Report of Anton Maksimov (antonma, 16-952-137), Task 5 "Image segmentation" on ETHZ course "Computer Vision".

8.1

Figure 1: Original image.

Figure 2: Blurred with Gaussion with $\sigma = 5$ image .

Figure 3: Converted to L*a*b blurred image. This color scheme shows better how human perceives the image (distances in this color space are better correlated with how we see difference of colors), so it's more beneficial than RGB scheme for image segmentation as human do it. (ref. exercise session, Wikipedia)

8.2 We implement mean-shift algorithm using different thresholds and radii r for moving of the feasible peak location. After each step we create new peak if there are no peaks nearby (less that $r/2$). Otherwise, we merge it with the first found and change the position of the peak using votes (weights to find new center of mass of all pixels which were assigned to this peak).
Algorithm works good (though rather slow, so we used resized smaller images), we can distinguish zebra and cow with black and white spots.
In order to achieve results which are possible to compare using different rescaling factors, we use corresp. rescaled σ of Gaussian filter applied before segmentation.
8.3 We initialize μ as random points in the range between minimal and maximal values of pixels in the image in every axis. Covariance matrix is initialized as diagonal matrix with values equal to lengths of ranges of pixel values in every corresponding axis.

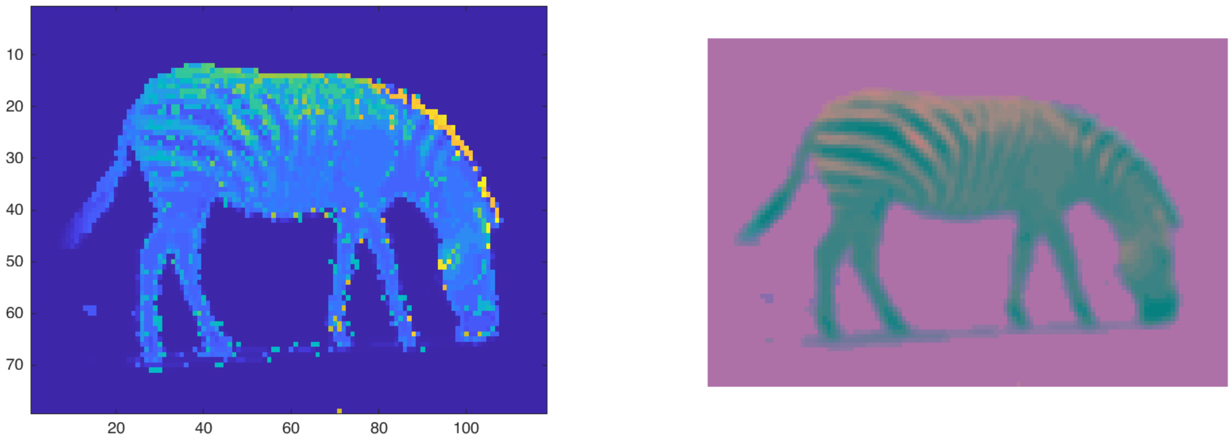

Figure 4: Mean shift algorithm, $r = 3$, `threshold` $= 0.01$, resized to 0.2 image.

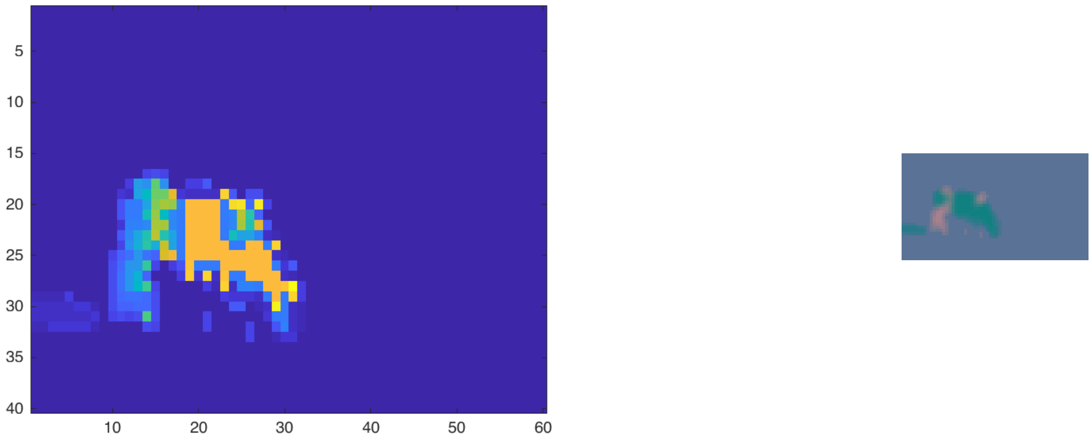

Figure 5: Mean shift algorithm, $r = 3$, `threshold` $= 0.01$, resized to 0.2 image.

