5/24/2017

Image Segmentation

CV Assignment 1

Chinmay Dhawan

I6171015

CONTENTS

OBJECTIVE

The following report discusses the process of image segmentation in colored images by employing various color spaces including RGB, HSV and LAB by using K-means clustering. Further it discusses how spatial features, Elbow method, feature normalization and smoothing filters influence the result.

BACKGROUND

Three types of color features are employed i.e., RGB, HSV and LAB with and without including coordinate/spatial information. All the features are normalized in the range 0-1 using min-max normalization. For some images, extra weight is given to the color features for better segmentation and wherever that is the case, it has been explicitly mentioned. The L2(Euclidean distance) norm is used for calculating distance between the features for K-means clustering.

EFFECT OF SPATIAL FEATURES

It is observed that by including spatial features, the larger objects with similar color information, such as image background gets divided among different clusters as shown in Figure 2. This is because although the color information is almost the same, the positional value varies greatly between the two ends of the images leading to a larger Euclidian distance and hence, more chances of being divided into different cluster. Although, this should not be of concern in many images since background isn't the object of focus.

The advantage of using spatial features is that it binds pixels closer to each other within one cluster, especially if the number of clusters are overloaded. The same can be observed by comparing Figure 1 with Figure 2. In Figure 2, the clusters are more coherent and binded because spatial features are also employed along with color features.

One more advantage of using spatial features is that it is able to distinguish objects located far from each other. For example, two different objects located at a distance from each other but with similar color values won't be associated in different clusters if spatial information is not employed.

In regards with noisy images, it is also observed that by including spatial features the effects of noise like gaussian or salt and pepper noise is somewhat reduced. This can be accounted to the fact that although these types of noises do affect the color features but the relative spatial information for individual pixels remain intact. The same is shown in Figure 3 which shows various versions of a segmented image with gaussian noise.

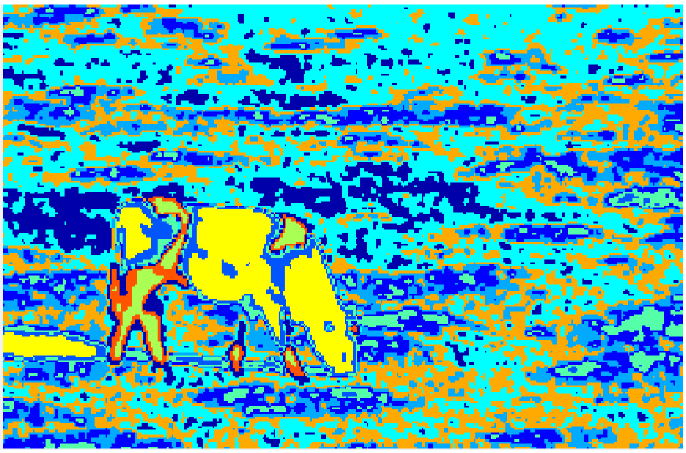

Figure 1: Segmented image with RGB features and k = 10.

Figure 2: Segmented image with RGB and spatial features, k = 10.

