Computer Vision, 8th assignment

Mathis Lamarre

Segmentation

Image Preparation

In order to prepare the image for segmentation, we smooth it with a Gaussian filter. Then, we convert it to the $L^*a^*b^*$ color space. This is more suited than the RGB color space because areas with the same color but different shading will have closer values, and will more likely be segmented together. For example, on the cow's leg, the different shades of white will have similar a^* and b^* values but slightly different L^* and will therefore be merged together, whereas there would have had very different RGB values for all 3 dimensions.

Fig. 1: Steps of the image preparation. Left shows the original cow image. Middle shows the image after Gaussian smoothing with a 5x5 matrix and $\sigma = 5$. Right shows the image in the $L^*a^*b^*$ color space.

Mean-Shift Segmentation

To do Mean-Shift segmentation, first we create the X density matrix by taking the value of all the pixels for each dimension of the $L^*a^*b^*$ space, putting them into a vector and concatenating these vectors into a matrix. Then we find the node of each pixel. To do so, for each pixel, we compute its distances in the $L^*a^*b^*$ space with all the other points, and select all those for the which it is smaller than a certain r. Then, we shift the center of mass to the mean of all these selected points and start over with the new center of mass, until convergence (shift between the point and center of mass smaller than a certain *tolerance*).
After, we calculate the distances between the computed mode and all the previous peaks we found. If it is smaller than $r/2$, we don't create a new peak and set the associate the pixel with the ID of the peak. Otherwise we create a new peak and set it to a new ID.

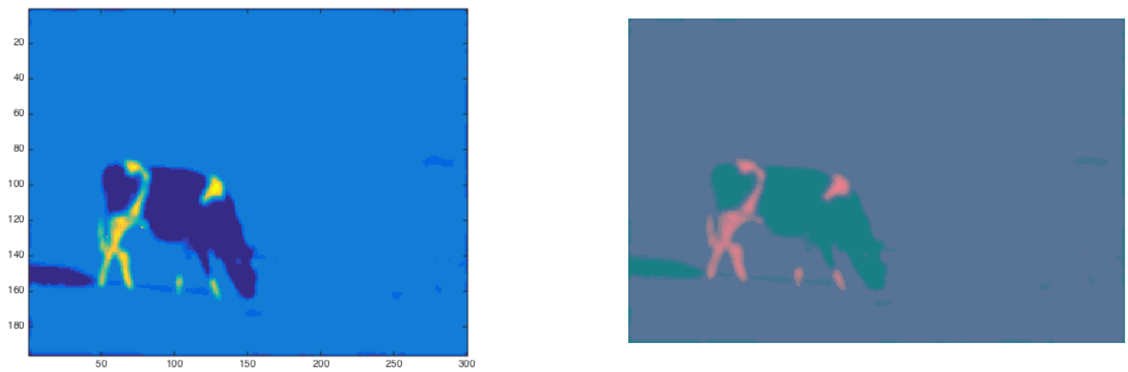

Fig. 2: Mean-Shift segmentation of the cow image. Left shows each cluster with a different color. Right shows the reconstruction by assigning its peak's color to each color. $r = 3$.

After trying with several parameters, we selected the results shown on Fig.2. Almost the entire background is in only one cluster which is good. The cow has several clusters which correspond to the different patches. One drawback is that a part of the leg that is in the shadow was merged with the background's cluster. However a smaller r results in way more clusters, particularly in the background.

EM Segmentation

First, we need to initialize the values of θ and α. For α, we set each weight to be equal, that is to say $1/K$. For the mean μ, we spread them equally in the values in the $L^*a^*b^*$ space taken in X. For the covariance matrices Σ, we set them as diagonal matrices with elements corresponding to the range of $L^*a^*b^*$ values taken in X.
After, in the expectation step, we compute for all the pixels its probability to be in each segment from μ, Σ and α. To do so, we use a gaussian model and make sure to normalize it.
Then, in the maximization step, we update the parameters μ, Σ and α that maximize the log likelihood under the current probabilities. We do so using the formulas derived and trying to minimize the use of for loops, since we need to iterate over all the pixels and all the clusters.
Finally, we repeat these two steps until convergence, that is to say when values in μ don't change too much between two steps (the clusters we defined are now relevant).

In the following figures, the μ vectors are displayed as lines and the Σ and Δ_μ are in the normalized version. The results more precise than with the Mean-Shift algorithm, particularly in the legs regions, even with fewer clusters.

