Segmentation - Lab 6 - Computer Vision

Yesith Juez
University of the Andes
jy.juez@uniandes.edu.co

Nicolas Florez
University of the Andes
n.florez228@uniandes.edu.co

Abstract

Segmentation of objects present in images is a important computer vision objective as we can use it to identify objects in the images and then interpreting its content. BSDS or Berkeley segmentation datasets is an approach of a database that represents a set of natural images to test methods against human anotations to observe and evaluate the performance of the machines and its implemented algorithms. We used this database with an implementation of kmeans gmm and watersheds with different colorspaces as RGB, LAB, HSV and we mix spaces of color with spatial information to use dimensions as RGB+XY, LAB+XY, HSV+XY. The best algorithm was the watershed in both time and accuracy but is important to always select a good height (H) as the algorithm is very sensitive to the change.

1. Introduction

In computer vision, image segmentation is one of the most widely studied problems [1]. In short, segmentation is looking to divide an image into several regions, where ideally each region corresponds to an object [2]. This area of interest has led to the development of a large number of methods for image segmentation. In this laboratory, 4 clustering methods (k-means, GMM, Hierarchical and watershed) will be applied in different color spaces, in order to evaluate which of the methods is better in a set of selected images.

2. Materials and methods

2.1. Description of the Database

The database used for the realization of this laboratory is a portion of the Berkeley Segmentation Dataset and Benchmark (BSDS500) is a database made by the Berkeley Computer Vision Group specificaly to be used in the cmputervision problems of Contour Detection and Image Segmentation.

Its a second version of a previous created one BSDS300 they used the 300 images of that database for training and validation and they added 200 more images with human annotations for testing.

Each image of the database is human annotated by 5 subjects in average.

The images are colored formated as JPEG that could be in landscape or portrait 481 x 381px.

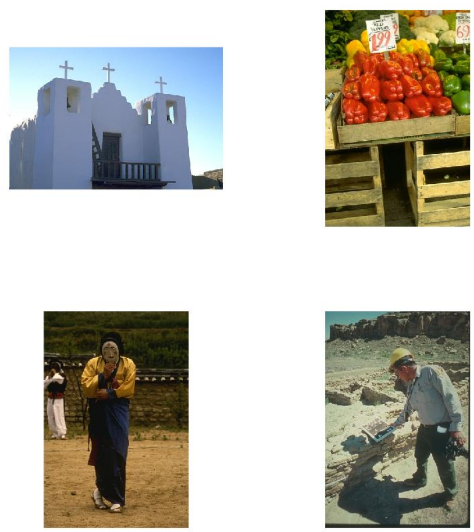

Figure 1. 4 random example images of the database

2.2. Clustering algorithms

In the present laboratory, four clustering methods were used to perform the segmentation in each of the colored spaces: k-means, GMM, hierarchical and watershed.

2.2.1 K-means

Also called Lloyd's algorithm its an unsupervised machine learning algorithm that iteratively cluster the data to the group where mean value is closer. The cluster in every iteration recalculates its center and the new cluster grouping of the data with the cluster center evolution.

The steps required by the algorithm are:

1. Choose the desired amount of clusters k.

2. Calculate every distance of the data observation to each cluster centroid.

3. Assign the data to its closest centroid.

4. Calculate the average of distances of each cluster to get the new centroid location.

5. Repeat until the centroids has minimum movement.

2.2.2 GMM

GMM is an algorithm that behaves almost equally as k-means but it can use a distribution of gausians that could make grouping with different distributions that as expected will classify better as the clusters are not expected to distribute uniformly as k-mean expects.

2.2.3 Hierarchical Clustering

Is an algorithm that groups data over different scales by creating a dendogram or a tree of clusters. The resulting tree represent a multilevel hierarchy, clusters at one level are joined as clusters at the next one.

The steps required by the algorithm are:

1. Find the similarity or dissimilarity between every pair of objects in the data set.

2. Group the objects into a binary, hierarchical cluster tree.

3. Determine where to cut the hierarchical tree into clusters.

2.2.4 Watershed

Its a grouping algorithm based on the river systems in the nature and carry the name of a ridge that divides areas drained by different river systems. Applying the same natural concept to a grayscale image it could be say that the image is a surface where the rivers and the mountains are opposite gray tones so it could be found watershed when a portion of the image abruptly changes its components that could inform the segmentation of different objects in it.

