SACLANTCEN REPORT
serial no: SR-307

SACLANT UNDERSEA RESEARCH CENTRE REPORT

SINGLE PING CLUTTER REDUCTION: SEGMENTATION USING MARKOV RANDOM FIELDS

R. Laterveer

March 1999

The SACLANT Undersea Research Centre provides the Supreme Allied Commander Atlantic (SACLANT) with scientific and technical assistance under the terms of its NATO charter, which entered into force on 1 February 1963. Without prejudice to this main task — and under the policy direction of SACLANT — the Centre also renders scientific and technical assistance to the individual NATO nations.

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

SACLANT Undersea Research Centre
Viale San Bartolomeo 400
19138 San Bartolomeo (SP), Italy

tel: +39-0187-5271
fax: +39-0187-524.600

e-mail: library@saclantc.nato.int

NORTH ATLANTIC TREATY ORGANIZATION

Single ping clutter reduction: segmentation using Markov random fields

R. Laterveer

The content of this document pertains to work performed under Project 041-3 of the SACLANTCEN Programme of Work. The document has been approved for release by The Director, SACLANTCEN.

Jan L. Spoelstra
Director

intentionally blank page

Single ping clutter reduction: segmentation using Markov random fields

R. Laterveer

Executive Summary:

Low frequency active sonar has been highlighted by a number of NATO nations as an important component of the next generation of undersea defence systems. The use of low frequencies in a shallow water environment, however, is known to result in a high false alarm rate due to the large number of clutter like returns. This high false alarm rate can overload automatic tracking and classification algorithms. The SACLANT Undersea Research Centre is currently investigating techniques to aid in the reduction of these false alarms without a reduction in detection probability.

This report describes an automatic method of image segmentation based on Markov random field modelling to reduce clutter. The method looks at detections over both range and bearing. It removes small objects which do not exhibit the right signature over beams. Separate detections corresponding to one large object are clumped together to form one single object. Objects too large to be a submarine can then be removed.

The Markov random field model used is based on the *a priori* physical and probabilistic knowledge of the sonar picture. It assumes that, statistically, a target has, on average, a large signal to noise ratio (SNR) and that, on a local scale, the sonar display exhibits homogeneity. The model is tuned to remove as much clutter as possible while retaining the target.

The algorithm was tested over a large number of sonar images. The performance is determined using several measures of performance including probability of detection and number of false alarm objects, i.e. connected detections. The number of false alarm objects was reduced by almost an order of magnitude while retaining more than 90% of the target detections.

With simulated sonar pictures the influence of target SNR on the segmentation is investigated. The algorithm performs well for high SNR targets but detection probability drops for low SNR targets. This is not a big problem as usually targets have sufficient SNR.

Future work includes evaluating the algorithm for other data sets and coupling it to automatic tracking and classification algorithms.

intentionally blank page

Single ping clutter reduction: segmentation using Markov random fields

R. Laterveer

Abstract:

The use of low frequency active sonar in shallow water leads to large numbers of clutter detections. This high false alarm rate can overload automatic tracking and classification algorithms. Traditional detection algorithms operate on each beam output individually searching for targets at all ranges. However, the target echo and bottom features may extend over several beams, either because a reflector is extended over space or because of the sidelobe structure of the beamformer. This suggest to associate detections over bearing, e.g. apply image processing to the range-bearing sonar data.

This report describes an automatic method of image segmentation based on a Markov random field model to reduce clutter. The segmentation is treated as a labelling problem, i.e. assign to each range-bearing cell either a target- or background label. It removes small objects which do not exhibit the right signature over beams. Separate detections corresponding to one large reflector are clumped together to form one single object. Objects too large to be a submarine can then be removed.

The Markov random field model uses *a priori* physical and probabilistic knowledge of the range-bearing sonar picture. It assumes that the background is Rayleigh distributed, a target has, on average, a large signal to noise ratio (SNR) and that the label at each range-bearing cell is influenced by its neighbouring cells. An iterative algorithm is used to produce a *maximum a posteriori* estimate for the labelling. The model parameters are tuned to remove as much clutter as possible without substantially reducing the probability of detection.

The algorithm was tested over a large number of sonar images. The performance is determined using several measures of performance including probability of detection and number of false alarm objects, i.e. connected detections. The number of false alarm objects was reduced by almost an order of magnitude while retaining more than 90% of the target detections.

To test the influence of target SNR on the segmentation, simulated sonar pictures, containing a target of known SNR, were produced. The algorithm performs well for high SNR targets but those with low SNR have lower detection probability. This is not a big problem as usually targets have sufficient SNR.

Keywords: clutter reduction ○ detection ○ classification ○ segmentation ○ low frequency active sonar ○ Markov random fields ○ image processing

Contents

1

Introduction

Low frequency active sonar has been highlighted by a number of NATO nations as an important component of the next generation of undersea defence systems. The use of low frequencies in a shallow water environment, however, is known to result in a high false alarm rate due to the large number of clutter like returns. The SACLANT Undersea Research Centre is currently investigating automatic techniques to aid in the reduction of these false alarms without a reduction in detection probability.

Detecting a target is usually not a problem, the SNR is usually sufficient. The large number of false detections makes finding the target difficult. Since present thinking is that the next stage of processing is likely to be some kind of target/non-target classification using algorithms such as tracking and fixed-feature removal the large number of false targets may make the computational load of such an approach prohibitive.

Automatic processing becomes especially important for broadband sonar . Due to the higher resolution of a broadband sonar there is much more information in the sonar picture, making it difficult or even impossible to present it all to a sonar operator. Therefore automatic processing is necessary to select the important features of the sonar picture.

This report discusses what is perceived to be a crucial first step in categorisation: the separation of regions of extended background reverberation. To this end we describe a segmentation type of algorithm which may be considered as a first stage classifier which will allow the estimation of size and position of extended objects and consequently facilitate their removal from the sonar image, and/or subsequent classification analysis.

In [1] an image segmentation algorithm based on a Markov random field (MRF) model was proposed. In this report the algorithm will be refined and its performance will be evaluated on a set of sonar pictures.

The segmentation is treated as a labelling problem, i.e. assign to each range-bearing cell either a target- or background label. The object is then, given a range-bearing sonar picture to determine the optimal labelling. The MRF model allows the optimisation to be put into a Bayes framework, the optimal labelling can be determined

by a *maximum a postiori* (MAP) estimate obtained through an iterative algorithm.

The Markov random field model uses *a priori* physical and probabilistic knowledge of the range-bearing sonar picture. The assumptions are:

- a possible target has a certain minimum size.
- the background follows a Rayleigh probability density.
- a target echo has, on average, a large signal to noise ratio (SNR).
- on a local scale, the sonar display exhibits homogeneity.

The model parameters are tuned to remove as much clutter as possible without substantially reducing the probability of detection.

In Sect. 2 the construction of the sonar image is discussed. It presents the signal processing prior to the segmentation and discusses the choices made. Section 3 gives a brief review of MRF modelling while Sect. 4 introduces the MRF model we use for active sonar. Section 5 shows the results of the algorithm applied to a data set. The model parameters are tuned for best performance. The influence of target SNR is investigated. Conclusions and prospects are presented in Sect. 6.

2

Sonar image description

The data set used in this study is known to have a lot of clutter and we have ground truth knowledge about range and bearing of up to two targets. Therefore it is highly suitable to study detection performance and false alarm reduction.

The data were collected using an active source and the Centre's low frequency towed line array as receiver. The transmission was a 2.29 s hyperbolic frequency modulated (HFM) waveform sweeping from 460 to 565 Hz. Assuming 750 m/s for the two-way sound speed this results in a range resolution of slightly less than 7.5 m.

The detection processing is illustrated in Fig. 1.

The data were beamformed using 128 hydrophones at one metre spacing using Hann shading over the array. The beams were equally spaced in cosine-space, and overlapped 3 dB down from their main response axis at 700 Hz, resulting in 86 beams spanning from forward to aft endfire.

Each beam was matched filtered and basebanded so that the centre of the waveform, 512.5 Hz, was shifted to zero Hz, and downsampled to a sampling frequency of 250 Hz so that each resolution cell contained about 2 samples.

