

Segmentación de imágenes por métodos de clustering en la base de datos BSDS500

Rubén Darío Bohórquez Cortázar
Universidad de los Andes
Bogota, Colombia
rd.bohorquez10@uniandes.edu.co

Javier Felipe Cifuentes
Universidad de los Andes
Bogota, Colombia
jf.cifuentes10@uniandes.edu.co

Abstract

En este laboratorio se evaluaron los métodos de segmentación por clustering correspondientes a k means y Gaussian Mixture Model (GMM) en la base de datos BSDS500. Adicionalmente, se realizó un estudio detallado atinente a la relevancia del hiperparámetro k referente al número de clusters en los resultados de la segmentación para ambos métodos. Para esto, se corrió el algoritmo de k means en el espacio de representación Lab, utilizando los siguientes valores de k [3 6 9 12 15 18]. De igual forma, se corrió el algoritmo de Gaussian Mixture Model en el espacio de representación rgb, con los mismos valores de k especificados anteriormente. Para evaluar el desempeño de estos métodos, se realizó una curva de precisión y cobertura y se calculó el F máximo, esto se realizó con ayuda de un script en Matlab proporcionado por el profesor. Se obtuvieron los siguientes resultados, para todos los valores de k evaluados, el método GMM presentó mejores resultados que el método de k means. Adicionalmente, se determinó que al aumentar el número de clusters para ambos métodos, el valor F máximo en las curvas de precisión y cobertura decrece, además, el procesamiento para un k mayor es más costoso computacionalmente, lo que implica un mayor tiempo de procesamiento. Finalmente, se concluyó que con los métodos de segmentación por clustering analizados no se pueden obtener mejores resultados que al implementar el método de mapas de contornos ultramétricos (UCM), debido a que segmentar con base al color no es criterio suficiente para lograr una separación óptima de los objetos o instancias presentes en una imagen.

1. Introducción

La segmentación de una imagen se define como el proceso de dividir una imagen en regiones diferentes de modo que cada región sea homogénea, es decir, dividir las imágenes en objetos o instancias que posean características

similares [1][2]. Para el estudio de la segmentación se han desarrollado diversos métodos basados en parámetros tales como el color, la textura, forma, luminancia, entre otros [1]. Kmeans y Gaussian Mixture Model son dos algoritmos de clustering, que han sido implementados en gran medida para el estudio del problema de segmentación en diversos contextos. Debido a esto, el objetivo principal de este laboratorio es evaluar el desempeño de estos algoritmos para la realización de una segmentación basada en color, sobre el consolidado de imágenes presentadas en la base de datos BSDS500. Adicionalmente, se busca comparar los resultados obtenidos a través de la segmentación con estos métodos con los resultados obtenidos en la implementación del método de mapas de contornos ultramétricos (UCM), con el fin de identificar las limitaciones de los algoritmos de segmentación basados únicamente en el parámetro color.

2. Materiales y métodos

2.1. BSDS500

La base de datos BSDS500 está compuesta por un total de 500 imágenes naturales, un conjunto de anotaciones de segmentación realizadas por humanos para cada imagen y el código de benchmarking. Los datos están separados de forma explícita en tres subconjuntos correspondientes a entrenamiento, validación y prueba [3]. En la figura 1 se pueden observar algunos ejemplos de las imágenes presentes en la base de datos.

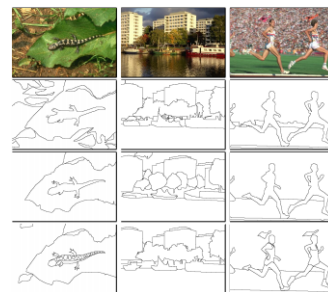


Figure 1. Imágenes naturales y anotaciones BSDS500

Para el desarrollo del laboratorio se realizó el proceso de segmentación sobre cada una de las imágenes de la base de datos. Para este fin se utilizó el código generado en el laboratorio anterior. Los métodos de segmentación utilizados fueron los dos mejores del laboratorio pasado, siendo el primero k means en el espacio de color Lab y el segundo mezcla de gaussianas en RGB. Fue necesario realizar una mejora sobre el método de gaussianas, debido a que este método no converge en todas las imágenes, sobre todo para un número de clúster grande. La mejora realizada fue alterar la propiedad *SharedCovariance*, a verdadero. Esta mejora permitió obtener una segmentación sobre cada imagen de la base de datos.

