DICTIONARY BASED AUTOMATIC TARGET RECOGNITION

Sathya Narayanan N*

Godwin Joe M**

Abstract—
There were different methods were developed for moving and stationary target acquisition and recognition (MSTAR) such as Non Negative Matrix Approximation (NNMA), Principal Component analysis (PCA), Linear Discriminant Analysis (LDA). But they have their lagging in the accuracy level. This paper explains that the recognition of multi view target automatically. To recognise such target without knowing the pose, a method used named Content Based Image Retrieval. In this method we can use the correlation of multiple views for the same target. It provides a greater accuracy for automatic target recognition (ATR) and convenient for different angles and configurations. It can out the neighbour of the new test image which is not found in the dictionary.

Keywords— **Automatic target recognition, Content Based Image Retrieval, Accuracy.**

* *PG Scholar* ME-Avionics, *PSN College of Engineering and Technology, Tirunelveli-152, India*

** *Associate Professor*, Department of Aeronautical Engineering, *PSN College of Engineering and Technology, Tirunelveli-152, India*

I. Introduction

Automated target tracking and recognition (ATR) is an important capability in many military and civilian applications. In this work, we mainly focus on tracking and recognition techniques for Synthetic Aperture Radar (SAR) imagery, which is a preferred imaging modality for most military applications. A major challenge in vision-based ATR is how to cope with the variations of target appearances due to different viewpoints and underlying 3D structures. Both factors, identity in particular, are usually represented by discrete variables in practical existing ATR algorithms. In this we will account for both factors in a continuous manner by using view and identity manifolds. Coupling the two manifolds for target representation facilitates the ATR process by allowing us to meaningfully synthesize new target appearances to deal with previously unknown targets as well as both known and unknown targets under previously unseen viewpoints. Common SAR target representations are non-parametric in nature, including templates, histograms, edge features etc. The target is represented by intensity and shape features and a self-organizing map are used for classification. Histogram-based representations were shown to be simple yet robust under difficult tracking conditions, but such representations cannot effectively discriminate among different target types due to the lack of higher order structure. The shape variability due to different structures and poses is characterized explicitly using a deformable and parametric model that must be optimized for localization and recognition. This method requires high-resolution images where salient edges of a target can be detected, and may not be appropriate for ATR in practical SAR imagery. On the other hand, some ATR approaches depend on the use of multi-view exemplar templates to train a classifier. Such methods normally require a dense set of training views for successful ATR tasks and they are often limited in dealing with unknown targets.

Due to its lateral and all weather view taking capabilities, SAR is an image-taking sensor particularly suited to the problem of battlefield surveillance, especially to meet a need for urgent programming. Within the context of photo-interpretation support activities, we have worked on shape recognition support algorithms for this type of imaging based, in particular, on the contribution of similarity measurements between a real SAR image and a simulated image.

Due to the unique characteristics of SAR image formation process, such as specular reflection, multiple bounces, typical low-resolution and non literal natures of the data, it is very difficult to extract linear features for recognition as used in the visible light image-based recognition tasks e.g., [1]. Templates or correlation based methods such as minimum noise and correlation energy (MINACE) filters have been developed in the past for SAR ATR [2, 3], which are designed via minimizing the correlation between the desired filters and the spectral envelope of the training images and noise at each frequency. A learning vector quantization (LVQ) method has also been introduced in [4] for SAR ATR to improve the template based SAR ATR. In that paper, the LVQ algorithm is used for learning the templates for classification. However, template based methods are not efficient for SAR image based target recognition, because the SAR image features can change dramatically with pose variations. Moreover, SAR images are very sensitive to different articulations or configurations
of the targets, which may lead to images with large
differences.

SAR images are sensitive to the variation in the pose of the targets, thus the classifiers trained with images within a specific pose interval may not perform well if the test image falls out of this interval. To alleviate this problem, several researchers have proposed to train an ensemble of classifiers on many different pose intervals [5—7]. For a test image, they first estimate the pose of the target and then perform the recognition with a specific classifier selected from the classifier ensemble via the estimated poses. However, it is known that pose estimation from SAR image itself is a very challenging problem, due to the enormous variability of the scattering phenomenology created by man-made objects [5]. Another work in [8] extended single-perspective based classifier such as K-nearest neighbors (K-NN) to handle multiple perspectives by first applying single-perspective based classifier to each view and then fusing the multiple outputs via majority-voting to get a single classification decision.