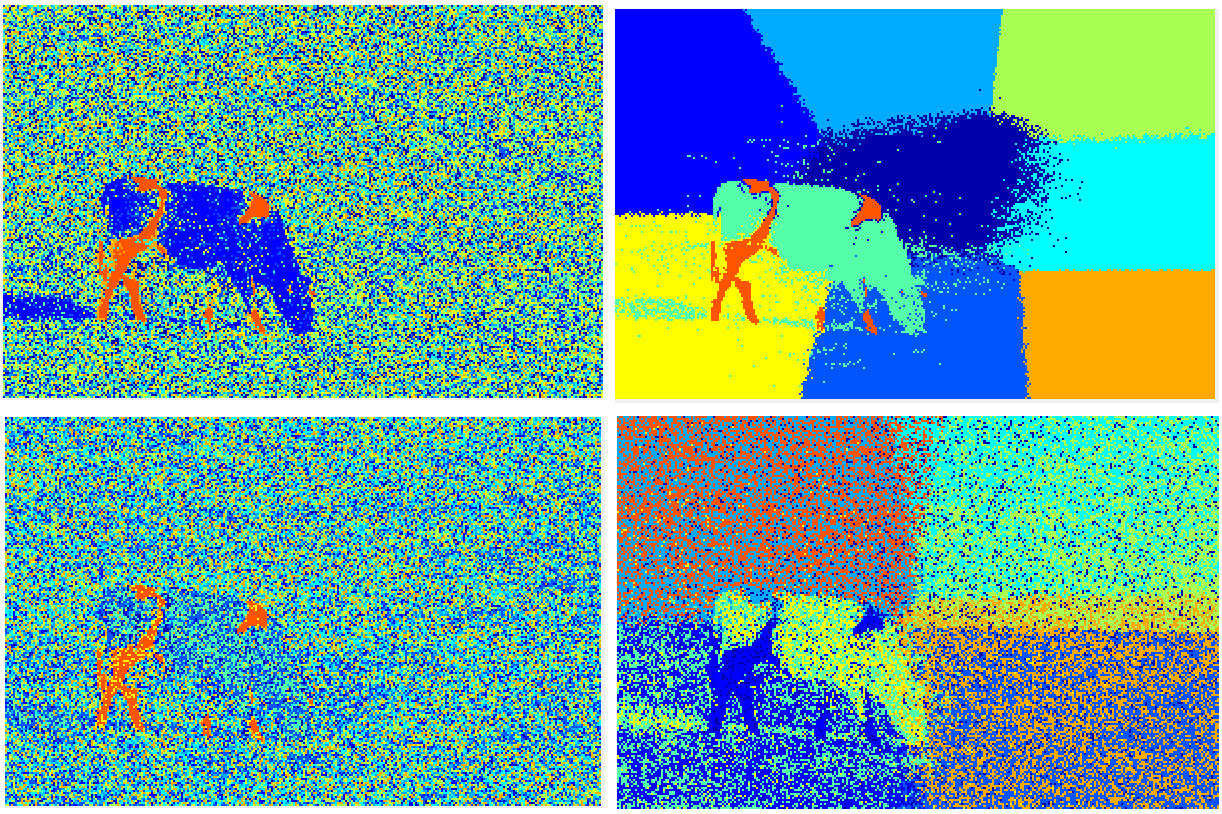

Figure 3: Left: Segmented image with RGB and HSV color features, Right: Segemented image with RGB and HSV color features including spatial features.

A major disadvantage of using spatial features is that they can easily overpower color information if not used carefully. For example, in the image of cow there are 3 major color values: Black, Green and White. So accordingly, if we do segmentation by using k = 3 and giving equal weight to all the features, the resulting image is overpowered with positional information. The results are shown in Figure 4. To improve the results on right, a greater weight needs to be given to the color space features.

Figure 4: Segmented image with LAB features and k = 3 including spatial features (right) and not including spatial features (left).

EFFECT OF DIFFERENT COLOR SPACES: RGB, HSV AND LAB

Each of the color space offers its own characteristics. And for the purpose of segmentation in the scope of this assignment it can not be concluded that one is better than the other. However, the various advantages one color space offer over another can be analyzed from the Figure 5. It shows, in clockwise direction, the original image at top left, and then the segmented images employing RGB, LAB and HSV color space features respectively, clockwise.

As you can observe from the figure, segmented image using RGB color features distinguish clusters solely based on color. The shadow of cow on the grass is almost black in color and hence assigned the same cluster as the black patch on cow. Whereas the segmented image using HSV features does a much better job, although not perfect, in seprating the cow from its shadow on the grass. This is because the hue component of the grass would still be the same, the saturation and intensity although would be different, making the HSV color space less sensitive to lightning variations. Also, it can be observed, that white legs of cow are classified to the same cluster as light portion of the grass. This is because as HSV can extract more information than just the three colors a higher value of k is needed for this segmentation. Finally, you can observe the segmented image using LAB features is much more defined than the other two, but this may not necessarily mean that it is better. LAB color features much more closely resemble to human perception of color than its counterparts. Hence, the image is much more smoothly segmented. RGB is a machine dependent color scheme, so LAB is usually a better choice than RGB. Although, as it is seen, there is a tradeoff between LAB and HSV and the choice should depend on the type and purpose of segmentation. For a specific image, If the segmentation can be done based on only color info LAB would be a good choice, otherwise for light sensitive images, HSV would prove better.

