An Adaptive Unified Algorithm for Both Detection and Recognition

Keith D Copsey and Chris J S Webber
QinetiQ
Malvern Technology Centre
St Andrews Road, Malvern
Worcestershire, WR14 3PS
United Kingdom

K.Copsey@signal.QinetiQ.com and C.Webber@signal.QinetiQ.com

ABSTRACT

An adaptive neural-network approach to target- and clutter-modelling is introduced. A key novelty of this approach is that both targets and clutter can be modelled within the same neural network, so that detection and recognition can take place simultaneously within an integrated framework. The approach can therefore be applied across the spectrum of ATR discrimination levels, e.g.: detection of unknown targets in clutter; detection of specific designated targets in clutter; recognition of target subclass post-detection. The approach is designed to be generically applicable, to data from a variety of sensors, including HRRPs, SAR intensity imagery, complex SAR imagery, visible and EO imagery, and burst-illumination LIDAR. This generic applicability is attributable to the fact that the algorithm adaptively models training-exemplar data of arbitrary type and dimensionality. Unlike many current approaches to target detection, this approach can exploit a wide range of cues for discriminating targets from clutter objects, including detailed grey-level shape information and, for RF sensors, complex/phase information. Furthermore, the approach is quick to use in operation, and has been designed with hardware implementations in mind. Successful results are presented for a target (designated building) detection and identification problem using real SAR imagery. The approach has been designed to have the future potential to offer other very significant new capabilities, e.g. the potential for reducing false-alarm rate in urban clutter and improving robustness to extended operating conditions.

1.0 INTRODUCTION

1.1 Integrated approach to target detection and recognition

This paper concerns an adaptive neural-network approach to target- and clutter-modelling. A key novelty of the approach is that both targets and clutter can be modelled within the same neural network, so that detection and recognition can take place simultaneously within an integrated framework. The approach can therefore be applied across the spectrum of ATR discrimination levels, e.g.:

- Detection of unknown targets in clutter;
- Detection and recognition of specific designated targets in clutter;
- Recognition of target subclass post-detection.

The proposed approach is designed to be generically applicable to data from a variety of sensors, including HRRPs, SAR intensity imagery, complex SAR imagery, visible and EO imagery, and burst-illumination LIDAR. This generic applicability is attributable to the fact that the algorithm adaptively models training-exemplar data of arbitrary type and dimensionality.

Paper presented at the RTO SET Symposium on "Target Identification and Recognition Using RF Systems", held in Oslo, Norway, 11-13 October 2004, and published in RTO-MP-SET-080.

When used for target detection, the approach exploits the kind of knowledge of the *signature* of the target that would conventionally be used only at the post detection identification stage, in order to influence the *detection* decision. This allows potentially crucial signature information (e.g. detailed grey-level shape information and, for RF sensors, complex/phase information) to be used at the earliest stages of target detection. This offers the potential to mitigate the false-alarm rate of the detection process very significantly. The resulting "identify-for-detect" principle is analogous to the "track-before-detect" principle well known in the context of tracking algorithms [8]. The key to this unified approach to automatic target detection and recognition lies in the use of a technique with the ability to:

- Model targets with varying degrees of specificity and generalisation, depending on the scenario;
- Model varied and complicated clutter with a high degree of generalisation.

Many potential approaches for modelling both targets and clutter cannot be used for simultaneous target detection and recognition, because of the computational expense involved in conducting the comparisons at each location. In contrast, the proposed algorithm is quick to use in operation, and has been developed with hardware implementations in mind.

1.2 Background

The developed approach is the result of neural network research into transformation-robust pattern recognition. The motivating aim has been to develop techniques that enable target detection and recognition to be robust to both continuous and discontinuous complex transformations (where complex transformations are defined to be transformations of individual objects within the sensor image, as opposed to simple perspective transformations of the whole image). In the ATR context, continuous complex transformations include 3-D rotation and articulation (e.g. the rotation of a tank's turret with respect to its hull). Examples of discontinuous complex transformations include the replacement of one type of vehicle with another of similar type, partial occlusion, differences in equipment fit, and change in operational configuration.

Complex transformations are intimately associated with ill-defined properties of the objects to be recognised. It is likely that characterisation of these properties will require reference to a complex, nonlinear, hierarchical pattern recognition technique, with the ability to adapt to training data that is characteristic of those ill-defined properties. A neural network is such a technique.

In the neural network literature, the term "invariance" is used to refer to what the military user or systems designer would term "robustness"; neural network papers generally adopt the mathematical terminology of group theory rather than the terminology of the ATR application domain. A neural network (or some part of it) is said to be invariant under a transformation of the data if its recognition output response does not change (or changes only gradually) as the transformation is applied to the data. This property of the recognition response is clearly what is desired if one wishes to build a neural network based ATR system that is robust with respect to variation in the sensor data.

1.3 Outline

The structure of this paper is as follows. Section 2 outlines the concepts behind the adaptive neural network algorithm, by discussing neural network approaches to transformation invariance. Section 3 briefly discusses the approach implemented within algorithms to date. Section 4 presents experimental results for detection and recognition of a particular building within SAR imagery of an urban area. Conclusions and future work are in Section 5.