The summary of the methods applied specifically in the computer vision problem of segmentation is:

The K-means algorithm allows the segmentation of an image by means of the realization of several clusters (minimum 2), where each cluster is assigned the pixels that have more similarity as they will be points that groups together in the hyperspace used. Having each group of clusters, the segmented signal is obtained using a ”reshape”.

the GMM algorithm is similar to the k-means algorithm. The only difference is the use of Gaussian distributions for the realization of each cluster.

The hierarchical algorithm allows the segmentation of images through a hierarchical grouping that groups the data in a variety of scales through the creation of a dendrogram, a tree shaped data that groups leafs into categories .

The watershed algorithm allows the segmentation of the image taking into account the magnitude of the gradient. In order not to generate over-segmentation, the minima imposition technique is used for the creation of markers.

2.3. Segmentation parameter tunings

For the segmentation with the k-mean and GMM method, the parameter to be established is the number of clusters. In the case of the watershed method it is necessary to establish the number of markers.

For the k number of clusters we selected 5 clusters as we wanted to differentiate sky from wall and from wood surfaces and observe if the algorithms were able to make that differentiation that we wanted.

In the case of the number of markers for the watershed method we selected h = 10. This was done by trial and error observing the behavior of the algorithm as we variate the hiperparameter. we fixed 10 as the number that for us segmentate the best the image.

2.4. Image preprocessing

For the image preprocesing we admited rgb images but the idea is to use different segmentation algorithms an observe their performance in different color spaces. So the function created do a image preprocesing of the color spaces as the user request it. So admiting RGB images we can use transformed LAB, HSV or the original RGB as color spaces but as segmentation algorithms could improve using dimensions of spatial distribution we also used RGB+XY, LAB+XY and HSV+XY.

To achieve the previous transformation we used the matlab functions

```
rgb2lab
```

```
rgb2hsv
```

And in the case of the spaces where also have spatial distribution we concatenate to the image matrix two more dimentions of the position in x and in y of the pixels as follows:

```
for i=1:size(rgbImage,1)
            for j=1:size(rgbImage,2)
                he(i,j,4)=j;
                he(i,j,5)=i;
            end
        end
```

2.5. Evaluation

In order to perform the evaluation of a general segmentation problem it is necessary to have a groundtruth made by humans. Currently, there is a database that allows the evaluation of different segmentation methods (the BSDS500 database). This contains a large amount of groundtruth made by humans of natural images. The evaluation metrics are made from the precision-recall curves and Average Classification Accuracy (ACA).

The evaluation strategy used for this particular problem was to use the Jacard index as an evaluation metric. The procedure was the following:

- To represent the segmented image and its Groundtruth in different binary masks (each mask refers to a contour of the image).
- To each binary mask find the Jaccard index:

$$J = A \cup B / A \cap B \tag{1}$$

  Where A is a mask of the segmented image and B is a mask of the Groundtruth.
- Add the Jaccard indices found of each mask.

Figure 2. Original image to be tested

The Jaccard index was chosen as an evaluation measure because only two images need to be compared (segmentation of the original image and its Groundtruth). being this metric adequate to evaluate the segmentation methods developed.

3. Results

3.1. Kmeans

The segmented images of the kmeans method can be found in the Figures 4, 5, 6, 7 ,8 and 9. They were used all the mentioned spaces just color and with spatial information.