The beams were normalised using a split window trimmed-mean normaliser [2]. The normalising window consisted of leading and lagging windows of 100 resolution cells each with a guard band of 5 resolution cells before and after the sample to be normalised. The samples in the normalising window were ordered and the lower and upper quarter were removed. The remaining samples were used to estimate the power in the auxiliary window. After normalisation, the sampling frequency was again reduced by a factor of two to 125 Hz, by taking the maximum in a window of two samples. This will change the statistics of the background somewhat but will increase the probability of detection for low SNR targets (Sect. 4.2). This procedure results in a timeseries with almost independent samples.

Other normalising methods were tried as well, including a recursive approach. After each MRF iteration the mean power is calculated of all the samples in the normalising window which are, at this iteration, classified as background. This estimate for the background power is then used in the next iteration. This approach uses the knowledge build up during the segmentation to improve the next iteration step. The

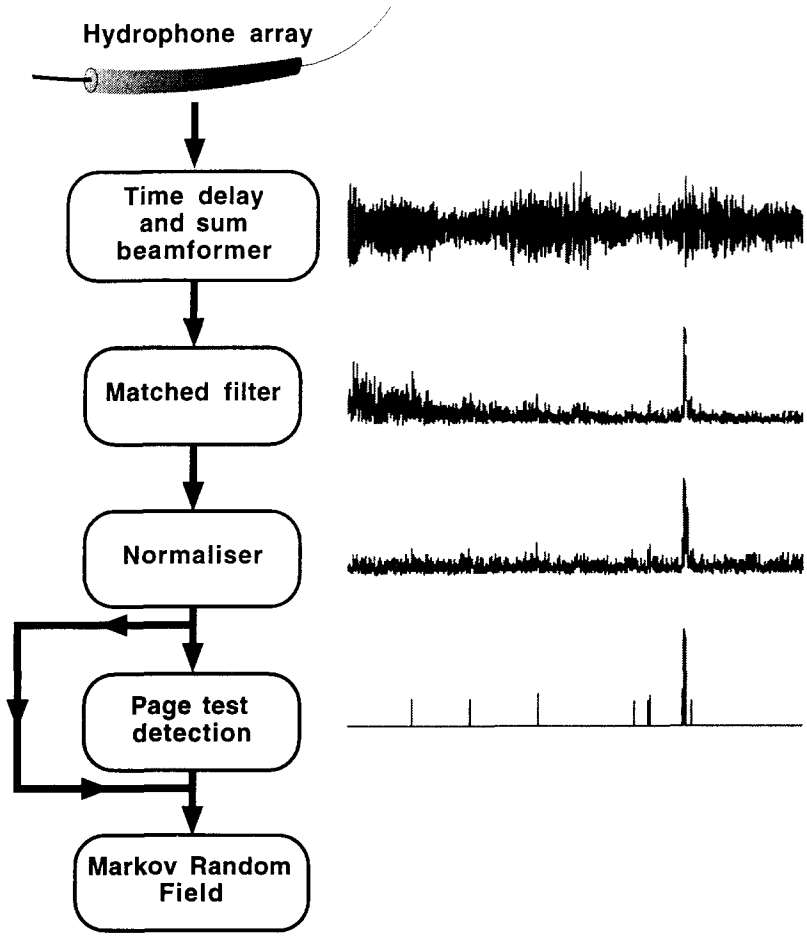


Figure 1 *Detection Processing chain.*

trimmed-mean normaliser resulted in less false alarms than the recursive approach. An explanation might be that it was designed to reduce clutter [2].

The normalised beam were put through a Page test detector [3]. This will result in a initialisation for the MRF segmentation. An initial segmentation close to the optimum one is crucial [1]. The Page test detector had a low threshold to allow for a high probability of detection. Also a simple threshold detector was tried, resulting in a higher number of false alarms. Also the value of the threshold was found to be critical, a too low value resulted in a large number of false alarms, while a high value reduced the probability of detection.

3
Markov random fields

In this section we will give a short review of the theory of Markov random fields. The problem is stated as a labelling problem, i.e. assigning a set of labels to image pixels or features. For a more detailed description see [4] and [5].

3.1 Sites and labels

The labelling problem is stated in terms of a set of sites and a set of labels.

A site can represent a pixel in an image or an image-feature such as a corner point. In the active sonar case the set of sites consists of all range/bearing cells

$$S = \{s = \{i, j\} \mid 1 \leq i \leq n_r, 1 \leq j \leq n_b\}, \tag{1}$$

where n_r is the number of ranges and n_b is the number of beams. In the general case the set of sites can be indexed by a single number, $S = \{s_1, ..., s_N\}$. A two dimensional image can be transformed into this representation by re-indexing the pixels, and then $N = n_r \times n_b$.

A label is an event that may occur at a site. Let L be a set of labels. A set of labels may be continuous or discrete. An example of a continuous label set is the dynamic range for an analogue pixel intensity. In the discrete case, a label assumes a discrete value

$$L = \{\lambda_1, \ldots, \lambda_M\}, \tag{2}$$

where M is the number of labels.

The labelling problem is to assign a label from L to each of the sites in S. In the segmentation of a sonar image, for example, the task is to assign to each range/bearing cell a label from the set $L = \{background, object\}$.

The set

$$\Lambda = \{\lambda_1, \ldots, \lambda_N\} \tag{3}$$

is called a labelling of the sites in S in terms of the labels in L, $\lambda_i \in L$. When all the sites have the same label set L, the set of all possible labellings, called the configuration space, is the following Cartesian product

$$\Omega = \underbrace{L \times L \cdots \times L}_{N\text{times}} = L^N, \tag{4}$$

where N is the number of sites.

3.2 Neighbourhood system and cliques

A neighbourhood system is used to describe the relations between sites in S. A neighbourhood system is defined as

$$V = \{v_s \mid s \in S\} \tag{5}$$

where v_s is the set of sites neighbouring site s. The neighbouring relationship has the following properties:

- a site is not a neighbour to itself: $s \notin V_s$
- the neighbouring relationship is mutual: $s \in V_r \Leftrightarrow r \in V_s$.

Some neighbourhood systems are shown in Fig. 2(a)-2(b).

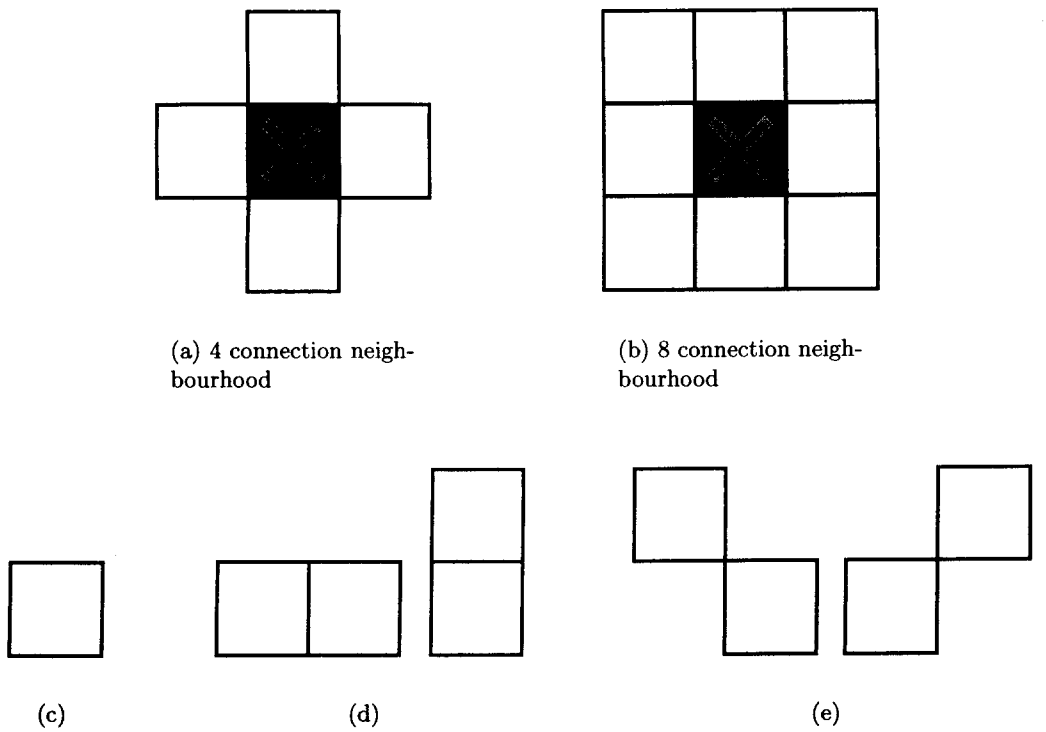
For a set of sites S with neighbourhood system V a clique is defined as a subset $C \subseteq S$ if every pair of distinct sites in C are neighbours. Also a single site $c \in S$ is a clique.