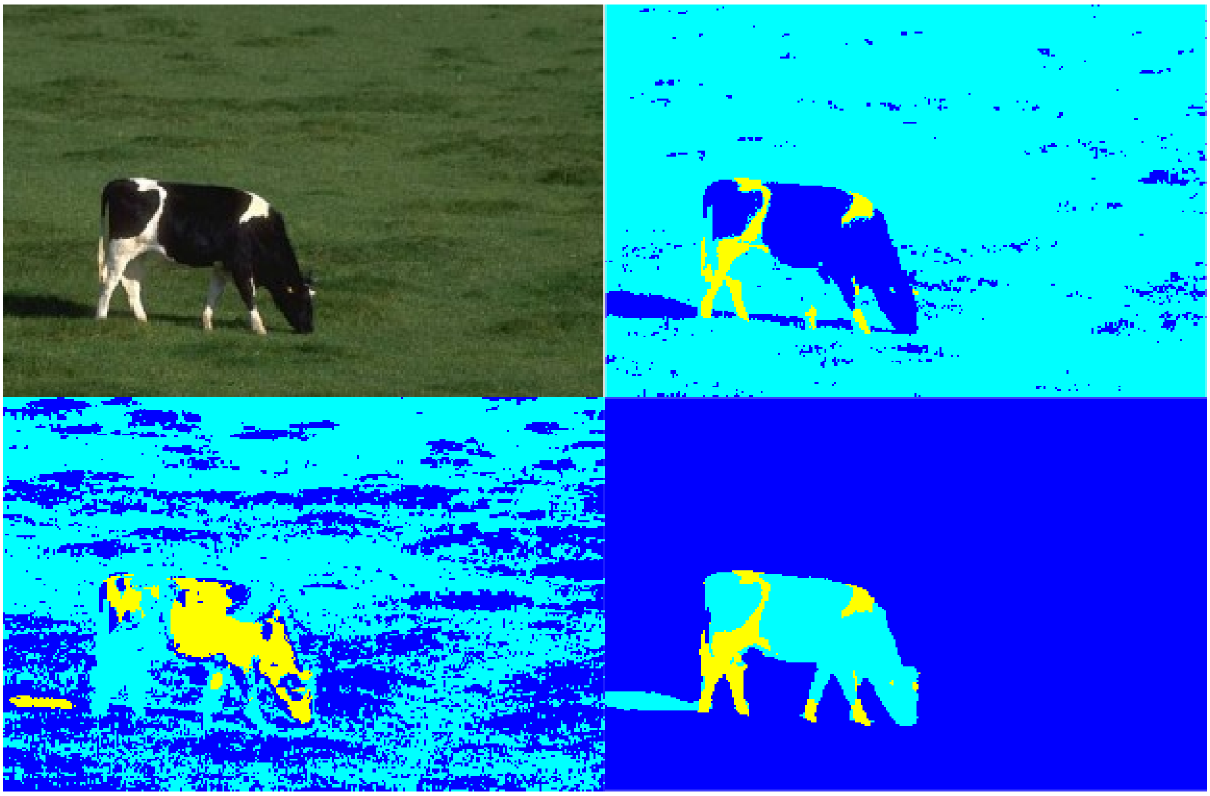

Figure 5: Image segmentation using different color features. (Clockwise from top left: Original, RGB, LAB and HSV)

DEPENDENCE ON INITIALIZATION

K-means only provides a locally optimum solution; The process of finding global optimum for this clustering is NP hard. For this assignment, Forgy method of initialization is implemented which works by randomly choosing k observations (pixels) as initial means. So, based on different initializations, different converging results are obtained. The same is shown in Figure 6 where for the same value of k and same features different segmentation results are obtained because of different initialization of k centers. For best results, the initial centers chosen should be as far away from each other as possible.

The results can be made more stable by running the algorithm several times with different initializations and then assigning the pixels to the cluster by majority voting concept.

Figure 6: Segmented image using HSV and spatial features with k = 8 using different k centers as initialization points

DISTANCE METRIC AND CALCULATION

The L2 norm is used for the distance among different pixel values. Each pixel is represented by its normalized features and the Euclidian distance is calculated between them according to the following formula.

Let Pixel a be represented by its features: a_1, a_2, a_3, a_4, a_5;

and Pixel b be represented as follows: b_1, b_2, b_3, b_4, b_5;

then distance is given by: $\{ (a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 + (a_4 - b_4)^2 + (a_5 - b_5)^2 \}^{1/2}$

A cluster center is also nothing but the average of all the observations belonging to one cluster and hence is calculated by the same formula as above.

Other forms of distance such as L1 or L3 norm may also be employed but L2 is norm is the most widely used and have certain significance since it represents the shortest distance between two points in the coordinate system, also satisfying the Pythagoras theorem.

ELBOW METHOD

Elbow method does seem to greatly improve the results of segmentation as it provides a way to balance the number of cluster and SSE. It is a fact that as the number of clusters will increase, the total intra cluster variance (also denoted by SSE) will decrease and eventually be zero when number of clusters is equal to the number of observations/pixels. By plotting a graph between the two, elbow method provides a way to observe the point (elbow point) where the decrease in gradient changes abruptly i.e., it's not as sharp as the from the previous points and that point gives a suitable number of clusters.

SSE is calculated by summing the square of all the distances of each data point from the center of the cluster for a specific value of k.

