Laboratory 6 - SuperPixels

Guillaume Jeanneret
Universidad de los Andes
Cra 1 #18a-12, Bogotá, Colombia
g.jeanneret10@uniandes.edu.co

Jorge Madrid
Universidad de los Andes
Cra 1 #18a-12, Bogotá, Colombia
ja.madrid2714@uniandes.edu.co

Abstract

Various segmentations methods were used to create various regions depending on chromatic and spatial characteristics of pixels. The methods used were k-means, gaussian mixture model, hierarchical and watershed segmentation.

1. Introduction

Segmenting an image is completing the task of finding a group of pixels which share a similar characteristic [1]. For example, in figure 2.3, the group of pixels in the structure *Lion* share the characteristic of being in the tiger. The programmers in this field have the task of generating an algorithm capable of grouping a pixel with its cluster. Various algorithms were created but the solution given by them may be further improved.

Some characteristics of the segmentation include a local pattern or a global one. In this paper, four methods are used: k-means, Gaussian Mixture Model (GMM), hierarchical clustering and watersheds.

2. Materials and Methods

2.1. Database

The database is composed of 8 pictures with their annotations, where each image is segmented with different amounts of clusters.

2.2. Methods

Given a color image in RGB format of dimensions $N \times M$, a matrix of dimension $NM \times k$, $k = \{3, 5\}$ is created in such way that each row represents a pixel. If the location is wanted to be used in the feature of the pixel, the dimension of each pixel is 5, otherwise the dimension is 3. The next function is implemented in MatLab in order to segment each image:

```
segmentByClustering(rgbImage,...
        featureSpace,method,clusters);
```

Figure 1. Tiger segmented from the background.

Where *rgbImage* is the color image, *featureSpace* is the 'rgb', 'lab' or 'hsv' color space with or without '+xy', meaning the incorporation of the spatial components, *method* -which is self explanatory- with possibles values of 'k-means', 'gmm', 'hierarchical' and 'watershed' and, finally, the number of clusters.

2.3. Algorithm

K-means

K-means unsupervised clustering method is an algorithm that given the number of clusters K and a array of vectors (training data) it extract the center vectors that represent each category. This central vectors are generated randomly taking K vectors from the array. Then, the algorithm assign each vector a label depending on the minimum distance to each cluster. Then, within each group, a new center is calculated. Iterating this process, when the centers don't move -greatly- anymore, those are the clusters of the algorithm.

GMM

This algorithm tries to approximate a each group with a multivariable-normal distribution. This task requires approximating the covariance matrix for each group and assigning each vector a probability of being in a cluster, then getting classified each vector does not give a label. This classification of the vectors is nominated *soft classification* [1]. The classification of a vector is given by the label of the i^{th} group where the vector has mayor probability.

Hierarchical

Hierarchical segmentation algorithm consist in comparing each vector and grouping each pair using a distance -like the euclidean or the taxi distance- one pair at time. This pairing is done with the minimum distance. Iterating this process, when the number of groups is achieved the labels are assigned depending on with branch of the tree each vector is. Because this classification takes to much time (up to 1 hour using 35% of the image), each image was reduced to a total of 80% of the original resolution.

Watershed

This algorithm attempts to segment the image by finding division lines between regions. These regions emulate the ridges that divide the valleys between adjacent valleys. Such divisory lines may be found by artificially flooding the intensity landscape of an image. This method on its own, however, usually yields an oversegmentation of the image due to the local variation of the intensity, even in homogeneous background regions. This problem may be overcome by imposing minima to begin the flooding. This, however, has the inconvenience of demanding interaction with the user or additional a priori information. Another solution is to only consider as initial minima, regions with a contrast difference above a given threshold. The value that this threshold takes must be fixed arbitrarily. Moreover, the code may not guarantee to segment the image in a given number of regions.

3. Results

Some results are shown in the images section. Selecting the effect of the method can be seen in images 3 for k-means method, 4 for GMM, 5 for hierarchical and 6 for watersheds, where 2 is the original.

The effect of the color space can be seen in the pictures 7 (original), 8 (RGB effect), 9 (HSV effect) and 10 (La*b* effect).

The effect of adding the space lieux component is shown in the following figures: 11 -original-, 12 -RGB feature space- and 13 -RGB+XY as feature space-.

Respect the time theme, in table 1, those are the times that our computer took to segment only one picture using *HSV+XY* and a total number of 4 clusters. The total time it took to segment every image using every clustering method with various clusters (between 10 and 30 clusters per image) was 21.72 hours.