The order of a clique corresponds to the number of elements contained in the clique:

$$\begin{aligned} C_1 &= \{i \mid i \in S\} \\ C_2 &= \{\{i,j\} \mid i,j \in S \quad \text{are neighbours}\} \\ C_3 &= \{\{i,j,k\} \mid i,j,k \in S \quad \text{are neighbours to one other}\} \end{aligned}$$

Fig. 2(c)-2(e) show first and second order cliques for the neighbourhood systems in Fig. 2(a)-2(b). The 4 connection neighbourhood 2(a) has only second order cliques 2(d). The other two neighbourhood systems also have diagonal second order cliques 2(e).

As the order of the neighbourhood system increases, the number of cliques grow rapidly and so do the computational expenses. We will limit ourselves to first and second order cliques.



(a) 4 connection neighbourhood

(b) 8 connection neighbourhood

(c) (d) (e)

Figure 2 *Neighbourhoods and cliques*

3.3 Markov random fields

Let $\Lambda = \Lambda_1, \ldots, \Lambda_N$ be a family of random variables defined on the set of sites S, where each variable Λ_i takes a value λ_i in the set of labels L. Λ is called a random field. A realisation $\lambda = \lambda_1, \ldots, \lambda_N \in \Omega$ is called a configuration of Λ. For a discrete label set L, the probability that random variable X_i takes value λ_i is denoted $P(\Lambda_i = \lambda_i)$, and the joint probability is denoted $P(\Lambda = \lambda) = P(\Lambda_1 = \lambda_1, \ldots, \Lambda_N = \lambda_N)$.

The random field Λ is called a Markov random field on S with respect to a neighbourhood system V if and only if:

$$\begin{array}{ll} P(\Lambda = \lambda) > 0, \forall \lambda \in \Omega & \text{(positivity)} \\ P(\lambda_i | \lambda_{S \setminus \{i\}}) = P(\Lambda = \lambda_i | \lambda_{v_i}) & \text{(Markovianity),} \end{array} \tag{6}$$

where $\lambda_{S \setminus \{i\}}$ is set of labels at the sites in S except i, and $\lambda_{v_i} = \{\lambda_j | j \in v_i\}$ are the labels of the neighbours of i.

The positivity assumption can usually be satisfied in practice. The Markovianity describes the local characteristics of Λ, the label at a site is dependent only on those

at the neighbouring sites. It is always possible to select a neighbourhood such that the Markovian property holds, the entire set of sites. Any field Λ is an MRF with respect to such a neighbourhood system.

3.4 Gibbs Random Fields

A set of random variables W is said to be a Gibbs random field (GRF) on S with respect to V if and only if its configurations follow a Gibbs distribution:

$$P(w) = \frac{1}{Z} e^{-U(w)/T}, \tag{7}$$

where $Z = \sum_{w \in \Omega} \left(\exp -U(w)/T \right)$ is called the partition function, T is constant called the temperature, and $U(w)$ is the energy function. The energy function is a sum of clique potentials over all possible cliques C

$$U(w) = \sum_{c \in C} V_c(w). \tag{8}$$

3.5 Markov-Gibbs equivalence

An MRF is characterised by its local properties (the Markovianity) whereas a GRF is characterised by its global properties (the Gibbs distribution). The Hammersley-Clifford theorem establishes the equivalence of these two types of properties. The theorem states that W is a MRF on S with respect to neighbourhood V if and only if W is a GRF. A proof can be found in e.g. [6].

The Markov-Gibbs equivalence makes it possible to determine the optimal labelling for a set of sites, a global property, by only considering neighbour interactions, a local property.

3.6 Bayes labelling

The Bayes rule states that

$$P(\Lambda = \lambda | Z = z) = \frac{P(Z = z | \Lambda = \lambda) P(\Lambda = \lambda)}{P(Z = z)} \tag{9}$$

where $P(\Lambda = \lambda | Z = z)$ is the *a posteriori* probability of labelling λ given sonar image z, $P(Z = z | \Lambda = \lambda)$ is the conditional probability, $P(\Lambda = \lambda)$ is the *a priory*

probability and $P(Z = z)$ is a constant for a given sonar image $Z = z$. Using the definition of the MRF we have

$$P(\Lambda_i = \lambda_i | \lambda_{S\setminus\{i\}}) = P(\Lambda_i = \lambda_i | \lambda_{v_i}). \tag{10}$$

The conditional probability $P(Z = z|\Lambda = \lambda)$ models the probability of the sonar image given a certain labelling. The *a priory* probability $P(\Lambda = \lambda)$ allows us to model the physical knowledge of the sonar image, i.e. the mapping of the sonar image onto physical space. This includes the relation of labels of neighbouring pixels.

The goal is, given a sonar image $z = \{z_1, \ldots, z_N\}$, to determine the optimum labelling of the range/bearing cells. The maximum *a posteriori* (MAP) estimate is the estimate that maximises the *a posteriori* probability. Using the Bayes rule this leads to

$$\begin{aligned} \hat{\lambda}_{\text{opt}} &= \arg\max_{\hat{\lambda}\in\Omega} P(\Lambda = \hat{\lambda}|Z = z) \\ &= \arg\max_{\lambda\in\Omega} P(Z = z|\Lambda = \lambda)P(\Lambda = \lambda) \end{aligned} \tag{11}$$

Because of the Markov-Gibbs equivalence the probabilities can be modelled by a Gibbs distribution. The maximisation of the product of the two probabilities in (11) is then equivalent to the minimisation of sum of two energy terms, the conditional energy $U(Z = z|\Lambda = \lambda)$, and the prior energy $U(\Lambda = \lambda|\lambda_{v_i})$.

$$\hat{\lambda}_{\text{opt}} = \arg\min_{\hat{\lambda}\in\Omega} \big(U(Z = z|\Lambda = \lambda) + U(\Lambda = \lambda)\big). \tag{12}$$

3.7 Optimisation algorithm

For small images it would be feasible to calculate the energy (12) for each possible configuration and choosing the one with minimal energy. For realistic images the number of configurations makes it impractical to do an exhaustive search over all possible configurations. Therefore we have to use an optimisation algorithm. These algorithms can be divided into two classes, global methods and local methods.

Many combinatorial minimisation problems like the discrete labelling problem considered here are frustrated systems, i.e. the energy surface has many local minima with almost equal energy. This makes finding the true global minimum hard, but a close local minimum might be good enough.

Global optimisation methods can escape from local minima by sometimes allowing an increase in energy. These global algorithms include genetic algorithms [7] and

simulated annealing [8]. A disadvantage of these methods is that they are computational intensive.

We will use a local method called iterated conditional modes (ICM) [9]. The main advantages of this method is that it is guaranteed to converge and its speed. Being a local method it can become trapped in a local minimum. This should not be a problem as long as we can start at a good initial point fairly close to the global minimum. In our case such an initialisation is available. The algorithm consists of the following steps:

- Determine an initial configuration $\lambda^{(0)}$ close to the optimal configuration.
- Scan the set of sites S. Each site s, at iteration k, is characterised by the pixel value z_s, its label $\lambda_s^{(k)}$ and the labels $\lambda_{v_s}^{(k)}$ of its neighbours. For each possible label $\lambda \in L$ determine the conditional energy $U(\lambda|z_s)$ and choose the one with the smallest energy.
- Continue scanning until a stop criterium is reached.

If the energy (12) contains some adjustable parameters it is possible to change them at each iteration, thereby allowing the modelling to change as the solutions is moving towards the energy minimum.

Another possibility is to estimate model parameters during ICM [9]. For example a parameter ϕ in the *a priory* probability can be estimated by the maximum likelihood method, i.e., after each iteration use the current labelling λ_i to obtain a new estimate $\phi = \hat{\phi}$ which maximises $\prod_{i \in S} P(\lambda_i; \phi)$. An example is given in Sec. 4.3.1.

4

Markov random fields for Active Sonar

4.1 Introduction

For the MAP optimisation of the MRF two probabilities must be defined, the conditional probability $P(Z = z|\Lambda = \lambda)$ and the *a priori* probability $P(\Lambda = \lambda)$. The different terms in the energy will compete to determine the new label at the site.

