Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information

S. Reed, Y. Petillot and J. Bell

Abstract: The majority of existing automatic mine detection algorithms which have been developed are robust at detecting mine-like objects (MLOs) at the expense of detecting many false alarms. These objects must later be classified as mine or not-mine. The authors present a model based technique using Dempster–Shafer information theory to extend the standard mine/not-mine classification procedure to provide both shape and size information on the object. A sonar simulator is used to produce synthetic realisations of mine-like object shadow regions which are compared to those of the unknown object using the Hausdorff distance. This measurement is fused with other available information from the object's shadow and highlight regions to produce a membership function for each of the considered object classes. Dempster–Shafer information theory is used to classify the objects using both mono-view and multi-view analysis. In both cases, results are presented on real data.

1 Introduction

Sidescan sonar imagery can provide high resolution images of the seafloor, making it useful for a variety of civilian and military applications. One such application is in the field of mine-countermeasures (MCM), where the advances in autonomous underwater vehicle (AUV) technology have made it necessary to research automated analysis techniques. Complete automated MCM systems generally operate using a two-tier process where the first stage detects all possible mine-like objects (MLOs). The second part of the process classifies each of the objects as mine or not-mine. This paper focuses on the classification section of this process, extending the basic mine/not-mine discrimination to include shape and dimension information on the detected object.

A common approach used to classify the objects is to compare an extracted set of features from the MLO [1–4] to a set of training data. This approach works well when the test data are similar to the training data set but can provide poor results when this criterion is not met [5]. Sidescan sonar imagery is very dependent on the sensor to target azimuth, making images of the same underwater scene look very different depending on the particular conditions. This creates problems when attempting to analyse feature sets to determine if an MLO is a mine or not [6]. Other classification techniques use the more specific knowledge that objects in sidescan sonar leave a recognisable signature. This is characterised by a paired highlight and shadow region [7]. Man-made objects such as mines often have a regular form which generates a recognisable shadow region. The appearance of this shadow region is more constant than the highlight region to change in the sonar conditions. This has led to active research being carried out in shadow extraction techniques [7–9]. Classification models have been developed which then attempt to match the extracted shadow to a template approximation of the shadow expected from a man-made object [10–12]. While this approach has been shown to give good results, it is completely removed from the sonar shadow formation process. Therefore the plausibility of the tested templates is generally not considered during the classification process.

This paper will present a novel classification model that considers information from both the highlight and shadow region of the object. These regions are first extracted from the sidescan sonar image using a co-operating statistical snake model (CSS) [13]. A sonar simulator is then used iteratively to generate shadow regions from objects in different scenarios (size, orientation, depth). These synthetic shadows are compared to the real objects shadow using the Hausdorff distance [14, 15]. This information is fused with additional highlight and size information to produce a membership function for each of the considered object classes. Dempster–Shafer (D–S) information theory then uses these membership functions to carry out mono- and multi-view classification analysis [16]. Both options are useful for sidescan data due to the 'lawn-mower' nature of surveys, which ensures that the same object often appears in multiple images. The ability to consider multi-view analysis allows the classification system to use more of the available information before providing a classification result. The current system classifies the object as a cylinder, sphere, truncated cone or clutter (not-mine), although new classes can easily be included. To avoid confusion, a clutter object is defined as any object which is not a mine. Results for the mono- and multi-view analyses are show on real data sets.

2 Extracting the object's shadow

A three-stage solution to the automated CAD/CAD (computer aided detection/classification) problem was proposed by Reed *et al.* in [13].

IEE Proceedings online no. 20040117

doi: 10.1049/ip-rsn:20040117

Paper first received 21st February and in revised form 1st September 2003

The authors are with the Ocean Systems Laboratory, School of Engineering and Physical Sciences, Heriot-Watt University, Riccarton, Edinburgh, EH14-4AS, UK

The first stage detects possible mine-like objects (MLOs) using a Markov random field (MRF) model. Once detected, the highlight and shadow of the object are extracted in the second stage using a co-operating statistical snake model. Other shadow extraction techniques have been developed using template models [l2, 10], fuzzy logic [8] and statistical models [17, 11]. These provide good results on simple backgrounds but often fail in the presence of complex backgrounds such as sand ripples. The CSS model obtains accurate shadow segmentations on different seabed types by using available *a priori* information on the relationship between the highlight and shadow region of the object. Examples of the CSS segmentation results can be seen in Fig. 1. All the images seen in Fig. 1 are from real sidescan data except the middle-centre and middle-right images, which have been generated using a sonar simulator model [18]. The real data shown were collected using an EdgeTech dual frequency DF-1000 Side Scan Sonar at frequency 390 kHz. The resolution of these images is 10 cm.

The detection and co-operating statistical snake phases of the proposed CAD/CAC approach have been tested on over 200 real sidescan sonar images taken from the BP'02 test experiments at the NATO Saclant Centre, La Spezia, Italy [19]. Over 80% of MLOs were successfully detected with their highlight and shadow regions accurately extracted. This was a shallow water survey, demonstrating the model's ability to cope with multi-path effects. However, it should also be noted that these effects were minimised by using a high frequency sonar under controlled conditions.

The co-operating statistical snake model is also useful for classification purposes as both the highlight and shadow regions are extracted. Both these regions contain information which can be used when classifying an object.

3 Data information

All the results relating to the classification model presented in this paper were obtained using real sidescan sonar data. The data sets were obtained from GESMA, France and the DRDC-Atlantic/Saclantcen Maple 2001 trials. The data from GESMA were taken using an EdgeTech DF 1000 sonar mounted on an AUV. The DRDC-Atlantic data used a Klein 5500 sonar with a towfish. Table 1 contains relevant information regarding the two sonars.

As detailed before, the GESMA data were collected using an AUV. The AUV maintained a constant altitude by carrying out altitude following manoeuvres using bottom bathymetry data [20]. Regular GPS fixes were obtained to minimise the AUV positioning error by regularly resurfacing the AUV.

The DRDC-Atlantic sonar used a towfish and was equipped with an altimeter and pressure sensors to ensure that the sonar depth and height information were accurate.

4 Generation and comparison of synthetic object shadow

When an MLO object has been detected, information such as the image resolution ρ, sonar fish height ϵ and the slant range to the object s are readily available (see Section 3). This allows synthetic images containing objects to be generated under the same sonar conditions that the MLO was detected. For the experiments detailed here, the use of regular GPS fixes and a towfish for the GESMA and DRDC-Atlantic data, respectively, ensured that this information was reasonably accurate. Further improvements in inertial navigational systems (INS) would allow this synthetic simulation to be improved.

