Segmentation

Johana M. Ramirez Borda
Los Andes University
jm.ramirez11@uniandes.edu.co

Francisco A. Rozo Forero
Los Andes University
fa.rozo1843@uniandes.edu.co

Abstract

On of the most studied problems in the image processing field is the segmentation. This problem is based on the need to cluster o regroup the pixels of an image that belongs to the same object or region. In this project, four clustering methods (kmeans,gmm,hierarchical and watershed) were tested in order to segmentate an RGB image. Also, LAB ans HSV colorspaces were taken into account. Moreover, spatial information such as the (x,y) coordinates of each pixel were used. Overall, the hierarchical method had the best performance while kmeans and gmm followed it. In contrast, watersheds didn't show a good performance because it oversegmentated or separated specific details of the image instead of objects themselves. These methods were evaluated using the Jaccard Index, being hierarchical with an HSV+xy colorspace the best, having an average jaccard index of 0.6109 *in all the images of the dataset.*

1. Introduction

The segmentation problem in image processing is based on the clustering of pixels that belongs to the same regions or objects. Therefore, segmentation can be used for objects recognition, occlusion boundary estimation, image editing, etc [4].

In this project, several segmentation strategies were tested by varying between four clustering methods: kmeans, gmm, hierarchical and watershed. Also, RGB, LAB and HSV feature spaces were used. However, not only color information was used in the clustering methods, but also the spatial information by taking into account the *(x,y)* coordinates of each pixel: RGB+xy, LAB+xy, HSV+xy.

First, kmeans clustering performs partitions of the color channels of the image into k clusters and return the cluster index of each pixel. This information allows to classify each pixel within a region or object and then display it on screen with a specific color.

Second, the gmm (gaussian mixture distribution) method, fits a gaussian mixture distribution with k components to the color channels of the image. Then, k clusters are created and this information is reshaped into an image.

On the other hand, hierarchical clustering is an strategy to built a hierarchy of groups. This method has two main approches: in the agglomerative approach, each cluster starts by beeing one single pixel and then pairs of clusters are merged; in the divisive approach, all the pixels are contained in one cluster and then splits are performed recursively[4].

Finally, he watershed clustering method was used to ”cluster” histograms of the image. The main idea is to consider the gray level version of the image and consider it as a topographic relief, in which actuating an immersion process[1].

2. Materials and methods

For this lab, the software implemented was MATLAB R2016b.

2.1. Database

The database used in this project is a small group of images with their respective annotations from the Berkeley segmentation dataset (BSDS). The creation of this database was supported by the University of Berkeley to promote the study of computational segmentation and boundary detection on images.

This sub-database contains 24 images in RGB, .jpg format and have 321x481 or 481x321 pixels. Some random examples of images are shown in Figure 1.

Figure 1. Random examples of images used for this project.

As it can be seen, there are images with landscapes, humans and animals, all showing different changes in brightness, color and texture. Moreover, each image has as annotations a *.mat* file that contains 5 human annotations consisting of one binary matrix of boundaries and another matrix with the segmented image. These can be seen in Figure 2.

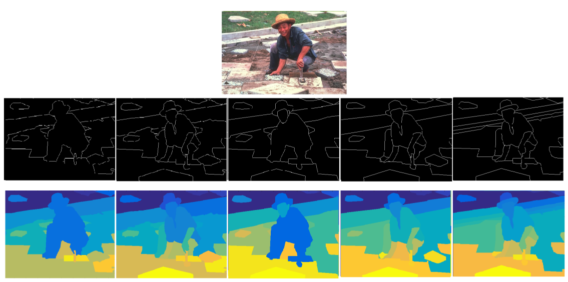


Figure 2. Example of human annotations for one image.

2.2. Image data pre-processing

Before processing, the pixels of the image were rearranged into a vector, starting in the upper left position and moving up to down and left to right. The first three columns corresponded to each color channel RGB/HSV/LAB in that order, the last two columns were used only when the coordinates were also included, being x and y position the 4th and 5th column respectively. However, these arrange only applies for the kmeans, gmm and hierarchichal methods. Additionally, all the data was normalized to prevent data such as the position or the RGB values to have more weight when measuring distances.

