Big Cat (Tiger) Detection – Final Project Report

Namrata Deka
Indraprastha Institute of Information Technology, Delhi
namrata13065@iiitd.ac.in

Avikal Kohli
Indraprastha Institute of Information Technology, Delhi
avikal13130@iiitd.ac.in

Abstract

This is the final report for the work done as part of the course project on Computer Vision, 2016. All work done and results will be summarized.

1. Introduction

Our aim was to be able to segment out tigers from forest camera-trap images. This problem is challenging due to the highly textured nature of tiger skin as well as forest backgrounds due to which usual intensity gradient-based segmentation methods tend to get stuck in the textures and as a result are not able to segment the actual object boundaries. To tackle this problem we took some inspiration from [1] and based our segmentation on edge probability maps obtained from a modified version of geodesic active contours and statistical prior energy term which makes use of Gaussian Mixture Models. We then followed up with a few morphological actions to obtain the final result. Section 2 will describe the statistical prior energy term followed by the edge probability map in section 3. Section 4 covers the morphology operations and final results.

2. Statistical Prior Energy

We have used Gaussian Mixture Model as a segmentation technique to separate the foreground and background of the image. A Gaussian Mixture Model is the weighted sum of component Gaussian densities. The parameters used by the model are weight, mean and co-variance of the components. The parameters are collectively represented by the notation $\lambda = \{w_i , \mu_i , \Sigma_i\}$ $i = 1, . . . , M$

To get the proper results the same image is used iteratively as training data. We have used the Maximum A Posteriori (MAP) Parameter Estimation. The first step of this algorithm is that it estimates the statistics from the first model. The statistics here are the weight, mean and variance. Then the statistics are adapted and the new statistics and then combined with the old statistics from the prior parameters using a mixing coefficient. The final parameters are then used to determine the foreground and background.

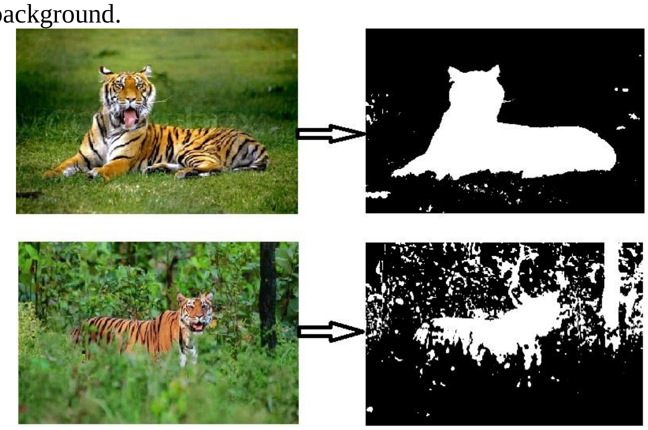

Figure 1: Results of foreground-background separation using Gaussian Mixture Models

3. Edge Based Energy term

The edge-based active contour model based on gradient edge indicator function is not applicable for natural scene images rich in texture and clutter like tiger images in the wild are expected to be. Here we calculate an edge indicator based on the local distribution of color class labels. The color class labels are generated by extracting the dominant color mode of the image and assigning each pixel the label of the corresponding dominant color. After that the pixels are clustered into the dominant color modes as per their labels. This dominant color mode image is used to construct the edge indicator function such that the indicator is high near boundaries of color-texture homogenous regions and low elsewhere.

3.1. Extracting Dominant Color

As mentioned in [1], this part is performed in the CIE Lab color space due it being perceptually uniform. The procedure for this is shown in figure 1.

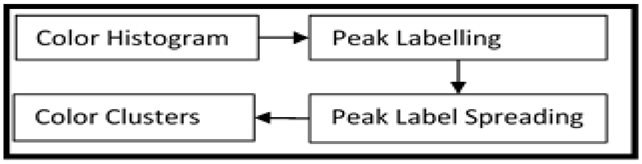

Figure 1: Procedures involved in Dominant Color Extraction

Peak Labelling is done by finding all the peaks in the color histogram and assigning each with a unique label.

The remaining bins are clustered iteratively with a label spreading procedure where labels are spread simultaneously from the peaks such that labels of the peak closer to a bin is likely to arrive first. If multiple peak labels arrive at the same time then the peak with higher intensity value is considered for labelling. This method seeks the shortest ascending route to a local maximum.

All bins with the same label belong to the same cluster. Next, we calculate the weighted mean dominant color for each cluster and the dominant color map is obtained by assigning the pixels in the original image the value of the dominant mean color intensity of the cluster to which that pixel belongs. The results are shown in figure 2.