2.0 ALGORITHMIC CONCEPTS

2.1 Discussion of neural network approaches to transformation invariance

Whether explicitly acknowledged or not, many neural networks capable of transformation invariance are symmetry networks [10]. These are neural networks in which some synaptic weights are equal (or approximately equal) to others. Put in another way, these are networks whose configuration of synaptic weights is symmetric (invariant) under certain groups of permutation transformations. This is illustrated graphically in Figure 1. The recognition outputs of neurons that possess this symmetry property are consequently invariant under corresponding groups of transformations on their input data. This conclusion is related to the group-invariance theorem of Minsky & Papert [5], and is explained more fully by Webber [16].

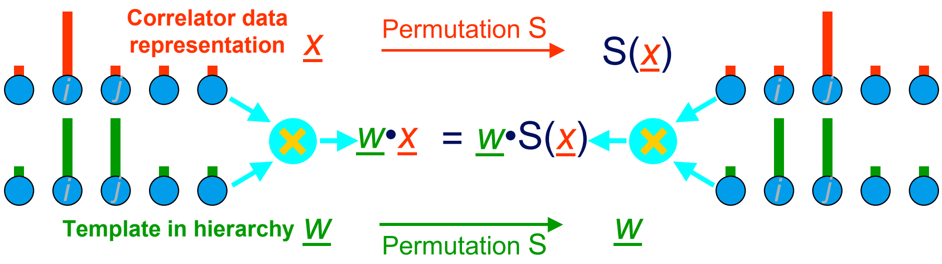

Figure 1: Illustration of a symmetry network. The permutation transformation in the data shown in the top half of the diagram is matched by a permutation in the synaptic weights (the templates in the hierarchy W) shown in the bottom half.

"Weight-sharing" [5][10][11] provides a trivial example of synaptic-weight symmetry in a symmetry network, although this does not begin to exemplify the full potential of synaptic-weight symmetry as a means of achieving transformation robustness. In a weight-sharing architecture, an array of neural network recognition nodes ("neurons") is forced to generate an output recognition response that is invariant with respect to global 2-D translations of the whole image. This is achieved by constructing the architecture of the network so that neighbouring neurons in the array share neighbouring input connection weights (neighbouring "synaptic weights") in common, *i.e.* the architecture is constructed so that neighbouring synaptic weights of neighbouring neurons are equal. The result is that as objects are translated across the sensor image, the recognition responses of the neurons are correspondingly just translated across their array rather than being changed entirely beyond recognition. This trivial example of a symmetry network therefore exhibits the same translation invariance property as the standard template correlator. More sophisticated forms of synaptic weight symmetries than simple globally translation-invariant weight-sharing can be designed into neural network architectures. The rotation-invariant network of Fukumi et al [1] (applied to coin recognition), the handwritten zip-code recogniser of Le Cun et. al. [4] and the Neocognitron [2][3] all use built-in weight equalities implicitly to achieve their transformation robustness.

In all the examples in the previous paragraph, the weight symmetries are designed into the architecture at the outset (*a priori*) and are not derived by learning. Such static or "hard-wired" weight symmetries are rarely sufficient for handling transformations more sophisticated than simple global transformations of the whole sensor image such as 2-D translation, rotation and scale magnification. This is because prior knowledge of the transformation properties of individual objects is generally insufficient or too ill-defined to allow the neural network's designer to understand how to hard-wire appropriately sophisticated and complex weight symmetries into the network's architecture; in other words, *object/model-dependent* weight symmetries will be necessary for true robustness to articulation and changes in operational configuration.

Some improvement on hard-wiring fixed symmetries in this way may be obtained by constraining the network learning algorithm's equations to enforce and maintain particular groups of simple weight symmetries throughout the learning process. An example is provided by Rumelhart et al. [7], in which weight sharing constraints on the learning equations were used to distinguish T and C shapes in simple synthetic images, independently of 2-D translation and 90^{o} rotation. This approach still has the disadvantage that the network designer must understand exactly what groups of symmetry transformations are to be imposed on the learning equations, and how to do so.

In each of the approaches outlined above, the algorithm's designer imposed onto the neural network a particular, known, well-defined group of transformation invariances. In the case of ill-defined, complex, object-dependent transformations such as articulation or changes in operational configuration, the designer generally has insufficient understanding of how to impose the ill-defined invariances onto the network in the form of supervisory knowledge. To address this problem, one needs a data-driven means of acquiring robustness/invariance under these transformations, through unsupervised learning of real data representative of the particular objects that undergo these object-dependent transformations.

2.2 Symmetry-preserving networks

Our solution is based on Webber's discovery that a new class of unsupervised neural networks ("symmetry-preserving" networks) can detect invariances/symmetries in the probability distribution function (PDF) of their training data, and exploit that new functionality to develop robust response with respect to precisely those transformations. In other words, this new class of algorithms can *preserve* the symmetries of the data's PDF, in the form of matching symmetries in the trained configuration of synaptic weights and consequently in the form of matching invariances in the network's recognition output. In principle, such networks can acquire robustness with respect to all manner of ill-defined complex transformations of objects in their training data, both continuous and discontinuous discrete transformations, simply through exposure to training data containing sufficient (but far from exhaustive) numbers of exemplars of the transformed objects. These claims are proved algebraically in Webber [16]. That paper also demonstrates this new functionality using real images having the full statistical complexity of natural scenes and shows that, through exposure to natural scenes, symmetry-preserving networks can derive and explain the translation-invariance properties of the complex cells of the visual cortex. Webber [17] goes on to demonstrate that, through exposure to natural scenes, symmetry-preserving networks can generate fully translation-, rotation- and scale-invariant codes for natural images. Symmetry-preserving network algorithms tend to be algorithmically simple, with the advantage that fast hardware implementations are feasible.