Method (4 clusters only)	Time(second)
K-Means	3-5
GMM	6-10
Hierarchical	20-30 (can vary greatly)
Watershed	3-5

Table 1. Time results segmenting one image using only

4. Discussion

- The comparison of images 3 , 4 , 5 , and 6 shows that, of all methods, segmentation by watersheds with contrast minima yielded the best results as it did not oversegment homogeneous regions such as the sky in the background or the mountains themselves.
- The segmentation of all the images took too much time, specially with the hierarchical segmentation. The time can vary from 10 to 20 minutes per image using this method and a number of clusters big enough.
- As seen in the image 14, segmentation using K-Means method with La*b*+XY as feature space and comparing it his segmentation without the spacial component, figure 15, the result can vary, specially using a lower amount of clusters. This phenomena is caused by the impulse given by the coordinates. Improvements can be made by normalizing every dimension of each vector. By doing this, the weight of each component can be compared.
- GMM algorithm gives great results, as seen in 4, but it can fail giving certain segmentations sometimes. This is because the algorithm will drop an ill-conditioned covariance matrix. The error was fixed using a *Regularization Value*, but sometimes the algorithm will not give any answer.
- Colorspaces alone are not always sufficient to perform appropriate segmentations. This is exemplified by figures 8 , 9 , and 10. Note that the original image 7 exhibits a large number of pixels with very low saturation and similar value, such as those in the lower section on the sky and those in that make up the church. No chromatic channel can effectively discriminate all these pixels into large, uninterrupted regions. The oversegmenatiton that we appreciate are caused by very small variations in the colorspcae features, thus the blunt edges. Other features must be incorporated into the segmentation criterion to obtain better results. This calls for more robust segmentation criteria. For instance, considering chromatic and spatial information yields a much better segmentation, as in 16.
- These methods have limitations. When the number of cluster is high, various artifacts can be extracted without giving any important result. For example, every segmentation of 7, the church got partitioned into several regions. In the other hand, when the number of clusters is too low, a effect similar to figure 15 will be

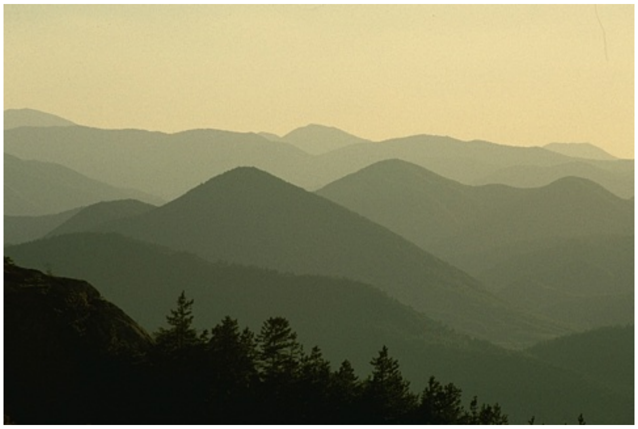

Figure 2. Original landscape photo

generated. This partition of the image mixes various regions, like the background and the clothes of the human. But, as shown in figure 16, picture segmented using the exact same parameters as 15, the church is almost segmented perfectly.

5. Conclusion

- The adequate number of clusters to correctly segment an image is image-dependent. This poses a dificulty to elaborate more robust segmenting algorithms.
- Because CPU and RAM memory are well known constraints, the hierarchical segmentation is not recommended because, at least using this methodology, the result is *pixelated* as shown in 5.
- Segmentation based solely on color cues yields limited results in umages with homogeneous chromatic distributions or low saturations. Incorporating spatial cues yields improved results.
- Although watersheds may result in oversegmentation, imposing contrast minima enhances its performance tu to point of surpassing pixel grouping techniques.
- Using spatial location for segmenting a group of adjacent structures can be a simple way to get them in the same region. But this method only will give a good result if the number of clusters is high enough; otherwise the result can be very poor.

References

[1] R. Szeliski. *Computer Vision: Algorithm and Application*. Springer, 2010.

