Compact Descriptors for Automatic Target Identification

A Gammerman, A Chervonenkis, I Nouretdinov,
Royal Holloway, University of London, Egham, Surrey
J Nothard and K Smart
QinetiQ, Cody Technology Park,
Iveley Road, Farnborough, GU14 0LX

Abstract

This paper reports on the development of invariant algorithms for automatic target identification. The aim was to build invariance to target pose and other variable parameters by deriving special kernels for the support vector machine classifier. Additional invariance to line of sight and in-plane target rotations were added to the pre-processing stages, and integrated into QinetiQ's existing real-time feature-based automatic target identification processing chain. The new algorithms were evaluated using a mix of high quality synthetic electro-optic imagery and real imagery taken with a Wescam MX-15 turret flown on a Britten Norman Islander aircraft. This paper describes the derivation of the invariant algorithms, discusses the tests carried out and evaluates the performance and military benefits of the new techniques.

Keywords: invariant, automatic target recognition, identification, kernel, PATRICIA

Introduction

Automatic target identification (ATI) is a key technology for future UK airborne platforms, required by systems in theatre airspace, deep target attack and ISTAR. One of the difficulties facing ATI algorithms is that targets can present a different appearance as a function of articulation, orientation and time of day. Typically, the orientation (pose) and articulation of the target is unknown, and current ATI algorithms have to be trained to recognise these vehicles using many different views of the target, covering all possible orientations and articulations. The resulting "target descriptor" is based on large quantities of real imagery, and requires a long-winded off-line algorithm training process.

Existing systems (like DEC TA's PATRICIA ATI system) are achieving reasonable levels of performance, but not quite good enough for operational deployment. Incremental improvements have been made under previous MOD contracts by optimising the first stages of the ATI process (image processing and feature extraction). For example, the existing PATRICA ATI system has some built in invariance to target size and shift by the use of orthogonal Fourier Mellin moments (OFMM) to decompose the target image. To realise any significant improvements in performance the algorithms must be made more robust to changes in the target state or the sensing environment.

The approach taken in this project is to develop a more compact target descriptor by building invariance into the ATI process e.g. to target orientation, allowing many target views to be potentially characterised with a sparse data set. The sparse data sets need only to contain enough information to distinguish the required targets and hence form a more compact descriptor.

Improvements in the ATI processing chain

Several different methods of finding shape invariance were investigated, using a set of four non-military synthetic targets; cow, coach, Land Rover and tractor. The algorithms were then implemented and tested on additional real and high quality synthetic image sets, both containing the same nine military targets. These consisted of two large tracked vehicles, four smaller armoured personnel carriers (APC) and three other distinct vehicles including a Land Rover.

The baseline ATI system used for comparison was the feature-based automatic target identification (FATI) processing chain developed by QinetiQ [1] for DEC TA's passive automatic target recognition and identification capability integration assessment (PATRICIA) programme.

Three new invariant algorithms were developed to insert into the existing FATI processing chain. These were:

- a de-projection stage to correct for distortion due to line of sight;
- invariance to target rotation within the sensor plane by a small modification to the OFMM algorithm; and
- a special kernel for the SVM classifier invariant to small changes in target pose.

Figure 1 compares the original FATI processing chain to the new, invariant processing chain. Note that the first stage, common to both new and old processing schemes is to segment the target from the image. This simply means removing the background, by finding the target outline and preserving intensity values only for pixels within the target.

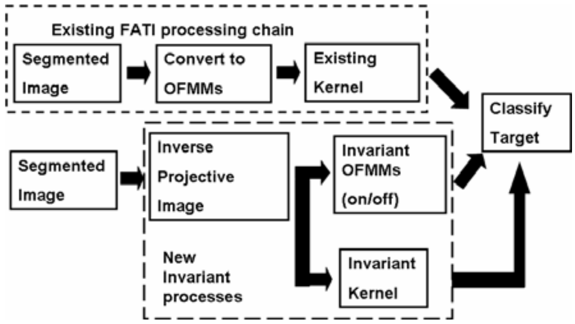

Figure 1 - Original FATI processing chain from PATRICIA ATI system (top) compared to the new, invariant process (bottom)

In practice segmentation of cluttered electro-optic imagery can be very difficult. The PATRICIA segmentation algorithm is based on the watershed technique. Where it is necessary to retrieve a very accurate target outline, e.g. for off-line training of the ATI algorithms, segmentation may be done by hand. Note the on-line automated system would not require hand segmentation; it is only necessary for these kinds of research experiments to gain a good understanding of system performance through the whole processing chain.