2.3 Componential coding, aka combinatorial, multiple-cause or factorial coding

The unsupervised neural network learning algorithms applied here to integrated target detection and recognition are capable of adaptively deriving *componential codes* to encode/model their training data. Componential coding has been alternatively called constituent coding, multiple-cause coding, combinatorial coding and factorial coding by various papers that have illustrated the concept using simplified synthetic data, e.g. [13][9]. It has since been applied to real data, e.g. for modelling handwriting [14] and for modelling sensor signals for the purpose of machine condition monitoring [15][6]. The idea is that sensor data having enormous variability may nevertheless be modelled effectively, by factorizing the variability down into its constituent building blocks, or components. Thus, one attempts to model the various exemplars of data as variable combinations of the building-block components, in the same way that the many tens of thousands of words in the English language may be represented as various combinations of 26 letters. The adaptive learning property of the algorithm is needed to derive the appropriate building blocks from the data, because these are generally not known a priori.

Componential coding is the approach used here to model the combinatorial complexity of urban clutter. We introduce a classifier that can distinguish designated targets from clutter by comparing the sensor image against a clutter model that *combines* building block components according to a statistical framework. (Figure 4 will show a few examples of the building blocks we derive from an urban clutter training set.) This is a much more sophisticated approach than simply attempting to compare the sensor image directly against a library of target exemplars and a library of clutter exemplars, by measuring which library fits the sensor image best. Indeed, attempting simply to compare the sensor image directly against a library of clutter exemplars could never produce robust performance, because one would need an effectively infinite number of exemplars to catalogue the enormously variable complexity of possible urban clutter scenes. This is the main reason why conventional correlation-filter or template-matching methods have not been effective in discriminating targets from urban clutter. In contrast, an adaptive clutter model that can *interpolate* between training exemplars by attempting to model them as combinations of building-block components has a far better chance of being able to model the enormous variability of urban clutter sufficiently well to be useful as part of an effective classifier.

3.0 USE OF THE APPROACH

3.1 Single-layer implementation

Learning algorithms capable of symmetry preservation have been incorporated into a Bayesian density estimation framework, and can be used to produce a likelihood distribution for the training data. To date only a single-layer implementation of these algorithms has been developed and assessed for radar target detection applications. A multi-layer implementation that will perform hierarchical feature extraction is under development. Such a multi-layer implementation is likely to be necessary for ATR that is robust to extended operating conditions (EOCs).

3.2 Clutter and target likelihoods

For the problem of recognition of target subclass post-detection, these algorithms are trained on exemplars of each target subclass. This would allow any subsequent unseen training data to be compared against each trained subclass model in order to generate a separate likelihood for each target subclass. After specifying prior subclass probabilities, Bayes' theorem can then be used to classify new objects by means of posterior subclass probabilities in the conventional Bayesian manner, e.g. [12].

For the problem of integrated target detection and recognition in clutter (identify-for-detect), a likelihood distribution is learned for the clutter as well as the target. This allows target detection in clutter to be performed, by comparing the relative likelihoods with which the image chip around any given location in a SAR scene matches the target and clutter models. In such cases, training the algorithm produces:

- A set of parameters θ_c that define the likelihood map for the clutter;
- A set of parameters θ_t that define the likelihood map for the target.

For a scene x, these can be used to produce a clutter likelihood map $l_c(x|\theta_c)$, and a target likelihood map $l_t(x|\theta_t)$. Using Bayes' theorem these can be combined to provide a map of the posterior target class probability:

$$p(t|x) = \frac{\pi_t l_t(x|\theta_t)}{\pi_t l_t(x|\theta_t) + \pi_c l_c(x|\theta_c)} = \frac{l_t(x|\theta_t)}{l_t(x|\theta_t) + \frac{\pi_c}{\pi_t} l_c(x|\theta_c)} \tag{1}$$

where $\pi_t > 0$ and $\pi_c > 0$ are the prior probabilities for target and clutter respectively, subject to $\pi_c + \pi_t = 1$. Target-versus-clutter decisions can then made by comparing the posterior target class probability against a pre-specified threshold, and the value of this threshold may be varied in order to trace out a receiver-operator characteristic (ROC) curve. Note that an identical ROC curve may be produced by fixing the threshold for the posterior probability at some nominal value (say 1/2) and instead varying the ratio π_t / π_c, i.e. the ratio of the prior probability for targets against the prior probability for clutter.