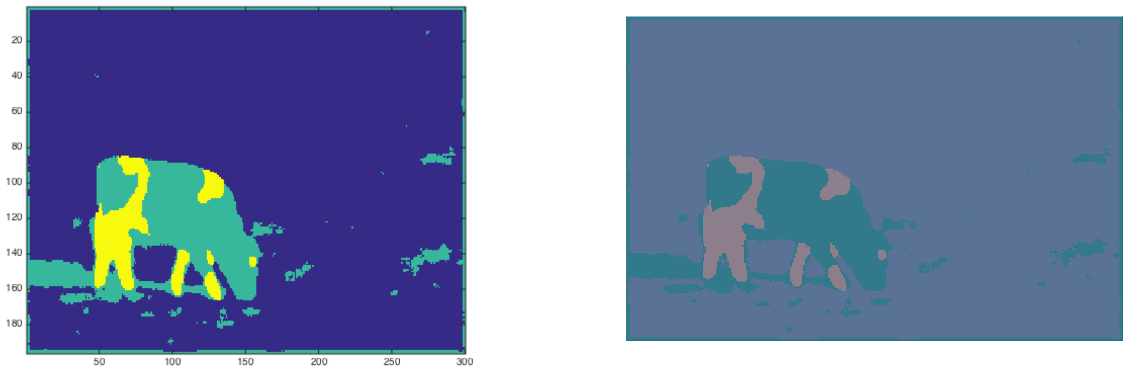

Fig. 3: EM segmentation and reconstruction. $K = 3$ and $\Delta_\mu = 0.3$

$$\mu = \begin{array}{ccc} 89.3311 & 114.4394 & 149.0918 \\ \hline 49.4438 & 121.4665 & 139.4818 \\ \hline 141.1582 & 125.5591 & 140.4582 \end{array}$$

$$\Sigma_1 = \begin{bmatrix} 0.7622 & 0.0013 & 0.0073 \\ 0.0013 & 0.0111 & -0.0023 \\ 0.0073 & -0.0023 & 0.0213 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 10.5424 & -1.9122 & 3.0484 \\ -1.9122 & 0.4863 & -0.6553 \\ 3.0484 & -0.6553 & 1.0125 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} 31.6837 & 0.7651 & 0.6701 \\ 0.7651 & 0.1361 & -0.0889 \\ 0.6701 & -0.0889 & 0.3467 \end{bmatrix}$$

$$\alpha = \begin{pmatrix} 0.8117 & 0.1527 & 0.0356 \end{pmatrix}$$

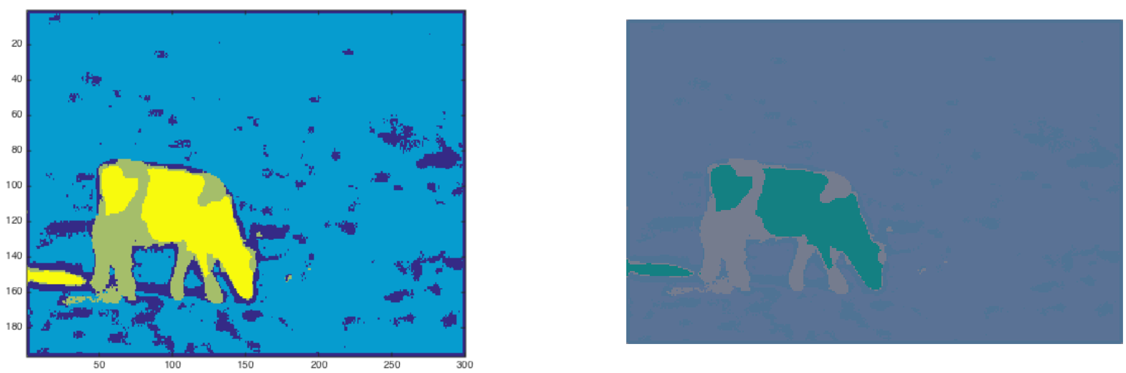

Fig. 4: EM segmentation and reconstruction. $K = 4$ and $\Delta_\mu = 0.3$

$$\mu = \begin{array}{ccc} 78.5808 & 115.4814 & 147.9853 \\ \hline 89.7275 & 114.4122 & 149.0209 \\ \hline 116.7977 & 123.7949 & 141.5904 \\ \hline 19.6305 & 128.1852 & 129.8127 \end{array}$$

$$\Sigma_1 = \begin{bmatrix} 3.1847 & -0.2709 & 0.5432 \\ -0.2709 & 0.0590 & -0.0747 \\ 0.5432 & -0.0747 & 0.1503 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 0.6754 & 0.0110 & -0.0033 \\ 0.0110 & 0.0102 & -0.0006 \\ -0.0033 & -0.0006 & 0.0182 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} 40.0417 & 1.6159 & 0.1289 \\ 1.6159 & 0.2050 & -0.1638 \\ 0.1289 & -0.1638 & 0.4274 \end{bmatrix}$$

$$\Sigma_4 = \begin{bmatrix} 1.4429 & -0.0030 & 0.2409 \\ -0.0030 & 0.0651 & -0.0439 \\ 0.2409 & -0.0439 & 0.1154 \end{bmatrix}$$

