Lab Assignment - Image segmentation

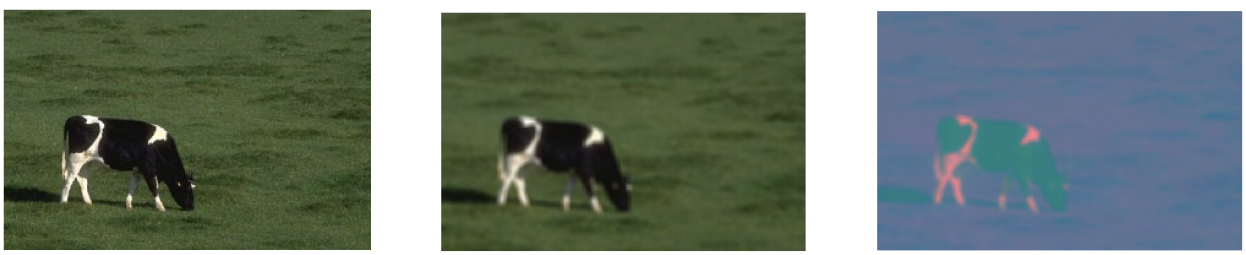
Kaiyue Shen

December 11, 2020

1 Image Preprocessing

Before we implement image segmentation methods, we first do a two-step image preprocessing. First, we use a 5*5 gaussian filter with $\sigma = 5.0$ to smooth the image. Then we use the build-in functions to convert the image from RGB to L^*a^*b color space. The original, smoothed and converted image are shown in figure 1.

The advantage of doing segmentation in L^*a^*b color space is: we know L is for perceptual lightness, and a and b are for the four unique colors of human vision: red, green, blue, and yellow. Pixels with similar colors but different lightness should be considered as the same segment. In L^*a^*b color space, their difference mainly lies in L channel while in RGB space, all three channels are different, making these pixels hard to be gathered together. So, using L^*a^*b color space to do image segmentation makes more sense.



(a) Original image (b) Smoothed image (c) L*a*b

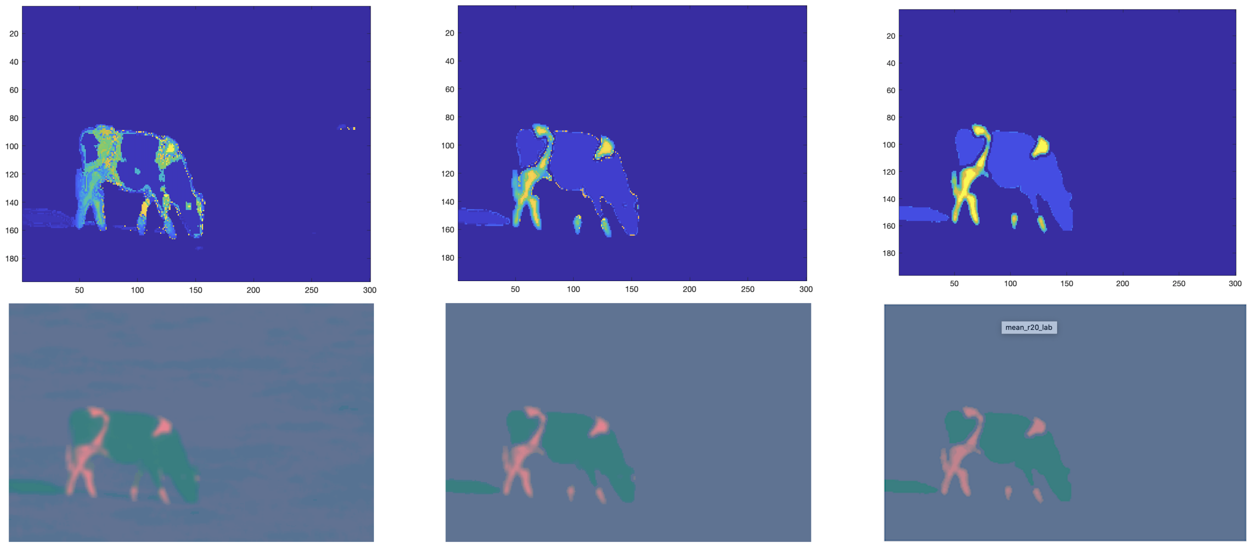
Figure 1: The original, smoothed and converted image of "cow"