However, there are limitations to this method. As it analyses the SSE for several values of k, it becomes computationally expensive as the resolution of the image or the maximum size of k increases. Also, one more major limitation of this method is that the elbow point is not always visible, and sometimes there are more than one number of elbows. This can be more easily explained by seeing the plots in Figure 7, 8, 9 and 10. From Figure 7, it can be argued that k = 5 can prove to be a good elbow point. In Figure 8, all the clusterings from Figure 7 are plotted and for k = 5, the cow objected is clearly segmented from the background without adding any more segments in the background. However, in Figure 9, the elbow method falls short as there is no clear elbow defined as the gradient decreases smoothly. For Figure 10, one can argue that k = 4 or k = 8 might prove to be good values for k and can be checked by viewing the segmentation for both values.

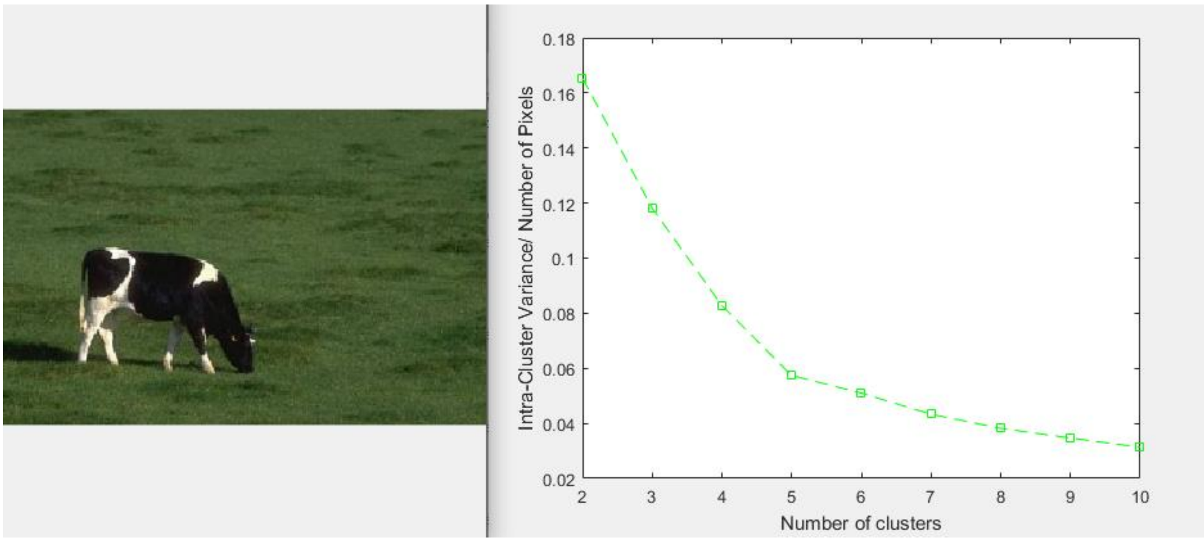

Figure 7: Intra-cluster variance vs number of clusters for the Cow image; LAB features with spatial information.