II. CONTENT BASED IMAGE RETRIEVAL

CBIR differs from classical information retrieval in that image databases are essentially unstructured, since digitized images consist purely of arrays of pixel intensities, with no inherent meaning columns is larger than the number of rows. CBIR draws many of its methods from the field of image processing and computer vision, and is regarded by some as a subset of that field. It differs from these fields principally through its emphasis on the retrieval of images with desired characteristics from a collection of significant size. Image processing covers a much wider field, including image enhancement, compression, transmission, and interpretation. While there are grey areas (such as object recognition by feature analysis), the distinction between mainstream image analysis and CBIR is usually fairly clear-cut.

Color, texture and shape features have been used for describing image content. Different CBIR systems have adopted different techniques. Few of the techniques have used global color and texture features [9],[10],[11] where as few others have used local color and texture features.

III. CONTENT BASED IMAGE RETRIEVAL IN MULTIVIEW ATR

A. *Steps for Retrieval*

Grid:

An image is partitioned into 24 (4 x 6 or 6 x 4) non overlapping. These tiles will serve as local color and texture descriptors for the image. Features drawn from conditional co-occurrence histograms between image tiles and the corresponding complement tiles are used for color and texture similarity. With the Corel dataset used for experimentation (comprising of images of size either 256 x 384 or 384 x 256), with 6 x 4 (or 4 x 6) partitioning, the size of individual tile will be 64 x 64. The choice of smaller sized tiles than 64 x 64 leads to degradation in the performance. Most of the texture analysis techniques make use of 64 x 64
blocks. This tiling structure is extended to second level decomposition of the image. The image is decomposed into size M/2 x N/2, where M and N are number of rows and columns in the original image respectively.

Integrated Image Matching:

Since at any given level of decomposition the number of tiles remains the same for all the images all the tiles will have equal significance. A similar tiled approach is proposed, but the matching is done by comparing tiles of query image with tiles of target image in the corresponding positions. In our method, a tile from query image is allowed to be matched to any tile in the target image. However, a tile may participate in the matching process only once. A bipartite graph of tiles for the query image and the target image is built The labeled edges of the bipartite graph indicate the distances between tiles. A minimum cost matching is done for this graph. Since, this process involves too many comparisons, the method has to be implemented efficiently. To this effect, we have designed an algorithm for finding the minimum cost matching based on most similar highest priority (MSHP) principle using the adjacency matrix of the bipartite graph. Here in, the distance matrix is computed as an adjacency matrix. The minimum distance d_{ij} of this matrix is found between tiles i of query and j of target. The distance is recorded and the row corresponding to tile i and column corresponding to tile j, are blocked (replaced by some high value, say 999). This will prevent tile i of query image and tile j of target image from further participating in the matching process. The distances, between i and other tiles of target image and, the distances between j and other tiles of query image, are ignored (because every tile is allowed to participate in the matching process only once). This process is repeated till every tile finds a matching. The complexity of the matching procedure is reduced from $O(n_2)$ to O(n), where n is the number of tiles involved. The integrated minimum cost match distance between images is now defined as:

$$D_{ij}=\sum\sum d_{ij}$$

where d_{ij} is the best-match distance between tile i of query image q and tile j of target image t and *qt D* is the distance between images q and t.

Shape Size

Shape information is captured in terms of the edge image of the gray scale equivalent of every image in the database. We have used gradient vector flow (GVF) fields to obtain the edge image.

The GVF field gives excellent results on concavities supporting the edge pixels with opposite pair of forces, obeying force balance condition, in one of the four directions (horizontal, vertical and diagonal) unlike the traditional external forces which support either in the horizontal or vertical directions only. The algorithm for edge image computation is given below:

Algorithm: (edge image computation)

1. Read the image and convert it to gray scale.
2. Blur the image using a Gaussian filter.
3. Compute the gradient map of the blurred image.
4. Compute GVF. (100 iterations and $\mu = 0.2$)
5. Filter out only strong edge responses using *k*s , where s is the standard deviation of the GVF. (k – value used is 2.5).
6. Converge onto edge pixels satisfying the force balance condition yielding edge image.

B. Experimental Setup

Data set: The images are of the size 256 x 384 or 384 x 256.

