabor filters, to segment Sonar images. To calculate the radial frequencies of the Gabor filters a new approach is proposed. To confirm the obtained results a large of Sonar images to segment is necessary. Furthermore, we plan to study the influence of the number of Gabor filters on detection objects on sea floor followed with a ROC analysis (False alarm,....). Finally, comparison of the obtained results to those of existing methods and particularly those based one the Markovian model (Mignotte et al., 1999) in terms of time complexity and False alarm is necessary.

ACKNOWLEDGMENTS

The authors would like to thank GESMA (Groupe d'Étude Sous Marine de l'Atlantique), Brest, France for providing real Sonar images.

REFERENCES

Collet, C., Thourel, P., Perez, P., and Bouthemy, P. (1996). Hierarchical mrf modeling for sonar picture segmentation. In *Proc. IEEE ICIP*, pages 979–982.

Daugman, J. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Amer.*, 2:1160–1169.

Gabor, D. (1946). Theory of communication. *J. Inst. Elect. Eng.*, 93:429–457.

Guo, G., Li, S., Chan, K., and Pan, H. (2000). Texture image segmentation using reduced gabor filter set and mean shift clustering. In *Proc. of the Fouth Asian Conference on Computer Vision*.

Jain, A. and Dubes, R. (1988). *Algorithms for clustering data*. Prentice Hall, New Jersey.

Jain, A. and Farrokhnia, F. (1991). Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12):1167–1186.

Jolliffe, I. (1986). *Principal Component Analysis*. Springer-Verlag, New York.

Manjunath, B. and Ma, W. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:937–842.

Mignotte, M., Collet, C., Perez, P., and Bouthemy, P. (1999). Three class markovian segmentation of high resolution sonar images. *Computer Vision Image Understanding*, 76(3):191–204.

Murino, V. (2001). Reconstruction and segmentation of underwater acoustic images combining confidence information in mrf models. *Pattern Recognition*, 34(5):981–997.

Stitt, J., Tutwiler, R., and Lewis, A. (2001). Fuzzy c-means image segmentation of side scan sonar images. In *Proc. IASTED*.

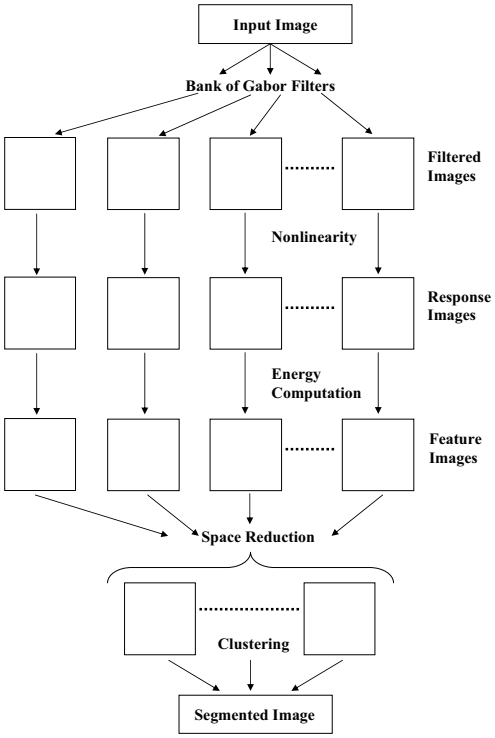

Figure 1: Overview of the segmentation process.

Figure 2: A real Sonar image involving sea floor and a man-made object (trolley) (a:left image). Extracted cluster "Shadow" (b: middle image). Complement of the cluster "Shadow" (c: right image).

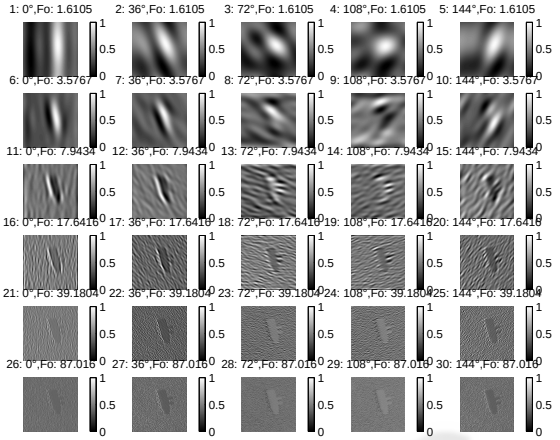

Figure 3: Example of 20 filtered images for Trolley image.

