Robust Automatic Target Detection/Recognition System For Second Generation FLIR Imagery

Huaibin Zhao, Shishir Shah, Jae Hun Choi, Dinesh Nair, J.K.Aggarwal
Computer & Vision Research Center,
Dept. of Electrical & Computer Engineering, ENS 522
The University of Texas at Austin, Austin, TX 78712-1084
{hzhao, shishir, jchoi}@rhine.ece.utexas.edu, aggarwaljk@mail.utexas.edu

1. Abstract

Automatic target detection and recognition (ATD/R) is of crucial interest to the defense community. We present a robust ATD/R system developed at the CVRC at UT-Austin for recognition in second generation forward looking infrared (FLIR) images. An experiment conducted on 1,930 FLIR images shows that this ATR system can achieve recognition with a high degree of accuracy and a low false alarm rate. This demo first presents a brief overview of the whole methodology, then shows the detailed procedures and temporary outputs step by step, by running this ATR system on typical low-contrast FLIR images. Results and examples are presented at the end of the demonstration.

2. Description

This ATD/R system is designed for FLIR imagery, applying a general methodology for automatic target detection, segmentation, and recognition in complex environments. The input to the system is an outdoor field FLIR image, in which the military target is at a distance of 2,000 meters to 3,500 meters from the sensor, occupying fewer than 5% of the pixels in the whole image. The system will determine if the image contains any targets, and then output the class of each target (tank, truck, missile launching vehicle, etc), the specific type of the target (e.g. M60 tank) and its pose or aspect angle.

The ATD/R system is composed of four sequential modules, including preprocessing, detection, segmentation, and recognition, shown in Fig 1.

The quality of the input FLIR image is enhanced in the preprocessing stage by removing or reducing the sensor noise. Then a detection system searches for targets by finding the regions of interest(ROIs) that are likely to include tactical targets. The focus in this stage is to achieve a high detection rate while maintaining a low false alarm rate. Except for the rough distance from the target to the sensor, no significant prior knowledge about the scene and objects is used. Edge information and heuristic knowledge such as region complexity, saliency and regularity are analyzed and used as criteria to find all the candidate ROIs[5].

The objective of the segmentor is to get a good representation of the target boundary that is used for the feature extraction stage and the recognition process. A focused analysis of each candidate target location found by the detector is performed[3]. By modeling the background using Gaussian and Weibull functions, most of the pixels belonging to the background are removed from the detected regions. Furthermore, a region-growing procedure that uses a diffusion process, driven by the underlying probability distribution of the background and modulated by local shape changes of the target, combined with salient edge information, is used repeatedly to get an estimate of the target shape. A Bayesian classification scheme is used here to combine region and edge information[4]. To further reduce the false alarm rate, a higher level interpretation module classifies the segmented areas as manmade or natural objects using geometric and FLIR-intensity based features extracted from the target.

The target classification is achieved by a hierarchical, modular recognition strategy based on shape representation and parts analysis[1]. This hierarchical structure gets better recognition results in the presence of occlusion and poor segmentation. Zernike moments are used to represent the object shape knowledge, while several modular units are created for different object parts. Each part-module is actually an independent expert trained to recognize one specific part, which is
modeled as a mixture density of multivariate Gaussian distributions. The mixture density parameters are obtained using the Expectation-Minimization (EM) algorithm. Each module has a weight associated with it, which relates to the importance of the part in the recognition process. The relation between the target parts is encoded using a Markov Random Field framework. A sequential Bayesian approach is taken to incorporate all the decisions made by the part experts, first discriminating different classes of target, then specific target type, and finally the pose, thus achieving recognition.

The entire system was tested on 1,930 FLIR images from 28 different datasets, ranging in quality from poor to good to excellent. All the images were obtained outdoors under various ambient scene and weather conditions. 89% of the targets are correctly located in the detection stage, with a false alarm rate of less than 5%. 90% of the detected regions can be correctly segmented. For most datasets, 70% of the targets’ types were correctly reported, with an 80% pose recognition rate.

Typical results of the developed ATD/R system on FLIR images are shown in Fig 2 & Fig 3. The detection, segmentation and recognition results are shown as overlay on the input image. The ROIs located by the detector are indicated by white blocks, while the segmentation of these regions are shown in a black background. The recognized type, specific class and pose are presented as text.

See http://rhine.ece.utexas.edu/~zhao/atr.htm or the following papers for more information, results, and complete references.

<table>
<thead>
<tr>
<th>Table 1. Performance results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input NO.</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>1930</td>
</tr>
</tbody>
</table>

Reference: