Correntropy Based Matched Filtering for Classification in Sidescan Sonar Imagery

Erion Hasanbelliu, Jose Principe, Clint Slatton
Department of Electrical and Computer Engineering
University of Florida
Gainesville, FL

Abstract—This paper presents an automated way of classifying mines in sidescan sonar imagery. A nonlinear extension to the matched filter is introduced using a new metric called correntropy. This method features high order moments in the decision statistic showing improvements in classification especially in the presence of noise. Templates have been designed using prior knowledge about the objects in the dataset. During classification, these templates are linearly transformed to accommodate for the shape variability in the observation. The template resulting in the largest correntropy cost function is chosen as the object category. The method is tested on real sonar images producing promising results considering the low number of images required to design the templates.

Keywords—correntropy, matched filtering, classification, sidescan sonar imagery

I. INTRODUCTION

Improvements in sonar technology have allowed for large portions of the sea bottom to be scanned quickly. Nevertheless, identifying and classifying objects lying on the sea-bed is still a daunting task for a human operator especially when the amount of data to inspect is large and the possible targets that need to be identified are small in size – such as mines. Automated techniques are required to replace the human operator and thus speed up the process.

Special interest is shown in automating these techniques in the field of mine-countermeasures where the process of mine-hunting involves two major steps: first, identifying possible mine-like objects, and second, classifying them as mines or not and also the type of mine. Our work involves an additional step of fusing images coming from the same object. During the survey of the sea bottom, the sensor(s) may detect the same object multiple times due to overlapping tracks – the zig-zag nature of the surveys. Thus, over the full course, the same object may be detected many times resulting in numerous observations that need to be fused into a single target location label. Navigation knowledge, characteristic target features, and texture information are utilized to fuse the data via probabilistic graphical models. In this paper, we focus only on the classification of the objects assuming that they are already detected and provided in the form of snippets.

Various approaches have been taken to solve the classification problem. A common method is to extract a set of features representing geometrical and statistical information of the training data [1], [2], [3]. These features are then used as input to supervised classifiers such as K-nearest neighbors neural networks, but any discriminatory filters can be used to classify the objects. This approach, though, requires a large amount of training data and assumes that the testing data will have similar characteristics.

Since man-made objects have regular geometrical shapes, usually characterized by straight or uniform edges, they generate recognizable shadows. This characteristic has been used to extract and classify mines based on their shadow features as in [4], [5] or using other knowledge about the shape such as straight lines to locate possible targets [6], [7]. Other classification models match the extracted contours to templates approximating the shadows of these objects as in [8], [9]. All these methods rely on quality of the extracted features/contours which may cause problems in noisy images, complex sea-beds, and partly occluded objects. Another approach is to work directly on images, such as using associative memories to find the instances in the database that are similar to the observation [10]. This approach is very simple to implement and use, however it does require that the database be large and diverse enough to provide a good representation for the encountered data.

Our approach uses template matching to work with the extracted contours of the object highlight and shadow. We define a set of prototype templates to represent the possible mine categories found in the dataset and a set of possible transformations that they can take to accommodate for the variability in the images due to object position and orientation with respect to the sonar sensor. Our classification uses a nonlinear version of matched filter based on a new similarity measure called correntropy [11].

Correlation and MSE are second order statistics which fully quantify random variables that are Gaussian distributed. Correntropy generalizes correlation; it quantifies higher order moments of the PDF yielding solutions that are more accurate in non-Gaussian and nonlinear signal processing. This is important in our data where there is high noise level, the object contours may have partially occluded areas, and the boundaries may not have been extracted correctly due to the structure of the sea bottom or other imperfections from the sonar sensor.

This work was supported by the Office of Naval Research under Grant N00014-07-1-0698.
The idea is to treat these occlusions or imperfections as outliers. The correntropy matched filter is robust to impulsive noise [12], so there is no need to extract any landmarks or other features from the objects. Instead the algorithm uses the contour directly and tries to match it to one of the known template shapes.

The remainder of the paper is organized as follows. Section 2 describes the dataset, the image segmentation and the template generation procedures. Section 3 describes matched filters, the correntropy metric, and the template matching method. Section 4 presents current results of the work. Finally, section 5 summarizes our study.

II. IMAGE ANALYSIS AND TEMPLATE GENERATION

A. Data Description

The original sonar images have already been segmented into snippets containing an object highlight and its shadow, which arise from the presence or lack of acoustic reverberations, along with some of the background. There are 450 images belonging to one of the four classes: cube, cylinder, cone, and rectangular cuboid. The images are taken in a range of 10 to 106 meters, aspect angle of 0 to 360 degrees, and they are 8 bit grey level. Figure 1. shows an example of the snippets provided.