2 Mean-Shift Segmentation

Mean-Shift Segmentation mainly consists of three steps:

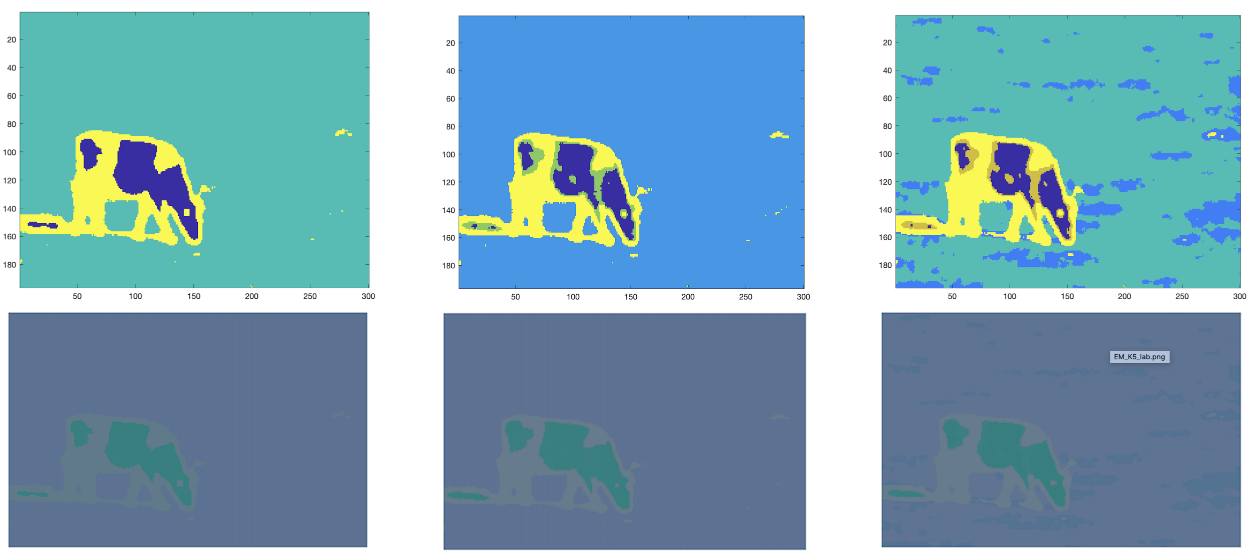
1. Create the density function in L^*a^*b color space. We vectorize the pixel values in each channel and concatenate them to build a L*3 matrix, where L is the number of pixels.
2. For each pixel, we initialize the peak with pixel itself. Then we find all other pixels within the spherical window and update the peak with the mean of inliers. We keep updating until the peak does not shift anymore.
3. For each new peak, we search the closest peak. If their distance is smaller than the threshold r/2, we merge them. Otherwise, we add the new peak.

We vary the threshold r, and show the results in figure 2. Comparing three columns, we can see with r increasing, newly-find peak is more likely to merge with old ones, so fewer peaks are found finally and the segmentation will be coarse. Comparing the bottom row with figure 1c, we can see the segmentation results are good enough to reconstruct the image in L*a*b color space.



(a) r=5, peak number=268 (b) r=10, peak number=30 (c) r=20, peak number=10

Figure 2: Mean-shift segmentation results using different thresholds. The top row shows the segment in label space, the bottom row shows the segment in L*a*b space.



(a) K=3, iteration number=14 (b) K=4, iteration number=40 (c) K=5, iteration number=47

Figure 3: EM segmentation results using different number of segments K. The top row shows the segment in label space, the bottom row shows the segment in L*a*b space.

3 EM Segmentation

EM Segmentation mainly consists of four steps:

1. Create the density function same as step 1 in Mean-shift.

2. Initialization. Since we assume the segments is modeled as a Gaussian mixture model with parameters $\Theta = (\alpha_1, \ldots, \alpha_K, \mu_1, \ldots, \mu_k, \Sigma_1, \ldots, \Sigma_k)$, we first need to initialize the parameters. For α, we use

uniform weights. For μ, we use "linspace" to spread them equally in L*a*b space. For Σ, we build a diagonal matrix using L, a, b range.