4.0 EXPERIMENTAL RESULTS

4.1 Introduction

The proposed approach has been applied to the problem of recognising a particular building within SAR imagery of an urban area. Detection of specific building types could have many military uses, such as:

- Detection and identification of terrorist training facilities.
- Detection and identification of complexes associated with the development and storage of Weapons of Mass Destruction (WMD).
- Detection and identification of civilian buildings that define exclusion areas for weapon engagement, such as hospitals and schools.

The main purpose of the example is, however, to provide an initial proof of principle for the application of these new techniques to the generic problem of target detection and identification in urban clutter using SAR imagery. These techniques will in future be applied to the more specific problem of detecting military vehicles in urban clutter. In the mean-time, several factors must be taken into account when considering the relevance of the building identification results to the problem of military vehicle detection and identification. Firstly, the selected target building is larger than most military vehicles; thus, more "pixels on target" are available than is usual, which it could be argued may make the task of detecting and identifying a particular building type easier than the task of detecting and identifying a particular military vehicle type. However, a counter-argument is that size information cannot be exploited as a discriminant between the target and the clutter when the target is a building of similar size to the surrounding clutter buildings; this factor makes detecting building types more challenging than detecting vehicle types for methods that (unlike this method) rely only on size discriminants. Another counter-argument is that the target building is likely to have more features in common with the surrounding clutter buildings than a military vehicle would have in common with the surrounding clutter buildings. Thus, some features that could be used to distinguish a vehicle from a building (perhaps related to different radar cross-sections from different types of material) can no longer be exploited as discriminants when distinguishing buildings from other buildings.

4.2 Experiments

The dataset used for this demonstration consists of high-resolution SAR imagery of a built-up area. Specifically, the dataset consists of 18 SAR images of the urban scene, from a range of different aspect angles. All 18 images were 512×512-pixels in size. 10 of these were used for training, with the 10 successive training images separated by 4-degree intervals in aspect angle, and the other 8 were used for testing, with the 8 successive test images separated by 4-degree intervals in aspect angle. Alternate aspect angles were used for training and testing to ensure a proper test of generalisation over aspect angle, i.e. no test image was closer than 2 degrees in aspect angle to any training image.

Examples of the SAR imagery are shown in Figure 2, with the designated (target) building ringed. It is clear that there are many clutter objects that have returns of similar size and intensity to those from the selected target. Figure 3 displays extracted (64×64-pixel) image chips, centred on the target.

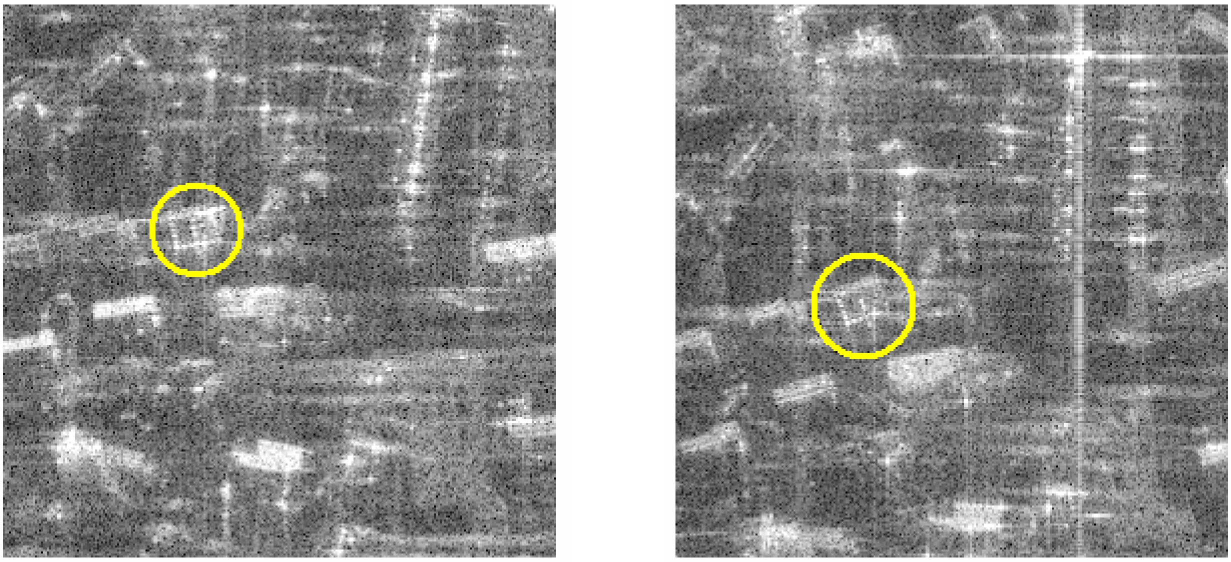

Figure 2: Examples of the SAR imagery of an urban scene, with the designated (target) building ringed in yellow. The left-hand image is view A and the right-hand image is view B.

