Accurate Image Segmentation using Gaussian Mixture Model with Saliency Map

Navya and Saksham Rathi

April 2024

1 Introduction

GMM is a flexible tool for image segmentation and image classification. But a major drawback in it is that it doesn't consider spatial information present in the image. The research paper on the basis of which this report has been generated has tried to come up with a method so that we can incorporate spatial information as well.

2 Saliency Map

The saliency map reflects the regions of an image, which can present an interest in the sense of visual per-ception. It highlights the pixels, which can potentially contain information to be used in a more complex image classification scheme.

This is how Saliency Map is calculated: L(f) is log spectrum of image

$$h_n(f) = \boxed{\begin{matrix} \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \\ \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \end{matrix}}$$

A(f) - Average Spectrum, S(x) - Saliency Map, G(x) - Gaussian Blur, F - Fourier Transform

$$A(f) = h_n(f) * L(f)$$

$$R(f) = A(f) - L(f)$$

$$S(x) = G(x) * F^{-1} exp[R(f) + P(f)]^2$$

The following images will show that why saliency map indeed capture salient features.

As we can see that those locations are captured where there is gradient in the image. The pixel intensities at the edge location are very high as compare to other pixels. Informally saying, the saliency map has captured the outline of different objects in the original image.

We also did some experiments on the method of calculating saliency maps. By varying the parameters we were able to exaggerate the boundaries. All this helps at times when we want minute details to be reflected in segmented image. The images below are to illustrate a few of such experiments.

Figure 1: Original Saliency Map

Figure 2: After Exponentiation of Features

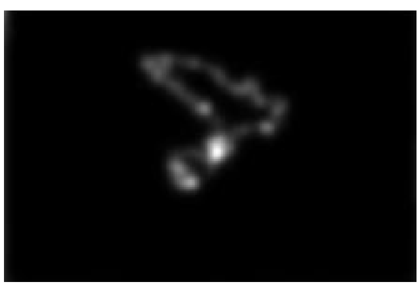

Figure 3: Original Saliency Map

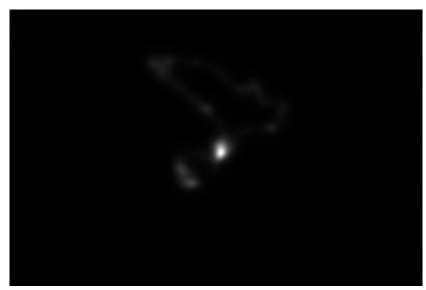

Figure 4: After Exponentiation of Features

3 Improvised Likelihood Function

The new likelihood for the data point x_i to belong to cluster j is given by:

$$\phi(x_i|\theta_j) = \frac{1}{2\pi|\sum_j|} exp[-\frac{1}{2}(x_i - \mu_j)\sum_j^{-1}(x_i - \mu_j)]$$

In a conventional GMM, the pixel value distribution can be discribed by the following equation

$$f(x_i|\Pi, \Theta) = \sum_{j=1}^{L} \pi_j \phi(x_i|\theta_j)$$

Π are parameters constituting prior probabilites and Θ consitutes the parameters of the gaussian distributions.

In our GMM we will incorporate the spatial information by using a weighted neighborhood template for calculating conditional probability by using neighborhood probabilities.

$$f(y_i|\omega) = \sum_{j=1}^{L} \pi_{ij} [\sum_{m \in N_i} \frac{S(x_m)}{R_i} p(y_m|\theta_j)]$$

This is the likelihood for our Saliency Weighted GMM.

One can see that in structure the formula is similiar to conventional GMM in the sense that this likelihood also constitutes summation over different exponentials. That said, we can actually use EM method for finding optimal/approximate values for parameters for which the likelihood for the given image is maximized.

4 Flow Chart Of GMM-SMSI

Step 1. Salience Map Extraction.
1. The image is converted in the spectral domain using FFT. This gives the amplitude spectrum F(f) and the phase spectrum P(f) of the image.
2. The log spectrum representation L(f) is given by the logarithm of F(f).
3. The estimation of the average spectrum A(f).
4. The calculation of the residual value R(f).
5. The generation of the saliency map S(x).

