Report: Image Segmentation

Rishabh Singh: 19-953-793

November 8, 2019

1 Image Preprocessing

The given image was smoothed using Gaussian filter of standard deviation $\sigma = 5$. This is necessary to make the color uniform and remove the noisy pixels which may give spurious clusters in the segmented image. Fig 1 (b). The resultant image was converted from RGB to L*a*b color space (Fig 1(c)). The Lab color space is the most exact means of representing color and is device independent. Working with the Lab color space includes all of colors in the spectrum, as well as colors outside of human perception. Since the segmented output should be close to what human eyes perceive, Lab color space is a better choice. Also, Lab space deals with noise better than the RGB space where the illumination effects can be found in each channel. It is also well known, that the perceptual difference of RGB colors does not correspond to the Euclidean distance within the RGB vector space. L*a*b* is designed so that the perceptual difference of individual colors is proportional to the distance in the corresponding three-dimensional vector space. Thus, L*a*b is preferred to RGB for image segmentation tasks.

(a) (b) (c)

Figure 1: Smoothing and Lab conversion

2 Mean-Shift segmentation

The result shown in Fig.2 were obtained for a radius value of 2 and the threshold value of 0.1. Here threshold means that when the distance between the successive means is lesser than the value of 0.1, the algorithm for finding peaks is assumed to have converged.

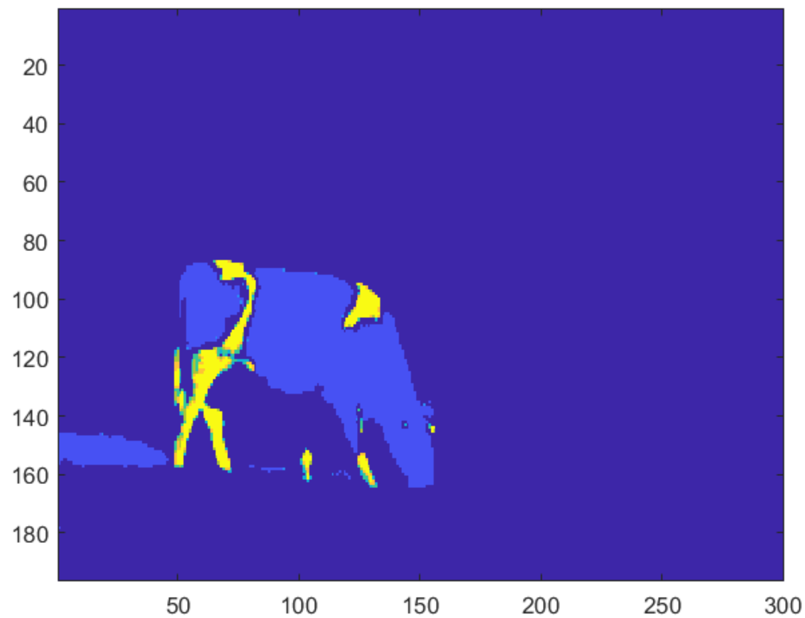


Figure 2: Segmentation map for Mean shift.

3 EM Segmentation

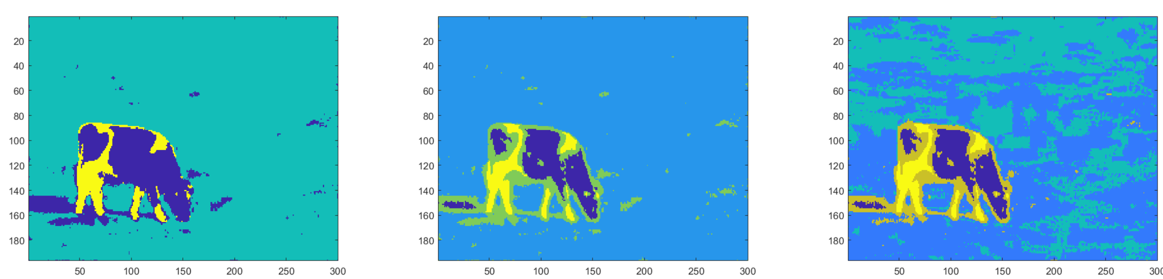


Figure 3: Segmentation map from EM algorithm

3.1 Parameters on convergence

3.1.1 K=3

$$\alpha = \begin{bmatrix} 0.11 \\ 0.86 \\ 0.03 \end{bmatrix} \mu_1 = \begin{bmatrix} 40.09 \\ 124.02 \\ 136.75 \end{bmatrix}, \mu_2 = \begin{bmatrix} 89.03 \\ 114.43 \\ 149.08 \end{bmatrix}, \mu_3 = \begin{bmatrix} 147.34 \\ 125.07 \\ 140.65 \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} 836.18 & -126.49 & 240.53 \\ -126.49 & 31.97 & -47.13 \\ 240.53 & -47.13 & 83.72 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 85.77 & 0.49 & 0.30 \\ 0.49 & 1.01 & -0.30 \\ 0.30 & -0.30 & 1.81 \end{bmatrix}, \Sigma_3 = \begin{bmatrix} 3779.31 & 130.85 & 10.07 \\ 130.85 & 13.54 & -4.02 \\ 10.07 & -4.02 & 24.94 \end{bmatrix}$$