On the other hand, when taking into account the *(x,y)* coordinates, it was found an inconsistency because the pixels of the image were in a range of $[0, 1]$ and the coordinates were much larger than 1. To overcome this, each vector of *(x,y)* coordinates was normalized: the vector of *x* coordinates between the number of rows and the *y* coordinates between the number of columns.

2.3. Kmeans Segmentation

Kmeans is a method that based on given vectors containing data of categories, iteratively finds the best centroids that represent each of them based on feature similarity. This similarity is found based on metric distance between vectors. The matrix of the pixels with their respective data values was used on the kmeans function to separate them in *k* clusters. The distance metric implemented was euclidean.

2.4. GMM Segmentation

In this method, the function *fitgmdist* was firstly used in order to study the matrix that contained the three channels (and the coordinates when necessary) of the image and to fit a Gaussian mixture distribution with k components. Then, the function *cluster* was used in order to create the k clusters that were regrouped to form the segmented image.

2.5. Hierarchical Segmentation

This method takes every pixel as a single class and then, by measuring with the euclidean/cityblock metric, it pairs those that are closer until they all become one unique cluster. For this method, the function linkage revalued the pixels to position them on the tree based on their distance, then the dendrogram function displays the dendrogram and gives the assigned group for each pixel.

2.6. Watershed Segmentation

Watersheds method consist on ”flooding” through the regional minimums and barricade where these floods join together; being these barriers highest peaks between minimums. This way, the objects are segmentated based on their gradient. This method requires imposed minima and height of flooding to segmentate, as minimums is where the flooding start, while the height correspond to the threshold where flooding stops. Hence, at higher values of height, lesser regions will appear.

2.7. Evaluation Method

Using our own MATLAB function JaccardInd, by introducing the segmentated image and the groundtruth images, it returns the corresponding Jaccard Index. First, it creates a matrix where the results of the jaccard index between each region from the segmented image, is calculated with each region of the groundtruth image. Then, it picks the best pairs of matching regions based on their jaccards' index, where the number of pairs is equal to the minimum number of regions between the two images. Finally, it calculates the average to give the final jaccard index.

3. Results and discussion

The function was tested for several images; by varying the different clustering methods and the color space, the results of the average Jaccard index and the average processing time are in Table 1.

Table 1. Jaccard index and time spent when testing the segmentation methods with all colorspaces.

Method	Colorspace	Time [s]	Jaccard Index
Kmeans	RGB	12.17	0.2933
Kmeans	RGB+xy	6.15	0.3419
Kmeans	HSV+xy	3.86	**0.5723**
Kmeans	HSV	9.84	**0.5644**
Kmeans	Lab	7.93	0.4709
Kmeans	Lab+xy	7.02	0.4725
Hierarchical	RGB	866.03	0.3752
Hierarchical	RGB+xy	1352.58	0.3845
Hierarchical	HSV	1223.67	**0.5974**
Hierarchical	HSV+xy	1043.19	**0.6109**
Hierarchical	Lab	2056.09	0.5347
Hierarchical	Lab+xy	1190.02	0.5563
GMM	RGB	20.24	0.4673
GMM	RGB+xy	14.128	**0.5207**
GMM	HSV	22.31	0.5012
GMM	HSV+xy	36.77	0.4337
GMM	Lab	17.75	0.414
GMM	Lab+xy	22.98	**0.5531**
Watersheds	-	3.41	0

As seen in Table 1, hierarchical was the method that performed the best regarding similarity of segmentation. It means this method had the highest values of Jaccard index: 0.5974 for HSV space and 0.6109 for HSV+xy colorspace. Excepting for the RGB colorspace, its overall jaccard indices are above the other methods, being HSV colorspaces the best of all. This method is efficient because it clusters the pixels one by one instead of centroids. However, the time required to perform the hierarchical method is high because it requires minutes to process one image (from 14 to 34 minutes depending on the space), while the other methods only take seconds. Yet, once the hierarchy tree is calculated, all the scale segmentation's are fast to obtain.

