Object Detection - Overview

This document outlines the high-level algorithm employed by the CUAV project that identifies regions of the image that differs significantly from its background.  The algorithm is statistical that processes each image independent of the previous results. In general, the algorithm identifies regions of interest based on spectral and spatial analysis.  Future improvements would take into account textural analysis.

The algorithm can be thought of as a sequence of filters where data is pruned out based on a set of parameters:

*HistogramCountThreshold (int): a threshold to delineate background pixels where pixel-count above which have a high-degree of spectral repetition*

*MetersPerPixel (float): the size of each pixel in meters, which depends on the sensor resolution, lens focal length and altitude from the ground*

*MaxRegionArea (float): the maximum area of interest in meter-squared*

*MinRegionArea (float): the minimum area of interest in meter-squared*

*MaxRegionSize (float): the maximum length of target in meter*

*MinRegionSize (float): the minimum length of the target in meter*

*MinRegionScore (int): indicates the rarity of a region with a value from 1 - 1000, where a higher value indicates a region that is significantly more distinct from the background*

Spectral Analysis

At the very onset, the image is quantized from a 24-bit RGB (8-bit per color) down to 9-bit RGB (3-bit per color) to generate a histogram where like-color pixels are grouped in the same quantization bin. While 9-bit RGB works well for detecting regions/targets with a distinct or high-degree of spectral variation as compared to the background, a higher bit value may otherwise be better suited for regions/targets with subtle or low-degree of spectral variation as compared to the background. As of now, all work has been done using the 9-bit RGB.  Future improvements could use other bit-values given the region/target spectral range as compared to the background terrain types.

Once the histogram is generated, the *HistogramCountThreshold* is used to remove the background by setting pixels, whose quantized group higher than the threshold, to zero.  In a data-driven model, only non-zero pixels are processed. A higher threshold value will include more pixels and potentially result in more false-positive regions/targets detected, while a lower-value will exclude pixels that may result in more false-negative missing the region/target of interest.

Spatial Analysis

With the background pixels removed, the remaining foreground pixels now represent potential regions of interest. The next step groups all neighboring foreground pixels as a 2-dimensional regions and adjoining regions are merged to form a larger region. Depending on the complexity of the scene, there may be hundreds to several thousands of regions that make up the foreground pixels. Most, if not all, of the regions may represent false-positive.  In an effort to maximize relevant results, the final step is to eliminate all regions from the result that lie outside of the specified MaxRegionSize, MinRegionSize, MaxRegionArea and MinRegionArea configuration parameters.  These parameters depend on a priori knowledge of dimensions of regions/targets of interest.  For best results, it is advised to give a buffer zone when settings these parameters in the event the HistogramCountThreshold may include/exclude pixels that impact the final dimensions of the merged region.

Scoring Regions

The result of the spatial analysis may still contain many dozen to several hundred regions.  To further reduce false-positive, the final step involves generating a score for each region.  The MinScoreRegion is used to further filter out regions below which are removed from the result set.

A region score is the average of the rarity for all the pixels in the given region.  The rarity of each pixel is the delta between the HistogramCountThreshold and the histogram count of the quantized pixel, normalized to 1000. The rationale behind the pixel rarity indicator is that the lower the quantized pixel count, the more distinct it is from the background. While a higher MinScoreRegion always guarantees finding the most distinct regions in an image, but also consider that often times targets of interest may be contained within a much larger merged region hence potentially lowering the averaged region score.



