

Intelligent selection of useful features for optimal feature-based classification

Tan Hwee Pink, Umaiyal Ramanathan
Defence Science Organisation National Laboratories
20 Science Park Drive Singapore 118230

Telephone 65 8712305/8712294 / Fax 65 8724366 / email: thweepin@dso.org.sg , rumaiyal@dso.org.sg

Abstract – In feature-based classification, each target class is characterised by a reference feature vector that comprises a combination of physical and statistical attributes. Different combinations of features are useful to distinguish amongst different target classes. In this study, an intelligent features selection method is proposed which selects features with minimum intra-class variance/inter-class variance. Classification results obtained with MSTAR data for Tanks, APCs and Trucks have shown a significant improvement in classification performance over using all measured features.

FEATURE-BASED CLASSIFICATION

When posed with an unknown target signature, a feature-based target classification algorithm matches this unknown target to one of M target classes based on minimum Euclidean distance computed over some physical or statistical features. It is assumed that each target class exhibits certain attributes / combination of attributes that are more representative of this class than other classes. Examples of physical attributes are length, width and average RCS of the target signature, while statistical attributes may include Fractal Dimension and Log Standard Deviation. Feature based classification comprises training and testing phases.

During the training phase, images comprising known targets are collected. Features are then measured from these target signatures and stored as a feature vector, F , as illustrated in Equation (1), for N features.

$$\tilde{X} = [X_1 \quad X_2 \quad \dots \quad X_N] \quad (1)$$

Each feature varies widely across data from different target classes. For example, the length measured could range from 4-5m for cars, 6-8m for tanks and >9m for trucks. In addition, different features exhibit different range of values for different targets. For example, length ranges from 4-10m while average RCS ranges between 10-40dB. In order to represent these features collectively in a feature-space, they have to be normalized to a common range of values e.g. [0,1].

Normalization is performed for each feature independently of other features. Measured values of a feature e.g. length for all training data is put in a vector and normalized to the maximum value such that all normalized values will be [0,1]. This is performed for all features based on all data belonging to candidate target classes.

For each target class, the mean feature vector is then computed by averaging over all feature vectors for this target class. Each target class is represented by its mean feature vector.

During the testing phase, features are measured from the unknown target signature and normalised to the corresponding maximum feature values obtained from training phase normalization. Classification is then performed based on the minimum Euclidean distance between this vector with the mean feature vector of each target class.

2. CONCEPT FOR INTELLIGENT FEATURE SELECTION

The classification procedure described in Section 1 uses all measured features since it assumes that they are equally useful for all classification scenarios. This is usually sub-optimal because different feature combinations are useful for discriminating against different groups of target classes. In addition, each feature within the selected combination exhibits different levels of importance.

To be useful, a feature must exhibit small intra-class variance as well as large inter-class variance. Intra-class variance translates to variability of the feature within the target class, and a small value would indicate better representation of the feature for this class. On the other hand, inter-class variance measures the separability of different target classes using this feature, and hence, a large value is desirable.

For each input data, a set of features are extracted and placed in a vector as shown in (1). This is repeated for all N_{train} training data for all classes to form the matrix as shown in (2):

$$X = \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(N_{train}) \end{bmatrix} \quad (2)$$

Each feature is normalized across all the data (i.e. column-wise in matrix X) such that all the entries of X are [0,1]. The mean feature vector, X_{bar} , is extracted by averaging along each column, and each column of X is subtracted from X_{bar} as shown in (3):

$$\tilde{X} = X - \begin{bmatrix} X_{bar} \\ \vdots \\ X_{bar} \end{bmatrix} \quad (3)$$

The covariance matrix is computed from X, where each diagonal element represents the spread of each feature across the training data i.e. it gives the interclass variance of this feature, as illustrated in (4).

$$\Sigma = X^T X = \begin{bmatrix} \sigma^2_{f_1} & \dots & \dots & \dots \\ \dots & \sigma^2_{f_2} & \dots & \dots \\ \dots & \dots & \sigma^2_{f_3} & \dots \\ \dots & \dots & \dots & \sigma^2_{f_N} \end{bmatrix};$$

$$\sigma^2_{INTER} = [\sigma^2_{f_1} \quad \sigma^2_{f_2} \quad \dots \quad \sigma^2_{f_N}] \quad (4)$$

The intra-class variance of each feature can be similarly obtained by considering training data for each target class at a time. These are averaged over the number of target classes to obtain σ^2_{INTRA} . Features with small $\sigma^2_{INTER} / \sigma^2_{INTRA}$ are selected in order to satisfy the criteria stated above.

3. DESCRIPTION AND EXTRACTION OF FEATURES

In feature-based classification, different target classes are distinguished based on a combination of physical and statistical features that are extracted from the target signature. This is illustrated in Figure 1 below, where $I(x,y)$ indicates the intensity level of pixel (x,y) and (\bar{x}, \bar{y}) is the intensity-weighted centroid of the target, which comprises N pixels. It is assumed that the aspect angle, θ , as well as major and minor axes have been determined prior to extraction of features.

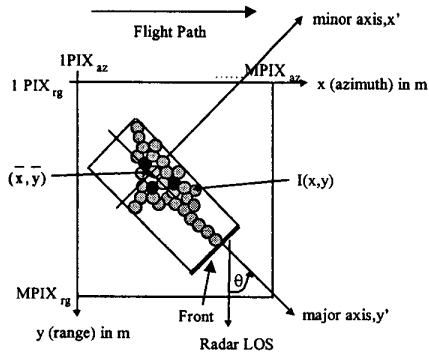


Fig. 1 : Illustration of target signature

In this study, the following features are studied.