3. Expectation. For each pixel, for each segment, we compute the probability that it is in that segment.
4. Maximization. With the above probability matrix, we can update the parameters Θ by maximizing the expectation of the complete log likelihood, which can be done in the analytic form.

We iterate between step 3 and 4 until the change of cluster centers is smaller than the threshold, in our case, we set the threshold to be 1. We vary the number of segments K, display the results in figure 3 and results of Θ as follows.

Results of Θ when $K = 3$:

$$\Sigma_1 = 10^4 \cdot \begin{bmatrix} 0.0015 & 0.0003 & -0.0001 \\ 0.0003 & 1.2908 & 1.2890 \\ -0.0001 & 1.2890 & 1.2879 \end{bmatrix} \quad \Sigma_2 = 10^3 \cdot \begin{bmatrix} 0.0586 & 0.0003 & 0.0005 \\ 0.0003 & 0.6510 & 1.5342 \\ 0.0005 & 1.5342 & 3.6229 \end{bmatrix}$$

$$\mu_1 = [15.0684 \quad 128.6718 \cdot 128.5405] \quad \mu_2 = [88.9472 \quad 114.4442 \quad 149.1241]$$

$$\Sigma_3 = 10^3 \cdot \begin{bmatrix} 3.0856 & 0.0708 & 0.0922 \\ 0.0708 & 1.0872 & 1.6046 \\ 0.0922 & 1.6046 & 2.4933 \end{bmatrix} \quad \mu_3 = [90.5892 \quad 123.2520 \quad 140.2095]$$

$$\alpha = [0.0423 \quad 0.8728 \quad 0.0849]$$

Results of Θ when $K = 4$:

$$\Sigma_1 = 10^4 \cdot \begin{bmatrix} 0.0005 & 0.0001 & -0.0001 \\ 0.0001 & 1.3210 & 1.3196 \\ -0.0001 & 1.3196 & 1.3187 \end{bmatrix} \quad \Sigma_2 = 10^3 \cdot \begin{bmatrix} 0.0582 & 0.0003 & 0.0005 \\ 0.0003 & 0.6498 & 1.5323 \\ 0.0005 & 1.5323 & 3.6205 \end{bmatrix}$$

$$\mu_1 = [13.6088 \quad 128.5382 \quad 128.4352] \quad \mu_2 = [88.9686 \quad 114.4425 \quad 149.1257]$$

$$\Sigma_3 = 10^4 \cdot \begin{bmatrix} 0.0053 & 0.0004 & 0.0007 \\ 0.0004 & 1.0909 & 1.1136 \\ 0.0007 & 1.1136 & 1.1393 \end{bmatrix} \quad \Sigma_4 = 10^3 \cdot \begin{bmatrix} 2.8896 & 0.1121 & 0.0369 \\ 0.1121 & 0.6584 & 1.0939 \\ 0.0369 & 1.0939 & 1.9512 \end{bmatrix}$$

$$\mu_3 = [23.8983 \quad 128.3084 \quad 130.5934] \quad \mu_4 = [97.3710 \quad 122.6326 \quad 141.2379]$$

$$\alpha = [0.0326 \quad 0.8713 \quad 0.0185 \quad 0.0776]$$

Results of Θ when $K = 5$:

$$\Sigma_1 = 10^4 \cdot \begin{bmatrix} 0.0005 & 0.0001 & -0.0001 \\ 0.0001 & 1.3223 & 1.3208 \\ -0.0001 & 1.3208 & 1.3197 \end{bmatrix} \quad \Sigma_2 = 10^3 \cdot \begin{bmatrix} 0.0304 & -0.0019 & 0.0027 \\ -0.0019 & 1.4686 & 2.8173 \\ 0.0027 & 2.8173 & 5.4147 \end{bmatrix}$$

$$\mu_1 = [13.5530 \quad 128.5395 \quad 128.4232] \quad \mu_2 = [75.6217 \quad 113.9187 \quad 149.1945]$$