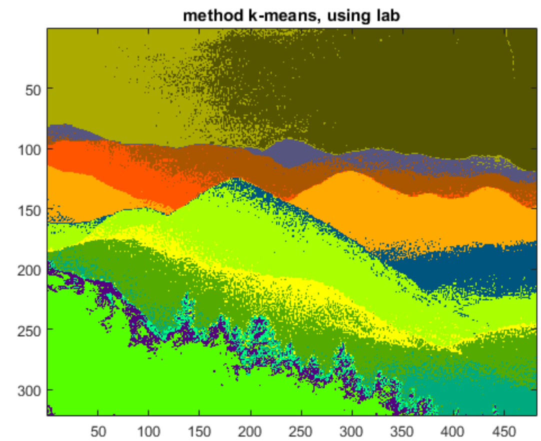

Figure 3. Segmentation by K-Means with 14 clusters using La*b* color space

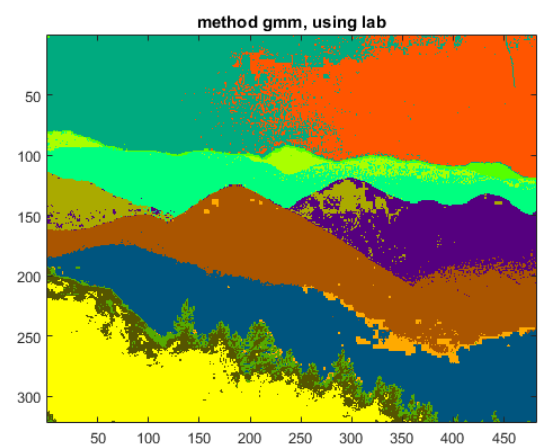

Figure 4. Segmentation by GMM with 13 clusters using La*b* color space

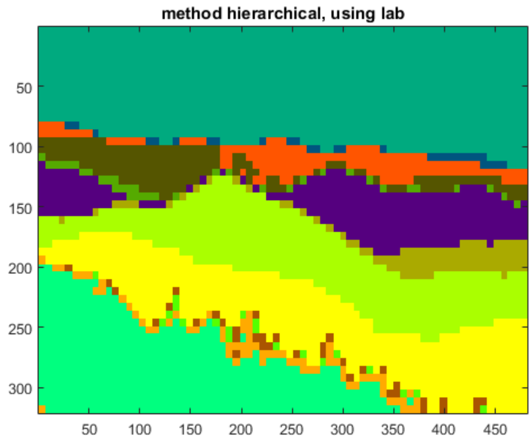

Figure 5. Segmentation by Hierarchical algorithms with 13 clusters using La*b* color space

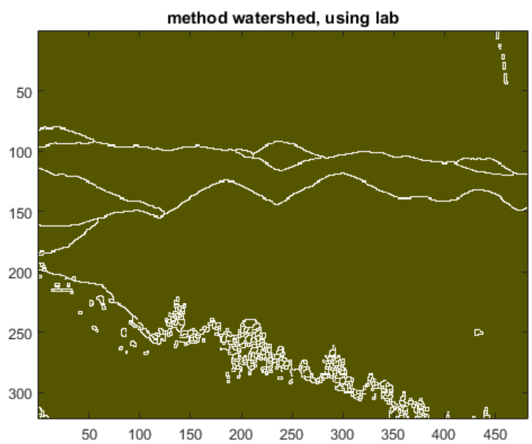

Figure 6. Segmentation by watershed with 13 (approximately) clusters using La*b* color space