Feature set: The feature set comprises color, texture and shape descriptors computed as follows:

Color and Texture: Conditional co-occurrence histograms between image and its complement in RGB color space provide the feature set for color and texture

Our proposed method is an extension of the co-occurrence histogram method to multispectral images i.e. images represented using n channels. Co-occurrence histograms are constructed, for inter-channel and intra-channel information coding using image and its complement. The .complement of a color image $I = (R,G, B)$ in the RGB space is defined by $I = (255 - R,255 - G,255 - B)$ º (R,G, B) . The nine combinations considered in RGB color space are: (R , G), (G , B), (B , R), (R , G), (G , B), (B, R), (R, R), (G,G) & (B, B) , where R, G & B represent the Red Green and Blue channels of the input image and R,G & B represent the corresponding channels in the complement image. The translation vector is t[d,a] where d is distance and a is direction. In our experiments we have considered a distance of 1 (d=1) and eight angles (a=00, 450, 900, 1350, 1800, 2250, 2700, 3150). Two co-occurrence histograms for each channel pair, for each of the eight angles, are constructed using a max-min composition rule, yielding a total of 16 histograms per channel pair. Then the histograms corresponding to opponent angles are merged yielding a total of 8 histograms per channel pair i.e. 00 with 1800, 450 with 2250, 900 with 2700 and 1350 with 3150. The feature set comprises of 216 features in all with 3 features each computed from the normalized cumulative histogram i.e. 9 pairs x 8 histograms x 3 features.

Method of Obtaining Histogram:

1. A pixel r in R plane and a pixel g in the corresponding location in G plane are shown above with their immediate eight neighbors. The neighboring pixels of r and g considered for co occurrence computation.

2. Consider two histograms H1 and H2 for R based on the maxmin composition rule stated below:

Let _ = max(min(r,g1), min(g,r1))
Then, r Î H1 if _ = min(r,g1)
and r Î H2 if _ = min(g,r1)

It yields 16 histograms per pair, 2 for each direction.

Feature computation:

The features considered are:

a. The slope of the line of regression for the data corresponding to the normalized cumulative histograms .

b. The mean bin height of the cumulative histogram.

c. The mean deviation of the bins.

A total of 216 features are computed for every image tile (per resolution).

Shape: Translation, rotation, and scale invariant one-dimensional normalized contour sequence moments are computed on the edge image.

The z(i) is the set of Euclidian distances between centroid and all N boundary pixels of the digitized shape.

A total of 12 features result from the above computations. In addition, moment invariant to ranslation, rotation and scale is taken on R, G and B planes individually considering all the pixels.

IV. RESULT ANF DISCUSSION

From the Figure1 the extraction of feature for a training image. After extracting it is added to the dictionary.

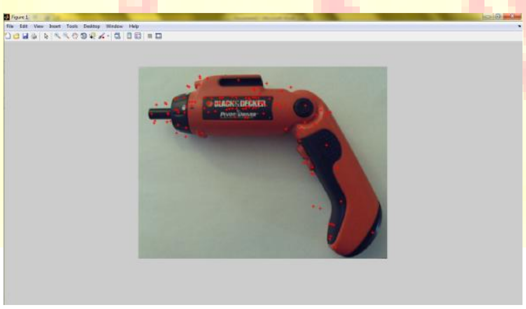

Figure 1

From the Figure2, it explains that the extraction of features of test image.

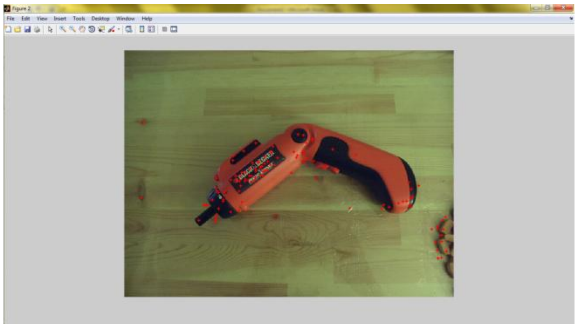

Figure 2

From the Figure3 and Figure4, comparing the test and trained images by using the extracted features respectively.