Moreover, as seen in figures 3 to 6, varying the colorspace changed the segmentation results for each clustering method. This observation has also been found in several studies [2] [3]. More important, it has also been found that combining the information provided by more than one colorspace is a useful way to improve the segmentation results [2]. For these reasons it makes sense to try and test different spaces until finding the best for the method being implemented. Also, observing Figure 5 it is possible to see that adding the position xy allows to cluster based on the object's position in the image instead of only their color information, improving the performance. This information agrees with Table 1 in which it is noticed that the Jaccard index always is better when the xy coordinates are used. Therefore, adding spatial information is useful when trying to segmentate objects that share colors with others but are far from them. Furthermore, each colorspace defines colors and intensities with different values, making each one of them unique and useful for different applications.

On the other hand, looking at Table 1. and Figures 3 to 6 it was noticed that in most cases the hue channel from the HSV colorspace is the most discriminative between the other channels as it contains information about all the colors. All the channels from the RGB colorspace represent an specific color with their different saturation scales and need to be combined to form other colors. Likewise, happens with the ab channels from the Lab colorspace, as those two channels represent 2 different colors each and need to be combined to form other colors. Finally, the L in Lab and the V in HSV work the same way and only give information of darkness of the colors, making them less representative than the hue channel.

Finally, the evaluation method calculates the jaccard index for all combinations between the segmentated regions from the image and the groundtruth, then it chooses the best values among that matrix and then it computes the average value. The results obtained when testing the function to evaluate the methods are shown in Table 1.
While calculating this index, it was noticed that when the number of regions between the two images differs, the jaccard index lowers as well. Additionally, sometimes when picking the maximum values in the matrix, doesn't bring the best combination of paired regions. Hence, a graph's optimization model can solve it.

3.1. Limitations

Segmentating just with one colorspace might not provide enough information about an image or the objects that are going to be segmentated. Also, not all images respond the same way to each colorspace; some have a better representation when segmentated with one colorspace or another. Same happens with the number of clusters, all images require a different number of clusters to be represented with.
Fureheremore, none of the methods had a reliable performance. The information given by only three or five parameters is not enough information to separate objects. Also, details from the objects tend to be segmentated separately from them, specially when the xy position is not provided. Moreover, the watersheds method requires to evaluate several values for the extended min method to acquire the desired segmentation without separating combining objects or over-segmentating (see figure 5). Additionally, watersheds

was susceptible to strong edges of the

3.2. Further work

In order to improve the performance of the algorithm, adding more information of each pixel such as combining the information of two colorspaces or three of them may give a more reliable segmentation. Also, trying to segmentate using the previous information with hierarchical, kmeans, gmm, mean-shift and normalized cuts, all with the same number of clusters. Then, using the intersection between the different segmentations, it can weight how much each pixel correspond to each region and then assign them based on this data.

4. Conclusions

According to the results, the watershed method didn't care about the position of the pixels, because the results were the same regardless of the position data. However, this method is imprecise when clustering specific areas because it does based on the topography of the image instead of the information of the pixels. In contrast, the hierarchical method was the best method regardless of the time it spends, as it compares every pixel instead of trying to cluster them by centroids. This make it less susceptible to noise or outliers. The opposite happens with kmeans, which is susceptible to outliers and gmm presents noise in the contours of the objects. Nevertheless, kmeans with HSV colorspace and gmm with Lab+xy colorspace had a close approach to the hierarchical method and did it in a more efficient time.

References

[1] M. Bicego, M. Cristani, A. Fusiello, and V. Murino. Watershed-based unsupervised clustering. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 83–94. Springer, 2003.

[2] D. J. Bora, A. K. Gupta, and F. A. Khan. Comparing the performance of l* a* b* and hsv color spaces with respect to color image segmentation. *arXiv preprint arXiv:1506.01472*, 2015.

[3] D. Girish, V. Singh, and A. Ralescu. Extended pixel representation for image segmentation. 2016.