Para el método K means, se procede a hacer un proceso de clustering utilizando la función *kmeans* de Matlab. Esta recibe como parámetros, los datos, en forma de matriz bidimensional y el número de clúster, siendo estos los centroides de los datos, a los cuales se da la mínima distancia entre cada punto y su centroide. Estos se escogen arbitrariamente. Esta devuelve un mapa y un vector de centroides. Posteriormente se generó un *reshape* sobre el mapa de manera que se muestre como imagen. Esta es la imagen resultante.

Para el proceso de mezcla de gaussianas, se sigue un procedimiento similar al de k means con los mismos parámetros. Se utiliza la función *fitgmdist* para ajustar funciones gaussianas a los datos. A partir de este resultado, se utiliza la función *cluster*. Para asignar cada uno de los puntos a una función y así, darles un valor. Posteriormente a este nuevo vector se le altera la forma con un *reshape*, para obtener la imagen segmentada.

Se corrió la segmentación a un número de clúster de 3, 6, 9, 12, 15 y 18, para los dos métodos. Cada imagen tiene como resultado un cell de segmentaciones en el que cada valor del cell, representa una segmentación a un determinado número de clúster. Luego se dio la evaluación de los métodos de segmentación, utilizando la función *testbenchsfast* de la base de datos. Esta da como resultado valores numéricos, siendo los más relevantes el F máximo y el área bajo la curva. Con la ayuda de la función *ploteval* se muestra la gráfica de precisión-cobertura obtenida. Se realiza este procedimiento para cada método, tanto en training como validación. Adicionalmente se corre la evaluación sobre segmentaciones a un único número de clúster, para obtener el F máximo de cada número de clúster. Finalmente se corre la evaluación sobre los datos, con los dos métodos, a los mejores tres clústers. Se comparan estos resultados con el método UCM.

3. Resultados

Al realizar las evaluaciones de los conjuntos de datos se obtuvieron diversos resultados. Dentro de estos se resaltan los valores de área bajo la curva y el máxima medida de F.

Adicionalmente se obtiene una curva de precisión cobertura para la segmentación. Se muestra una de las gráficas resultantes en la figura 2. El total de los valores calculados se encuentran en la tabla 1.

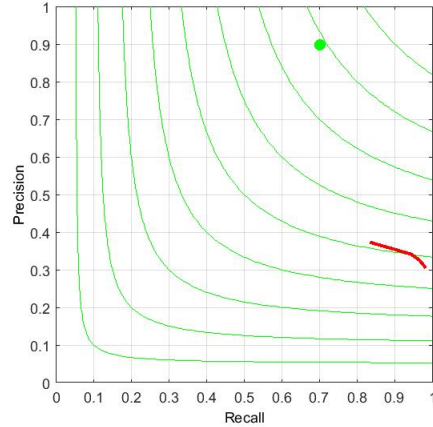


Figure 2. Evaluación del entrenamiento para el metodo K means.

Metodo	F máximo	Area bajo la curva
Entrenamiento		
K means	0.47	0.05
Gmm	0.52	0.09
Validación		
K means	0.44	0.05
Gmm	0.49	0.1

Table 1. Resultados numericos obtenidos para el training y la validacion

Posteriormente se obtuvieron de los valores de cada número de clusters por separado. Los resultados se muestran en la tabla 2. se obtuvieron gráficas como resultado, se muestra un ejemplo en la figura 3.

