Segmentation

Federico Ulloa
Universidad De Los Andes
Cra. 1 18a-12, Bogotá, Colombia
f.ulloa10@uniandes.edu.co

Esteban Vargas
Universidad De Los Andes
Cra. 1 18a-12, Bogotá, Colombia
e.vargas11@uniandes.edu.co

Abstract

Image segmentation is one of the main areas of study in computer vision. It consists in assigning each pixel of an image to a group. In this laboratory, we built a function that performs segmentation to a given image with k (number of clusters or groups), clustering method and color space as parameters. Clustering can be done by K-means, GMM, hierarchical or watersheds; while color spaces can be RGB, Lab, HSV, RGB+XY, Lab+XY or HSV+XY. After performing segmentation to the test images used with each of the different parameters, we determined that the best performance was obtained with GMM and K-means clustering in Lab color space.

1. Introduction

Segmentation is one of the main branches of computer vision. It consists of determining the pixels in an image that belong to an object of interest. A wide variety of methods exist for this purpose, and the main idea of all of them is to do segmentation by clustering the pixels of the image. There are also several methods to cluster data, so the ones used in this work will be briefly explained.

1.1. K-means

The first one is K-means. This is the most basic non-supervised clustering method. The main idea is that, given a k number of desired clusters, the algorithm puts k centroids randomly in the data. Then, elements in each cluster are assigned by finding the closest centroid with a defined distance. Then, the centroid of each resultant cluster is calculated and the process iterated until it reaches a non-changing result.

1.2. GMM

The second method is the mixture of gaussians (GMM). This methods is a generalized version of k-means. In GMM clusters are formed by representing the probability density function of observed variables as a mixture of normal densities.

1.3. Hierarchical

Another method is the hierarchical clustering. As its name suggests, this type of clustering produces a hierarchy of clusters. This can be made by top-down or bottom up methods. Top-down starts with all data in one cluster and starts dividing them until every data is in a different cluster. Bottom-up is the opposite, starting with each data in a different cluster and then grouping similar clusters.

1.4. Watersheds

Lastly, watersheds is a method based on contours. The algorithm finds the regional minima and starts "flooding" the image, fomring lakes. When two lakes are about to meet, a dam is constructed in so that lakes stay independent. When the image is completely submerged, the set of dams is the watershed lines.

2. Materials and methods

2.1. Database

The dataset used to test the segmentation function is a very small subset of only 2 images of the BSDS. This color images are in JPG format and have sizes of 481x321 and 321x481.

2.2. Segmentation function

A simple segmentation function was created with an image, a color space, a clustering method and a number of clusters as parameters. For this, a subset of the BSDS500 database was used. Depending on the color space entered as parameter, the image is converted to that requested space. Then, the image is segmented by the clustering method and number of clusters entered as parameters. For the watersheds method, some preprocessing was made so that we could obtain markers to run the algorithm with k clusters and better results.

In the RGBXY space, the X and Y channels had to be rescaled, so that all 5 channels were between 0 and 255. Also, in the hierarchical clustering, the images had to be rescaled because this algorithm takes a very long time to run.

Determining the number of clusters in segmentation depends on the application and method. This requires previous knowledge in the type of images that are being analyzed. Factors as the number of different objects in an image, the distribution of the data, shape and scale are things that can affect the "optimal" number of clusters. In this cases, experiments with different number of clusters can be done and from there, the k with better results can be chosen.

The evaluation discussed in this work is merely qualitative, as we could not implement a good working method to objectively evaluate the results in the given time. Nevertheless, possible metrics to evaluate this kind of problems are discussed later.

3. Results

Figures 3 through 10 show the obtained segmentation for the different clustering methods in all color spaces for images 1 and 2. It is relevant to try different color spaces than RGB because GMM and K-means uses euclidean distance to cluster the image, so color spaces as Lab that is perceptual uniform and distance between colors is euclidian or HSV wich gives color, saturation and brigtness information separately and also we can add spatial information to the representation space adding two features more, one with the x coordinate of the pixel, and another one with y coordinate ending with a 5 dimensional representation of the image.

As shown in the images, the channels LAB and HSV show the best results in general. Also, another determining factor is the computation time that these methods represent. Hierarchical clustering takes the longest by far, as this algorithm is an agglomerative algorithm, starting with as many clusters as data. Then, K-means and GMM take approximately the same time as they are quite similar and, lastly, watersheds takes the shortest time, as it is not a random and iterative algorithm as the others. For all this, we can say that GMM and K-means are the best options, as watersheds produce big errors with this method.