![Figure 1. Example of a snippet provided showing object highlight and shadow of a cube.](image1)

B. Image Segmentation

Extracting an object highlight and its shadow in sonar imagery may be difficult depending on the texture of the seabed and the resolution of the images. Various methods have been applied to accurately extract objects such as co-operating statistical snakes [13] and Markovian segmentation [14]. Since our focus is mainly on template matching and the algorithm is robust to irregularities, the segmentation process is kept simple.

First, some basic morphological algorithms such as opening and closing and region filling are applied to simplify the segmentation process. Then, region growing is used to segment the object highlight and shadow from the background as shown in Figure 2. Notice that the segmentation process does not get rid of the outliers in the highlight or the noise due to the background texture in the shadow. In the next section, we show how the system is capable of ignoring most of this noise.

![Figure 2. Example of segmenting object highlight and shadow. Object highlight is shown in red, shadow contour is shown in blue.](image2)

C. Template Generation

To generate the category templates, we used two different approaches. Since the dataset is classified and thus no specific information or templates were provided, we used images of the object taken at four strategic angles to create the contour. In sonar imagery, the object is visible only on the side that is hit by the sonar. As a result, we chose images close to the angles 45, 135, 225, and 315 degrees to provide more than one visible side. The images were scaled, rotated, and aligned to create a presentation of the object of interest. The boundary of the final object was segmented out and used as a template. Figure 3. shows the first set of templates, which were used for object highlight matching.

![Figure 3. Template contours of the four categories.](image3)

The templates shown above do not represent all the possible instances of the object highlight shape. To account for the variability of the objects based on its relative position and orientation to the sonar sensor linear transformations based on rotation and scale are introduced (the details are provided in the next section).

In addition to the object highlight, the shadow contains a lot of useful information about the object that should be utilized in shape matching. However, there is no linear transformation of template A in Figure 3. that will match the shadow contour of the same object shown in Figure 2. This is due to the shadow being created from a 3D object whereas the template is only 2D. To overcome this problem, we designed 3D models for each of the categories. The shadow contour was then generated by projecting the object onto the ground based on a point of view as depicted in Figure 4. The point of view projection follows the actual synthesis of sonar images. Sonar images are synthesized one line at a time as the sonar fish follows the path. Similarly, we use a line rather than a single point as our projection viewpoint.
III. CURRENTROPY BASED MATCHED FILTERING

A. Matched Filtering

A matched filter is an optimal linear filter for maximizing the signal to noise ratio (SNR) in the presence of added white Gaussian noise. It is obtained by correlating a known signal (in our case, a category template contour) to the unknown signal (object contour) to detect the presence of the template in the unknown signal. Consider the linear model

\[ Y = AX_m + E, \]

where \( Y \) is a 2xN matrix of the object’s contour points, \( X_m \) a 2xN matrix of contour points of the \( m \)th template, \( A \) is 2x2 transformation matrix, and \( E \) is 2xN error matrix. We consider the object of interest to simply be a transformed state of the template corrupted by some white Gaussian noise.

B. Currentropy

Currentropy is a generalized correlation measure involving high-order statistics which measures a nonlinear similarity between random variables [11]. For our purpose, currentropy is useful because it is directly related to the probability of how similar two random variables are in a neighborhood of the joint space controlled by the kernel bandwidth [15]. Controlling the kernel bandwidth allows us to adjust the ‘observation window’ in which similarity is assessed resulting in an effective mechanism to eliminate the detrimental effect of outliers; thus, making this technique robust against impulsive noise, which may occur in the form of occlusions or other imperfections on the images.

Cross currentropy between two random variables \( X \) and \( Y \) is defined as

\[ V_{\sigma}(X, Y) = E[\kappa_{\sigma}(X - Y)], \]

where \( \kappa_{\sigma}(X - Y) \) can be any non-negative definite kernel. In our case, the Gaussian kernel is used:

\[ \kappa_{\sigma}(X - Y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(X - Y)^2}{2\sigma^2}\right). \]

In practice, the joint PDF is unknown and only a finite number of data \( \{(x_i, y_i)\}_{i=1}^N \) are available, leading to the sample estimator of currentropy

\[ \hat{V}_{N,\sigma}(X, Y) = \frac{1}{N} \sum_{i=1}^N \kappa_{\sigma}(x_i - y_i). \]