Consider for example the schematic example in Fig. 3. We are trying to determine the new label for the central pixel. It has three neighbours with an object label and 11 neighbours with a background label. Therefore it is likely to become a background pixel, unless its intensity is too high, in which case it will remain an object pixel.

4.2 Likelihood energy

The likelihood energy $U(Z = z|\Lambda = \lambda)$ corresponds to the conditional probability $P(Z = z|\Lambda = \lambda)$. It allows us to model the influence of the measured pixel value on the pixel label. We assume that background and object pixels have well separated distributions. Then the conditional probability can be implemented by a simple threshold crossing [1]. A pixel value below the threshold, z_t, will have a fixed probability p that it is a background pixel, while a pixel value above the threshold will have fixed probability q that it is an object pixel

$$\text{If } z < z_t \text{ then} \qquad P(z|0) = p \tag{13}$$

$$P(z|1) = 1 - p \tag{14}$$

$$\text{If } z \geq z_t \text{ then} \qquad P(z|0) = 1 - q \tag{15}$$

$$P(z|1) = q. \tag{16}$$

If we ignore constant terms the likelihood energy is

$$\text{If } z < z_t \text{ then} \qquad U(z|0) = -\ln(p) \tag{17}$$

$$U(z|1) = -\ln(1 - p) \tag{18}$$

$$\text{If } z \geq z_t \text{ then} \qquad U(z|0) = -\ln(1 - q) \tag{19}$$

$$U(z|1) = -\ln(q). \tag{20}$$

Note that it would be more correct to use $P(z|1) = 1 - q$ and $P(z|0) = 1 - p$ (see (21) and (22)). Since the final results are not sensitive to either choice we decided to keep the conditional probability from [1].

Figure 4 shows the histograms of background pixels and Page detections. The background pixels are all those not detected by the Page test. In Fig. 4(a) the timeseries is sampled at 250 Hz, while in Fig. 4(b) the sampling frequency was reduced by a factor of two to 125 Hz, by taking the maximum in a window of two samples, resulting in a timeseries with almost independent samples. The solid lines shows a Rayleigh fit through the background histogram and a noncentral Rayleigh pdf with 8.5 dB SNR.

In Fig. 4(a) he Rayleigh distribution fits the background well. This is confirmed by both the χ^2-test and the Kolmogorov-Smirnov test[10]. The thinned out data (Fig. 4(b)) does not fit the Rayleigh distribution as well, so taking the maximum of every two samples changes the statistics of the timeseries. This is still preferred over taking every second sample which will decrease the probability of detection for low SNR targets.

The background and object distributions are well separated, so the likelihood energy (13-16) makes sense. From Fig. 4 we see that p and q can be determined from the cumulative probability densities (CDFs):

$$p = \frac{\int_0^{z_t} f_0(z)dz}{\int_0^{z_t} \big(f_0(z) + f_1(z)\big)dz} \tag{21}$$

$$q = \frac{\int_{z_t}^{\infty} f_1(z)dz}{\int_{z_t}^{\infty} \big(f_0(z) + f_1(z)\big)dz}, \tag{22}$$

where $f_0(z)$ is the background pdf and $f_1(z)$ is the object pdf.

We model the object pdf with a noncentral Rayleigh distribution. Note that this is just an approximation, the object pixels correspond to different returns with different SNR. Even so the approximation is appropriate. A reasonable fit is a noncentral Rayleigh pdf with 8.5 dB SNR, a good choice for the threshold is $z_t = 2$. With these distributions we get $p = 0.8$ and $q = 0.97$. This is confirmed by the cumulative histograms of the data.

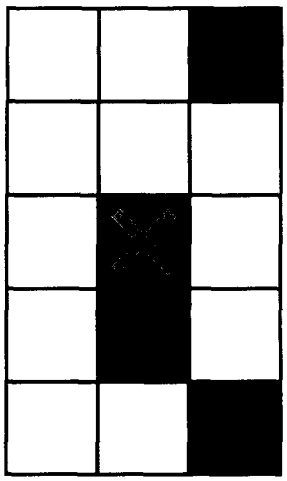
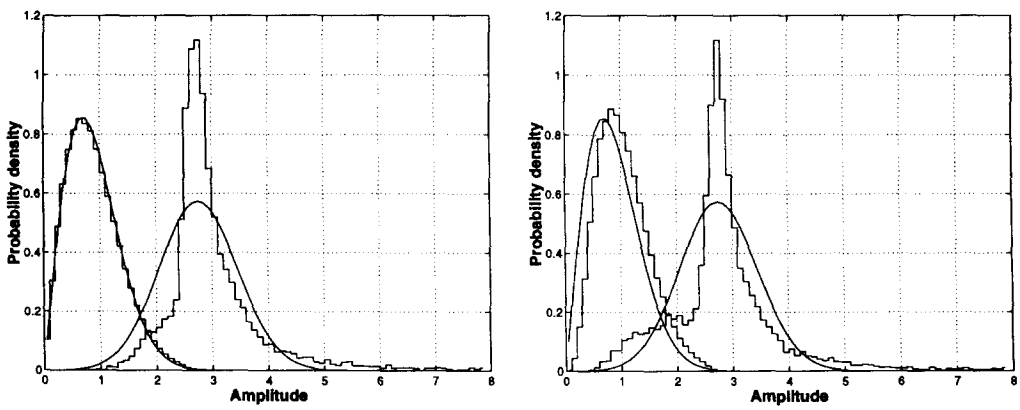


Figure 3 *The influence of neighbours*



(a) Sampled at 250 Hz

(b) Sampled at 125 Hz

Figure 4 *Histograms of background and Page detections, (a) sampled at 250 Hz, (b) sampled at 125 Hz, by taking the maximum over two samples. The solid lines show a Rayleigh pdf and a noncentral Rayleigh pdf with 8.5 dB SNR.*

Figure 5 *Flipping of pixel from iteration to iteration*

4.3 Prior energy

The prior energy $U(\Lambda = \lambda)$ allows the modelling of the physical knowledge of the sonar image. It includes two terms:

$$U(\Lambda = \lambda) = -\alpha \ln(c_s) - \beta \sum_{j \in v_s} c_{t,j} \delta(\lambda_j, \lambda), \tag{23}$$

where λ is 0 for background pixel or 1 for an object pixel, α and β are homogenising constants, c_s is a weight of the first order clique, $c_{t,j}$ are weights of the second order cliques and $\delta(a, b) = 1$ if $a = b$ and 0 otherwise.

The homogenising constants are increased every iteration. This allows acceleration of regularisation without removing small objects too quickly, slowly moving from a conditional probability dominant model to an *a priori* dominant model [9]. We choose $\alpha = k\beta$, where k is a constant, and let β increase every iteration from 0.5 in steps of 0.1 to 1. The optimal value for k is determined in Sec. 5.

The neighbourhood chosen is a window of five range cells by three beams. The five range cells span a length of around 30 m, which is about half the length of a typical target. The size of neighbourhood window determines the appropriate range of values for α and β.

4.3.1 Single order clique

The first order clique is implemented as a Markov chain from iteration to iteration (Fig. 5) [1]. This allows small objects surrounded by background pixels to behave differently from small clusters of background pixels surrounded by object pixels. The idea is that targets can correspond to a small number of range/bearing cells, so we don't want small objects to disappear too easily. On the other hand we want small holes in an object to be filled quite rapidly.