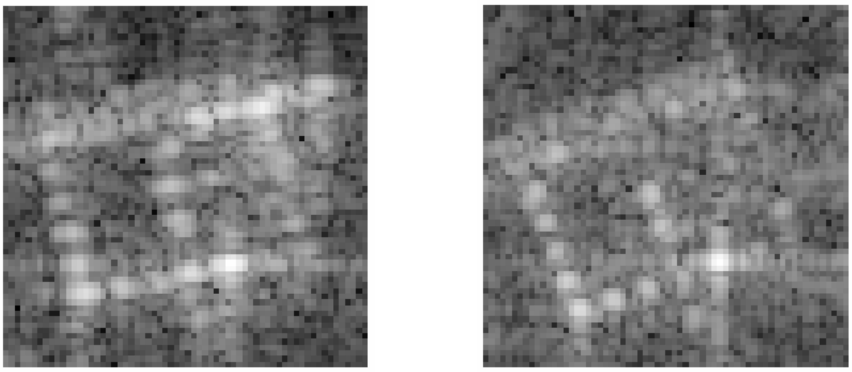

Figure 3: Examples of extracted target chips from the SAR imagery. The left-hand chip is extracted from view A and the right-hand chip is extracted from view B.

Training data for the target model were obtained by manually locating the target in each of the training images and, for each training image, extracting a 64×64-pixel image chip centred on the target. This produced 10 target-model training chips. Training data for the clutter model was obtained by sliding 64×64-pixel input windows over the entire set of training imagery. Some of the building blocks that constitute the trained clutter model are displayed in Figure 4.

Input windows that contained the target were not removed from the training data for the clutter model. The motivation for not removing the targets from the clutter training data is that it allows the clutter model to be trained on large areas of surveillance imagery, without human intervention to edit out targets. This introduces the potential to train the clutter model in situ, at the same time as the sensor platform surveys the area in which targets are to be detected. This mode of operation allows the clutter model to be refined so as to model the target's local environment optimally. The inclusion of targets in the clutter training data has minimal effect on the properties of the trained clutter model, and so does not cause significant degradation in the performance of the target-versus-clutter likelihood comparison. This is because the clutter model is trained in order to derive building blocks that best model the average properties of the bulk of the clutter training data, so the inclusion of a few targets amongst the clutter training data will bias this average insignificantly. This is in marked contrast with the effect that including targets among the clutter exemplars would have on traditional template-matching classifiers (correlation-filters); for template-matching classifiers, any target chips that pollute the clutter training data would match as would the exemplars of the target. This is another advantage of the componential coding approach over traditional template-matching classifiers, over and above the combinatorial complexity issue discussed earlier.

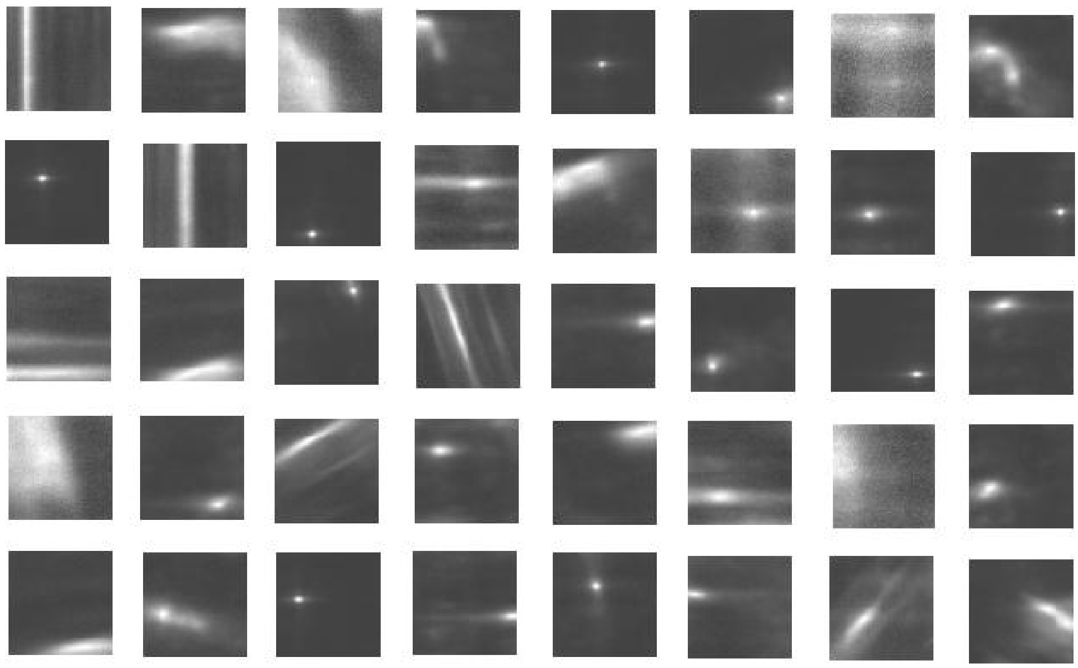

Figure 4: A subset of the building blocks that define the trained clutter model.

4.3 Results

The trained target and clutter models are used to produce target and clutter likelihood maps for each test scene. Once the target and clutter prior probabilities are specified, these likelihood maps can be used to produce a map of the posterior target class probabilities for each scene, using equation (1). Figure 5 displays the posterior class probabilities for views A and B, using a low ratio for the prior probability of a target versus the prior probability of clutter. Grey boxes have been centred on the locations for which the posterior class probabilities are higher than 0.5. For view A, the only detection box is centred on the target. For view B, there are two areas in which the posterior class probability is higher than 0.5. The upper area corresponds to the target, while the lower area is the result of a clutter false alarm.