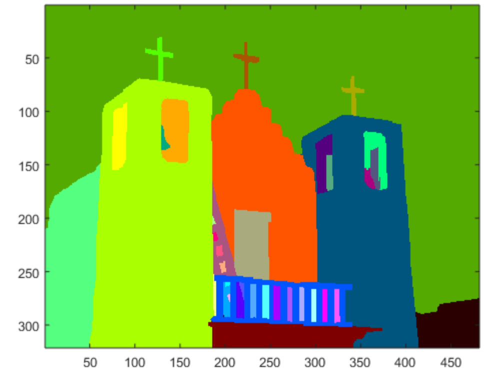

Figure 3. mean segmentation made by humans

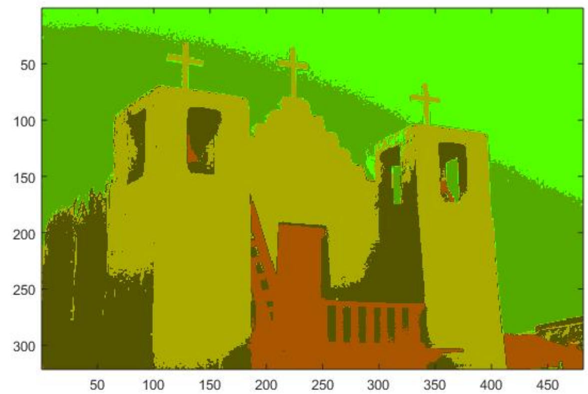

Figure 4. Image clustered by kmeans in rgb space of color

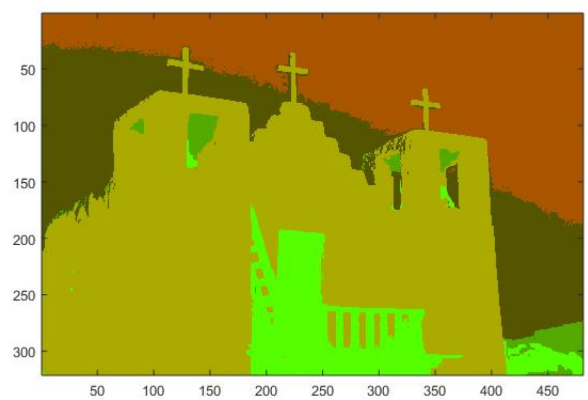

Figure 5. Image clustered by kmeans in hsv space of color

3.2. GMM

The segmented images of the gmm method can be found in the Figures 10, 11 and 12.