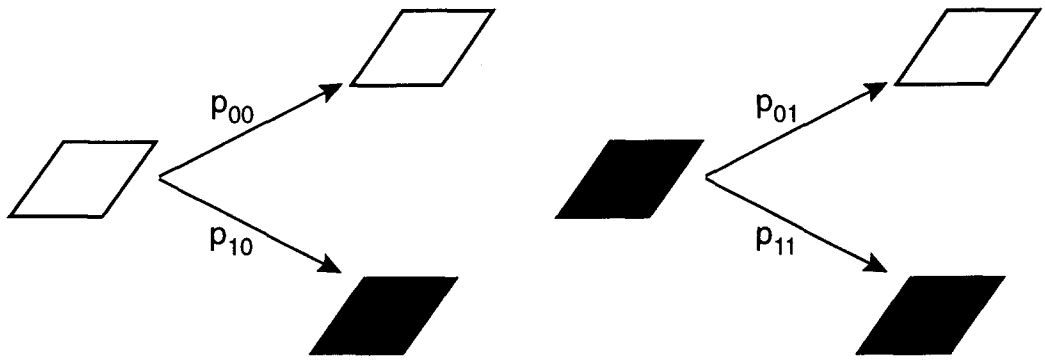


Figure 6 *Transition probabilities*

Going from iteration n to $n+1$ we can define transition probabilities (Fig. 6)

$$P(\lambda_s^{(n+1)} = 0 | \lambda_s^{(n)} = 0) = p_{00} \tag{24}$$
$$P(\lambda_s^{(n+1)} = 1 | \lambda_s^{(n)} = 0) = p_{10} = 1 - p_{00} \tag{25}$$
$$P(\lambda_s^{(n+1)} = 0 | \lambda_s^{(n)} = 1) = p_{01} = 1 - p_{11} \tag{26}$$
$$P(\lambda_s^{(n+1)} = 1 | \lambda_s^{(n)} = 1) = p_{11}. \tag{27}$$

The value of p_{00} should not be too big if we want fragmented objects to merge into one object. The size of the smallest detectable object is determined by p_{11}, it should be close to one. The weight coefficient for the first order clique is now given by

$$\begin{aligned} c_s &= \big(p_{00}\delta(\lambda_s,0) + p_{01}\delta(\lambda_s,1)\big)\delta(\lambda,0) + \\ &\quad \big(p_{10}\delta(\lambda_s,0) + p_{11}\delta(\lambda_s,1)\big)\delta(\lambda,1) \\ &= \big(p_{00}\delta(\lambda_s,0) + (1-p_{00})\delta(\lambda_s,1)\big)\delta(\lambda,0) + \\ &\quad \big((1-p_{11})\delta(\lambda_s,0) + p_{11}\delta(\lambda_s,1)\big)\delta(\lambda,1) \end{aligned} \tag{28}$$

where λ_s is the label of site s at iteration n, and λ is the argument of (23).

The performance of the algorithm is rather sensitive to the value of p_{11}. The optimum value might also depend on the data set used. A solution is to estimate it during the optimisation [9].

After each iteration we determine the maximum likelihood estimate $\hat{p}_{11}$ which maximises the joint probability $\prod_{i \in S} P(\lambda_i; p_{11})$

$$\hat{p}_{11} = \frac{n_{11}}{n_{11} + n_{01}}, \tag{29}$$

where n_{ij} the number of transitions from label j to label i. This estimate is then used for the following iteration. As the segmentation converges $n_{01} \to 0$ so that $p_{11} \to 1$. Because we do not want p_{11} to become too small in the early iterations the estimate is used only if it increases p_{11}.

A similar maximum likelihood estimate can be formed for p_{00}. For this estimate also $p_{00} \to 1$ as the segmentation converges. Because we want to keep p_{00} small, the estimate is not used and we keep it fixed at $p_{00} = 0.5$.

4.3.2 Second order cliques

The second order cliques allows the neighbourhood to influence the new label at a site. The sonar image contains two kinds of neighbours, cells in the same beam as the site we are considering and cells in different beams. With increasing range intra-beam neighbours remain a constant physical distance to the site while the centers inter-beam neighbours will increase the separation in physical space. Therefore in [1, 11] it was proposed to introduce a range and beam dependent weight

$$c_t = 1 \qquad \text{for intra-beam links} \tag{30a}$$

$$c_t = ((R/R_0)\sin(\theta/2))^{-\Gamma} \qquad \text{for inter-beam links} \tag{30b}$$

where θ is the angle between the beams in which the pixels of interest lie, R is the range of the central pixel and R_0 is a given minimum radius. The weight from [11] extends the work in [1] where $\Gamma = 0.5$.

To evaluate the notion of different weights for intra-beam and inter-beam neighbours, we analyse the probability for an object pixel to remain an object pixel. Let the neighbourhood be N range cells by M beams, and consist of $r \leq N-1$ intra-beam pixels and $s \leq N(M-1)$ inter-beam pixels with an object label. If the pixel value is above the threshold and, at iteration n, has an object label then the posterior energies are

$$U_{01} = -\ln(1-q) - \alpha\ln(1-p_{11}) - (N-r-1)\beta - ((N(M-1)-s)\beta c_t \tag{31}$$

$$U_{11} = -\ln q - \alpha\ln p_{11} - r\beta - s\beta c_t, \tag{32}$$

where c_t is according to (30b).

The new label remain an object label if $U_{11} < U_{01}$ or

$$\ln\frac{q}{1-q} + \alpha\ln\frac{p_{11}}{1-p_{11}} - \beta(N-2r-1+(N(M-1)-2s)c_t) > 0. \tag{33}$$

These equations allow us to determine, for fixed p_{11}, the range of values of α and β for which the new label will stay an object label. Because $0.5 < q, p_{11} < 1$ there is no restriction on α and β if

$$N-2r-1+(N(M-1)-2s)c_t \leq 0. \tag{34}$$

This means that, independent of the parameter values, the object will be stable.

The probability for a background pixel flipping to an object pixel gives the same result if $p_{00} \leq 0.5$.

As an example consider an 5×3 neighbourhood ($N = 5, M = 3$). If $c_t = 1$, i.e. all cells have the same weight, then $r + s \geq 7$, so an 8 pixel object is stable. The stability of smaller objects depends on the parameters. If $c_t = 0.5$, then $2r + s \geq 9$, so for example a 6 pixel object consisting of 5 intra-beam pixels and 1 inter-beam pixel is stable. For $c_t = 0.25$, then $4r + s \geq 13$, so an 5 pixel object consisting of 4 intra-beam pixels and 1 inter-beam pixel is stable. Reducing the weight to $c_t = 0$, i.e. the neighbourhood consists of 5 range cells, gives $r \geq 2$, so a 3 range pixel object is stable.

We see that reducing c_t, and thus the weight of inter-beam pixels, makes the size of unconditionally stable objects smaller. Therefore at increasing range, and thus decreasing c_t, the number of remaining small objects will increase. This is not desirable because we want the stability of objects to be controlled by p_{11}, α, and β. We conclude that it is better to let all the neighbours have the same weight, i.e. $\Gamma = 0$ in (30b) so that $c_t = 1$ for all neighbour cells.

Now choose $\alpha = k\beta$ with k fixed over all iterations, $c_t = 1$ and let there be r object neighbours, then the object will be stable if

$$\log \frac{q}{1-q} + \left(k \log \frac{p_{11}}{1-p_{11}} - (NM - 1 - 2r)\right)\beta > 0 \tag{35}$$

or

$$r > \frac{1}{2}(NM - 1 - k \log \frac{p_{11}}{1-p_{11}}) - \log \frac{q}{1-q}/\beta. \tag{36}$$

From (36) we see that the smallest size of objects increases with increasing β, increasing k and decreasing p_{11}. Because β is increased at every iteration ($0.5 \leq \beta \leq 1$) at the beginning of the segmentation small objects remain, while later on small objects will be removed. This allows a big object fragmented in several small pieces to merge into one big object and small isolated objects to be removed. The value of k and p_{11} allows us to tune the model to get a big reduction of the false alarm objects while retaining the target.

5
Results

5.1 Parameter optimisation

The algorithm was applied to a data set of 200 pings, each with a pulse repetition interval of 60 seconds. They contain areas of extended reverberation and clutter as well as areas of clear display, and thus highly suited for this study. Each ping has previously been investigated in order to locate significant targets of opportunity, 175 pings contain 1 or 2 such targets. This ground knowledge is used to test the performance of the algorithm. The data were processed using the detection processing method described in Sec. 2. The Page test detected all target opportunities, i.e. it had 100% probability of detection.

There are several ways to quantify how well the algorithm performs. We will use the following measures of performance (MOPs): probability of detection, probability of false alarm, number of false alarm objects and computation time.

In an operational situation it is important that the signal processing can be done in real time, the processing has to be completed before the next ping of data arrives. For the present data set this time is 60 seconds. The MRF algorithm is computationally intensive, it consists of relatively simple operations but have to be repeatedly performed over the entire sonar picture, 5001 range cells by 86 beams in our case. Even so, due to an efficient implementation, the algorithm runs in real time.

The probability of detection is calculated by counting the number of target detections which remain after the MRF segmentation and dividing by the total number of detection opportunities. The probability of false alarm can be estimated by counting the range/bearing cells which do not correspond to a known target and dividing by the total number of range/bearing cells.