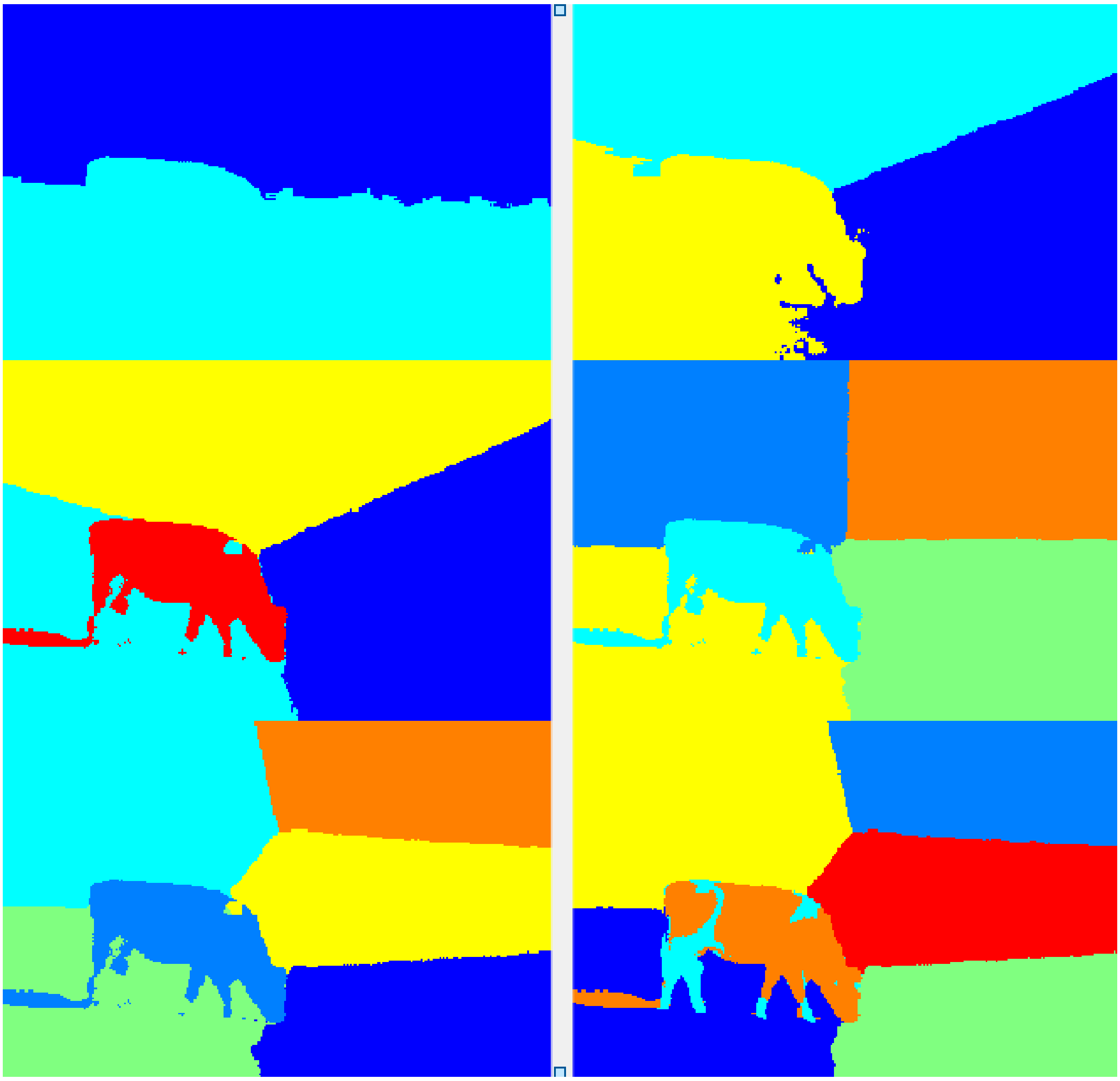

Figure 8: Segmented Images from k = 2 to k = 7 for the cow image with LAB features and spatial information corresponding to Figure 7.

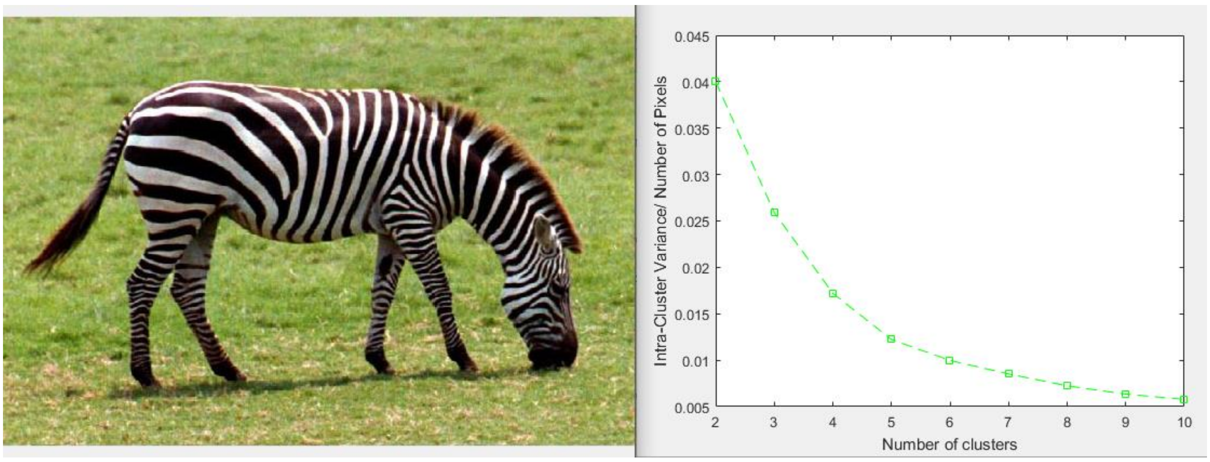

Figure 9: Intra-cluster variance vs number of clusters for the Zebra image; RGB Features.