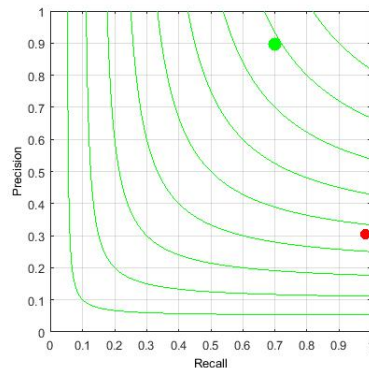


Figure 3. Resultado de correr evaluacion sobre un unico numero de clusters.

Numero de Clusters	Entrenamiento		Validación	
	K means	Gmm	K means	Gmm
3	0.40	0.44	0.39	0.42
6	0.37	0.42	0.35	0.4
9	0.33	0.40	0.31	0.38
12	0.29	0.37	0.27	0.35
15	0.27	0.36	0.26	0.33
18	0.25	0.34	0.23	0.32

Table 2. Resultados por cada numero de clusters

Finalmente se corrió la evaluación sobre los datos de test para los tres mejores clusters, para cada uno de los métodos. Los clusters escogidos fueron 3, 6 y 9. Finalmente se compara con el método UCM. La gráfica se muestra en la figura 4, mientras que los datos se muestran en la tabla 3.

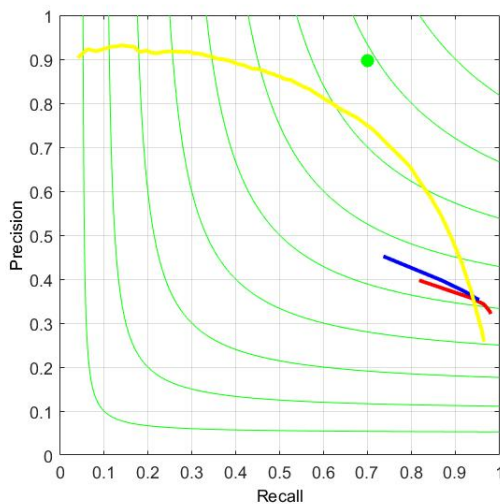


Figure 4. Gráfica que muestra los resultados para los tres métodos. La línea roja representa el método K means, la línea azul la mezcla de gaussianas y la línea amarilla el método UCM

Metodo	F máximo	Área bajo la curva
K means	0.46	0.06
Gmm	0.49	0.09
UCM	0.74	0.73

Table 3. Resultados numéricos obtenidos para el test

4. Análisis de resultados

Al evaluar los métodos correspondientes a k means y GMM en la base de datos BSDS500 se observó que para todos los k seleccionados, el método de GMM presentaba mejores resultados, es decir, este método expuso valores de F medida máxima mayores a los del método de k means. Se

debe resaltar que el método GMM fue mejor tanto en entrenamiento y validación como en test. Esto se puede observar en la tabla 2.

Ahora bien, estos resultados son iguales a los obtenidos en el laboratorio anterior y se deben a que el algoritmo de GMM es más flexible y adaptable a la distribución de los datos que entran como parámetro, lo que se ve reflejado en una mejor segmentación, mientras que el algoritmo de k means es menos flexible y tiende a agrupar los datos en clusters circulares y de tamaño similar, lo cual perjudica los resultados de la segmentación. Adicionalmente, se debe resaltar que una desventaja que presentan ambos métodos es la dependencia de la inicialización y el conocimiento previo del número de clusters. Por otro lado, se debe resaltar que aunque GMM presente mejores resultados que kmeans en términos de la calidad de la segmentación, el tiempo de computo que requiere correr el método de kmeans en la base de datos, es mucho menor que el requerido para el método de GMM. Aunque los resultados obtenidos tanto en este laboratorio como en el pasado, presentan a GMM como un método de segmentación superior a K means, se debe resaltar que en este laboratorio la diferencia en la calidad de los resultados expuestos por cada método, es más marcada, y eso se debe a que la metodología de evaluación implementada en este laboratorio es más adecuada para el análisis de los algoritmos de segmentación, y presenta menos limitaciones que la desarrollada en el laboratorio anterior.