Size

The size parameter is intuitively chosen because it can often serve as a first level coarse discrimination amongst small (e.g. Jeeps), medium (Tanks and Armoured Personnel Carrier (APC)s) and large (Trucks) land targets. This reduces the possible candidate classes in subsequent fine classification.

Log Standard Deviation

This measures the variation of RCS across the target signature. It is computed as shown in (5).

$$\sigma_{std} = \sqrt{\frac{\sum_{i=1}^N [10 \log_{10} I(x_i, y_i) - \bar{\sigma}_{dB}]^2}{N-1}} \quad (5)$$

Average Radar Cross-Section (RCS)

Average RCS gives an indication of the radar reflectivity of the target, and it is normalized against clutter level. Clutter level is measured on grass area in our study. It is computed as illustrated in (6).

$$\bar{\sigma}_{dB} = \frac{\sum_{i=1}^N 10 \log_{10} I(x_i, y_i)}{N} - clutter_{dB} \quad (6)$$

Objects with rough surface, e.g. tanks, or which contain edges are better radar reflectors than objects with smooth surfaces e.g. cars.

Fractal Dimension [1]

This parameter measures the spatial distribution of the brightest scatterers of the target.

Weighted Rank Fill Ratio (RFR) [1]

This parameter measures the % of total energy contained in the brightest scatterers of the target. Man-made targets typically exhibit large RFR as most of the energy comes from a small group of bright scatterers. Clutter, on the other hand, exhibits small RFR since the energy is more distributed over all the pixels. Hence, it is useful for discriminating between man-made target and natural clutter.

Surface level parameters

These parameters measure the compactness of the target. It is computed as shown in (7).

$$N^k \cdot PIX_{az} \cdot PIX_{rg}, \quad k=10, 50 \text{ or } 90$$

N^k : no. of target pixels whose intensity $\geq k\%$ of its peak value
 PIX_{rg} : pixel size in m in rg dir
 PIX_{az} : pixel size in m in az dir

A small variation in surface on level $\forall k$ implies a sharp response, which indicates a compact target and vice versa.

Volume level parameters

This parameter is an extension of the surface parameter. It is computed as shown in (8).

$$PIX_{az} \cdot PIX_{rg} \cdot \sum_{i=1}^L 10 \log_{10}[I(x_i^k, y_i^k)], \quad k = 10, 50 \text{ or } 90 \quad (8)$$

L: no. of target pixels whose intensity $\geq k\%$ of its peak value
 (x_i^k, y_i^k) : target pixels whose intensity $\geq k\%$ of its peak value

4. TEST DATA AND CLASSIFICATION APPROACH

The algorithms were tested using the MSTAR public domain database collected by Sandia National Laboratory and released in 1997. The images are collected at 17 degree grazing angle with 1ft resolution. The range pixel size is 0.2021m and azimuth pixel size is 0.2031m. The target classes used are the T72(Tank), BMP2(APC) and ZIL(Truck).

For training, known target signatures of each target class over an aspect angle range are trained to generate a reference feature set for this target class for the given aspect range. In our studies, training is performed independently over 6 angular sectors, namely 0-60, 60-120, 120-180, 180-240, 240-300 and 300-360°, and the respective reference feature sets are stored in a database.

For testing, to classify an unknown test image, the measured orientation is used to determine the angular sector for which the test image belongs to. Features are then extracted from the test image and matched against all potential candidate reference feature sets in this angular sector and is classified based on some distance measure.

Two experiments are conducted: The first uses all features with equal weighting when computing the reference feature vector for each target class. The second experiment uses the intelligent feature selection approach described in Section 2 to generate the reference feature vectors.

5. RESULTS

Classification results are shown in terms of confusion matrices. Figure 2 and 3 shows the results obtained with (a) all features used, (b) intelligent selection of features.

	Tank	APC	Truck
Tank	52.5	21	20
APC	35	73	10
Truck	12.5	6	70

Fig. 2 : Confusion matrix using all features

	Tank	APC	Truck
Tank	90	13	2.5
APC	10	87	0
Truck	0	0	97.5

Fig. 3 : Confusion matrix using intelligent feature selection

It is observed that average correct classification is improved from 65% to 92% when features are filtered based on their intra-class and inter-class variance for the groups of targets concerned. Table 1 shows the distribution of the features selected for all the angular sectors.

Table 1 : Distribution of selected features

Feature	Prob of Selection
Length	1
Width	1/3
LSD	1/6
RCS	1/2

Only four features were selected from a list of 15 features, which demonstrates significant filtering of useful features for optimal classification performance.

6. CONCLUSIONS

In feature-based classification, each target class is characterised by a reference feature vector that comprises a combination of physical and statistical attributes. Different combinations of features are useful to distinguish amongst different target classes, and the objective of this study is to derive an optimal way to select this set of useful features based on its variance characteristics.

Classification results have shown that optimal selection of features improves classification performance by more than 25% over using all features for the case of Tanks, APCs and Trucks.

REFERENCES

[1] L. M. Novak, G. J. Owirka and C.M. Netishen, "Performance of a High-Resolution Polarimetric SAR Automatic Target Recognition System", pp 11-24, Lincoln Laboratory Journal, Volume 6 Number 1, 1993.