Inverse projective image

The aim of the inverse projection is to remove the variation in apparent target shape and size just due to the range and viewing angle of the sensor. The segmented target is fitted into a rectangle with sides S_x and S_y. Two types of pre-processing were considered:

- distance between the observer and target unknown, rescale the image to square with a diagonal equal to 2; and
- distance between observer and target known, rescale the image to rectangle with sides S*x and S*y, such that the proportion S*x : S*y is corrected by a multiplier $\sqrt{(d^v + d^h)} / d^h$ and the whole image is rescaled so that $\sqrt{(S*2x + S*2y)}$ is equal to 1.

This allows a correct projection reconstruction.

The scaled segmented target is next fitted onto a disk by bicubic interpolation; Figure 2 shows an example of a synthetic, segmented image of a tank being de-projected. Notice the gun appears in the centre end for both de-projected images but not in the original sensor images.

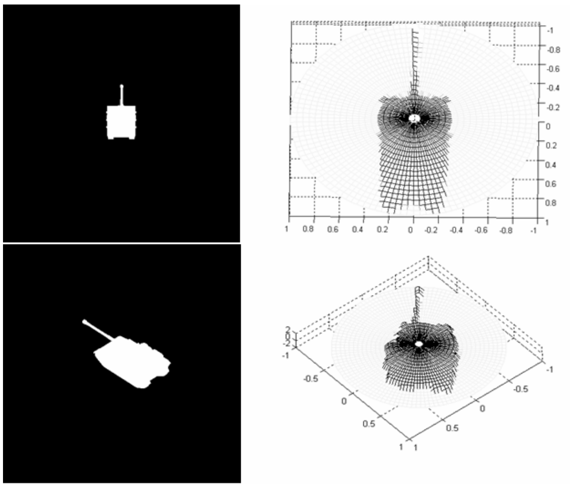

Figure 2 - de-projected synthetic tank. Original, segmented sensor images (left) are from the same range but a different viewing aspect. Corrected images (right) remove the projection effect

Invariant orthogonal Fourier-Mellin moments

Invariant OFMM [2] eliminate both scale and rotational changes from an image by simply using the absolute values or modulus of the moments, not the phase terms.

A number of experiments were performed on non-military synthetic targets to evaluate the benefit of the new algorithm. The images were split into training and test with aspects 0°-135° in the training and aspects 180°-315° in the test set.

Table 1 shows the correct classification rates using combinations of the two new invariant algorithms for the non-military synthetic target set. The classifier for these experiments was a simple 1 nearest neighbour (1NN) algorithm. Two different pre-processing methods were applied to the segmented target sets (a) altitude and slant range not used to scale (b) altitude and slant range used to scale. The best performance was found using images that were scaled using the altitude and slant range information and making use of the size of the target significantly improves performance.

Projection invariance	**Invariant OFMM**	**1NN (a) No size**	**1NN (b) No size**	**1NN (b) Size**
Off	Off	49.0	59.4	91.0
Off	On	48.0	49.9	71.9
On	Off	70.1	75.3	93.1
On	On	64.5	65.0	83.3

Table 1 – Classification rates for different combinations of invariant pre-processing algorithms

Table 2 shows the error rates for each target, and indicates which targets are most easily confused. Note the coach is very distinct but the tractor and Land Rover are occasionally confused, probably because they are similar in size.

	Coach	**Cow**	**Landrover**	**Tractor**
Coach	-	0%	0%	0%
Cow	0.8%	-	0.8%	0.8%
Landrover	0.4%	0.4%	-	0.8%
Tractor	5.4%	4.1%	24.6%	-

Table 2 - error rates for 2-class problem with 1NN algorithm

An experiment was undertaken to determine if the added scale invariance from the invariant OFMM allowed a reduction in the required amount of training imagery. Specifically, images at different ranges were removed from the training set, and the classification rate measured on the test set. Figure 3 shows the results. Most views of the target are required for robust training, indicating that the training set was already sparse in range.

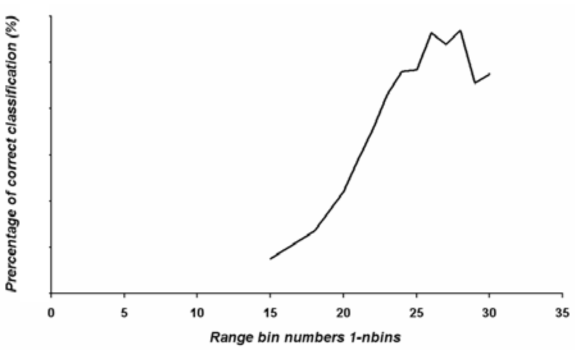

Figure 3 - Number of range bins in training set versus classification rate, for the synthetic non-military targets

To complete this experiment, only the long range images were removed from the training set, and the classification rates again measured on the test set. The results are summarised in Figure 4.