An object is a set of detections over range/bearing cells which are connected with an 8-connection window[12]. The number of false alarm objects are all the objects which do not correspond to a known target. Note that this is different from the number of false alarms. The notion of false alarm objects is more appropriate in classification because a classifier can consider a set of connected detections as one object entity. For example a tracker does not need to track all individual cells in a detected object, but can represent it by the cell with the largest pixel value.

The normalised data and the Page test detections were used as inputs to the MRF algorithm. Two types of MRF models were tried, one with p_{11} fixed over all iterations, and one with p_{11} adjusted according to the maximum likelihood estimate (Sect. 4.3.1). Both models had β starting from 0.5, increasing 0.1 at each iteration, up to 1.0, $\alpha = k\beta$, with $k = 1, 2, 3$. The starting value of p_{11} was from 0.8 up to 0.975. By running the MRF models for all these parameter values we can determine the optimum set of parameter values which gives the maximum reduction in false alarm objects while still retaining a high probability of detection.

Figure 7 show the results for p_{11} fixed. The graph shows the probability of detection against the number of false alarm objects. Note that this is not a ROC curve, as the horizontal axis shows number of false alarm objects as apposed to false alarm probability. The big circle shows the Page detections, with 100% probability of detection. The top axis shows the reduction of the number of false alarm objects compared to the Page test. The green triangles are for $k = \alpha/\beta = 1$, the red diamonds are for $k = 2$, while the blue squares are for $k = 3$. The different points on the lines are for different values of p_{11}. Table 2 shows the numbers, the columns show, respectively, the value of p_{11}, the value of $k = \alpha/\beta$, the number of targets detected, the probability of detection, the number of false alarm objects and the reduction of false alarm objects compared to the Page test.

Figure 8 and Table 3 show the results for adaptive p_{11}. Table 1 shows typical values of p_{11} at each iteration. At the first few iterations p_{11} is fixed at the starting value while at the later iterations it quickly goes toward 1.

iteration	p_{11}
1	0.900
2	0.900
3	0.952
4	0.972
5	0.990
6	0.997

Table 1

From the results we see that, as expected, both the probability of detection and the remaining number of false alarm objects increases for increasing k and p_{11}. For high k and p_{11} all the target opportunities are detected, resulting in a reduction in false alarm objects without any reduction in probability of detection.

When p_{11} is updated according to the maximum likelihood estimate there is little dependence of the results on the initial value of p_{11}, especially for $k = 1, 2$, making the algorithm more robust. It does result in a slightly lower reduction of false alarm objects.

The optimum set of parameter values is $k = 1$ and p_{11} starting at 0.9 and updated according to the maximum likelihood estimate.

Figures 9 and 10 show a typical example of the algorithm's results. Figure 9(a) shows part of the sonar returns. Strong echoes from large objects are visible. The sonar image has a spiky nature, resulting in a large number of clutter detections.

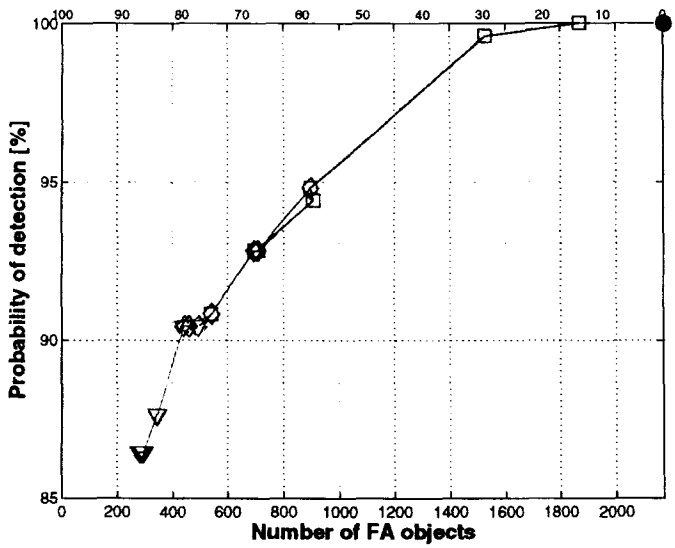


Figure 7 *Results for p_{11} fixed for all iterations. The big circle shows the Page detections, e.g. 100% probability of detections. The top axis shows the reduction of the number of false alarm objects compared to the Page test. The green triangles are for $k = \alpha/\beta = 1$, the red diamonds are for $k = 2$, while the blue squares are for $k = 3$. The different points on the lines are for different values of p_{11}.*

p_{11}	$k = \alpha/\beta$	detections	p_d	FA objects	reduction
0.800	1	217	86%	275	87%
0.825	1	217	86%	284	87%
0.850	1	217	86%	293	86%
0.875	1	217	86%	293	86%
0.900	1	220	88%	338	84%
0.925	1	220	88%	342	84%
0.950	1	227	90%	437	80%
0.975	1	227	90%	490	77%
0.800	2	227	90%	437	80%
0.825	2	227	90%	453	79%
0.850	2	227	90%	455	79%
0.875	2	227	90%	489	77%
0.900	2	228	91%	537	75%
0.925	2	233	93%	692	68%
0.950	2	233	93%	706	67%
0.975	2	238	95%	894	59%
0.800	3	228	91%	537	75%
0.825	3	233	93%	694	68%
0.850	3	233	93%	695	68%
0.875	3	233	93%	706	67%
0.900	3	237	94%	904	58%
0.925	3	238	95%	895	59%
0.950	3	250	100%	1525	30%
0.975	3	251	100%	1867	14%

Table 2 *Results for p_{11} fixed for all iterations. The columns show the value of p_{11}, the value of $k = \alpha/\beta$, the number of targets detected, the probability of detection, the number of false alarm objects and the reduction of false alarm objects compared to the Page test.*

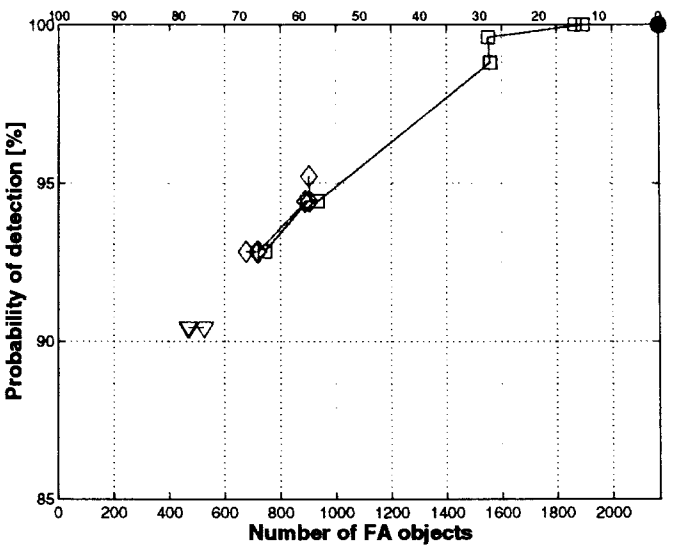


Figure 8 *Results for p_{11} increasing according to the maximum likelihood estimate. The symbols are as in Fig. 7*

Table 3 *Results for p_{11} increasing according to the maximum likelihood estimate. The first column shows the starting value of p_{11}, the other columns are as in Table 2*

p_{11}	$k = \alpha/\beta$	detections	p_d	FA objects	reduction
0.800	1	227	90%	468	78%
0.825	1	227	90%	468	78%
0.850	1	227	90%	468	78%
0.875	1	227	90%	468	78%
0.900	1	227	90%	468	78%
0.925	1	227	90%	474	78%
0.950	1	227	90%	474	78%
0.975	1	227	90%	527	76%
0.800	2	233	93%	682	69%
0.825	2	233	93%	722	67%
0.850	2	233	93%	726	67%
0.875	2	233	93%	726	67%
0.900	2	237	94%	890	59%
0.925	2	237	94%	888	59%
0.950	2	237	94%	904	58%
0.975	2	239	95%	902	58%
0.800	3	233	93%	746	66%
0.825	3	237	94%	898	59%
0.850	3	237	94%	935	57%
0.875	3	237	94%	934	57%
0.900	3	248	99%	1558	28%
0.925	3	250	100%	1550	29%
0.950	3	251	100%	1896	13%
0.975	3	251	100%	1867	14%

Figure 9(b) shows the detections, the red object in the centre is the detection of a surface ship. Some detections from large objects are present but are fragmented into several separate pieces. The detections include a lot of small clutter objects. In Fig. 9(c) the result of the MRF segmentation is shown. Most of the small clutter has been removed, separate fragments of large objects have been segmented into larger objects, the target is still detected. Over the entire sonar picture the number of false alarm objects is reduced by 85%.