Figure 7. Original church photo

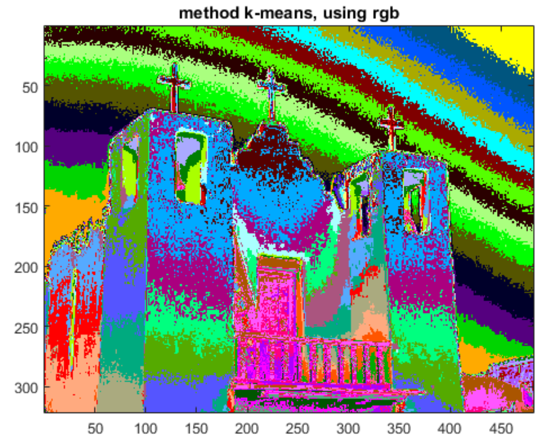

Figure 8. Segmentation using k-means and RGB space color

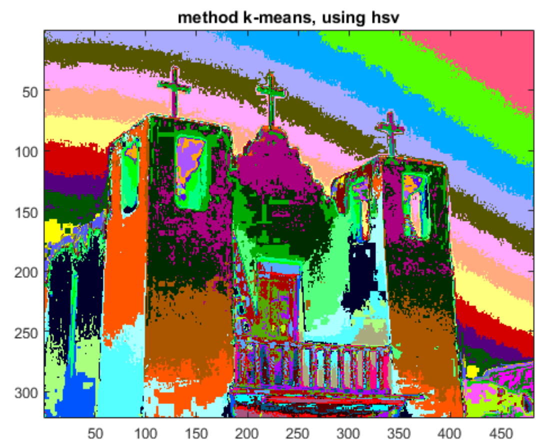

Figure 9. Segmentation using k-means and HSV space color

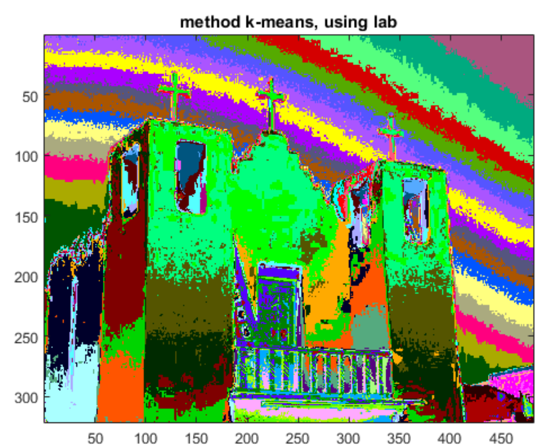

Figure 10. Segmentation using k-means and La*b* space color

Figure 11. Original pepper photo

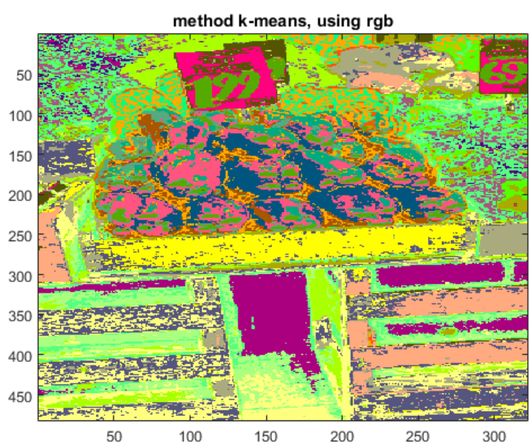

Figure 12. Segmented image using k-means and RGB as color space

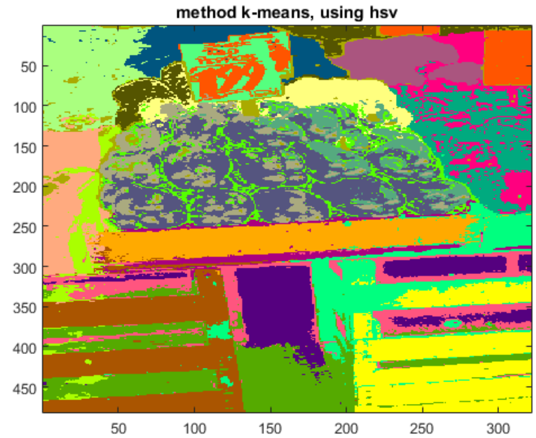

Figure 13. Segmented image using k-means and RGB as color space plus the coordinates

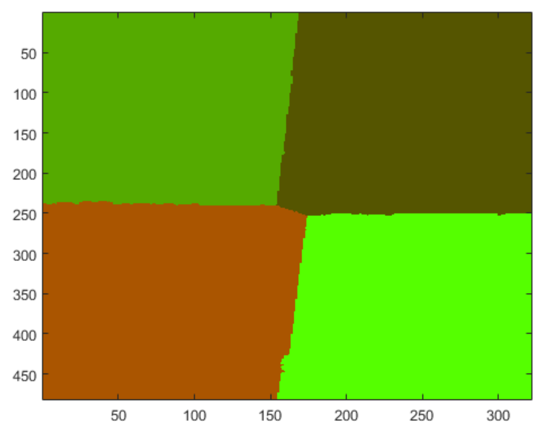

Figure 14. Segmentation using K-Means with $K = 4$, using La*b* feature space with spacial components.

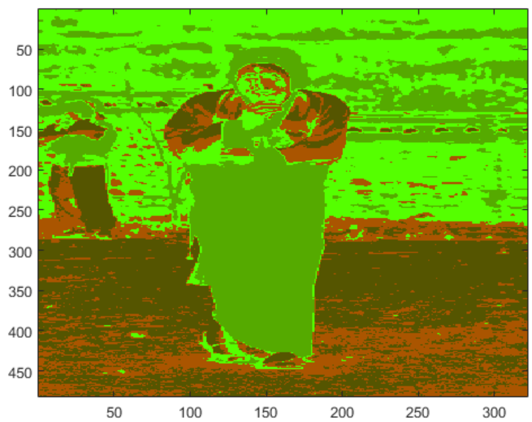

Figure 15. Segmentation using K-Means with $K = 4$, using La*b* feature space without spacial components.

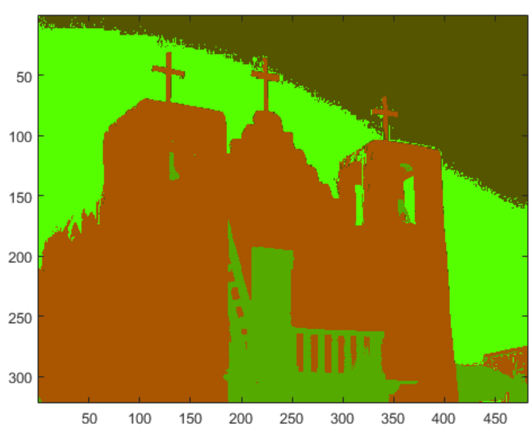

Figure 16. Segmentation using K-Means with $K = 4$, using La*b* feature space without spacial components.