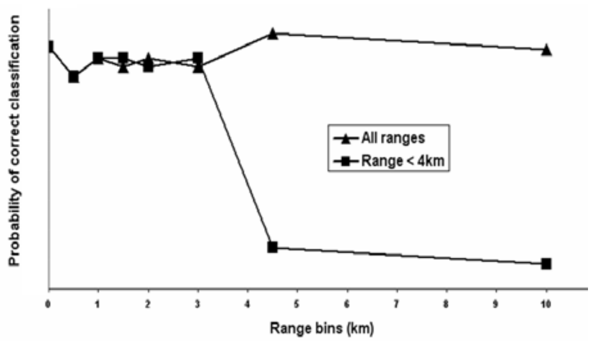

Figure 4 - Training with half aspects and testing on remaining half aspects for military synthetic targets using invariant OFMM. In the lower curve, long range images were removed from the training set

The classification performance in Figure 4 is very flat with range, implying that the misclassifications are not range dependant. The maintenance of performance with range is also exceptional, due to the target edges being well defined in the synthetic imagery.

Finally, the aspect invariance of the new invariant OFMM was compared to the original algorithm. Figure 5 shows the classification rates for aspect versus range, for a single target, with white indicating 100% correct classification and black 0%. The contour plot indicates that the drop in performance is due to different target aspects being confused.

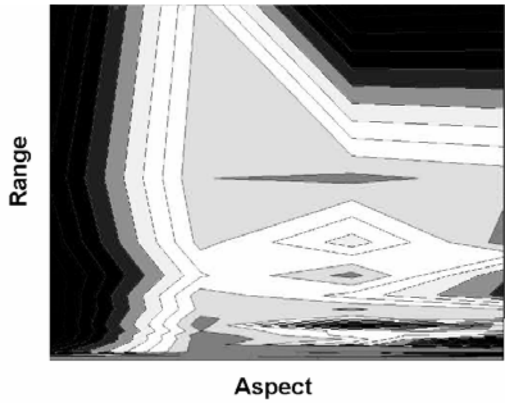

Figure 5 - Synthetic Warrior half aspect response with invariant OFMM

Figure 6 shows the same plot, but for the original phase-sensitive OFMM from PATRICIA. For this, high quality, synthetic test imagery, the response of the invariant OFMM is much more even across all angles, confirming they do provide extra robustness to target aspect.

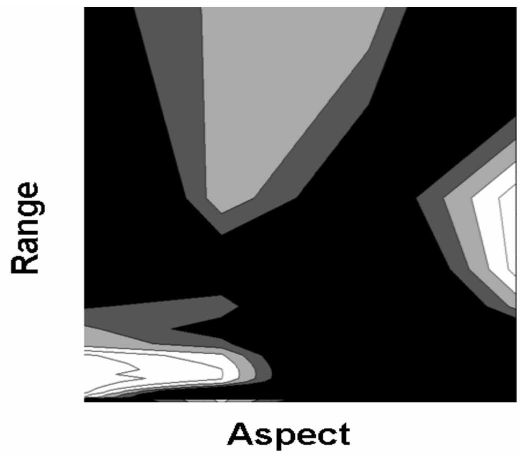

Figure 6 - Synthetic Warrior half aspect response with original OFMM

Special kernel for the SVM

The aim was to derive a new kernel with some invariance to target pose. The invariant kernel was developed using the synthetic non-military image set. The new kernel was designed to ensure all the redundant information regarding aspect and range is excluded from the target descriptor, leaving only invariant discriminative features.

The kernel was developed using several experimental criteria, including satisfying Mercer's condition [3]. Figure 7 shows an example of dividing vertical integration strips through a target. The kernel compares all integrated regions from all targets with aspect and range.