$$\Sigma_3 = 10^3 \cdot \begin{bmatrix} 0.0370 & -0.0006 & 0.0005 \\ -0.0006 & 0.5693 & 1.3934 \\ 0.0005 & 1.3934 & 3.4170 \end{bmatrix} \quad \Sigma_4 = 10^4 \cdot \begin{bmatrix} 0.0048 & 0.0005 & 0.0006 \\ 0.0005 & 1.1072 & 1.1262 \\ 0.0006 & 1.1262 & 1.1480 \end{bmatrix}$$

$$\mu_3 = [90.6703 \quad 114.5148 \quad 149.1112] \quad \mu_4 = [23.2555 \quad 128.4445 \quad 130.3597]$$

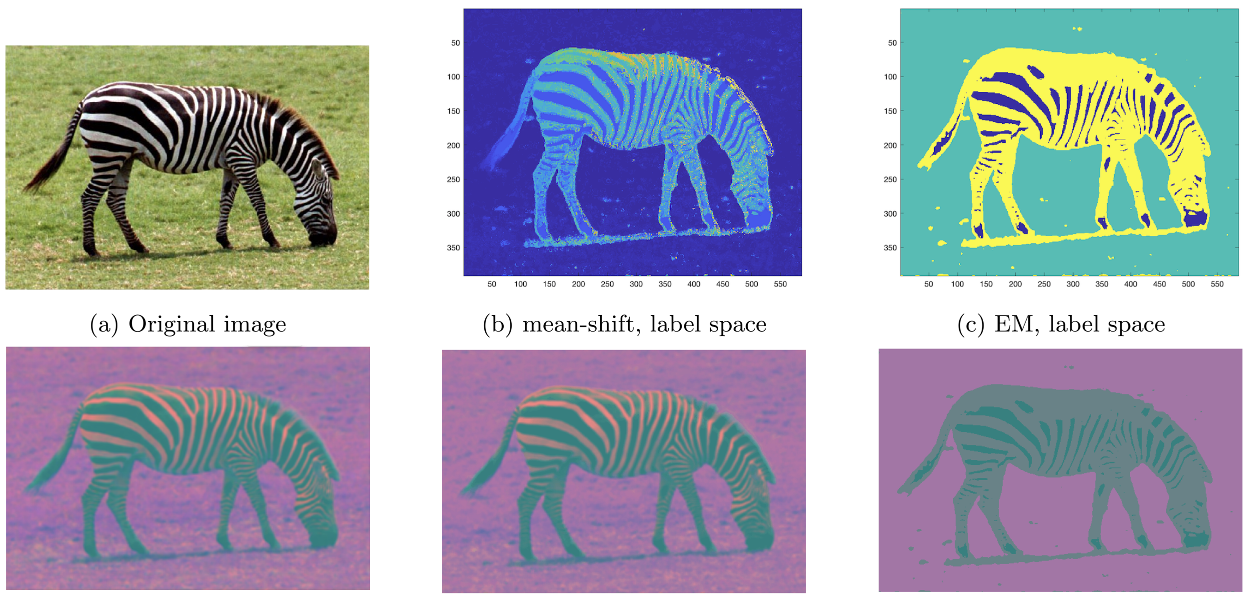
$$\Sigma_5 = 10^3 \cdot \begin{bmatrix} 2.9955 & 0.1025 & 0.0535 \\ 0.1025 & 0.6500 & 1.0674 \\ 0.0535 & 1.0674 & 1.8716 \end{bmatrix} \quad \mu_5 = [97.8859 \quad 123.0238 \quad 140.8444]$$

$$\alpha = [0.0321 \quad 0.1039 \quad 0.7707 \quad 0.0181 \quad 0.0752]$$

Comparing three columns in figure 3, we can see with more segments, the image is segmented more finely. Besides, we find in all cases, the maximum of α is much larger than the second maximum, e.g. when $K = 3$, $0.8728 >> 0.0849$.

4 Other Experiments

We display the results of both Mean-Shift Segmentation and EM Segmentation for "Zebra" image in figure 4.



(a) Original image

(b) mean-shift, label space

(c) EM, label space

(d) L*a*b

(e) mean-shift, L*a*b space

(f) EM, L*a*b space

Figure 4: Segmentation results using two methods.