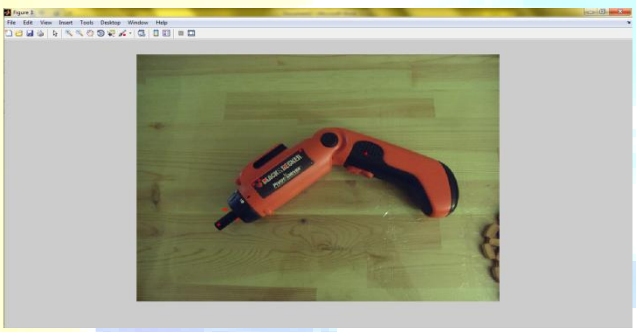

Figure 3

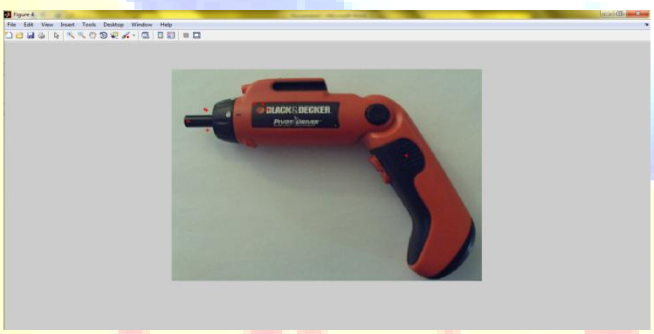

Figure 4

From the Figure5, matching the test and trained image from the dictionary

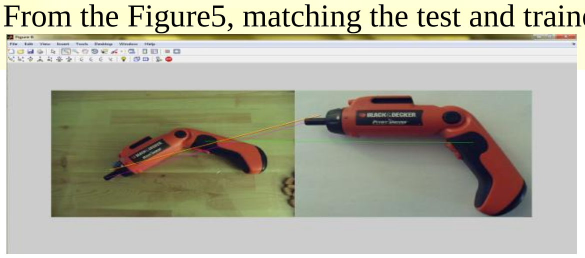

Figure 5

From the Figure6, the retrieval of new image and compared it with neighboured image.

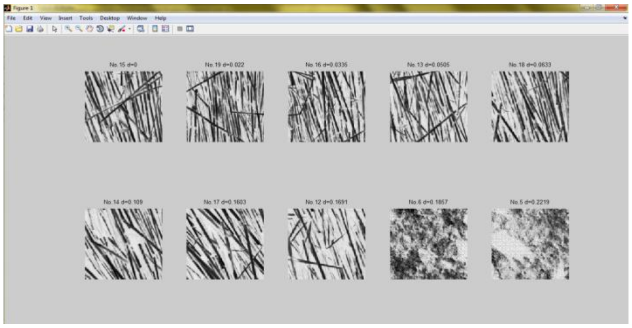

Figure 6

V CONCLUSIONS

The Content Based Image Retrieval for multi view and multi target recognition automatically is presented in this paper. This method can handle pose variation in the test images without requiring the poses to be known. Also, it can utilize multiple images of the same physical target from different views for a joint target recognition in order to improve the recognition performance.

In this proposed method, the dictionary is constructed as the training data directly. A better approach would be to learn a compact dictionary from the training data, which can give comparable or better performance, while with much fewer atoms.

And the retrieval of new image and compared with new image is integrated in this paper. The future is going on for the recognition and retrieval of video data.

REFERENCES

[1]. Jones, III, G. and Bhanu, B. Quasi-invariants for recognition of articulated and non-standard objects in SAR images. In *IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, 1999.

[2]. Patnaik, R. and Casasent, D. SAR classification and confuser and clutter rejection tests on MSTAR ten-class data using MINACE filters.

[3]. *Optical Pattern Recognition*, **6574** (2007), 657402.1—657402.15. Patnaik, R. and Casasent, D. MINACE filter classification algorithms for ATR using MSTAR data. In *Proceedings of SPIE*, vol. 5807, 2008, pp. 100—111.

[4]. Marinelli, A. M. P., Kaplan, L. M., and Nasrabadi, N. M. SAR ATR using learning vector quantization. In *Proceedings of SPIE*, vol. 3647, 1999.

[5]. Principe, J. C., Xu, D., and Fisher, III, J. W. Pose estimation in SAR using an information theoretic criterion. In *Proceedings of SPIE*, vol. 3370, 1998, pp. 218—229.

[6]. Zhao, Q., Xu, D., and Principe, J. C. Pose estimation for SAR automatic target recognition. In *Proceedings of the Image Understanding Workshop*, 1998, pp. 827—832.

[7]. Sun, Y., et al. Adaptive boosting for SAR automatic target recognition. *IEEE Transactions on Aerospace and Electronic Systems*, **43**, 1 (2007), 112—125.

[8]. Vespe, M., Baker, C. J., and Griffiths, H. D. Radar target classification using multiple perspectives. *IET Radar, Sonar & Navigation*, **1**, 4 (2007), 300—307.

[9]. W. Niblack *et al.*, "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape," in *Proc. SPIE*, vol. 1908, San Jose, A, pp. 173–187, Feb. 1993.

[10]. A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Content-based Manipulation of Image Databases," in *Proc. SPIE Storage and Retrieval for Image and Video Databases II*, San Jose, CA, pp. 34–47, Feb. 1994.

[11]. M. Stricker, and M. Orengo, "Similarity of Color Images," in Proc. SPIE Storage and Retrieval for Image and Video Databases, pp. 381-392, Feb. 1995.