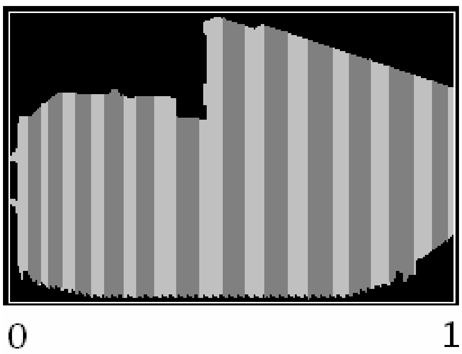

Figure 7 - Example of dividing integrating strips through a target

After several iterations the kernel was designed, based on this region integration idea, to be:

$$K(x,y) = \left[\sum_{i=1}^{k} \min\{S(X,i), S(Y,i)\} + 1\right]^3$$

The response of the kernel for the non-military targets had the best performance when not taking the size of the target into account, giving an overall performance of 90.8%, which is an improvement of 15.5 % over the standard kernel.

However, the kernel didn't perform as well on the synthetic military data set, where there are more vehicles and several are very similar. For example, two of the APCs share the same chassis and are very similar in size and shape. Figure 8 shows the confusion matrix achieved using the new kernel on the synthetic military target set. If all the targets were correctly classified there would be a strong white diagonal line and all the other squares would be black. Notice the confusion arises around vehicles of similar shape and size, such as between the small APCs and Land Rover.

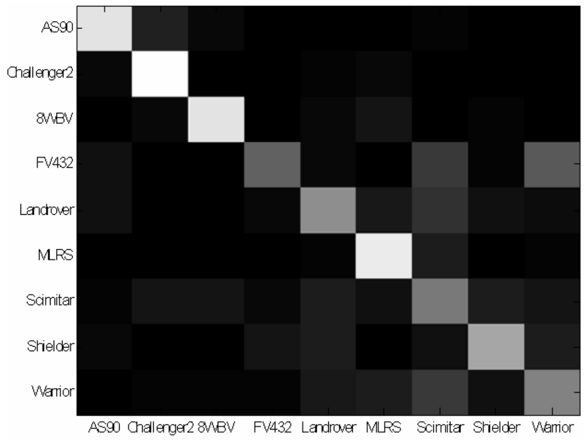

Figure 8 - Confusion matrix for synthetic military target data set and the new invariant kernel, showing that the confusion is between vehicles of similar shape and size

Finally, Figure 9 compares the performance of the new kernel, which works directly with the segmented image, to de-projecting and finding the invariant OFMM. The same improvement in performance is not seen on the military target set as was seen on the non-military set.

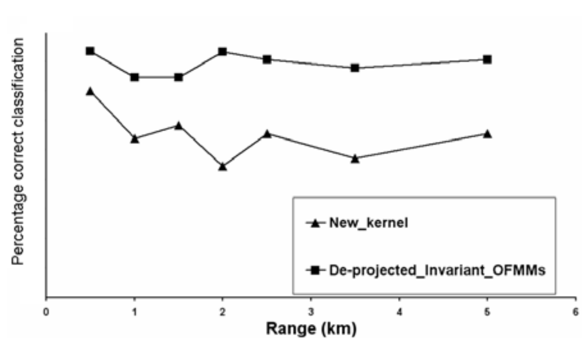

Figure 9 - Performance of the new kernel compared to the de-projected invariant OFMM

This is believed to be because the non-military targets are all quite distinct in shape and size, so the subtle information lost by building in pose invariance is not important. For the military targets, the changes due to vehicle type can be very subtle, and are removed by the invariant process as part of pose compensation.

Evaluation using real sensor data

The final part of this work examined the performance of the new invariant algorithms on real electro-optic imagery from the PATRICIA trials. The aim of the PATRICIA programme was to transition laboratory ATI algorithms into a real-time system and test them, in the air, against real military vehicles in a range of deployment scenarios.

The sensor selected was the Wescam MX-15 turret, currently in service on military platforms. For the airborne trials, it was mounted on a Britten Norman Islander aircraft, and the operator console was located in the cabin with the real-time ATI processor. Phase 1 of the flight trials gathered a complete set of training data for all nine military vehicles by imaging them at all aspects and from multiple ranges. Phases 2 and 3 gathered test imagery, of the same vehicles at a different time of year, in different weather conditions and employing a range of camouflage, concealment and deception (CC&D) measures.

The SVM classifiers trained with the synthetic data were tested on targets extracted from phase 3 PATRICIA imagery using the watershed segmentation algorithm. The response for the original FATI processing is shown in Figure 10, and for the new invariant algorithms (de-projection and invariant OFMM only) in Figure 11. Performance in both cases is low, due to the mis-match between very clean, perfectly segmented synthetic imagery and the real data.