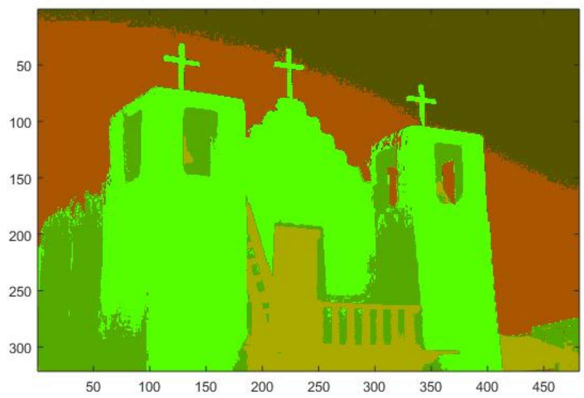

Figure 6. Image clustered by kmeans in lab space of color

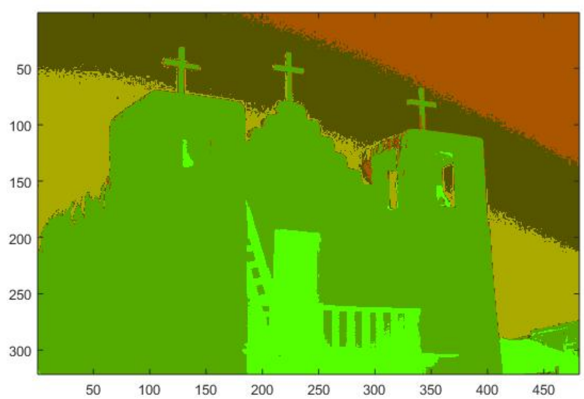

Figure 7. Image clustered by kmeans in rgb+xy space of color

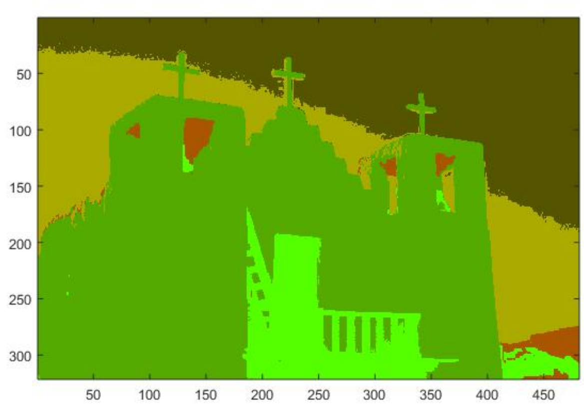

Figure 8. Image clustered by kmeans in hsv+xy space of color

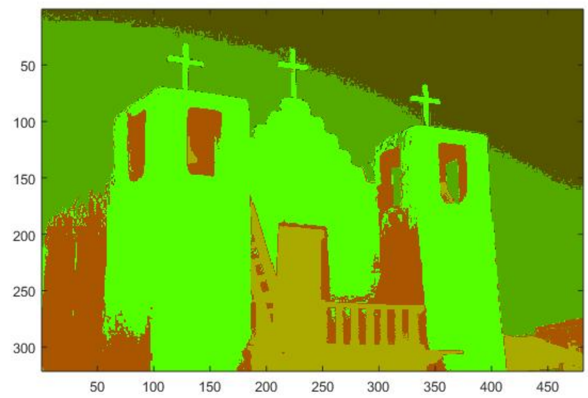

Figure 9. Image clustered by kmeans in lab+xy space of color

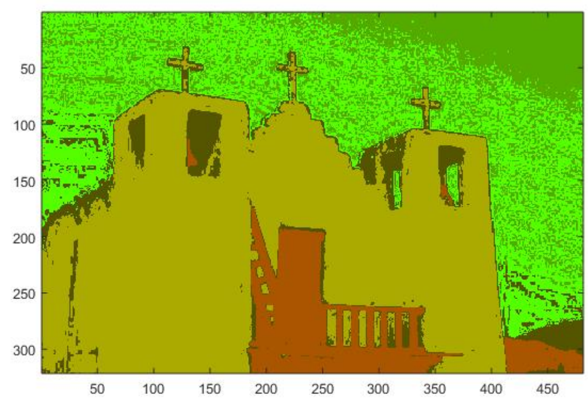

Figure 10. Image clustered by gmm in rgb space of color

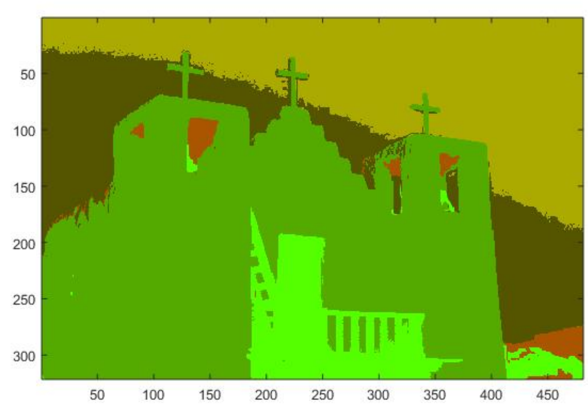

Figure 11. Image clustered by gmm in hsv space of color

3.3. Watersheds

The segmented images of the Watershed method can be found in the Figures 13, 14 and 15. They were only used in the color spaces without the spatial info. as the method as previously explained they use a gray image to apply the watersheds, and it already use spatial information.