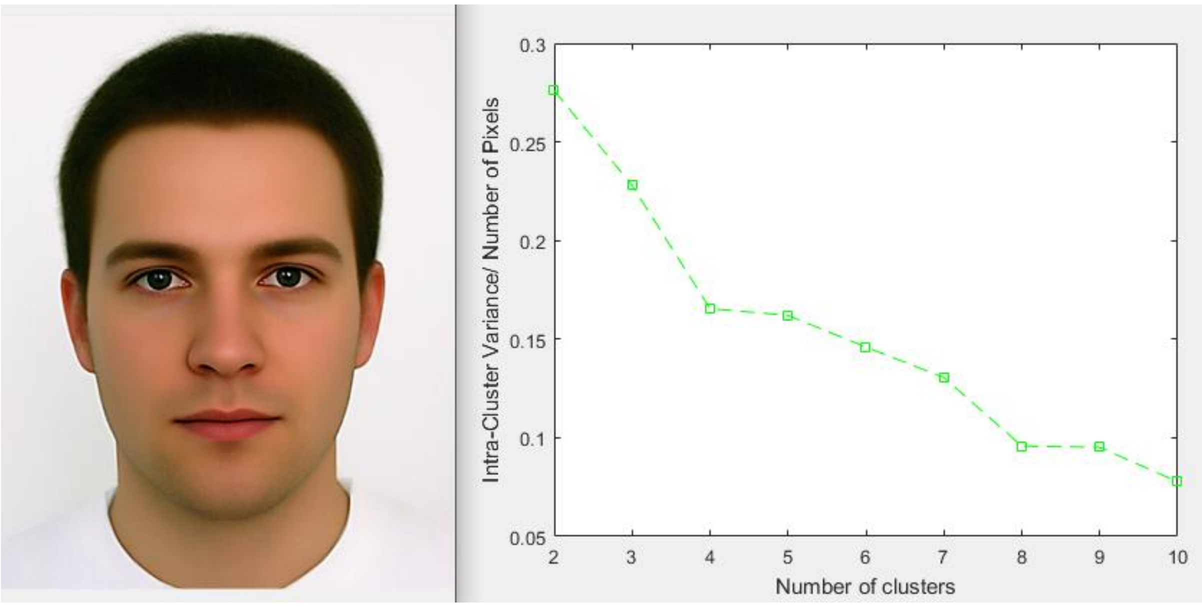

Figure 10: Intra-cluster variance vs number of clusters for the Face image; HSV features with spatial information.

PRE-PROCESSING: FEATURE NORMALIZATION AND IMAGE SMOOTHING

Feature normalization is an important step to assign equal weights to all the features. Otherwise, features with larger vales may dominate the features with smaller values. Further, once it is known that all features contribute equally, the weight of individual features can easily be modified to further analyze the contribution of each feature and tweak the weights manually.

For this report, Min-Max normalization has been employed on all the features.

Figure 8 shows the results of image segmentation with and without normalization, and the it can easily be observed that when the images are not normalized spatial features are overpowering the segmentation resulting in a not so useful segmentation.