C. Template Matching

We consider the object of interest as a rotated, scaled, and translated version of one of the category templates. So, through a fixed number of parameter updates (transformation and translation), given \( Y \) we can identify which \( X_m \) template is present. The parameters are updated based on maximizing the currentropy between the object of interest and the ‘transformed and translated’ template contours using as a cost function

\[ J(A, d) = \sum_{i=1}^N \left\| \frac{y_i - A X_m \cdot d}{2\sigma^2} \right\|^2, \]

where \( y_i \) and \( x_i \) are the \( i \)th column entries of \( Y \) and \( X_m \) respectively. The function is maximized with respect to the scaling and rotation matrix

\[ A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \]

and the translation vector \( d \). In order to update the matrix \( A \), we need to take into consideration the repetition of the elements \( a \) and \( b \) and differentiate with respect to the vector \([a, b]^T\) instead. Differentiating with respect to this vector and equating the gradient to zero results in the following fixed point update rule:

\[ \begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \frac{1}{N} \sum_{i=1}^N \left \langle \begin{bmatrix} x_i \cdot [y_i - A X_m \cdot d] \\ \| x_i \|^2 \end{bmatrix} \right \rangle_{2\sigma^2}, \]

with

\[ z_i = \begin{bmatrix} x_{i1}y_{i0} + x_{i2}y_{i1} \\ x_{i1}y_{i2} - x_{i2}y_{i1} \end{bmatrix}, \]

where \( x_{ik} \) and \( y_{ik} \) are the \( k \)th components of the \( x_i \) and \( y_i \) respectively. Similarly by differentiating with respect to \( d \) and equating the gradient to zero we get the update rule for \( d \):

\[ d \leftarrow \frac{1}{N} \sum_{i=1}^N \left \langle \frac{y_i - A X_m}{2\sigma^2} \right \rangle_{2\sigma^2}. \]
The object highlight depends not only on the shape of the object but also on the reflectivity properties of its material. In addition, in side-scan sonar images, at least half of the object contour is occluded. While correntropy can overcome noise due to partial occlusion, its performance deteriorates as the occluded part grows. The shadow contour, on the other hand, is less prone to imperfections. So, to have a better template matching, we need to consider separate cost functions for updating the rotation/scaling matrix and the point of view. To update the rotation/scaling matrix we need to match the shadow of the object to the simulated shadow contour from the template in addition to the highlight matching.

The shadow contour matching is more difficult than highlight matching because in order to create the correct shadow shape we need to consider 3D object templates and highlight matching because in order to create the correct contour from the template in addition to the highlight matching. The object highlight depends not only on the shape of the object but also on the reflectivity properties of its material. In side-scan sonar images, at least half of the object but also on the reflectivity properties of its material. In addition, in side-scan sonar images, at least half of the object but also on the reflectivity properties of its material. In side-scan sonar images, at least half of the object but also on the reflectivity properties of its material. In side-scan sonar images, at least half of the object but also on the reflectivity properties of its material.

The scaling/rotation matrix becomes

$$A = \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & r \end{bmatrix}$$

(11)

where we have an additional row/column counting for the third dimension. The element $r$ represents the scaling factor in $a$ and $b$. Again, differentiating with respect to the vector $[a, b]^T$ and equating the gradient to zero results in the following fixed point update rule:

$$\begin{array}{l}
\begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \frac{\sum_{i=1}^{N} z_i e^{-\frac{||\mathbf{sh}_i - (p_i, a, c_i)||^2}{2\sigma^2}}}{\sum_{i=1}^{N} p_i \sum_{i=1}^{N} z_i e^{-\frac{||\mathbf{sh}_i - (p_i, a, c_i)||^2}{2\sigma^2}}}, \\
\end{array}$$

(12)

with

$$z_i = \left[ (\mathbf{sh}_i \cdot c_i + \mathbf{sh}_i \cdot c_i - p_i \cdot c_i + r c_i \cdot (p_i \cdot c_i + p_i \cdot c_i)) - (\mathbf{sh}_i \cdot c_i - \mathbf{sh}_i \cdot c_i - p_i \cdot c_i + r c_i \cdot (p_i \cdot c_i - p_i \cdot c_i)) \right].$$

(13)

The $r$ is updated using $\sqrt{a^2 + b^2}$.

To update the point of view, the cost function then becomes

$$J(A, d) = \sum_{i=1}^{N} e^{-\frac{||\mathbf{sh}_i - (p_i, a, c_i)||^2}{2\sigma^2}},$$

(14)

where $\mathbf{sh}_i = [\mathbf{sh}_1, \mathbf{sh}_2]^T$ and $c_i = [c_1, c_2, c_3, 1]^T$ are the $i$th column entries of the shadow contour and object template respectively. Notice how the two vectors have different dimensions representing the two different spaces. In addition, the vector $c_i$ has a fourth dimension which is needed to homogenize the coordinates. The function is maximized with respect to the projection matrix

$$M = \begin{bmatrix} -p_x & 0 & 0 \\ 0 & -p_y & 0 \\ 0 & 0 & -p_z \end{bmatrix},$$

(15)

resulting in the fixed point update

$$p \leftarrow \sum_{i=1}^{N} z_i e^{-\frac{||\mathbf{sh}_i - (p_i, a, c_i)||^2}{2\sigma^2}},$$