Con base a los resultados obtenidos, se puede afirmar que seleccionar k puede mejorar o empeorar los resultados atinados a las segmentaciones obtenidas para cada método. Se determinó que con valores pequeños de k, el valor de la F medida máxima es más grande, mientras que al usar valores de k grandes, el valor de la F medida máxima es más pequeño. En otras palabras, se obtuvieron mejores resultados con los k más pequeños (mejor resultado con k=3). Ahora bien, se debe resaltar que aunque al utilizar valores más grandes de k los valores de F medida máxima eran menores, la cobertura obtenida era mayor que al utilizar valores de k más pequeños (máxima cobertura k=18).

5. Discusión

Las curvas de precisión y cobertura exponen la relación entre los parámetros precisión y cobertura para cada método de segmentación con cada valor de k determinado. En otras palabras, para obtener la curva total se deben correr los métodos con múltiples valores de k. La precisión y la cobertura se definen como: $P = \frac{TP}{TP+FP}$ $R = \frac{TP}{TP+FN}$ TP: verdaderos positivos, FP: falsos positivos, FN: falsos negativos, P=precisión, R=Recall.

La curva indica que tan buenos son los resultados obtenidos con los diferentes métodos de segmentación, en comparación con las anotaciones de segmentación utilizadas. Para evaluar los distintos métodos, se debe calcular el área

bajo la curva o reportar la F medida máxima. La F medida máxima se define como: $FMM = \frac{2PR}{P+R}$. Los resultados ideales en una curva de precisión y cobertura, es la obtención de un valor de F medida máxima igual a 1. El cual indicaría una precisión del cien por ciento y una cobertura del 100 por ciento, es decir, esto implicaría que las segmentaciones obtenidas con el método analizado son idénticas a las anotaciones de segmentación utilizadas.

Al comparar los resultados de los métodos, no fue posible superar los resultados obtenidos por el método UCM, desarrollado por el profesor Pablo Arbeláez. Esto se puede mostrar en los datos obtenidos para la F medida máxima. No se toma en cuenta la medida de área bajo la curva, esto se debe a que los resultados obtenidos por el método UCM2 toman en cuenta un gran número de K, mientras que los métodos desarrollados en este laboratorio solo toman en cuenta 6, generando que las curvas no abarquen la totalidad de los valores que debieran, haciendo que el área bajo la curva siempre sea menor. Tomando en cuenta los valores de F máximo, se puede denotar que el método de UCM presenta mejores resultados. Se atribuye este resultado al hecho de que el método UCM, utiliza más información de la imagen que los métodos desarrollados, información como los bordes obtenidos al realizar mapa de textones, watersheds de gradientes orientados, entre otros. El método desarrollado solo toma información respecto al color en la imagen.

Adicionalmente, se debe resaltar que las principales limitaciones que presentan los métodos de segmentación por clustering estudiados, k means y GMM, son las atinentes a la gran dependencia a la inicialización y a la selección previa del número de clusters. Además, los resultados obtenidos con estos métodos no son los mejores debido a que estos utilizan solo la información atinente al color en cada imagen. Información que como se observó en la comparación con el método UCM, es insuficiente para lograr una separación óptima de los objetos o instancias presentes en una imagen.