For the purposes of mine classification, the tested objects were limited to the cylinder, sphere and truncated cone classes, which closely represent the shape of many real mines. The cylinder class is assumed to be completely described by parameters $\Theta_{cyl} = \{r_{cyl}, l_{cyl}, d_{cyl}, \phi_{cyl}\}$, where r_{cyl} is the radius of the cylinder, l_{cyl} is the length of the cylinder and ϕ_{cyl} is the angle of the cylinder's major axis with respect to the along track direction. The parameter d_{cyl} describes how deeply the cylinder is sunk into the seafloor, with $d_{cyl} = 0$ relating to a cylinder sitting proud on the

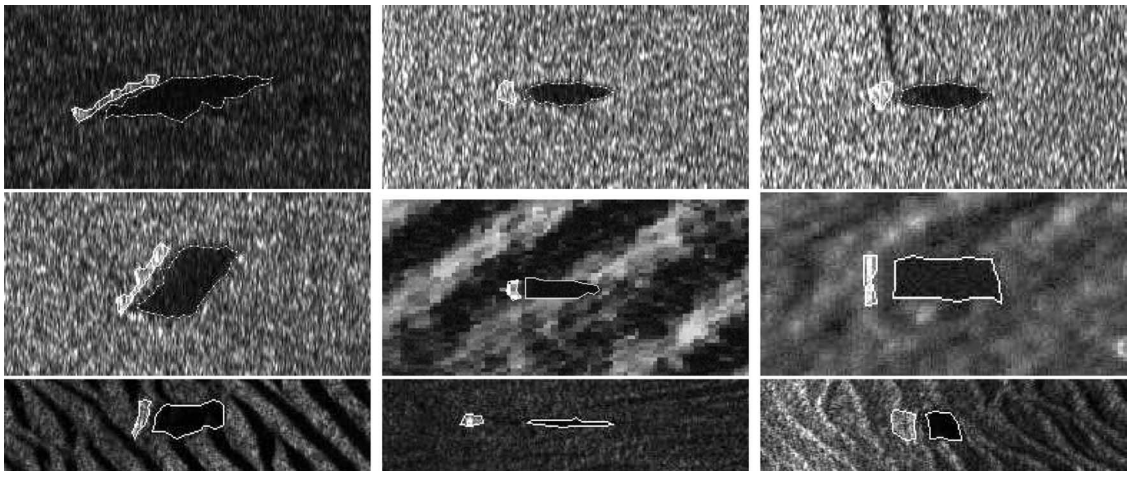

Fig. 1 *Examples on real and synthetic sidescan data of the co-operating statistical snake model's ability to segment both the highlight and shadow regions of the object*

The model successfully copes with a variety of seabed types

Table 1: Sonar information relating to the two different data sets used in this paper

Data set	Frequency, kHz	Sonar depth, m	Sonar height, m	Image resolution, cm
GESMA	390/105	0–30	5–10	3.3
DRDC-Atlantic	455	10–25	8–22	10.0

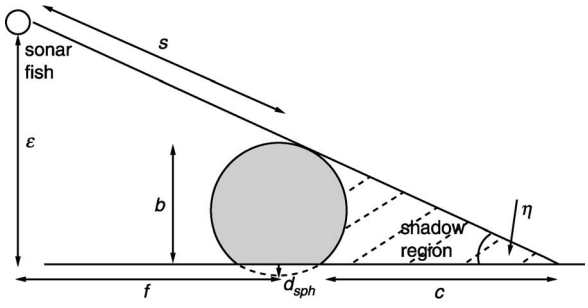

Fig. 2 *One row of a height map image of a partially sunk sphere where line-of-sight calculations are being used to determine the resulting shadow region*

seabed. The sphere class is described by parameters $\Theta_{sph} = \{r_{sph}, d_{sph}\}$ and the truncated cone class by $\Theta_{cone} = \{r_{cone}, d_{cone}\}$, where the representations used are the same as for the cylinder class. For simplicity, only the base-radius r_{cone} was used to describe the truncated cone class. The radius at the top of the truncated cone was assumed to be $0.5r_{cone}$ while the height was set at $0.8r_{cone}$. This is in good agreement with currently tested truncated cone objects [21] and allowed a generic truncated cone class to be specified rather than generating a specific class for each known mine type. Using these parameters, object height maps with the same resolution ρ as the real image can be generated easily.

4.1 Generating the synthetic shadow regions

The object shadow region is determined from the object's height map by considering a simple line-of-sight calculation. This is an approximation from an existing sonar simulator [18] which assumes isovelocity conditions and a simple point source receiver where the source and receiver are colocated. Effects such as beampattern are assumed to be compensated for. This reduction in the model is applicable when it is only the generated shadow which is of interest rather than the full backscattered signal. Figure 2 contains the cross-section of an object where simple geometry can be used to determine the shadow region.

For the spherical model, consideration was given to the underside region of the sphere, which can sometimes allow the sonar to reach regions behind the object, dependent on the sphere depth d_{sph}. Figure 3 contains example shadow regions from the three object classes where the object depth has been altered.

4.2 Comparing the shadow regions

As a consequence of the nonlinear shadow formation process, the same object at a different orientation will provide very different shadow regions. Standard pattern recognition techniques [22–24] are therefore difficult to use when matching the extracted real shadow regions to the generated synthetic shadows. The Hausdorff distance [14] is a technique which measures the similarity between two shapes. Defining $Y = \{y_1, \ldots, y_p\}$ and $V = \{v_1 \ldots, v_q\}$ as the points on the perimeter of the real and synthetic shadow regions, respectively, the Hausdorff distance is defined as

$$H(Y, V) = \max(h(Y, V), h(V, Y)) \tag{1}$$

where

$$h(Y, V) = \max_{y \in Y} \min_{v \in V} \|y - v\| \tag{2}$$

and $\|\cdot\|$ is some underlying norm on the points of Y and V.