3.1.2 K=4

$$\alpha = \begin{bmatrix} 0.04 \\ 0.84 \\ 0.08 \\ 0.02 \end{bmatrix} \mu_1 = \begin{bmatrix} 13.80 \\ 128.39 \\ 128.63 \end{bmatrix}, \mu_2 = \begin{bmatrix} 89.09 \\ 114.44 \\ 149.09 \end{bmatrix}, \mu_3 = \begin{bmatrix} 64.07 \\ 120.37 \\ 142.89 \end{bmatrix}, \mu_4 = \begin{bmatrix} 176.48 \\ 126.9 \\ 140.06 \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} 10.92 & 1.83 & -1.02 \\ 1.84 & 2.84 & -1.40 \\ -1.02 & -1.40 & 3.31 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 84.19 & 0.36 & 0.26 \\ 0.36 & 0.96 & -0.28 \\ 0.26 & -0.28 & 1.77 \end{bmatrix}, \Sigma_3 = \begin{bmatrix} 646.09 & -73.35 & 136.35 \\ -73.36 & 29.2 & -33.96 \\ 136.35 & -33.96 & 55.26 \end{bmatrix}$$

$$\Sigma_4 = \begin{bmatrix} 2807.73 & 28.19 & 56.87 \\ 28.19 & 4.85 & 0.04 \\ 56.87 & 0.04 & 26.35 \end{bmatrix}$$

3.1.3 K=5

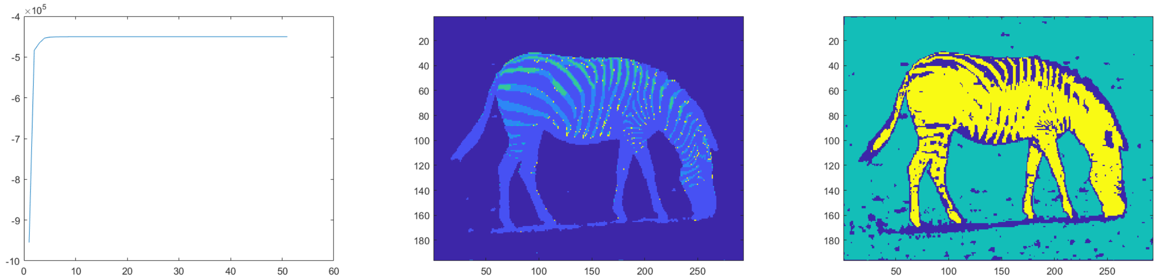
$$\alpha = \begin{bmatrix} 0.04 \\ 0.52 \\ 0.34 \\ 0.06 \\ 0.02 \end{bmatrix} \mu_1 = \begin{bmatrix} 13.58 \\ 128.35 \\ 128.61 \end{bmatrix}, \mu_2 = \begin{bmatrix} 85.50 \\ 114.35 \\ 149.55 \end{bmatrix}, \mu_3 = \begin{bmatrix} 93.82 \\ 114.58 \\ 148.41 \end{bmatrix}, \mu_4 = \begin{bmatrix} 60.78 \\ 122.08 \\ 140.56 \end{bmatrix}, \mu_5 = \begin{bmatrix} 179.96 \\ 127.01 \\ 140.16 \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} 9.39 & 1.52 & -0.98 \\ 1.52 & 2.66 & -1.35 \\ -0.98 & -1.35 & 3.15 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 106.66 & -0.32 & 4.91 \\ -0.32 & 1.49 & -0.31 \\ 4.91 & -0.31 & 2.17 \end{bmatrix}, \Sigma_3 = \begin{bmatrix} 23.83 & 0.09 & -0.64 \\ 0.09 & 0.37 & -0.02 \\ -0.64 & -0.02 & 0.52 \end{bmatrix}$$

$$\Sigma_4 = \begin{bmatrix} 778.30 & -79.70 & 146.99 \\ -79.70 & 26.11 & -30.74 \\ 146.99 & -30.74 & 52.8 \end{bmatrix}, \Sigma_5 = \begin{bmatrix} 2599.18 & 21.03 & 47.37 \\ 21.03 & 4.51 & -0.05 \\ 47.37 & -0.05 & 26.16 \end{bmatrix}$$

4 Discussion

- Segmentation of image is necessary for localisation unlike classification which only gives a global description.



(a) Log-lik vs iter

(b) Mean-shift

(c) EM

Figure 4: (a)Log-likelihood vs iteration number for K=3. (b) Results on zebra.jpg

- Smoothing image prior to segmentation reduces noisy assignments to clusters in the output.
- Lab colorspace is better than RGB because of its resemblance to human perception, better noise handling and a relation between perceptual difference and Euclidean distance in the Lab space.
- Low radius value in Mean Shift algorithm (say r=1) have heavy noise in the output as many background pixels were labelled as foreground i.e. the cow. Higher radius(r=50) led to eroded output since majority of pixels were labelled as background.
- Best results for Mean Shift were obtained for $r = 20, threshold = 0.1$
- The convergence of EM algorithm was assured when the difference in log-likelihood value from the previous iteration is lesser than a threshold.
- The best results for EM-based segmentation were obtained for $threshold = 1e - 3$.
- While EM results with K=3 is coarse as it misses out details like sharp boundaries, K=3 is too fine and it assigns unnecessary details in background as well. However, the foreground is well segmented in each case.
- The output of Mean Shift segmentation is finer than the EM counterpart, probably because it is based on hard assignment of individual pixels, while assignment is probabilistic in EM.
- The value of μ_2 doesn't show much change on varying the value of K. This component might be associated with a dominant color in the image as it is present in every case after convergence.
- Both the algorithms are computationally expensive.
- The log-likelihood increases and then saturates after certain iterations (Fig. 4)