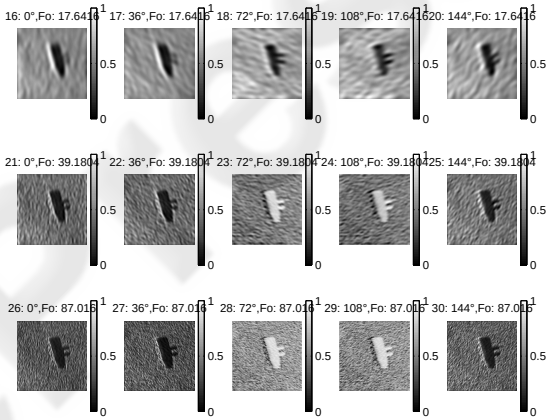

Figure 4: Example of 15 Feature images.

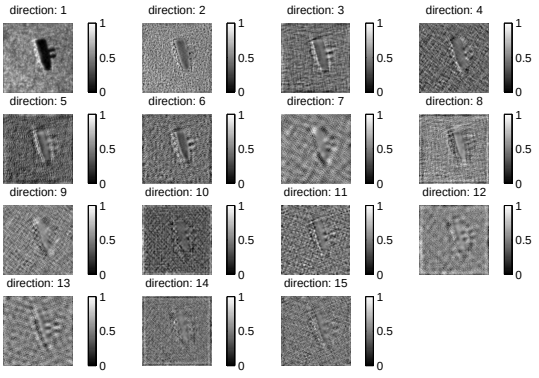

Figure 5: ACP analysis.

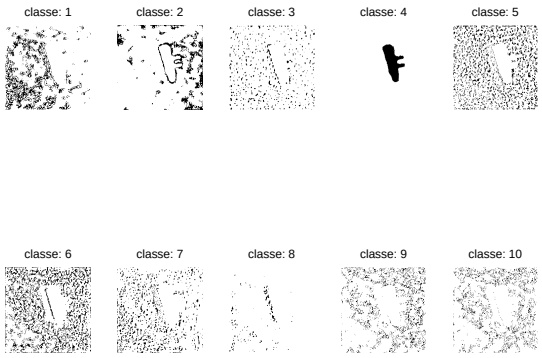

Figure 6: Clustering using 10 classes of Trolley image.

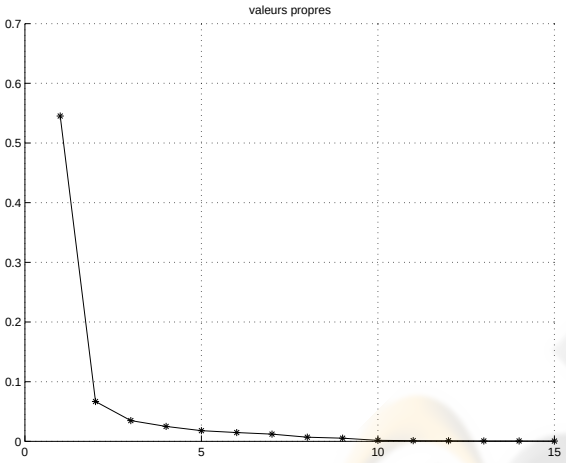

Figure 7: Eigenvalues.

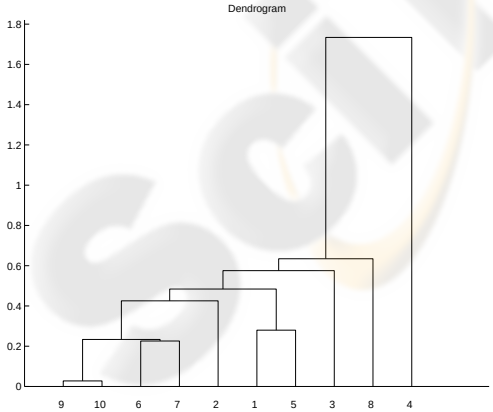

Figure 8: An example of dendrogram obtained using the cluster centers generated by the k-means.

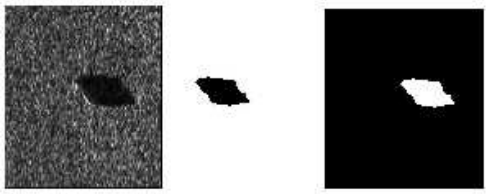

Figure 9: A real Sonar image of a sandy sea floor with the shadow of a man-made object (cylinder) (a:left image). Extracted cluster ”Shadow” (b: middle image). Complement of the cluster ”Shadow” (c: right image).

Figure 10: A real Sonar image involving an object and a rock shadows (a:left image). Extracted cluster ”Shadow” (b: middle image). Complement of the cluster ”Shadow” (c: right image).