Function $h(Y, V)$ is the directed Hausdorff distance and is computed by first computing the distance between each point in Y to its nearest neighbour in V. $h(Y,V)$ is the maximum of this set of values. Therefore if $h(Y, V) = u$, each point in Y must be within distance u of some point in V. A similar process is carried out to compute $h(V, Y)$, with the Hausdorff distance $H(Y, V)$ being designated the maximum value of the two directed distances. $H(Y, V)$ is therefore a measure of mismatch between Y and V. Unlike many other methods for comparing shapes, there is no explicit pairing of points in Y with points in V. The technique is also fast to compute, which is necessary when used within an iterative process. This technique would not cope well with outlier points on the shadow perimeters. However, the co-operating statistical snake model includes shape regulating priors which do not allow the extracted shadow to include jagged peaks [13], ensuring that this is not a problem.

4.3 Iterative matching process

An iterative process was used to search through $\Theta_{cyl}, \Theta_{sph}$ and Θ_{cone} to find the best match from each class. To limit the parameter space tested by the model, constraints can be imposed by considering the real object's extracted highlight and shadow regions. Moment analysis was used to define the image ellipse of the object highlight region [25]. This allowed estimates of the major and minor axis of the highlight to be obtained, which were used to determine initial values for $l_{cyl}, r_{cyl}, r_{sph}$ and r_{cone}. Angle ϕ_{cyl} was also initialised considering moment analysis. Defining ψ as the angle of the principal axis relative to the across-track direction and $\mu_{wz}\, w, z \in \{0, 1, 2\}$ as the centred moments where

$$\psi = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{3}$$

a good initial value for ϕ_{cyl} can be obtained [25]. Initial estimates for depth parameters d_{cyl}, d_{cone} and d_{sph} are found by considering Fig. 2 to obtain

$$b = \frac{c \cdot \epsilon}{f + c} \tag{4}$$

where b is the height of the object, c is the maximum shadow length over all the rows of the shadow region, ϵ is the sonar fish height and f is the distance to the object. As an estimate for the radius has already been obtained, a value for the depth can be computed by considering b.

Large margins were set on each of the initial parameter estimates, defining a discretised parameter space for each class which had to be searched through to find the best fit to the MLO's shadow. The minimum parameter increment allowed was set at the resolution of the image being tested. The parameter space for both Θ_{sph} and Θ_{cone} was two-dimensional, allowing a quick exhaustive search to be used to find the best Hausdorff distance. The parameter space for

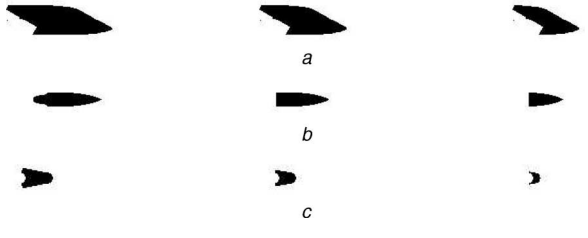

Fig. 3 *Example shadow regions*

a From the cylinder class
b From the sphere class
c From the truncated cone class
The object's burial depth increases from left to right

Θ_{cyl} was four-dimensional. A simple maximum likelihood search was carried out to find the lowest value for H_{cyl}. This has the disadvantage of being more time-consuming than for the other two classes (around 60 s to compute on a Pentium 4 1.3 GHz PC) but was necessary to guarantee a good estimate.

For cases where multiple parameter values gave the same Hausdorff distance H_j, $j \in \{cyl, sph, cone\}$, the most mine-like set of dimensions (those that maximise $\omega_j^{par}(\Theta_j)$ described later in Section 5.2) were used. This increases the false classification rate but also increases the chance of a mine being correctly classified.

This model-based approach also has the advantage of being relatively insensitive to noise. The co-operating statistical snake model has been shown [26] to be very robust to high levels of noise. Once an accurate shadow extraction has been achieved, the model-based nature of the classification ensures that noise does not further affect the model.

After the iterative matching process, each object class is allocated its lowest Hausdorff distance value H_j^b obtained with parameters Θ_j for $j \in \{cyl, sph, cone\}$. This information is then fused with additional information to produce a membership function for each of the classes.

5 Determining the class membership functions

Fuzzy membership functions are often used when a decision needs to be made by considering multiple sources of information [27]. These functions lie in the real, closed interval [0, 1]. For classification purposes, the overall membership function for each class is defined by

$$\omega_j^{final}(H_j^b, \Theta_j, \alpha) = \omega_j^{haus}(H_j^b) \times \omega_j^{par}(\Theta_j) \times \omega_j^{high}(\alpha) \quad (5)$$

Function $\omega_j^{haus}(H_j^b)$ considers the classes best Hausdorff distance values. $\omega_j^{par}(\Theta_j)$ considers the synthetic object parameter values (i.e length, radius etc.) used to obtain H_j^b, while $\omega_j^{high}(\alpha)$ considers the elongation (the ratio of the major/minor axis) α of the unknown object's extracted highlight region where $j \in \{cyl, sph, cone\}$. These three terms are now explained in more detail.

5.1 Function $\omega_j^{haus}(H_j^b)$

This function considers the best Hausdorff distance H_j^b obtained for each of the classes. The shape of the membership function was determined by training the model on real sonar data from known classes. If the sonar simulator exactly modelled the real sonar process, the correct object class would be expected to give a Hausdorff distance value $H_j^b \approx 0$. While accurate, such sonar models are computationally heavy, making them unsuitable for use within an iterative matching process. The line-of-sight simulator described in Section 4 produced Gaussian distributions around non-zero Hausdorff values. It should also be noted that the training and testing data were completely separate, with no overlap between the sets. The function was defined for each of the classes by

$$\omega_j^{haus}(H_j^b) = 1 \quad \text{if} \quad H_j^b \leq \bar{m}_j$$
$$= e^{-\frac{(H_j^b - \bar{m}_j)^2}{2\sigma_j^2}} \quad \text{if} \quad H_j > \bar{m}_j$$

where $\bar{m}_j$ and σ_j^2 are the mean and variance values of the class Gaussian distributions. The specific function values used were $\bar{m}_{cyl} = 5.17, \bar{m}_{sph} = 5.42, \bar{m}_{cone} = 5.3$, $\sigma_{cyl}^2 = 3.02, \sigma_{sph}^2 = 1.5$ and $\sigma_{cone}^2 = 3.9$. The training was carried out on different data-sets of objects provided by GESMA and DRDC-Atlantic.