Figure 1: Original image

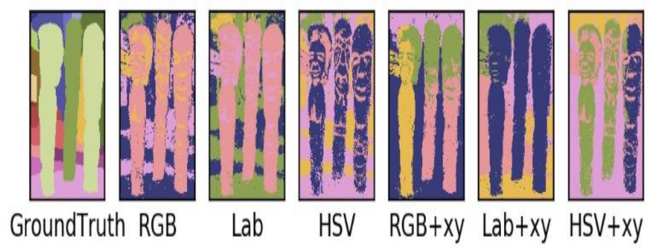

Figure 2: Results with GMM clustering

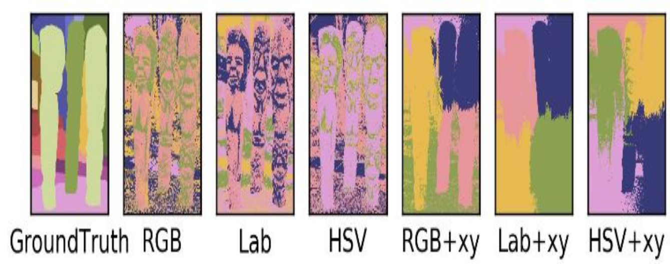

Figure 3: Results with K-means clustering

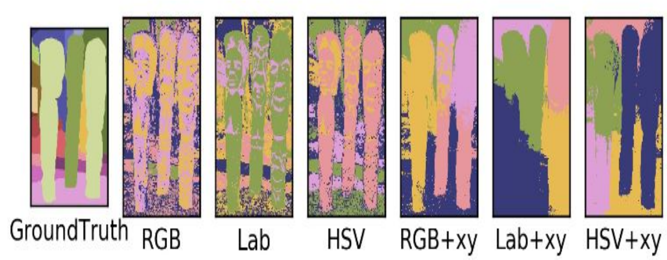

Figure 4: Results with Hierarchical clustering

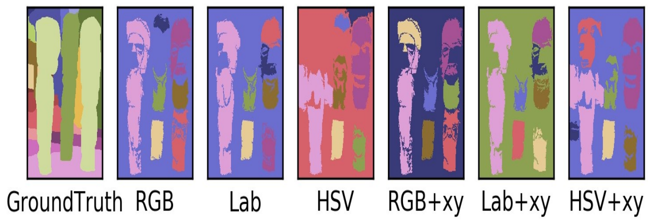

Figure 5: Results with Watershed clustering

As seen in all the methods when spatial information is added to the image representation (x,y coordinates of each pixel), the clustering method ends dividing the image in equal parts that have similar coordinates, this can be noticed in figure 6 where the image was divided by K-means algorithm in four equal parts.

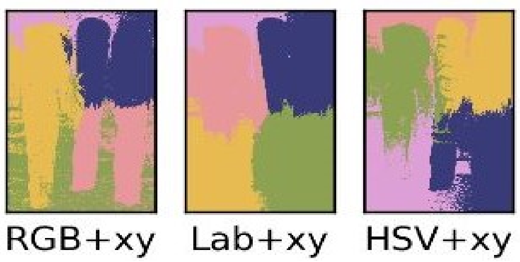

Figure 6: Problem with spatial information

The limitations of this function include the fact that a number of clusters needs to be entered as parameter, which implies that some good previous knowledge of the images is necessary. Furthermore, this segmentation method is a color based segmentation, meaning that segmentation is done with color as its only feature. This means that, as color is its only feature, errors in segmentation occur because different objects with similar colors can be separated into the same group when clustering is done. For this reason, to improve the method, other characteristics as shape could be features to include in the data to obtain a better performance in the segmentation.

Lastly, a quantitative evaluation method can be implemented by making several comparisons with the different ground truths of every image. This comparisons can include size, shape, error and other features to obtain a more liable metric. [1] [2] [3]

4. Conclusions

Adding spatial information with coordinates of the image is not a good idea because the clustering method will divide the image in equal parts.

In the function built in this work, the optimal methods were K-means and GMM, as they showed better results with an acceptable computation time. Two limitations of the method used are that a number of clusters is required meaning that previous information of the image is needed, and that it is a color based segmentation. With this, we observed that LAB and HSV are the color spaces that showed the best results.

References

[1] Crevier, D. (2008). Image segmentation algorithm development using ground truth image data sets. Computer Vision And Image Understanding, 112(2), 143-159. http://dx.doi.org/10.1016/j.cviu.2008.02.002

[2] Polak, M., Zhang, H., Pi, M. (2009). An evaluation metric for image segmentation of multiple objects. Image And Vision Computing, 27(8), 1223-1227. http://dx.doi.org/10.1016/j.imavis.2008.09.008

[3] Peng, B., Zhang, L. (2012). Evaluation of Image Segmentation Quality by Adaptive Ground Truth Composition. Computer Vision – ECCV 2012, 287-300. http://dx.doi.org/10.1007/978-3-642-33712-3$_2$1