$$\alpha = \begin{pmatrix} 0.1782 & 0.7101 & 0.0533 & 0.0585 \end{pmatrix}$$

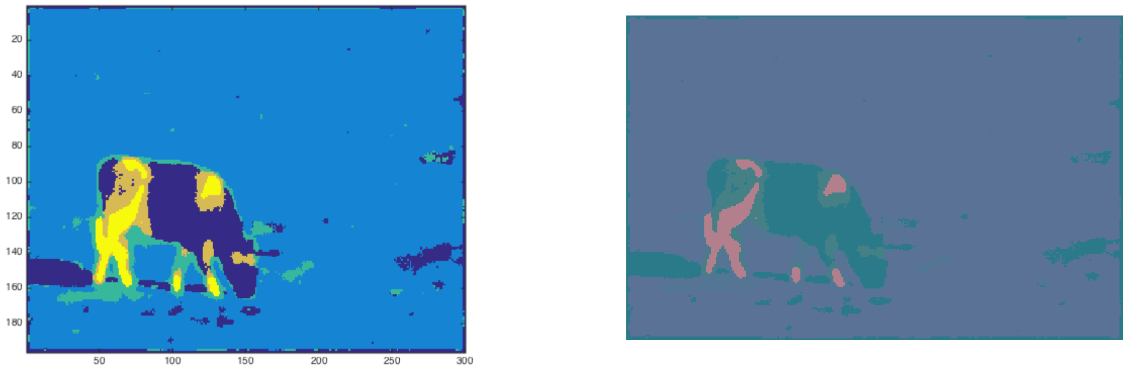

Fig. 5: EM segmentation and reconstruction. $K = 5$ and $\Delta_\mu = 0.3$.

$$\mu = \begin{array}{ccc} 42.0544 & 121.5194 & 138.0533 \\ \hline 89.5622 & 114.4372 & 149.0982 \\ \hline 81.6744 & 117.7979 & 146.9445 \\ \hline 68.2510 & 128.1831 & 134.1095 \\ \hline 178.7563 & 127.0330 & 141.4296 \end{array}$$

$$\Sigma_1 = \begin{bmatrix} 8.6716 & -1.9683 & 2.7776 \\ -1.9683 & 0.5180 & -0.6906 \\ 2.7776 & -0.6906 & 0.9733 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 0.7172 & 0.0001 & 0.0075 \\ 0.0001 & 0.0105 & -0.0023 \\ 0.0075 & -0.0023 & 0.0211 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} 2.4941 & 0.1764 & 0.0846 \\ 0.1764 & 0.0846 & -0.0764 \\ 0.0846 & -0.0764 & 0.2297 \end{bmatrix}$$

$$\Sigma_4 = \begin{bmatrix} 29.0982 - 1.3419 & 2.0191 \\ -1.34190.1105 & -0.1441 \\ 2.0191 - 0.1441 & 0.2656 \end{bmatrix}$$

$$\Sigma_5 = \begin{bmatrix} 15.2006 & 0.1302 & -0.1166 \\ 0.1302 & 0.0527 & -0.0342 \\ -0.1166 & -0.0342 & 0.2730 \end{bmatrix}$$

$$\alpha = \begin{pmatrix} 0.1083 & 0.8005 & 0.0510 & 0.0228 & 0.0174 \end{pmatrix}$$

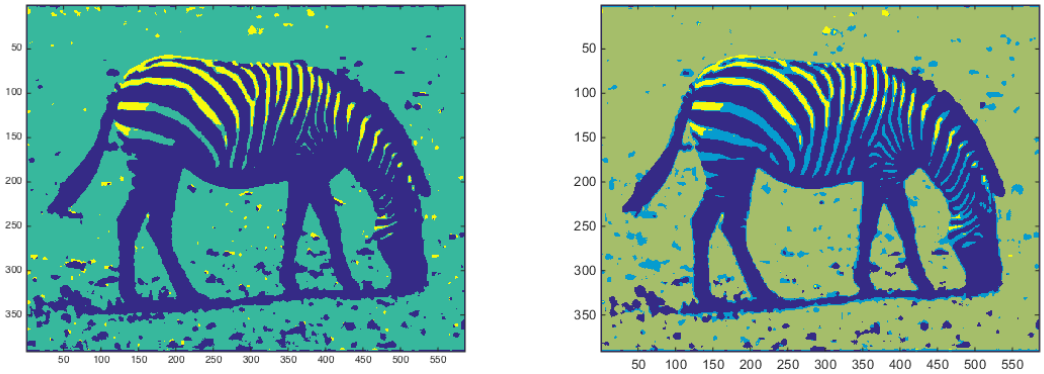

Fig. 6: EM segmentation on the zebra with 3 and 4 clusters.

The results are also good for the zebra, even if the stripe pattern is tricky.