The data sets were visually very different but as the model incorporates information such as the resolution ρ of the image, similar Hausdorff distance values were obtained on both. An interesting aside is that if the sonar simulator was made more complex, the membership functions could be made more severe, which would ensure larger separability between the different classes and so an improved classification rate.

5.2 Function $\omega_j^{par}(\Theta_j)$

The first membership function describes the degree of match between the synthetic shadow from each class and the real MLO's shadow. While the synthetic shadow may visually appear very similar to the real shadow, no consideration has yet been given to the parameters of the synthetic object. This membership function considers how closely parameters Θ_j match real-mine dimensions. The class membership functions are defined using functions $\Gamma(\cdot), \Psi(\cdot), \Omega(\cdot)$ and $\Pi(\cdot)$ as shown in Fig. 4, where $\Gamma(\cdot)$ relates to the length of the cylinder, $\Psi(\cdot)$ to the radius of the cylinder, $\Omega(\cdot)$ to the radius of the sphere and $\Pi(\cdot)$ to the base radius of the truncated cone. The form of these functions has been chosen for general mine-like dimensions but could be made case-specific.

The specific membership functions are defined by:

$$\begin{aligned}
\omega_{cyl}^{par}(\Theta_{cyl}) &= \max\{\Gamma(l_{cyl}),\ \Gamma(l_{cyl}-\rho),\ \Gamma(l_{cyl}+\rho)\} \\
&\quad \times \max\{\Psi(r_{cyl}),\ \Psi(r_{cyl}-\rho),\ \Psi(r_{cyl}+\rho)\} \\
\omega_{sph}^{par}(\Theta_{sph}) &= \max\{\Omega(r_{sph}),\ \Omega(r_{sph}-\rho),\ \Omega(r_{sph}+\rho)\} \\
\omega_{cone}^{par}(\Theta_{cone}) &= \max\{\Pi(r_{cone}),\ \Pi(r_{cone}-\rho),\ \Pi(r_{cone}+\rho)\}
\end{aligned} \quad (6)$$

where ρ is the resolution of the image being considered. The fuzzy t-conorm function $\max\{\cdot\}$ [28] was used to ensure that errors in the parameter measurement due to the image resolution were accounted for. T-conorms generalise union of sets where a t-conorm T follows the inequality

$$\forall(\nu, \tau) \in [0, 1]^2, T(\nu, \tau) \geq \max(\nu, \tau) \quad (7)$$

T-conorms are therefore disjunctive, with 'max' being the smallest possible function. Using a t-conorm (in our case the maximum of the range of measurements) ensures that a pessimistic standpoint is always taken when classifying an object, i.e. we always accept the most mine-like measurement. This should ensure that fewer real mines are disregarded as clutter objects.

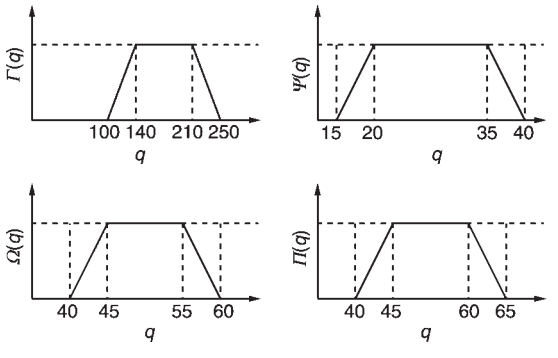

Fig. 4 *The four membership functions used to define* $\omega_j^{par}(\Theta_j)$
All dimensions are in centimetres

5.3 Function $\omega_j^{high}(\alpha)$

This function will consider the elongation α of the object's highlight region. Simply, objects from the cylinder class would be expected to have a high α value while objects from the other classes would not. The function can be seen in Fig. 5, where the crossover parameter β was set at 2.25 after considering the training data used to define $\omega_j^{haus}(H_j^b)$.

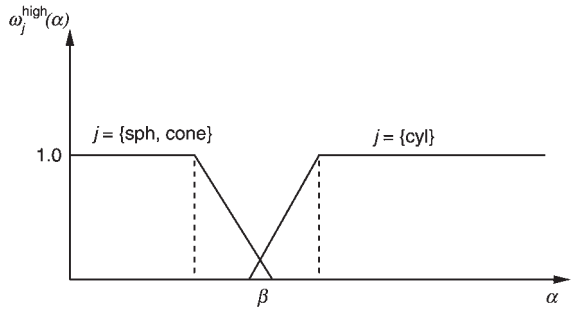

Fig. 5 *Form $\omega_j^{high}(\alpha)$ for $j = \{cyl\}$ and $j = \{sph, cone\}$*

5.4 Membership function results

The class membership functions were determined for 50 different objects. These data were taken from the GESMA and DRDC-Atlantic data sets. There was no overlap between the testing data and training data used to define the fuzzy functions. The testing data included 10 cylindrical objects, 10 spherical objects, 10 truncated cone objects and 20 clutter objects. All the tested images are from real Sidescan data. Examples of the different classes can be seen in Fig. 6. As can be seen, clutter can often appear very mine-like in sidescan imagery.

The final class membership functions for the 30 mine-like objects can be seen in Table 2.

As Table 2 shows, the cylinder and sphere objects are well classified. The truncated cone objects are not so clearly defined, with the sphere and cone classes often offering non-zero membership functions. However, these classes often provide similar shadow regions under certain sonar conditions. It should also be noted that the cylinder class has zero membership for all the cone and sphere images. The clutter image results are not shown but were dependent on the visual similarity to the considered classes. The very irregular clutter objects provided low membership functions for all three classes. For those that closely represent a specific object class (see Fig. 6), only the relevant membership function was non-zero.