The convergence of the algorithm is illustrated in Fig. 10, the circles show the number of objects left after each iteration, the triangles show the number of flipped pixels divided by the number of objects. The number of objects reaches a stable value and the number of flips becomes small compared to the number of objects, thus showing convergence.

5.2 SNR simulation

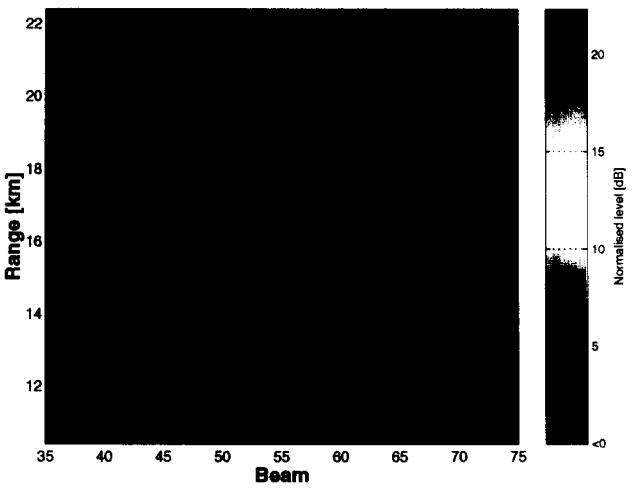
In this section we will investigate the influence of target SNR on the probability of detection after MRF segmentation. Curves are generated of mean number of false alarm objects against probability of detection, one curve for each target SNR, the points on the curve correspond to different values of the Page test detector threshold. Again note that these are not ROC curves.

We generate simulated sonar pictures based on real world data. An artificial target, with known SNR, is inserted into a real sonar picture at random range and bearing. The artificial target is taken from a ping with a high SNR target detection and scaled to obtain the desired SNR. By applying this procedure to all pings we obtain a set of sonar pictures with known targets under different types of background.

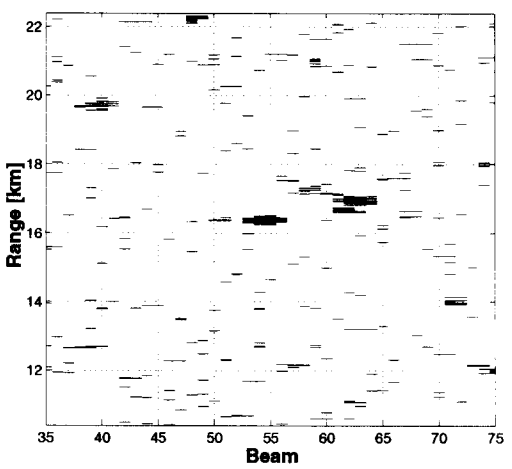
For the SNR estimation we use the modified 'double boxcar' averaging technique described in [13]. The average power is calculated in 1 s windows 1 s before and 1 s after the echo, the noise level (NL) is the minimum of the two values. The echo level (EL) is the peak value of the echo, the SNR then becomes $\mathrm{SNR} = \mathrm{EL} - \min(\mathrm{NL}_1, \mathrm{NL}_2)$. Note that this notion of SNR from real data should not be confused with the theoretical non-centrality parameter used in Fig. 4.

The MRF algorithm is then applied to the simulated sonar pictures, again with $k = 1$ and p_{11} starting at 0.9. The probability of detection for the given SNR is estimated by the number detections divided by the number of simulated sonar pictures.

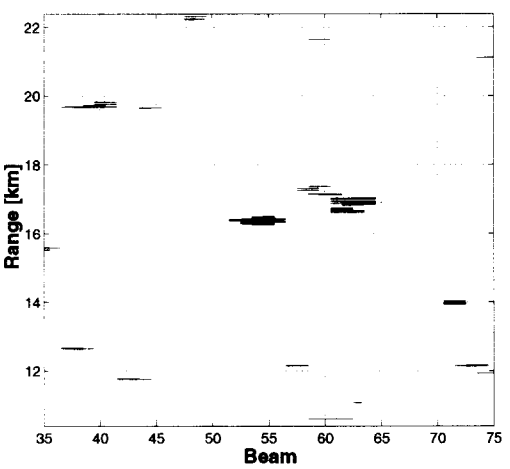
Figure 11 shows the results of the Page test detector, Fig 12 shows the results of the MRF segmentation. Targets stronger than about 14 dB have more than 90% probability of detection for the Page test detector with a low threshold. After MRF segmentation the probability of detection drops for targets weaker than 18 dB. Note



(a) Normalised MF output.



(b) Page test detections.



(c) MRF segmentation.

Figure 9 *Results for one ping.*

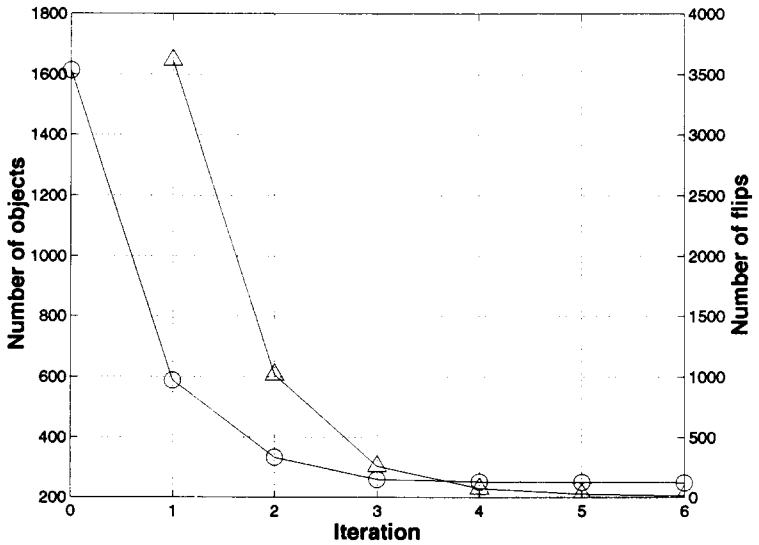


Figure 10 *Convergence of the algorithm, the circles show the number of objects left after each iteration, while the triangles show the number of flipped pixels divided by the number of objects.*

that the mean reduction in false alarm objects is almost 80%.