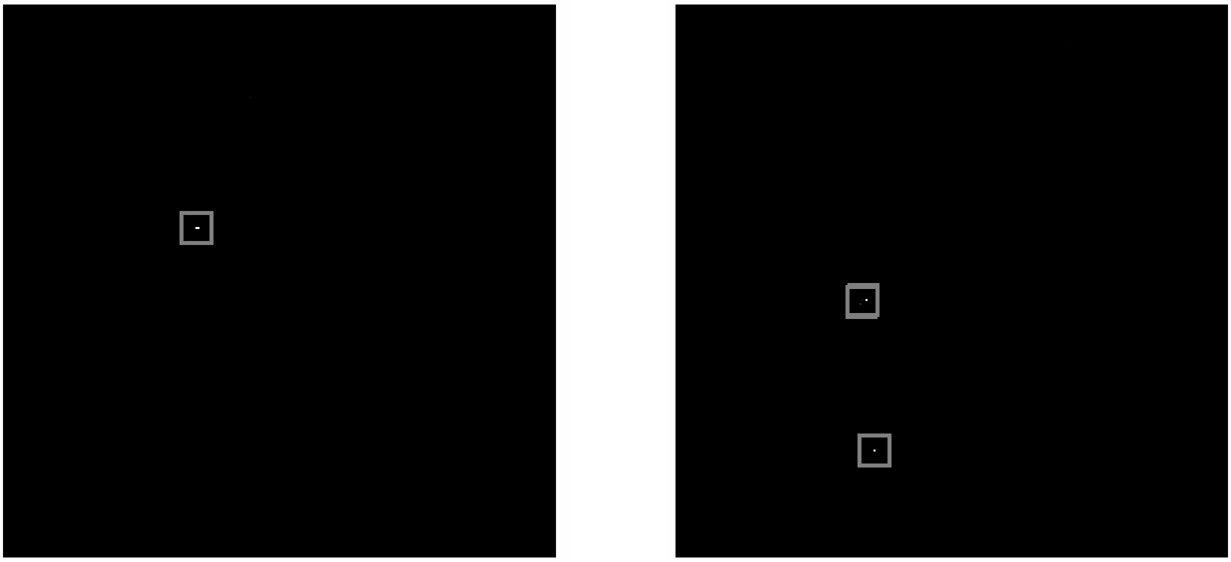

Figure 5: Posterior target class probability maps. The left-hand image is for view A and the right-hand image is for view B. Detections are surrounded by grey boxes.

Increasing the ratio of the prior probability for targets to the prior probability for clutter results in plots for the posterior target class probability that display more false alarms. (Increasing the ratio of the prior probabilities for a fixed value of the detection threshold on the posterior target probability is mathematically equivalent to reducing the detection threshold on the posterior target probability for a fixed value of the ratio of prior probabilities, as has been explained earlier.) Consequently, either varying the ratio of the prior probabilities for a fixed value of the detection threshold on the posterior target probability or varying the detection threshold on the posterior target probability for a fixed value of the ratio of prior probabilities will trace out an identical Receiver Operating Characteristic (ROC) curve for the target detection probability as a function of the false alarm rate. This ROC curve is provided in Section 4.5.

4.4 Baseline results

As an indicator of algorithm performance, two sets of baseline results are presented. Namely:

- Application of a correlation-filter (template matching).
- Application of a correlation-filter to the areas identified as anomalous by a Constant False Alarm Rate (CFAR) filter.

The templates for the correlation-filter were the same training target chips that were used to train the target model of the adaptive neural network algorithm. The same image chips as were used to train the clutter model of the neural network algorithm cannot be used as negative exemplars by the correlation-filter. This is because, for any correlator-filter classifier to function, the target would have to be manually edited out from the clutter training data, and this requirement would remove the operational potential for collecting the training data for the clutter model in-situ. More significantly, it would in general be impractical to use clutter chips as negative exemplars for a correlation-filter classifier, because an effectively infinite library of such negative exemplars would be required in order to provide robust generalisation to unseen clutter configurations. In contrast, the adaptive neural network algorithm presented in this paper avoids this problem, by using an adaptive interpolating clutter model to extract the building-block components of the clutter, and thus to model unseen clutter configurations as variable combinations of these building-block components.

The CFAR-filter assumed a Gaussian background noise distribution for the logarithms of the amplitudes of the complex-valued SAR returns.

Figure 6 displays the correlation-filter maps (i.e. the maximum correlation values over all target templates, as a function of image location) for views A and B. Both images are displayed using the same grey-scale map (i.e. the same relationship between correlator output value and pixel brightness). Comparison with the target locations in Figure 2 reveals that there is a (local) peak corresponding to the target in the correlation map for each of the two views; however, many clutter objects also give large correlation values.