Step 2. GMM incorporating the saliency map as spatial information.
1. The k-means algorithm is first used to initialize the parameters set $\Psi^{(0)}$.
2. Using the saliency map S(x) the EM algorithm is applied for the parameters estimation until convergence. At the end, we get the parameters set $\Psi(c)$.
3. The image pixels are then classified (labeled) based on the highest posterior probability.

Figure 5: Eagle

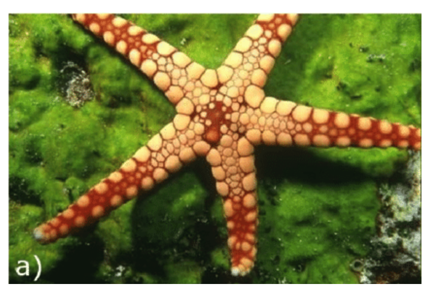

Figure 6: Starfish

5 Dataset

We have performed experiments from on a set of real images from the Berkeley Image Dataset. Some of the images from the dataset are shown in Figure 5 and 6.

6 Experiments

In this section we will showcase the results of the experiments mentioned in the research paper. Out of four experiments, we have found good results in two of them, in the later two experiments we have found results not much better than conventional GMM. But the reason for the same to whatever extent we could think of, has been mentioned in the respective sections.

6.1 Eagle - Tiny Object In Large Background

The original image is shown in Figure 5. The image below is of segmented image outputted by conventional GMM (Image obtained from the research paper).

Figure 7: Conventional GMM

The images below represent our results obatained using GMM-SMSI. On the left is the saliency map and on the right is the segmented result of the GMM-SMSI.

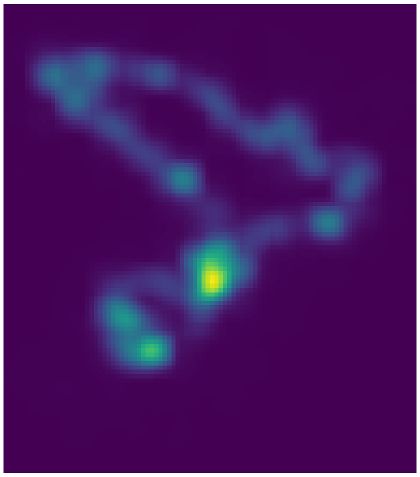

Figure 8: Saliency Map

Figure 9: Segmented Image

Here the result of GMM-SMSI is clearly better than conventional GMM.
1. The feathers are properly separated.
2. The outline of the birds resemble to that in the original image.

One can take a close look at the fourier transform to see that how fantastically it has captured the outline of the birds. Moreover, observe the feather section in the fourier transform.

But what result was research paper claimed? You can see below.

Figure 10: GMM-SMSI

Where are we lacking ? The middle thing that is properly segmented in there GMM-SMSI but not in ours. To get this we tried various initialization and changed the method to generate the fourier transform. The result of one such experiment is shown below. Although the result is not clean but it does to some extent preserve the middle part.

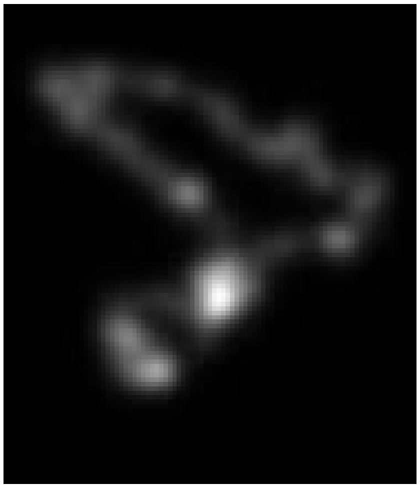

Figure 11: Saliency Map

Figure 12: Segmented Image