

Berkeley Segmentation Dataset and Benchmark (BSDS500)

José María Campo Viñas
Universidad de los Andes
Código: 201412002

jm.campo11@uniandes.edu.co

Mariana Franky
Universidad de los Andes
Código: 201313944

m.franky10@uniandes.edu.co

Abstract

El objetivo de este laboratorio es evaluar los dos mejores métodos de segmentación desarrollados en el laboratorio anterior, para esto se utilizó la base de datos de Berkely BSDS500 y el Benchmark, el cual se compone de 500 imágenes naturales, anotaciones humana y código de benchmarking. Se comparó estos dos métodos K-means y Watershed con el método de Pablo.

1.. Introducción

La segmentación de una imagen puede ser definida como la partición de una imagen en regiones constituyentes no solapadas, las cuales son homogéneas con respecto a alguna característica como intensidad o textura.[1]. La segmentación puede ser esencial en el análisis de imágenes, representación de objetos, visualización y muchas otras tareas de procesamiento de imágenes. Del mismo modo, es esencial tener un método de evaluación que permita la comparación con métodos anteriores. En este laboratorio se evaluaron dos métodos de segmentación, previamente desarrollados, en la base de datos BSDS500 de Berkley, estos métodos se evaluaron con el BSDS Benchmark desarrollado por el grupo de visión por computador de Berkley y se compararon con el método realizado por Pablo Arbeláez. Los métodos de segmentación k-means y Watershed fueron empleados en este laboratorio:

- *K-means* : el algoritmo de K-means se basa en el análisis de varianza, el objetivo es minimizar las distancias cuadradas entre elementos y centroides. Como resultado se obtiene el agrupamiento del conjunto de datos en un número predefinido de clases.

Las características que se debe tener en cuenta de este método son : Se asume que el número de clusters K se conoce, cada cluster es representado por su centro de masa. El algoritmo comienza con un conjunto aleatorio de k centroides, a cada pixel se le asigna un centroide, se calcula la distancia entre centroide y pixel, logrando

que cada pixel corresponda al centroide cuya distancia sea la mínima Se detiene hasta que converge al criterio de parada.[2]

- *Watersheds*: La analogía que hace watershed es que una imagen en escala de gris se puede interpretar como la imagen topográfica de un relieve terrestre; se puede decir que las intensidades de gris de mayor amplitud pueden corresponder a llanuras mientras que las que presentan menores intensidades pueden corresponder a valles y ríos. Se obtienen las líneas divisoras de agua mediante la inundación de cada uno de los valles (mínimo regional de la superficie topográfica), las aguas procedentes de la inundación de los valles forman lagos. Cuando dos lagos se encuentran, se construyen una represa evitando que las aguas de los dos lagos se mezclen. Cuando el conjunto está completamente sumergido, el conjunto de represas son las líneas divisoras.

Se compararon estos dos métodos con el método desarrollado por Pablo Arbeláez :

- *Ultrametric Contour Map (Mapa de contorno ultramétrico)*: Propone un algoritmo que construye una jerarquía de regiones a partir de cualquier detector de contorno. Este consta de dos pasos, inicialmente utiliza Oriented Watershed Transform (OWT) con el fin de formar regiones iniciales a partir de contornos, el segundo paso consta en la construcción de un mapa de contorno ultramétrico UCM que define una segmentación jerárquica. [3]

2.. Métodos y materiales

2.1.. Materiales

La base de datos que se utilizó en este laboratorio, fue la de BSDS500. Este conjunto de datos fue desarrollado por la universidad de Berkely , es una extensión de la base de datos de BSDS300, donde las 300 imágenes originales se

usan para entrenamiento y validación y 200 imágenes nuevas , junto con sus anotaciones humanas, son utilizadas para el test. Cada imagen fue segmentada por cinco temas diferentes en promedio. Además, esta base de datos cuenta con un BSDS benchmark code.