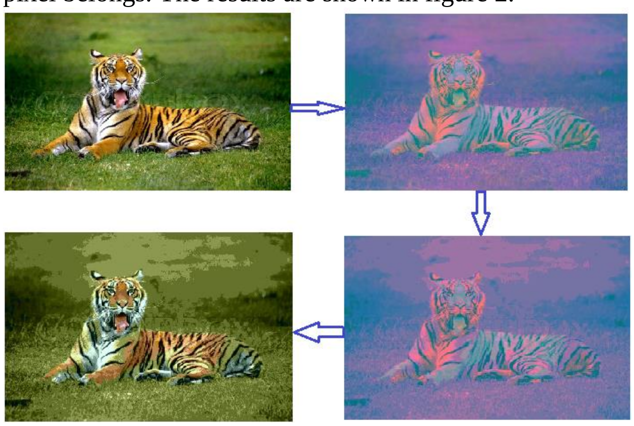

Figure 2: Top left - original image followed by image in CIE Lab space followed by dominant mode in Lab space followed by corresponding result in RGB space.

3.2. Constructing Edge Indicator Function

We used pixel variance to construct the indicator function as described in [1]. This is to incorporate the property that variance near edges tend to be higher than elsewhere, which is what we want in our edge indicator function. S_T is the variance image of the entire dominant color mode image and S_W is the total variance of pixels belonging to the same color mode. The edge indicator J is then defined as: $J = (S_T - S_W)/S_W$

The following figure shows that J is large near boundaries.

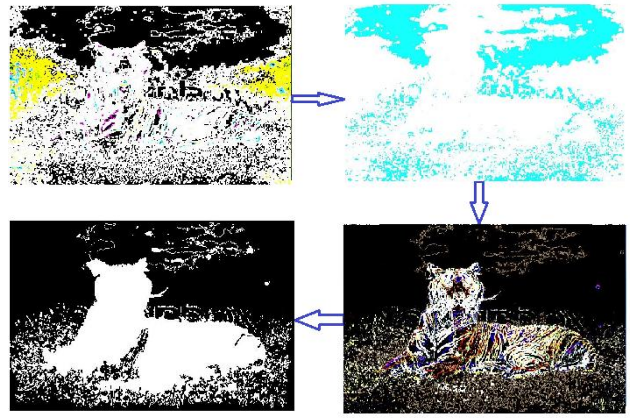

Figure 3: Top left - S_T followed by S_W follows by J. Bottom left - result of active contour on J

3.3. Geodesic Active Contour

Our final goal is to obtain E_e, which is the result of applying the geodesic model but modified by using the complement of J as the edge stopping function. We use the complement because we want the energy function to minimize at the boundaries and J was large near the boundaries. Traditionally, the stopping function is an inverse function of the gradient of the image, but as mentioned earlier that scheme will not work with highly textured images and so we used the variance of the dominant color mode image to construct the indicator and stopping functions. Figure 4 shows some results of this step.

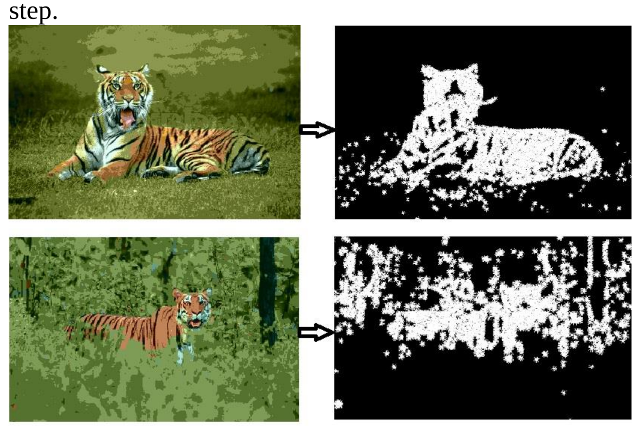

Figure 4

4. Morphology and Results

Finally the two results from the edge and the statistical energies where combined with more weight to the statistical term than edge. Results are illustrated in the figure 5.

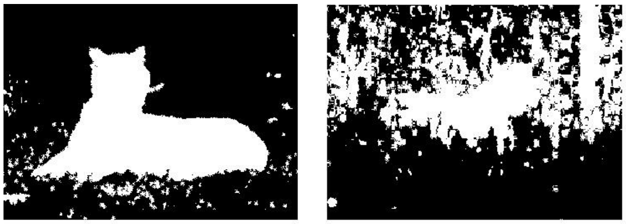

Figure 5

After that we applied the morphological operation of opening to get rid of some of the clutter from the background. Results are illustrated in figure 6.

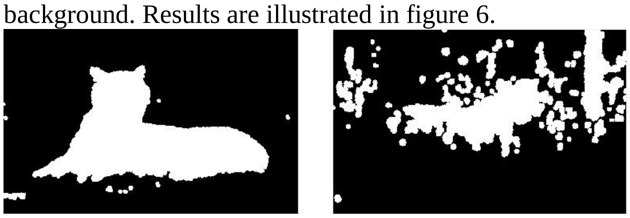

Figure 6

We applied on final morphology operation to extract the boundaries from the above images and overlaid them on the original images to check for accuracy. Results are shown in figure 7.

Figure 7: Final results

References

[1] Wang, Tinghuai, Bo Han, and John Collomosse. "TouchCut: Fast image and video segmentation using single-touch interaction." *Computer Vision and Image Understanding* 120 (2014): 14-30.