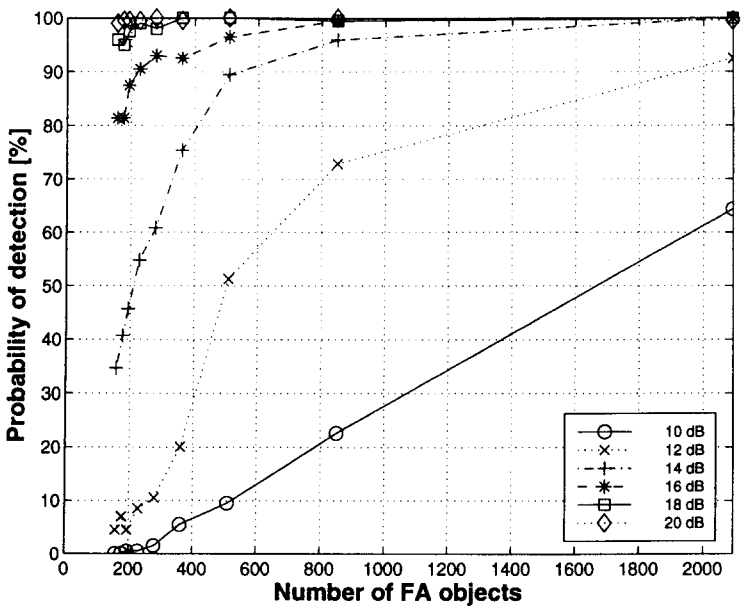


Figure 11 *SNR sensitivity study, Page test results.*

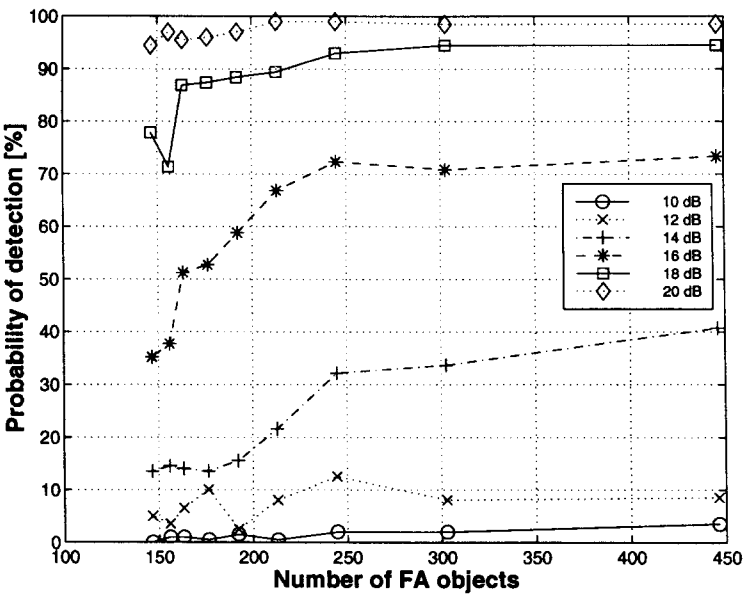


Figure 12 *SNR sensitivity study.*

6

Conclusions

This study demonstrates that image processing can reduce the number of false contacts considerably without substantial reduction in detection probability. The image processing technique is based on a Markov random field model which allows the physical and statistical knowledge of the sonar picture to be used in the modelling. The algorithm converges quickly and can be executed in real time or near real time. The model parameters were tuned for best performance.

The algorithm performs well, the number of false contacts reduces by about 80% while the probability of detection remains above 90%. The segmentation removes some of the weaker targets. This should not be a big problem as the target echo has typically a high SNR. Even if the target echo is weak it might still be detected over several pings.

Future work includes testing the algorithm on other data sets. Some of the parameters of the MRF model depend on the data used. Even so the segmentation results are rather robust against small variations of these parameters so their precise value is not important. Other possible enhancements include coupling the algorithm to tracking and classification algorithms. Target tracking can be implemented by extending the MRF model to multiple pings, introducing a neighbourhood over three dimensions, range, bearing and ping. The three dimensional objects obtained can then be classified according to their slope. Another possibility is a tracking algorithm not based on MRF but which uses the baricentres of the segmented objects, dramatically reducing the work compared with an algorithm which has to track every range-bearing cell.

The segmented objects can also be used to extract classification clues from a single ping sonar image, without looking of the ping history. Classification clues include echo shape and object size. The object size is an important clue, because the classification clusters together detections from the same physical object, size can be used to remove reverberation features which are to big to be from a target.

References

[1] S. Dugelay and D.A. Abraham. Reduction of low frequency active sonar clutter through image processing. Technical Report SR-272, SACLANTCEN, 1998.

[2] P.P. Ghandhi and S.A. Kassam. Analysis of CFAR processors in nonhomogeneous background. *IEEE Transactions on Aerospace and Electronic Systems*, 24:427–445, 1988.

[3] D.A. Abraham and P.K. Willett. Active signal detection in shallow water using the Page test. SACLANTCEN SR–252, 1996.

[4] S.Z. Li. *Markov Random Field Modeling in Computer Vision.* Springer-Verlag, 1995.

[5] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[6] J. Besag. Spatial interactions and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, Series B 36:192–236, 1974.

[7] D.E. Goldberg. *Genetic algorithms in search optimization, and machine learning.* Addison-Wesley, 1989.

[8] S. Kirkpatrick, C.D. Gellatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[9] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, Series B 48:259–302, 1986.

[10] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C, The Art of Scientific Computing, Second Edition.* Cambridge University Press, 1992.

[11] R. Laterveer, D. Hughes, and S. Dugelay. Markov random fields for target classification in low frequency active sonar. In *Oceans'98*, pages 1274–1278, 1998.

[12] A.K. Jain. *Fundamentals of Digital Image Processing.* Prentice-Hall, 1989.

[13] D. Sheldon and E. Capriulo. A comparison of active sonar frequency-monotonic and frequency-hop waveforms: theoretical background and shallow water results. SACLANTCEN, SR–248, 1996.

Document Data Sheet

NATO UNCLASSIFIED

Security Classification		Project No.
UNCLASSIFIED		041-3
Document Serial No.	**Date of Issue**	**Total Pages**
SR-307	March 1999	33 pp.

Author(s)

Laterveer, R.

Title

Single ping clutter reduction: segmentation using Markov random fields

Abstract

The use of low frequency active sonar in shallow water leads to large numbers of clutter detections. This high false alarm rate can overload automatic tracking and classification algorithms. Traditional detection algorithms operate on each beam output individually searching for targets at all ranges. However, the target echo and bottom features may extend over several beams, either because a reflector is extended over space or because of the sidelobe structure of the beamformer. This suggests to associate detections over bearing, e.g. apply image processing to the range-bearing sonar data.

This reports describes an automatic method of image segmentation based on a Markov random field model to reduce clutter. The segmentation is treated as a labelling problem, i.e. assign to each range-bearing cell either a target or background label. It removes small objects which do not exhibit the right signature over beams. Separate detections corresponding to one large reflector are clumped together to form one single object. Objects too large to be a submarine can then be removed.

The Markov random field model uses *a priori* physical and probabilistic knowledge of the range-bearing sonar picture. It assumes that the background is Rayleigh distributed, a target has, on average, a large signal to noise ratio (SNR) and that the label at each range-bearing cell is influenced by its neighbouring cells. An iterative algorithm is used to produce a maximum *a posteriori* estimate for the labelling. The model parameters are tuned to remove as much clutter as possible without substantially reducing the probability of detection.

The algorithm was tested over a large number of sonar images. The performance is determined using several measures of performance including probability of detection and number of false alarm objects, i.e. connected detections. The number of false alarm objects was reduced by almost an order of magnitude while retaining more than 90% of the target detections.

To test the influence of target SNR on the segmentation, simulated sonar pictures, containing a target of known SNR, were produced. The algorithm performs well for high SNR targets, but those with low SNR have lower detection probability. This is not a big problem as usually targets have sufficient SNR.

Keywords

Clutter reduction - detection - classification - segmentation - low frequency active sonar - Markov random fields - image processing

Issuing Organization

North Atlantic Treaty Organization
SACLANT Undersea Research Centre
Viale San Bartolomeo 400, 19138 La Spezia, Italy

[From N. America: SACLANTCEN (New York) APO AE 09613]

Tel: +39 0187 527 361
Fax:+39 0187 524 600

E-mail: library@saclantc.nato.int

Initial Distribution for Unclassified SR-307

Ministries of Defence

DND Canada	10
CHOD Denmark	8
DGA France	8
MOD Germany	15
HNDGS Greece	12
MARISTAT Italy	9
MOD (Navy) Netherlands	12
NDRE Norway	10
MOD Portugal	5
MDN Spain	2
TDKK and DNHO Turkey	5
MOD UK	20
ONR USA	34

NATO Commands and Agencies

NAMILCOM	2
SACLANT	3
CINCEASTLANT/ COMNAVNORTHWEST	1
CINCIBERLANT	1
CINCWESTLANT	1
COMASWSTRIKFOR	1
COMSTRIKFLTANT	1
COMSUBACLANT	1
SACLANTREPEUR	1
SACEUR	2
CINCNORTHWEST	1
CINCSOUTH	1
COMEDCENT	1
COMMARAIRMED	1
COMNAVSOUTH	1
COMSTRIKFORSOUTH	1
COMSUBMED	1
NC3A	1
PAT	1

Scientific Committee of National Representatives

SCNR Belgium	1
SCNR Canada	1
SCNR Denmark	1
SCNR Germany	1
SCNR Greece	1
SCNR Italy	1
SCNR Netherlands	2
SCNR Norway	1
SCNR Portugal	1
SCNR Spain	1
SCNR Turkey	1
SCNR UK	1
SCNR USA	2
French Delegate	1
SECGEN Rep. SCNR	1
NAMILCOM Rep. SCNR	1

National Liaison Officers

NLO Canada	1
NLO Denmark	1
NLO Germany	1
NLO Italy	1
NLO Netherlands	1
NLO Spain	1
NLO UK	1
NLO USA	1

Sub-total	199
SACLANTCEN	30
Total	229