Table 2: Final membership functions for all three considered object classes

Obj. number	Obj. class	ω_{cyl}^{final}	ω_{sph}^{final}	ω_{cone}^{final}
1	cyl	1	0	0
2	cyl	0.84	0	0
3	cyl	1	0	0
4	cyl	1	0	0
5	cyl	0.75	0	0
6	cyl	0.1	0	0
7	cyl	0.35	0	0
8	cyl	0.08	0	0
9	cyl	1	0	0
10	cyl	1	0	1
11	cone	0	0.13	0.16
12	cone	0	0	0.17
13	cone	0	1	1
14	cone	0	0	1
15	cone	0	1	1
16	cone	0	0.30	0.17
17	cone	0	1	1
18	cone	0	0	1
19	cone	0	1	1
20	cone	0	1	1
21	sphere	0	0.34	0
22	sphere	0	1	0
23	sphere	0	1	0.0
24	sphere	0	0.66	0.01
25	sphere	0	1	0
26	sphere	0	0.22	0
27	sphere	0	0.17	0
28	sphere	0	1	0
29	sphere	0	0.34	0
30	sphere	0	0.34	0

Results are shown for the cylinder, sphere and truncated-cone objects

A classification decision could be made on these membership functions by simply imposing a threshold. This is a 'hard' technique. The objects are classified simply as mine or not-mine. A more subtle approach is to allocate a probability or a 'belief' to each of the classes. This is particularly desirable for sidescan surveys where the same

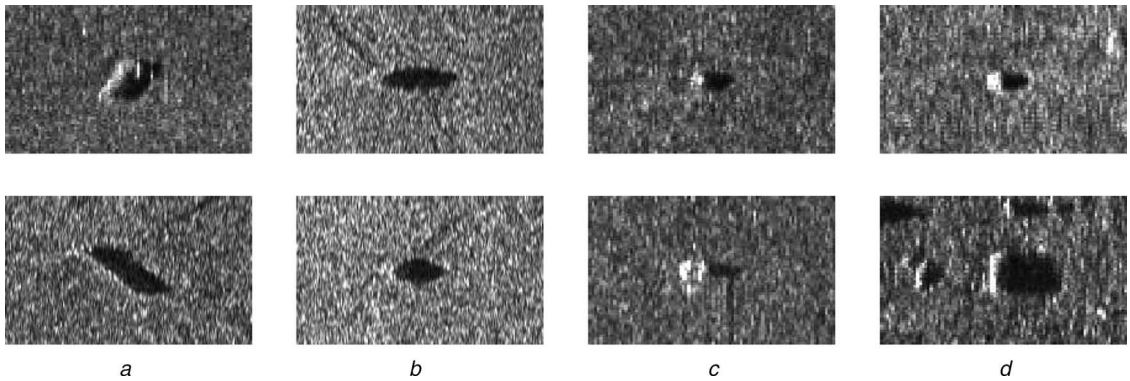

Fig. 6 *Examples of objects used for testing*

a Cylinder objects
b Sphere objects
c Truncated cone objects
d Clutter objects

object is often viewed multiple times offering the opportunity for multi-view analysis to be carried out on the object. This can be done by considering Dempster–Shafer (D–S) information theory.

6 Using Dempster–Shafer information theory

Data fusion collates information from different sources considering the same scene in an attempt to provide a more complete description. The most common numerical techniques used are Bayesian probability theory [29, 30], fuzzy systems [28] and Dempster–Shafer (D–S) theory [31, 32]. Fuzzy systems contain a wealth of possible fusion operators. However, the choice of operator is case dependent and many of the operators are non-associative, meaning the order in which the information is fused has an impact on the final result. Bayesian and Dempster–Shafer models have both been successfully applied but D–S theory provides some features which Bayesian theory cannot [33]. One of the most significant features is that D–S theory can consider union of classes. This feature is used to improve the separability of the cone and sphere classes. Results from the D–S classifier are presented in both a mono-view and multi-view context.

6.1 Dempster–Shafer model

Dempster–Shafer (D–S) theory allows the representation of imprecision and uncertainty through the definition of two functions: plausibility (*Pls*) and belief (*Bel*). These are derived from a mass function m which is analogous to the well known probability function. Mass functions are defined on the power set of the space of discernment D. For classification purposes, D may be the set of possible classes. Specific to D–S theory, D may also contain union of classes [33]. Denoting 2^D as the set of subsets of D, mass function $m(A)$ for every element A of 2^D is defined such that

$$m(\emptyset) = 0, \sum_{A \subset D} m(A) = 1 \tag{8}$$

Given a set of mass functions, the belief (*Bel*) and plausibility (*Pls*) can be determined by:

$$\begin{aligned} Bel(\emptyset) &= 0 \\ Bel(A) &= \sum_{B \subset A} m(B), \forall A \subset D, A \neq \emptyset \end{aligned} \tag{9}$$

$$\begin{aligned} Pls(\emptyset) &= 0 \\ Pls(A) &= \sum_{B \cap A \neq \emptyset} m(B), \forall A \subset D, A \neq \emptyset \end{aligned} \tag{10}$$

These functions have the following properties:

$$\begin{aligned} Bel(D) &= 1 \\ Pls(D) &= 1 \\ Bel(A) &\leq Pls(A), \forall A \subset D \\ Pls(A) &= 1 - Bel(A^C), \forall A \subset D \end{aligned} \tag{11}$$

where A^C is the complement of A. Different techniques have been implemented to make a decision based on the belief (*Bel*) and plausibility (*Pls*) values [33]. The most common, and the one used here, is the maximum belief over all the singleton classes (thus ensuring that the final classification result is not a union of classes).

For the mine classification model presented here, the allowed Dempster–Shafer classes were $A = \{clutter, cyl, sph, cone, sph \cup cone\}$. The nonsingleton class $A = sph \cup cone$ has been defined to remove the degeneracy between the sphere and cone classes on the truncated cone membership results seen in Table 2. It can be seen that the cone and sphere membership functions are only identical when $\omega_{cone}^{final} = \omega_{sph}^{final} = 1$. The mass functions are therefore defined:

$$\begin{aligned} m(cyl) &= \frac{\omega_{cyl}^{final}}{Z} \\ m(cone) &= \frac{\omega_{cone}^{final}}{Z} \\ m(sph \cup cone) &= \frac{G(\omega_{sph}^{final}, \omega_{cone}^{final})}{Z} \\ m(sph) &= \frac{\omega_{sph}^{final} - G(\omega_{sph}^{final}, \omega_{cone}^{final})}{Z} \\ m(clutter) &= \frac{F(x,t)}{Z} \end{aligned} \tag{12}$$

$$m(D) = \frac{0.1}{Z} \tag{13}$$

where $x = \max\{\omega_{cyl}^{final}, \omega_{sph}^{final}, \omega_{cone}^{final}\}$, t is a threshold parameter and Z is a normalising factor. Mass function $m(D)$ has simply been allocated a small, non-zero value and represents the ignorance between classes. $G(\cdot)$ is defined as

$$\begin{aligned} G\left(\omega_{sph}^{final}, \omega_{cone}^{final}\right) &= \xi \quad \text{if} \quad \omega_{cone}^{final} = \omega_{sph}^{final} = 1 \\ &= 0 \text{ otherwise} \end{aligned} \tag{14}$$

Constant ξ was arbitrarily set to $\frac{1}{3}$ to ensure that $m(cone) > m(sph)$ when $\omega_{sph}^{final} = \omega_{cone}^{final} = 1$.