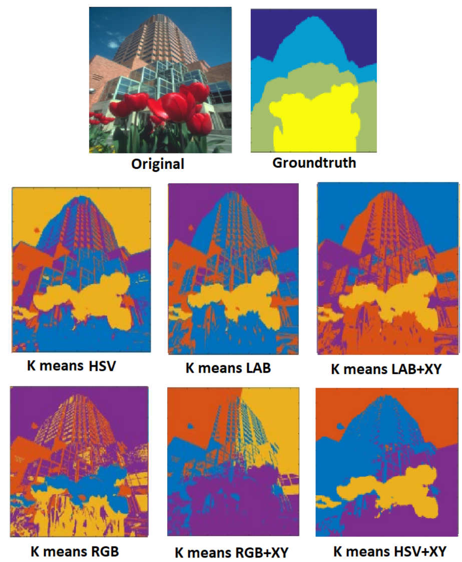
[4] A. Jepson and D. Fleet. *Reading on Segmentation*. 2007.

All the codes created and used are attached.

5. Annex

5.1. Examples of segmentation results

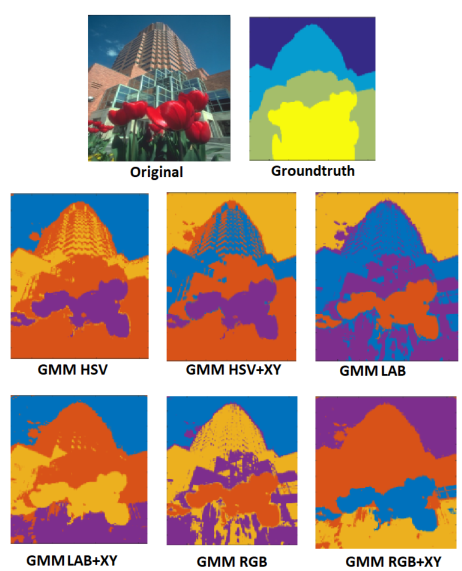
5.1.1 Kmeans



,

Figure 3. Segmentation through kmeans for K=4 in one image

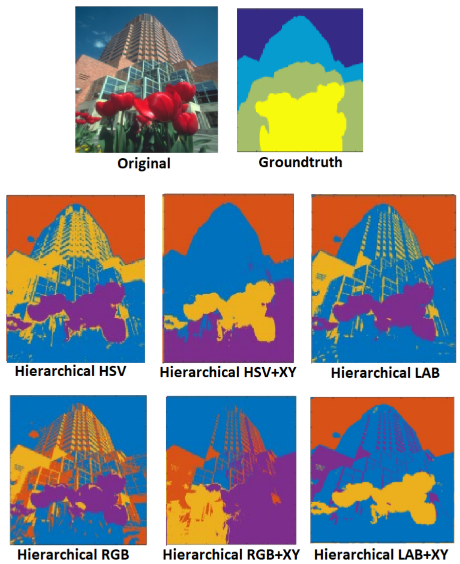
5.1.2 GMM



,

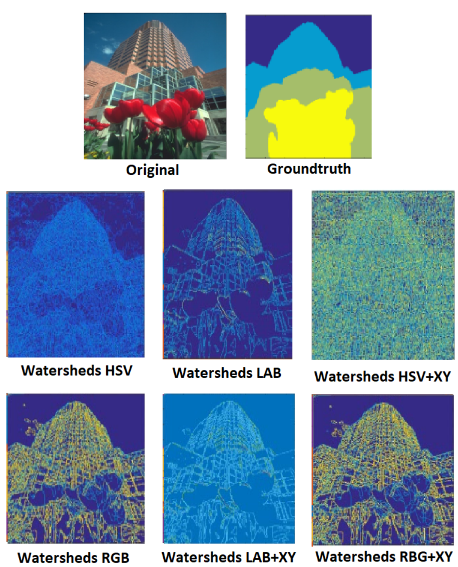
Figure 4. Segmentation through GMM for k=4 in one image

5.1.3 Hierarchical



,
Figure 5. Segmentation through Hierarchical for k=4 in one image

5.1.4 Watersheds



,
Figure 6. Segmentation through kmeans for k=4 in one image