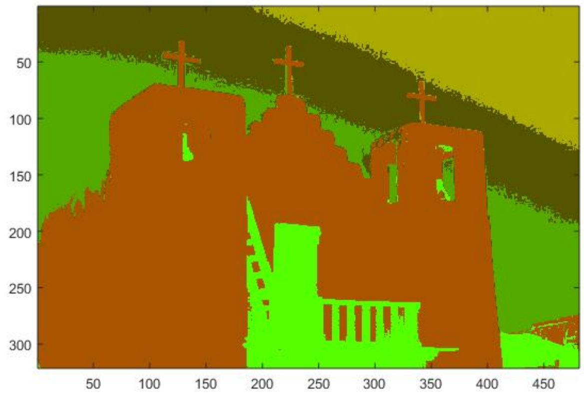

Figure 12. Image clustered by gmm in lab space of color

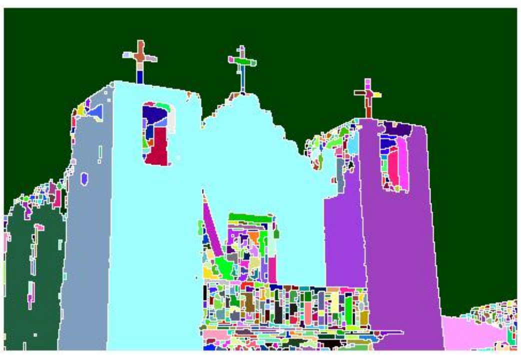

Figure 13. Image clustered by watershed in rgb space of color

Figure 14. Image clustered by watershed in hsv space of color

Figure 15. Image clustered by watershed in lab space of color

	RGB	RGB-XY	Lab	Lab-XY	HSV	HSV-XY
Kmeans (k= 5)	0,4249	0,4630	0,4741	0,5167	0,3696	0,6241
GMM (k=5)	0,4499	0,5256	0,4077	0,5246	0,3707	0,4517
Watershed (h=10)	0,6638	-----------	0,0780	----------	0,2768	------------

Figure 16. Evaluation of Segmentation Methods with Jaccard Index

4. Discussion

4.1. Kmeans

It can be seen that the RGB color space using the k-means method does not have much coverage it can be a lot noisy in the edges of the items within the images. This happens because the RGB space does not take into account the image to spatial level. It is observed that in the HSV and LAB color spaces an improvement in the k-means method were achieved as it presents less noise in the edges of the objects but still the whole 3 color spaces mistakes segments of the image as different objects as the sky dividing it in 2 because of the change of colors along the image of it. We expected that segmentation with the k-means method were better in the RGB-XY LAB-XY and HSV+XY spaces. but they also mistakes the sky as 2 separate objects, even they were used spatial information the sky is so big that it wasnt found a cluster that were shaped by the spatial information in this case. In all the cases the segmentation level is below to the desired level (see figure 2).

4.2. GMM

Using the GMM method, the segmentation improves considerably comparing them with the Kmeans algorithm in the spaces HSV and LAB. The color space that is not efficient is the RGB as it is the noisiest segmentation of all the analyzed. The noise can be seen clearly in the sky.(see figure 9, 10 and 11). This could be that the algorithm is so strong it can see the noise of the image in the sky captured by the camera an classify it as object. or simply that the algorithm is not so good. It also important to notice it wasn't used spatial information so it can support the theory of the camera noise that is irregular an not spatial located.

4.3. Watersheds

In the method watershed segmentation, results are very close to the segmentation image made by humans (see figure 2, 12 and 13), except for the LAB color space, where the results are not very positive. In the LAB space it is observed that an over-segmentation occurs (see figure 14). In the 2 spaces we can see a lot of noise but i we re threshold the small watersheds we can achieve the almost human segmentation as the edges are finner than Kmeans and gmm. The sky is also a single segment of the image so we can change a better H of the watershed method to remove this noise.

5. Conclusions

The watershed algorithm is overall better in time and in accuracy but is difficult to select the hiperparameter H-minima. Selecting a wrong parameter number can lead to bad results.

The algorithm for gmm is heavy time consuming but is a lot reliable because is less sensitive to the hiperparameters than watersheds.

The algorithm kmeans is the heaviest in time but its better when using space dimension along with the color space. it consumes less resources so it can handle better the images if time isnt a key factor in the usage.

The space color of the image is a key factor that would determine the well behavior of a clustering algorithm.

References

[1] R. Szeliski, "Computer Vision: Algorithms and Applications", Springer, 2011.
[2] P. Arbelez, "Computer Vision: Grouping 01", University of the Andes, 2018
[3] "K-means clustering - MATLAB kmeans - MathWorks America Latina", Matlab Mathworks, March 30 2018
[4] S. Lloyd, "Least Squeares Quantization in PCM", IEEE Transactions on Information Theory, 1982
[5] "Gaussian Mixture Distribution - MATLAB & Simulink - MathWorks America Latina", Matlab Mathworks, March 30 2018
[6] "The Watershed Transform: Strategies for Image Segmentation - MATLAB & Simulink", Matlab Mathworks, March 30 2018