The form of $F(x,t)$ can be seen in Fig. 7. This form ensures that when $x > t$, at least one of the object class belief functions is greater than the clutter belief, while for $x > t$ the clutter has the highest belief function.

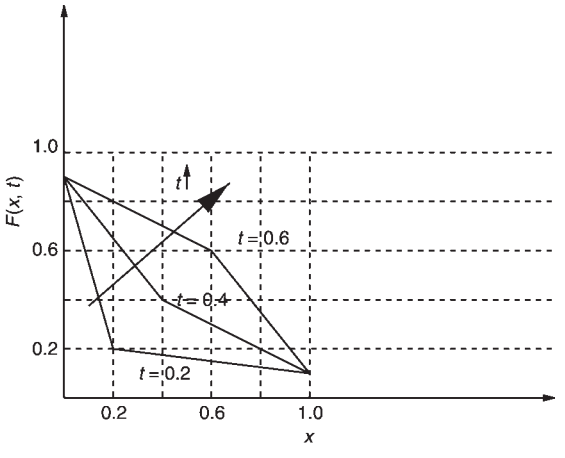

Fig. 7 *Form of F(x, t) used in the D–S mass function for clutter*
Form used depends on threshold t and x

6.2 Mono-view classification results

The tested data were classified using the D–S classifier. The results can be seen in Fig. 8 for a range of threshold t values. The graph contains two plots. The first considers the standard mine/not-mine classification, while the second considers the classification as correct only if the correct object class has been identified.

A threshold value $t = 0.18$ correctly classifies over 85% of the mines by identifying their specific object class.

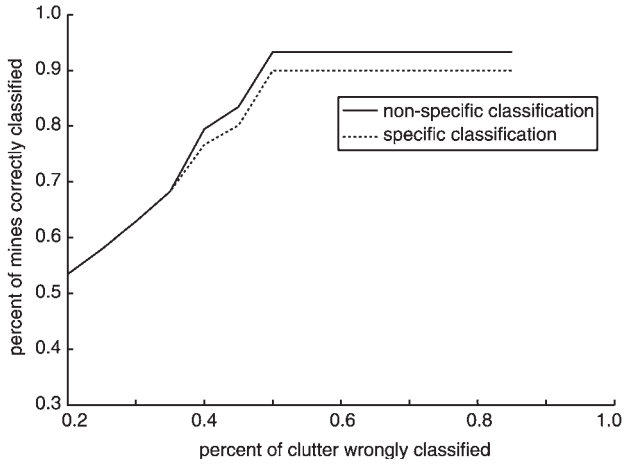

Fig. 8 *Results for Dempster–Shafer classification model*
Results obtained on 30 mine objects and 20 not-mine clutter objects

The same threshold correctly identifies over 50% of the clutter objects as false alarms where it is important to note that some of the clutter objects appear very 'mine-like' in the mono-view images. This result would allow the number of false alarms detected using any detection model to be dramatically reduced. The same t value allows $\sim$95% of the mines to be correctly labelled as simply mine. However, obtaining a specific shape classification (as well as the parameter information) allows the possibility of identifying the mine type and so perhaps affecting how the specific threat is dealt with. Further improvements to these results would be expected by again improving the sonar simulator used in generating the synthetic shadows.

6.3 Multi-image classification

When the same object has been viewed from multiple aspects, D–S theory allowed the mono-view mass functions to be combined. Considering mass function m_k from source $k, k = 1, \ldots, n$, where n is the number of sources, this rule is expressed by

$$(m_l \oplus m_2 \oplus \ldots \oplus m_n)(A) = \frac{\sum_{B_1 \cap \ldots \cap B_n = A} m_1(B_1) m_2(B_2) \ldots m_n(B_n)}{1 - \sum_{B_1 \cap \ldots \cap B_n = \emptyset} m_1(B_1) m_2(B_2) \ldots m_n(B_n)} \quad (15)$$

for all nonempty subsets A of D. The summation on the bottom line of (15) is often referred to as the *conflict* and is ≤ 1. A value of 1 means that the evidence from the sources is completely conflicting and so cannot be fused. Once the fused mass functions have been determined, the belief (*Bel*) functions can be determined as in (11) and the multi-view classification result obtained.

The multi-view classification model is demonstrated on three different objects. The first is a cylindrical object, where four views corresponding to objects 2, 3, 4 and 6 in Table 2 have been used. The second object is a truncated cone where the four views corresponding to objects 11, 12, 15 and 16 in Table 2 have been used. The final example considers a spherical object where objects 21, 22, 23 and 24 have been used. The different views can be seen in Fig. 9.

Figure 9 clearly shows the effect that direction, fish height and range can have on the appearance of the objects in sidescan sonar imagery, again demonstrating why it is important for these variables to be taken into consideration when classifying the object. Tables 3–5 show the mono- and multi-image classification results for the cylinder, truncated cone and spherical images respectively. A value of $t = 0.18$ was used throughout.

Tables 3–5 show all three objects being correctly classified with strong belief. Table 5 shows the fusion model simply confirming the mono-view classification results, offering a very high belief for the sphere class. The cone example is a more difficult example with the mono-image classifications offering conflicting results. However, the fusion model provides the correct classification result. Table 3 shows the cylinder being classified confidently even though object 6 is actually detrimental to the overall result (due to a low value for $\omega_{cyl}^{par}(\Theta_{cyl})$). Owing to a lack of prior information, the model currently carries out the fusion by blindly considering all considered views as equally important. A possible alternative offered by D–S theory is to only consider a view if, for a hypothesis $A \in D, Bel(A) \geq \max_{A' \in D, A' \neq A} Pls(A')$. This is a very strict condition called the *absolute decision rule* and would result in the model only fusing views with strong mono-view classification results. Another alternative offered by D–S theory is to consider the *conflict* (see Section 6.3) when deciding whether to consider a view. For information, the measured conflicts while fusing the cylinder, cone and sphere views were {0.23, 0.13, 0.81}, {0.41, 0.64, 0.62} and {0.37, 0.17, 0.31}, respectively. The final addition of object 6 in the cylinder example can be clearly seen to heavily conflict with the other views. D–S theory would also allow classification ignorance (certain sonar conditions can