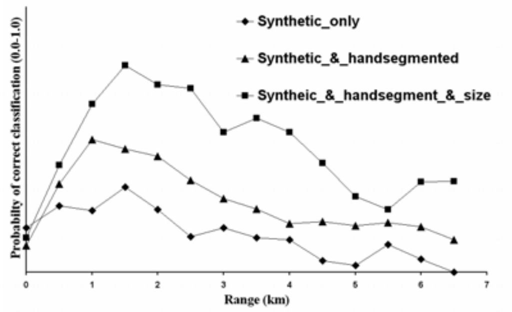

Figure 10 – Identification performance against watershed segmented targets from PATRICIA phase 3 using the existing FATI processing chain, trained on three different mixtures of synthetic and hand segmented real data

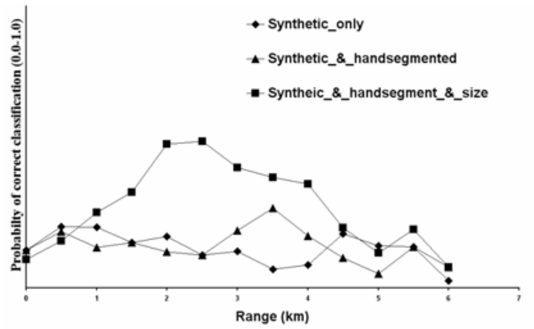

Figure 11 - Identification performance against watershed segmented targets from PATRICIA phase 3 using de-projection and the new invariant OFMM algorithm

Next the classifiers were trained with a mixture of synthetic and hand segmented data from PATRICIA phase 1. The classification performance of the original OFMM increases as shown in Figure 10 but no improvement is seen for the invariant OFMM in Figure 11.

Finally the classifier is trained with a size invariant term added to both the OFMM algorithms. The performance of the original OFMM in Figure 10 is now operationally credible. However, the invariant OFMM in Figure 11 improve but don't perform as well as the original OFMM. This is probably due to the following factors:

- automated (watershed) segmentation of the test imagery causes significant differences from the

synthetic or hand segmented training imagery;
- changes in target shape e.g. tanks have flat tracks in synthetic images, but real imaged tanks pick up the ground contours; and
- the real imagery is subjected to atmospheric variations, causing blurring and altering the thermal profile between phases 1 and 3.

The original OFMM had some robustness to these training/test mis-matches by using a slightly different implementation and retaining the phase terms. The new invariant algorithms do not have this robustness built in, so cannot realise their performance potential on this difficult real imagery.

Other programmes are looking at the segmentation of electro-optic imagery. When segmentation can be performed reliably, with no over or under segmentation and clean edges retained, the mis-match between training and test sets will be greatly reduced and the benefits of the new invariant algorithms could be realised.

Conclusions

The aim of this project was to develop a more compact target descriptor by building invariance into the ATI process e.g. to target orientation. Invariance has been added to: distortion caused by range and viewing angle; shift and in-plane rotations; and small changes in target pose by the derivation of a special kernel for the SVM classifier.

The new algorithms were extensively tested using a mix of high quality synthetic and real electro-optic imagery from the PATRICIA trials. The new kernel was very effective against dissimilar targets such as those in the synthetic non-military set, but was less effective against the more difficult military problem of distinguishing very similar vehicles such as APCs.

The de-projection and invariant OFMM algorithms gave a large performance benefit when training and testing on the military synthetic image set, indicating they had delivered extra robustness to target aspect. However, these performance benefits could not be realised on the real Wescam imagery, principally due to significant mis-matches between the training and test imagery. In particular, segmentation errors can significantly alter the target outline, and the invariant algorithms are less robust to these effects than the original FATI process.

However, segmentation is an ongoing area of research, and it is suggested that the invariant algorithms are re-evaluated when segmentation performance rises to acceptable levels.

Acknowledgements

The work reported in this paper was funded by the Electro-Magnetic Remote Sensing (EMRS) Defence Technology Centre, established by the UK Ministry of Defence and run by a consortium SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic.

References

1. Peter Randall, Jonathan Wood, Jo Nothard, Jason Cowell, Karen Smart and Dan Pearce, '*Demonstration of real-time automatic target identification from a surrogate UAV*', SPIE Vol. 6209, 62090B, 5th May, 2006
2. Yunlong Sheng and Lixin Shen, '*Orthogonal Fourier-Mellin moments for invariant pattern recognition'*, J. Opt. Soc. Am. A/Vol. 11, No.6/June 1994
3. Bernhard Schölkopf, Christopher J.C.Burges, Alexander J. Smola, '*Advances in kernel methods',* MIT Press 1998