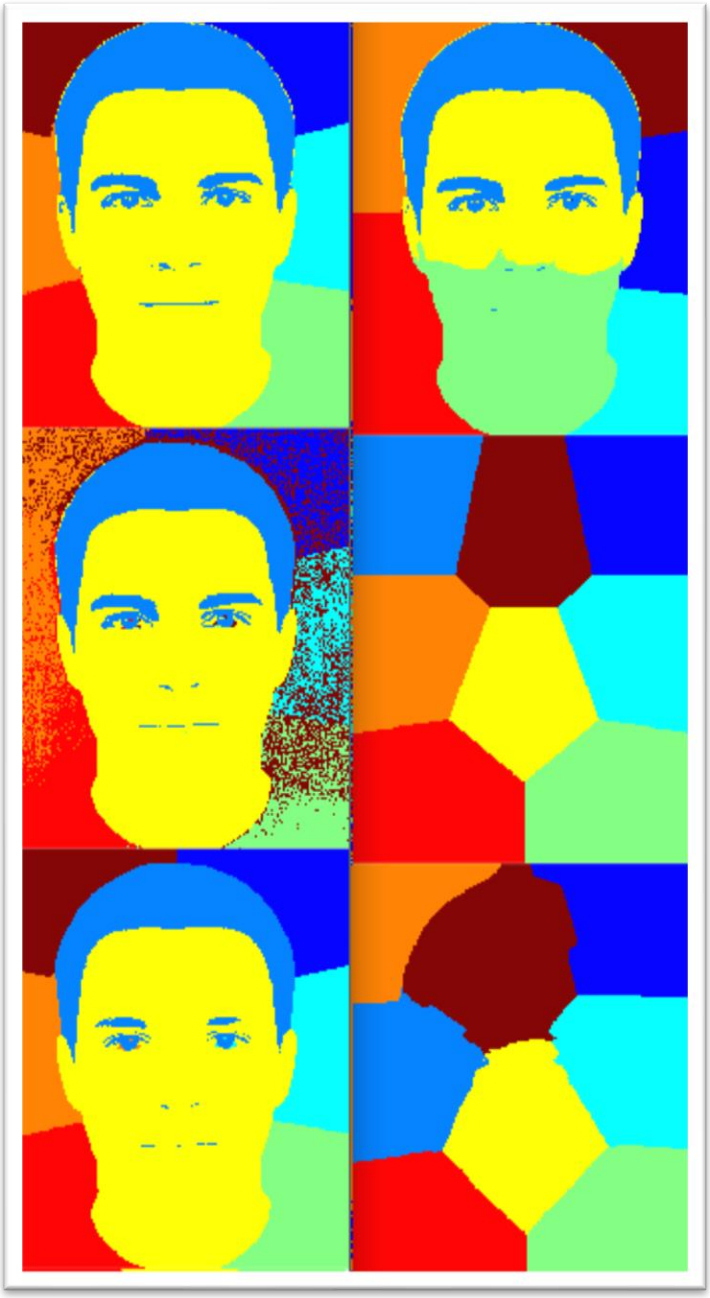

Figure 11: Segmented image with RGB, HSV and LAB features with spatial information (top to bottom). Left – normalized features, right – features not normalized.

Image smoothing is also an important step to reduce the effects of noise and redundant slopes in the images which may not necessarily provide significant information. For the purpose of smoothing a gaussian filter is applied with unit variance. The variance directly influences the scope of smoothing and can be varied depending upon the amount of noise present. Other filters that can be substituted for smoothing may include approximation filter such as box blur or moving average. The effect of smoothing is more prevalent when noisy images are used. Figure 12 to 15 shows segmentation of noisy images with and without applying smoothing filter. It can be clearly observed that the results are vastly improved as smoothing cancels out the noise by averaging.

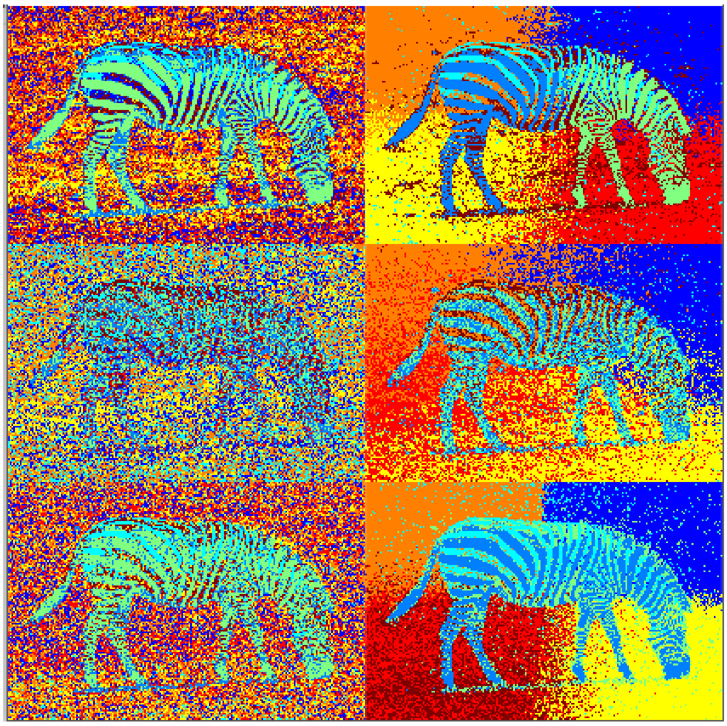

Figure 12: Different segmentations of Zebra image with Gaussian Noise. (No smoothing)