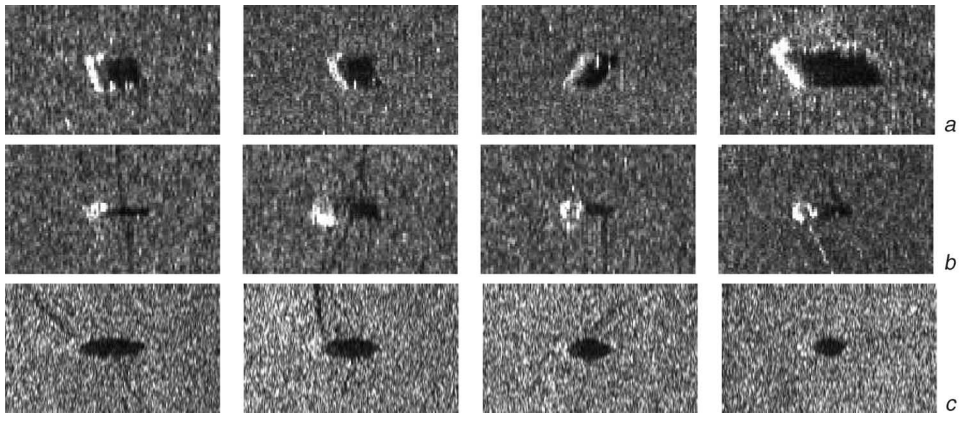

Fig. 9 *Different views of three objects*
a Four different views of the same cylinder
b Four different views of the same truncated cone
c Four different views of the same spherical object
These views are taken from different directions, fish heights and slant ranges

Table 3: Belief functions for the different classes for both individual images and overall fused result for cylindrical object

Mono-image belief					Fused belief				
Object	Cylinder	Sphere	Cone	Clutter	Objects fused	Cylinder	Sphere	Cone	Clutter
2	0.707	0.000	0.000	0.208	2	0.707	0.000	0.0000	0.208
3	0.833	0.000	0.000	0.083	2, 3	0.936	0.000	0.000	0.054
4	0.833	0.000	0.000	0.083	2, 3, 4	0.988	0.000	0.000	0.011
6	0.058	0.000	0.000	0.820	2, 3,4, 6	0.939	0.000	0.000	0.60

Fused belief functions change as a new image is included in multi-view classification

Table 4: Belief functions for the different classes for both individual images and overall fused result for truncated cone object

Mono-image belief					Fused belief				
Object	Cylinder	Sphere	Cone	Clutter	Objects fused	Cylinder	Sphere	Cone	Clutter
11	0.000	0.213	0.246	0.385	11	0.000	0.213	0.246	0.385
12	0.000	0.000	0.359	0.432	11, 12	0.000	0.076	0.334	0.535
15	0.000	0.303	0.454	0.045	11, 12, 15	0.000	0.152	0.675	0.142
16	0.000	0.358	0.199	0.323	11, 12, 15, 16	0.000	0.0223	0.592	0.174

Fused belief functions changes as a new image is included in multi-view classification

Table 5: Belief functions for the different classes for both individual images and overall fused result for spherical object

Mono-image belief					Fused belief				
Object	Cylinder	Sphere	Cone	Clutter	Objects fused	Cylinder	Sphere	Cone	Clutter
21	0.000	0.465	0.000	0.400	21	0.000	0.465	0.000	0.400
22	0.000	0.833	0.000	0.083	21, 22	0.000	0.858	0.000	0.124
23	0.000	0.832	0.001	0.083	21, 22, 23	0.000	0.971	0.001	0.027
24	0.000	0.597	0.012	0.300	21, 22, 23, 24	0.000	0.983	0.001	0.016

Fused belief functions changes as a new image is included in multi-view classification

provide images which are unsuitable for classification) to be modelled through the use of the dual hypothesis mass functions. Future research will focus on these points to create a more flexible fusion model.

7 Conclusions

This paper has presented a novel classification model which extends the standard mine/not-mine classification procedure to include shape and parameter information on the object. A Dempster–Shafer based model was implemented which allowed a classification to be obtained from either mono or multiple views of the object. This is important for sidescan surveys, which generally use 'lawnmower trajectories' and so often contain multiple views of the same object. The Dempster–Shafer mass functions were determined by using an iterative sonar simulator process to generate synthetic shadow regions from considered object classes. This allowed information such as the object's parameters to be included in the classification process. The approach also easily accommodates the inclusion of new object classes and allows the classification process to become mission-specific. Results were presented on real mine and clutter images. These results showed how the proposed approach could remove a significant amount of the false alarms picked up by any mine detection model while classifying a large percentage of the real mine objects. Identifying the mine's shape and size would allow the possibility of classifying the mine type, leading to an improved assessment of how best to neutralise the threat.

8 Acknowledgments

The authors would like to thank the Mine and Torpedo Defence group at DRDC-Atlantic, Canada, for providing much of the data used. These data were collected during the DRDC-Atlantic/Saclantcen Maple 2001 Trial. The authors are also grateful to Benoit Zerr at GESMA for providing both test data and helpful comments.

9 References

1 Dobeck, G.J., Hyland, J.C., and Smedley, L.: 'Automated detection/classification of sea mines in sonar imagery', *Proc. SPIE–Int. Soc. Opt. Eng.*, 1997, **3079**, pp. 90–110
2 Aridgides, T., Fernandez, M., and Dobeck, G.: 'Adaptive 3-dimensional range-crossrange-frequency filter processing string for sea mine classification in side-scan sonar imagery', *Proc. SPIE–Int. Soc. Opt. Eng.*, 1997, **3079**, pp. 111–122
3 Ciany, C.M., and Huang, J.: 'Computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines'. Proc. MTS/IEEE Oceans Conf. and Exhibition, 2000, vol. 1, pp. 277–284
4 Dobeck, G.J.: 'Algorithm fusion for automated sea mine detection and classification'. Proc. MTS/IEEE Oceans, Conf. and Exhibition, 2001, vol. 1, pp. 130–134
5 Ciany, C.M., and Zurawski, W.: 'Performance of computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines'. Presented at CAD/CAC Conf., Halifax, Novia Scotia, Canada, Nov. 2001
6 Fawcett, J.A.: 'Image-based classification of side-scan sonar detections'. Presented at CAD/CAC Conf., Halifax, Novia Scotia, Canada, Nov. 2001
7 Reed, S., Petiliot, Y., and Bell, J.: 'Unsupervised mine detection and analysis in side-scan sonar: A comparison of Markov random fields and statistical snakes'. Presented at CAD/CAC Conf., Halifax, Novia Scotia, Canada, Nov. 2001