(16)

where

$$z_i = \left[ (c_i, p_i, -\mathbf{sh}_i \cdot (p_i, c_i)) / c_i, \\
(c_i, p_i, -\mathbf{sh}_i \cdot (p_i, c_i)) / c_i, \\
(c_i, p_i, -\mathbf{sh}_i \cdot (p_i, c_i)) / c_i, \\
(c_i, p_i, -\mathbf{sh}_i \cdot (p_i, c_i)) / c_i \right].$$

(17)

IV. RESULTS

First, we compare the performance of correntropy against correlation on the object highlight with noise. Figure 5. shows the results of template matching using correlation (on the left) and correntropy (on the right). The template is shown in red, the object highlight along with noise is shown in blue circles, and the point correspondence in black. Notice that correlation attempts to accommodate all the samples whereas correntropy works on a window size and ignores points that are far away (outliers).

![Comparison of correntropy against correlation](image-url)
Next, we test the highlight and shadow template matching. The system requires no actual training for template matching. Prior knowledge about the geometrical shape of the objects is needed or some strategic images are used to design the templates. Each template is then transformed to match the extracted contour from the images. Since correntropy depends on scale, matching each template to the object contour also acts as a normalization. The template resulting in the largest correntropy value is selected as the class label.

Since the highlight shape is difficult to extract from the sidescan sonar images, we extract all the highlight pixels as shown in Figure 2. enclosed in red. These pixels are then weighed based on their intensity value during the match evaluation. The figure below shows the results of comparing an object against two templates: in (a) against the correct template and in (b) against a wrong template. The incorrect template deforms to adjust to the shadow shape but due to constraints in its shape, it can match well only on the sides resulting in poor match at the end. This brings an important point about shadow matching that the important information about the object is contained only on the end part of the shadow, the rest is just an elongation of the shadow depending on the relative height and distance of the sonar sensor from the object of interest. The poor match is noticed also on the highlight where to accommodate as much of the shape as possible the template rotates into a diagonal resulting in a suboptimal solution.

Since the highlight shape is difficult to extract from the sidescan sonar images, we extract all the highlight pixels as shown in Figure 2. enclosed in red. These pixels are then weighed based on their intensity value during the match evaluation. The figure below shows the results of comparing an object against two templates: in (a) against the correct template and in (b) against a wrong template. The incorrect template deforms to adjust to the shadow shape but due to constraints in its shape, it can match well only on the sides resulting in poor match at the end. This brings an important point about shadow matching that the important information about the object is contained only on the end part of the shadow, the rest is just an elongation of the shadow depending on the relative height and distance of the sonar sensor from the object of interest. The poor match is noticed also on the highlight where to accommodate as much of the shape as possible the template rotates into a diagonal resulting in a suboptimal solution.

Tables I and II show the classification results of the entire database for both the highlight and the shadow template matching. It can be seen that the shadow matching improves the classification due to its consistency in preserving the object shape. However, both cases suffer in distinguishing between classes B and D. This is due to the similarity of these two classes; class B is the cylinder category and class D is the rectangular cuboid category.

<table>
<thead>
<tr>
<th>TABLE I. HIGHLIGHT CONFUSION MATRIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II. SHADOW CONFUSION MATRIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
</tbody>
</table>

So far we have performed template matching separately on the highlight and the shadow. Figure 7. demonstrates the problem resulting from this approach where the template contour and its shadow are so different in size. The template shadow and highlight contour should have similar widths as well as orientation. Constraining the system to include these restrictions may improve its performance.

In addition, correntropy matched filter suffers from the point correspondence problem. The boundary points and their corresponding counterparts in the template have to be properly ordered since the kernel function is evaluated for each point in the boundary with one point in the template. In order to avoid the association problem, the template contains two to three times more points than the object contour and the Euclidean distance is used as a criterion to search for the corresponding points of each boundary point. This may lead to problems when the template drastically increases in size and all the corresponding points for the object contour come from only one region of the template and as a result the system is incapable of learning the correct transformation and thus recovering from this error.

V. CONCLUSION

In this paper, a new method to classify objects on sonar images was illustrated. The proposed method uses correntropy based matched filtering. This new metric based on kernel methods relies on an observation window to assess the similarity between the variable, which improves classification in template matching scenario where noise or occlusions occur. In addition, the matched filtering method requires no training.
except of prior information about the categories in the dataset in the form of templates. An object is classified based on the template that through some linear transformation provides the largest correntropy cost value. Besides the classification capabilities, this algorithm is useful in the full context of our project because it provides a set of features that are used for image fusion in our dynamic trees. The algorithm provides a degree of association with each template in addition to information about the rotation of the object and the possible sonar viewpoint. However, there still exist issues that need to be further investigated with the current algorithm such as constraining the updates to include information from both the highlight and shadow, as well as the point correspondence problem.

REFERENCES