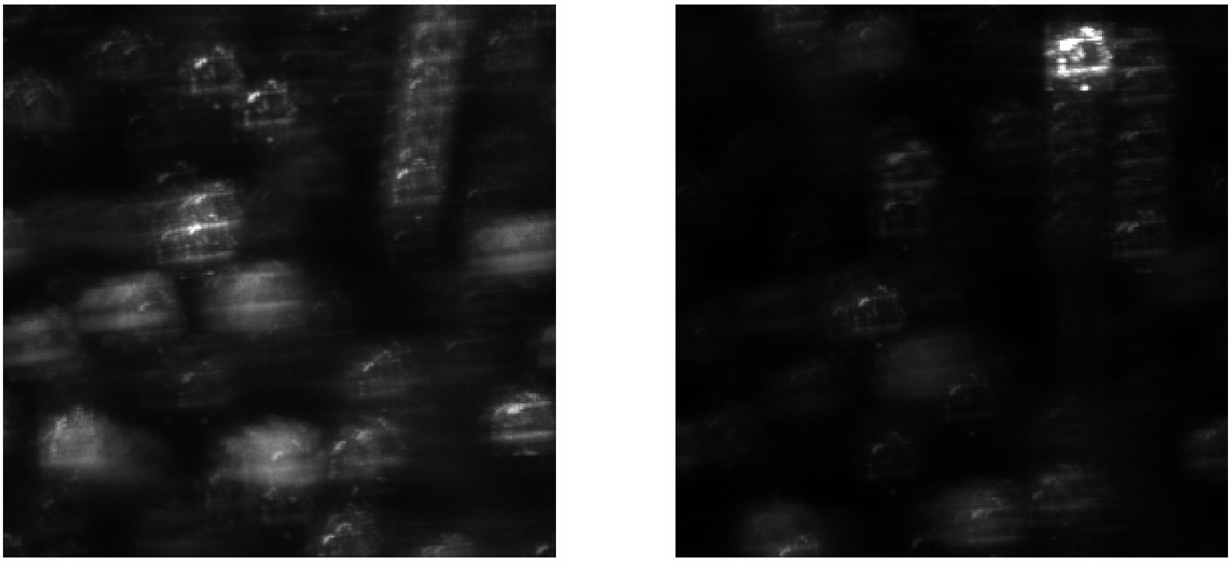

Figure 6: Maps of the outputs from a correlation-filter. The left-hand image is for view A and the right-hand image is for view B.

For the same views A and B, Figure 7 displays the outputs obtained by applying a correlation filter to the areas identified as anomalous by a CFAR filter; thus, the template-matching identification capability of the correlation filter is used to mitigate the initial false-alarm rate of the CFAR filter. Again, both images are displayed using the same grey-scale map. Comparison between the plots in Figure 7 with those in Figure 6 shows that one is able to remove many of the correlation peaks caused by clutter objects by combining the CFAR and correlation filters. Moreover, this does not seem to be at the expense of peaks at the target location. However, there are still considerably more false alarms than from the neural network algorithm (see Figure 5). The relative performances are quantified in ROC curves in Section 4.5.

4.5 ROC curve comparison

Visual comparison of Figure 5, Figure 6 and Figure 7 indicates that the adaptive neural network algorithm provides lower false alarm rates than the CFAR-filter/correlation-filter chain, which predictably provides lower false alarm rates than the correlation-filter alone. A more rigorous assessment of algorithm performance is possible by comparing ROC curves.

Figure 7: Maps of the output from a correlation-filter applied to the areas identified as anomalous by a CFAR filter. The left-hand image is for view A and the right-hand image is for view B.

To calculate these ROC curves, binary decisions between target and clutter are made for the pixels in a processed image, by applying a simple threshold to the image. If a pixel value lies above the threshold, then the pixel is declared to be a target, while if a pixel value lies below the threshold, the pixel is declared to be clutter. Targets are only counted once; exclusion zones around the targets have been applied in order to avoid counting the extremities of targets as clutter. By altering the threshold a ROC curve of target detection probability versus false alarm rate is obtained. For the adaptive neural network algorithm the processed image is the map of the posterior target class probabilities. Selection of the threshold therefore corresponds to a threshold on the posterior class probability. For the two baseline algorithms the processed images are the correlation filter outputs (applied to areas identified as anomalous by the CFAR-filter in the second case). The obtained ROC curves are display in Figure 8.

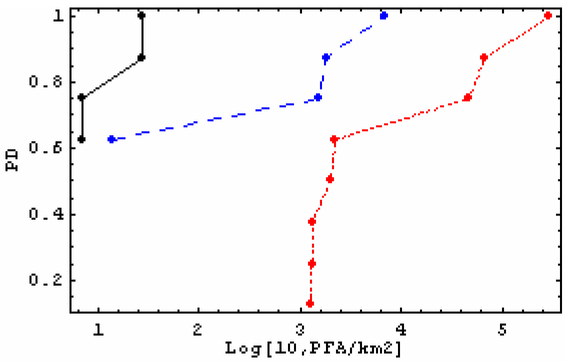

Figure 8: ROC curve comparison. Target detection probability along vertical axis, Log_{10} of false alarm rate per square km along horizontal axis. The black (solid) line is for the adaptive neural network algorithm, the blue (dashed) line is for a correlation-filter applied to the areas identified as anomalous by a CFAR filter, and the red (dotted) line is for a correlation-filter alone.