8 Myers, V.J.: 'Image segmentation using iteration and fuzzy logic'. Presented at CAD/CAC Conf., Halifax, Novia Scotia, Canada, Nov. 2001
9 Quidu, I., Malkasse, Ph., Burel, G., and Vilbe, P.: 'Mine classification using a hybrid set of descriptors'. Proc. Oceans MTS/IEEE Conf. and Exhibition, 2000, vol. 1, pp. 291–297
10 Mignotte, M., Collet, C., Perez, P., and Bouthemy, P.: 'Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (2), pp. 129–141
11 Quidu, I., Malkasse, Ph., Burel, G., and Vilbe, P.: 'Mine classification based on raw sonar data: An approach combining Fourier descriptors, statistical models and genetic algorithms'. Proc. Oceans MTS/IEEE Conf. and Exhibibition, 2000, vol. 1, pp. 285–290
12 Dura, E., Bell, J., and Lane, D.: 'Superellipse fitting for the classification of mine-like shapes in side-scan sonar images'. Proc. MTS/IEEE Oceans, Conf. and Exhibition, 2002, pp. 23–28
13 Reed, S., Petillot, Y., and Bell, J.: 'An automated approach to the detection and extraction of mine features in sidescan sonar', *IEEE J. Ocean. Eng.*, 2003, **28**, (1), pp. 1–16
14 Huttenlocher, D.P., Klanderman, G.A., and Rucklidge, W.J.: 'Comparing images using the Hausdorff distance', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1993, **15**, (9), pp. 850–863
15 Odone, F., Trucco, E., and Verri, A.: 'General purpose matching of grey level arbitrary images'. Proc. 4th International Visual Form IWvF4, 2001, pp. 573–582
16 Zerr, B., Bovio, E., and Stage, B.: 'Automatic mine classification approach based on AUV manoeuvrability and the COTS side scan sonar'. Proc. Autonomous Underwater Vehicle and Ocean Modelling Networks, GOATS 2000, SACLANTEN Conf. Proc. CP-46, La Spezia, Italy, Aug. 2001, pp. 315–322
17 Reed, S., Bell, J., and Petillot, Y.: 'Unsupervised segmentation of object shadow and highlight using statistical snakes'. Proc. Autonomous Underwater Vehicle and Ocean Modelling Networks: GOATS 2000, SACLANTEN Conf. Proc. CP-46, La Spezia, Italy, Aug. 2001, pp. 221–236
18 Bell, J.: 'A model for the simulation of sidescan of sonar'. PhD thesis, Heriot-Watt University, UK, August 1995
19 Reed, S., Petillot, Y., and Bell, J.: 'A model-based approach to the detection and classification of mines in sidescan sonar'. Presented at MREP'03 Conf., La Spezia, Italy, May 2003
20 An, E., Smith, S., Dunn, S., Pantelakis, T., Whitley, W., Frankenfield, J., Kuntz, J., Burns, A., and Christensen, R.: 'Ocean explorer AUV participation in the GOATS 2001 experiment'. Proc. Autonomous Underwater Vehicle and Ocean Modelling Networks, GOATS 2000, SACLANTCEN Conf. Proc. CP-46, Aug 2001, pp. 221–236
21 Myers, V.: 'A database of mine-like object'. Technical Report, Defence Research Establishment Atlantic, 2001
22 Gupta, L., and Srinath, M.D.: 'Contour sequence moments for the classification of closed planar shapes', *Pattern Recognit.*, 1987, **20**, (3), pp. 267–272
23 Wood, J.: 'Invariant pattern recognition: a review', *Pattern Recognit.*, 1996, **29**, (1), pp. 1–17
24 Flusser, J., and Suk, T.: 'Pattern recognition by affine moment invariants', *Pattern Recognit.*, 1993, **26**, (1), pp. 167–174
25 Prokop, R.J., and Reeves, A.P.: 'A survey of moment-based techniques for unoccluded object representation and recognition', *CVGIP, Graph. Models Image Process.*, 1992, **54**, (5), pp. 438–460
26 Chesnaud, C., Refregier, P., and Boulet, V.: 'Statistical region snake-based segmentation adapted to different physical noise models', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1999, **21**, (11), pp. 1145–1157
27 Seraji, H., Howard, A., and Tunstei, E.: 'Safe navigation on hazardous terrain'. Proc. IEEE Int. Conf. Robotics, and Automation, 2001, pp. 3084–3091
28 Bloch, I.: 'Information combination operators for data fusion: A comparative review with classification', *IEEE Trans. Syst. Man Cybern.*, **26**, (1), pp. 52–67
29 Solberg, A.H.S., Jain, A.K., and Taxt, T.: 'Multisource classification of remotely sensed data: fusion of landsat TM and SAR images', *IEEE Trans. Geosci. Remote Sens.*, 1994, **32**, (4), pp. 768–778
30 Buede, D.M., and Girardi, P.: 'A target identification comparision of Bayesian and Dempster–Shafer multisensor fusion', *IEEE Trans. Syst. Man Cybern. A, Syst. Humans*, 1997, **27**, (5), pp. 569–577
31 Le Hegart-Mascle, S., Bloch, I., and Vidal-Madjar, D.: 'Application of Dempster–Shafer evidence theory to unsupervised classification in multisource remote sensing', *IEEE Trans. Geosci. Remote Sens.*, 1997, **35**, (4), pp. 1018–1031
32 Van Cleynenbreugal, J., Osinga, S.A., Fierens, F., Suetens, P., and Oosterlinck, A.: 'Road extraction from multi-temporal satellite images by an evidential reasoning approach', *Pattern Recognit. Lett.*, 1991, **12**, pp. 371–380
33 Bloch, I.: 'Some aspects of Dempster–Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account', *Pattern Recognit. Lett.*, 1996, **17**, pp. 905–919