Se considera que la manera más adecuada de mejorar los algoritmos de segmentación desarrollados, es a través de un espacio de representación más preciso, es decir que tenga la mayor información posible acerca de cada punto o cada ventana. Puede ser de utilidad incluir información acerca de la textura, ya que es parte importante de lo que representa la diferencia entre objetos. No se considera el uso de espacio o de coordenadas espaciales, ya que como se vio en el laboratorio anterior, no mejoraban el resultado, solo lo empeora. Debe incluirse la información de bordes en la imagen, para la utilización de los mismos como delimitantes de las segmentaciones. Finalmente se considera que el uso de SIFT puede ser de gran utilidad en la segmentación de objetos, ya

que gran parte de los objetos tienen formas características, que pueden ser reconocidas con alguna facilidad. La forma óptima de unir esto, es a través de un histograma conjunto con toda la información recolectada. Para la parte de clustering, se pueden usar métodos de clustering como métodos de k means o mezcla de gaussianas.

6. Conclusiones

- El método de segmentación por clustering GMM presenta mejores resultados en la base de datos BSDS500 que el método k means.
- La segmentación basada únicamente en la información de color es insuficiente para lograr una separación óptima de los objetos o instancias presentes en una imagen. Lo que explica porque el método UCM expone resultados mucho mejores que los obtenidos con los métodos de clustering estudiados.
- La selección del parámetro k es de suma importancia para la obtención de mejores resultados en los algoritmos de segmentación por clustering k means y GMM. Con los valores más altos de k, se obtiene una mayor cobertura, no obstante, los valores de la F medida máxima son menores que los obtenidos con valores más bajos de k. Esto explica porque los mejores resultados se obtuvieron al correr el método GMM con valores más bajos de k, específicamente, el mejor resultado se obtuvo con k=3.

References

- [1] Palomino, N. L. S., Concha, U. N. R. (2009). Técnicas de segmentación en Procesamiento digital de imágenes. Revista de investigación de Sistemas e Informática, 6(2), 9-16.
- [2] Wu, Y., Yang, X., Chan, K. L. (2003, December). Un-supervised color image segmentation based on Gaussian mixture model. In Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on (Vol. 1, pp. 541-544). IEEE.
- [3] Contour Detection and Image Segmentation Resources. (n.d.). Retrieved March 30, 2017, from <https://www2.eecs.berkeley.edu/Research>

7. Anexos



Figure 5. Ejemplo: imagen de la base de datos BSDS500, entrenamiento.

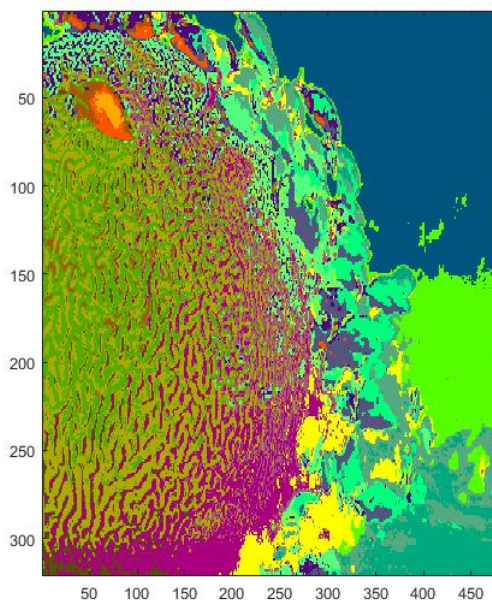


Figure 7. Ejemplo: segmentación imagen base con mezcla de gaussianas, con 18 cluster

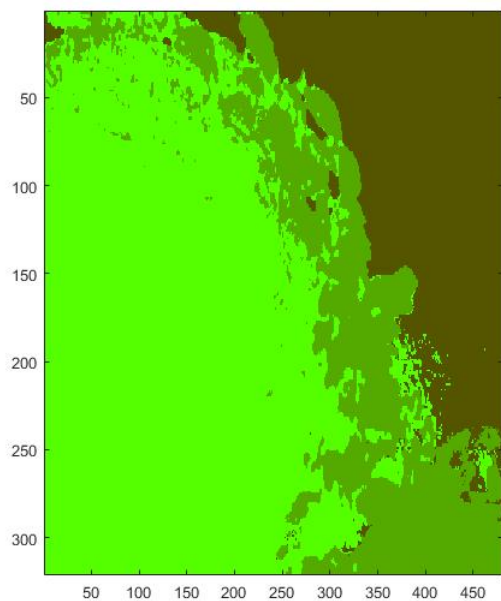


Figure 6. Ejemplo: segmentación imagen base con mezcla de gaussianas, con 3 cluster