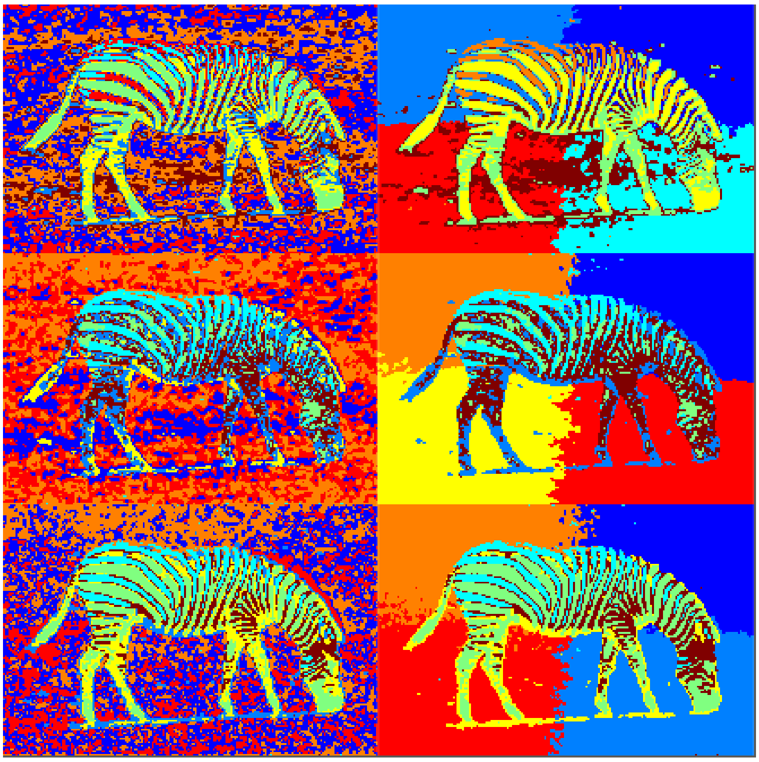

Figure 13: Different segmentations of Zebra image with Gaussian Noise (Gaussian smoothing is applied with variance = 2)

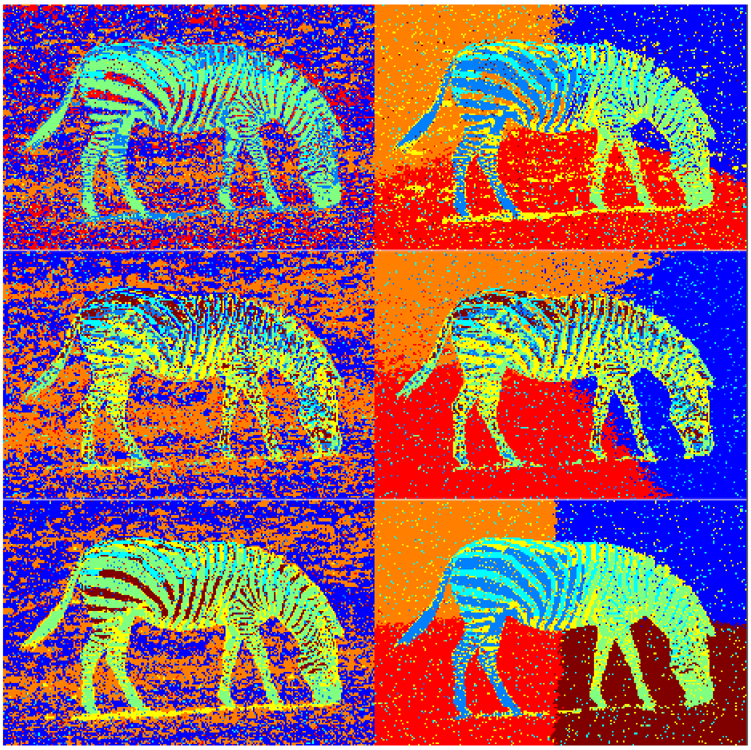

Figure 14: Different segmentations of Zebra image with Salt and Pepper Noise. (No smoothing)

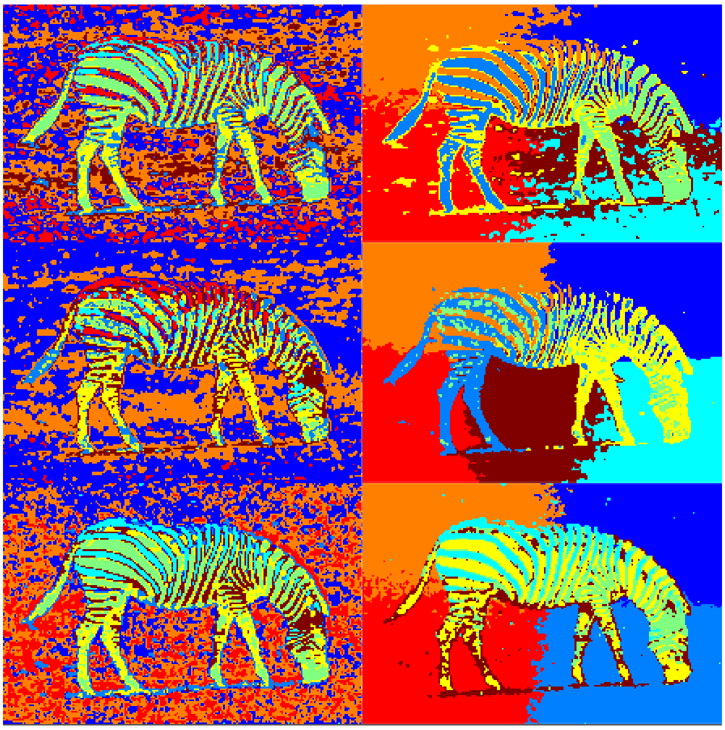

Figure 15: Different segmentations of Zebra image with Salt and Pepper Noise. (Gaussian smoothing is applied with variance = 2)