In further maximization steps we add $10^{-10} I_d$ to the estimated covariance matrices in order to avoid them to become negatively definite because of the rounding (such a small addition practically doesn't change work of the algorithm, but makes it robust).
As a stop condition we use working time (20 min) and iteration (200) constrains + check if log–likelihood $\sum_i \log \sum_k p_i * \mathcal{N}(X_i, \theta_k)$ changed less than for 10^{-4} during one step (sums over pixels (i) and over clusters (k). We use this criterion as it is proved that during EM negative log–likelihood decreases monotonically.
We can compare EM and mean shift algorithms by eye and conclude that the quality of mean shift is better, though it is also much slower. But also its advantage is that it doesn't need prior number of clusters, only parameters which characterize scale of the image but not the color distribution there. Though, is we now how many clusters we should have approximately (we know color composition of the image), then quality of EM could be comparable with than one of mean shift.
Obtained results in res.doc file, first 3 data for zebra images from report.

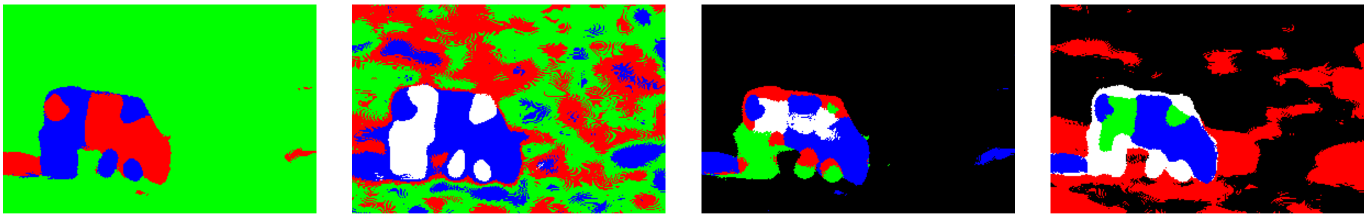

Figure 6: Result of EM-algorithm for K =3, 4, 5 and 5 and full cow image. We can see that for this image $K = 3$ is enough and probably the best in terms of the most adequate and simple model, because we have, in general, 3 colors at the picture: black, white and green. Also as it can be seen from all images and especially from comparison of two the most right that the result depends on the initial parameters and becomes more ambiguous with the increase of K. Here we can go either in more detailed segmentation of the cow or the grass. So for better result it could be beneficial to start EM several times and find the best for our purposes segmentation somehow.

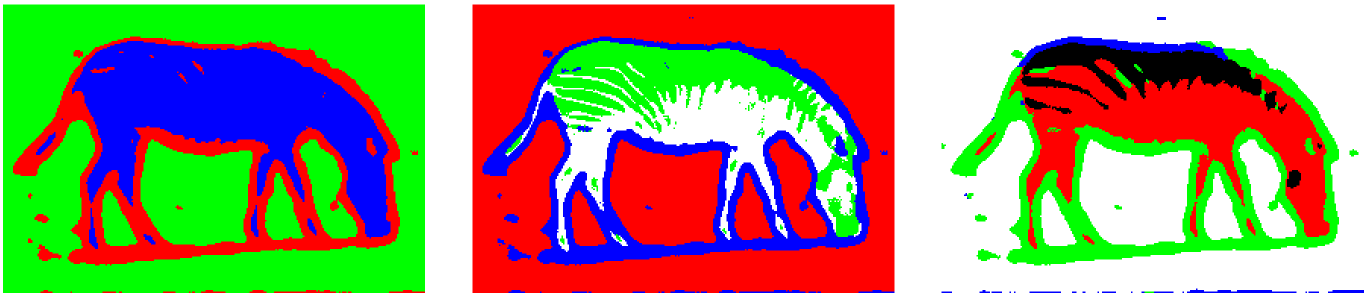

Figure 7: Result of EM-algorithm for $K = 3, 4, 5 and 5$ and zebra image at the 0.5 scale. Because of the more uniform background as it is at the cow image, here different initialization lead to more detailed segmentation of zebra and not grass.