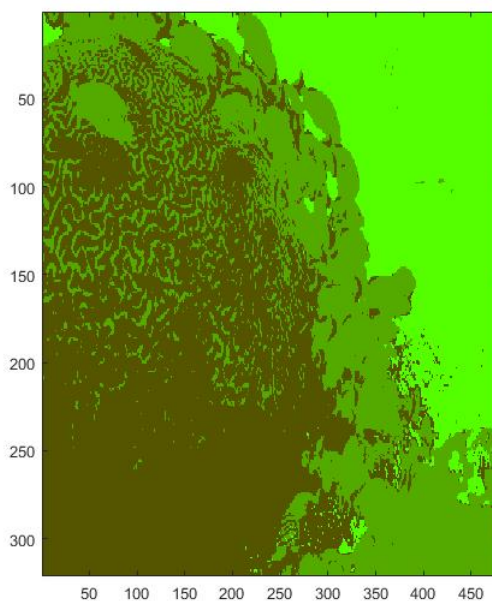


Figure 8. Ejemplo: segmentación imagen base con k means, con 3 cluster

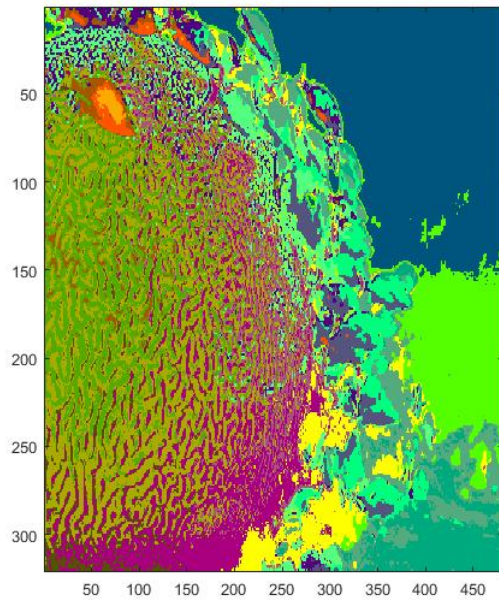


Figure 9. Ejemplo: segmentación imagen base con k means, con 18 cluster

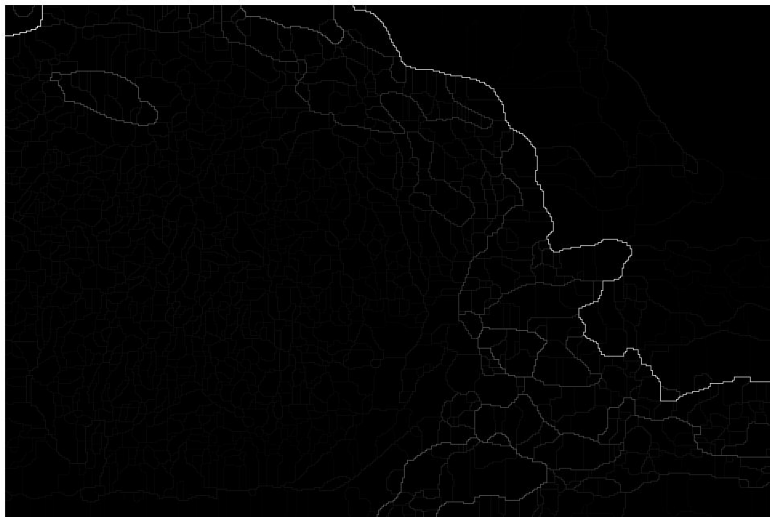


Figure 10. Ejemplo: segmentación imagen base con UCM