The curve for the adaptive neural network algorithm is closer to the top-left corner than the curves obtained using the two baseline techniques, indicating that better target detection and identification performance is being obtained. However, two caveats must be borne in mind when interpreting the results from this single experiment:

- Each target is only counted once, so the jumps in target detection probability are very coarse.
- No post-processing has been applied to cluster the detections, so there is some double counting of false alarms (visual inspection indicates that this double counting is likely to have more of an adverse effect for the correlation-filter baseline algorithm than the other algorithms). Clustering detections in urban clutter can be a very difficult exercise because, with so many clutter detections in the close vicinity of the target, it can be far from obvious how to design a clustering algorithm to deduce where the target detections end and the clutter detections begin.

5.0 CONCLUSIONS

This paper has introduced an adaptive neural-network approach to target- and clutter-modelling. The approach is such that both targets and clutter can be modelled within the same neural network, so that detection and recognition can take place simultaneously within an integrated framework. The algorithms can therefore be applied across the spectrum of ATR discrimination levels, e.g.: detection of unknown targets in clutter; detection of specific designated targets in clutter; recognition of target subclass post-detection. The approach builds on the unsupervised neural network principle of symmetry-preservation. Symmetry-preserving networks can detect invariances/symmetries (under complex transformations) in the PDF of their training data and exploit that functionality to develop robust responses with respect to precisely those transformations. We believe that such networks could ultimately offer the potential for target recognition with robustness to EOCs, although this has not been addressed in this paper. The paper provides an initial proof of principle for the application of componential coding to target detection and identification, using real SAR imagery in urban clutter; componential coding offers a new handle on modelling the combinatorial complexity of urban clutter. This demonstration concerned recognition of a particular building within SAR imagery of an urban area. Superior performance (in terms of target detection probability at a given false alarm rate) was obtained, compared to two baseline approaches based on a correlation-filter (template matching), one of which also exploits a CFAR filter as an initial detection stage. Planned future work will more fully assess the componential coding approach for target detection in urban clutter, and investigate the symmetry-preserving functionality with the aim of improving robustness to EOCs.

6.0 REFERENCES

[1] Fukumi M, Omatu S, Takeda F, Kosaka T, Rotation-invariant neural pattern recognition system with application to coin recognition, *IEEE Transactions on Neural Networks* 3, pp 272-279, 1992.

[2] Fukushima K, Miyake S, Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position, *Pattern Recognition* 15, pp 455-469, 1982.

[3] Himes G S, Inigo R M, Automatic Target Recognition using a Neocognitron, *IEEE Transactions on Knowledge and Data Engineering* 4(2), April 1992.

[4] Le Cun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L, Backpropagation applied to hand-written zip code recognition, *Neural Computation* 1, pp 541-551, 1989.

[5] Minsky M L, Papert S A, Perceptrons, MIT, 1969.

[6] Payne B, Gu F, Webber C J S & Ball A. Componential coding in the condition monitoring of electrical machines, Part II - Application to a conventional machine and a novel machine, *Proceedings of the Institution of Mechanical Engineers* 217(C8), pp 901-915, 2003.

[7] Rumelhart D E, Hinton G E, Williams R J, Learning internal representations by error propagation, *Parallel Distribution Processing: Explorations in the Microstructures of Cognition, ed. D E Rumelhart, J L McClelland* 1, pp 318-362, 1986.

[8] Salmond D J, Birch H, A particle filter for track-before-detect, *Proc. of American Control Conference, Arlington, VA, June 25-27, 2001*, Vol 5, pp 3755-3760, 2001.

[9] Saund E, A multiple cause mixture model for unsupervised learning, *Neural Computation* 7, pp 57-71, MIT Press, 1995.

[10] Shawe-Taylor J, Building symmetries into feedforward networks, *Proc. of First IEEE Conference on Artificial Neural Networks*, pp 158-162, 1989.

[11] Shawe-Taylor J, Symmetries and discriminability in feedforward network architectures, *IEEE Transactions on Neural Networks* 4, pp 816-826, 1993.

[12] Webb A R, Statistical Pattern Recognition, John Wiley & Sons, Chichester, 2nd edition, 2002.

[13] Webber C J S, Self-organisation of transformation-invariant detectors for constituents of perceptual patterns, *Network: Computation in Neural Systems* 5, pp 471-496, IOP Press, 1994.

[14] Webber C J S, Emergent componential coding of a handwriting-image database by neural self-organisation, *Network: Computation in Neural Systems* 9, pp 433-447, IOP Press, 1998.

[15] Webber C J S, Payne B, Gu F & Ball A. Componential coding in the condition monitoring of electrical machines, Part I - Principles and illustrations using simulated typical faults, *Proceedings of the Institution of Mechanical Engineers* 217(C8), pp 883-899, 2003.

[16] Webber C J S, Self-organisation of symmetry networks: Transformation invariance from the symmetry-breaking mechanism, *Neural Computation* 12(3), pp 565-596, MIT Press, 2000.

[17] Webber C J S, Predictions of the spontaneous symmetry-breaking theory for visual code completeness and spatial scaling in single-cell learning rules, *Neural Computation* 13(5), pp 1023-1043, MIT Press, 2001.

7.0 ACKNOWLEDGMENT

This research was sponsored by the UK MOD Corporate Research Programme.