2.2.. Métodos

Los métodos que se eligieron para correr la base de datos BSDS500 fueron K-means y Watershed, se escogieron estos dos métodos debido a que fueron los que mostraron un mejor resultado en el anterior laboratorio, resultado en cuanto al índice de Jacard y tiempo de procesamiento. Así mismo se escogió como espacio de color HSV +xy. Basados en los resultados obtenidos anteriormente no se realizó ninguna modificación de los dos métodos propuestos. Sin embargo, se realizó modificaciones en la función anterior para que funcionara en el BSDS Benchmark, la modificaciones se basaron en la forma de leer el conjunto completo de imágenes y ajustar los resultados de segmentación al mismo formato de benchfast.

Cabe destacar que se utilizó la función:

```
segmentation = segmentByClustering(rgbImage,featureSpace,...,clusteringMethod,numberOfClusters)
```

para la ejecución de esta laboratorio, esta función se desarrolló en el anterior laboratorio. Como featureSpace se utilizó hsv+xy , clusteringMethod: se utilizó K-means y Watershed.

K-means: Es un método de segmentación por agrupación, se utilizó la función de matlab k-means, esta recibe como parámetro de entrada recibe una matrix con las características de la imagen y el número de clusters.

Watershed: En éste método, inicialmente se obtienen los bordes más pronunciados de la imagen, empleando el algoritmo de Canny, para luego hacer conteo de componentes conexos de dicha imagen binaria resultante. El propósito de los componentes conexos es usar aquellos de mayor tamaño para imposición de mínimos en la imagen que será introducida a la función de Watershed, la ventaja de realizar esto, consiste en regular la cantidad de clusters que se deseen obtener para la imagen, sin tener en cuenta información previa.

Hiperparámetro K

- **En entrenamiento y validación** Con respecto al anterior laboratorio, en este se debe calcular la segmentación de la imágenes con diferentes K , este hiperparametro hace referencia al número de regiones en los que va estar segmentada la imagen.En la etapa de entrenamiento y validación se eligió un vector de 8 K, la elección de este hiper parámetro fue aleatoria, sin embargo se tuvo

en cuenta que la distancia entre los K no fuera tan cercana que hubiera espacio entre ellos. Después de realizar las segmentaciones , los resultados se guardaron en un arreglo de celdas, donde cada elemento es una segmentación de la imagen original calculada con un K diferente. Se evaluó las distintas segmentaciones de la base de datos con el benchmark y en los resultados se encontró la F-measure para cada K.

- **Test:** Los K que fueron utilizados para esta etapa, se escogieron basándose en la F-measure de cada K. Es decir, para cada método se escogió los 4 mejores K con F-measure. Teniendo los mejores K se realizó la segmentación en la base de datos de test.

2.3.. Metodología de evaluación:

La metodología de evaluación con la que se evaluó los métodos de segmentación propuestos es con la curva de precisión y cobertura, esta curva considera dos aspectos importantes. La precisión mide la fracción de verdaderos positivos en las regiones producidas por la segmentación. Cobertura mide la proporción de pixeles en los contornos en el ground truth que fueron detectados con éxito por la segmentación se obtiene una curva parametrizada. Además, proporciona una F-measure , que es definida como la media armónica de precisión y cobertura, esta medida proporciona una puntuación sobre el algoritmo. El resultado ideal sería que nuestro resultado de segmentación tuviera alta precisión y una alta cobertura, obteniendo un resultado similar al de los humanos, que en este caso la F-measure es 0.79.

3.. Resultados

En la Figura 1 se observan las curvas obtenidas por los métodos empleados: k-means (azul), watershed (amarillo) y el algoritmo de probabilidad de bordes hecho por Pablo Arbeláez [3]. En la Figura 2 se observan las curvas obtenidas por los métodos mencionados anteriormente y empleando los números de clusters cuyo umbral de cobertura era óptimo.

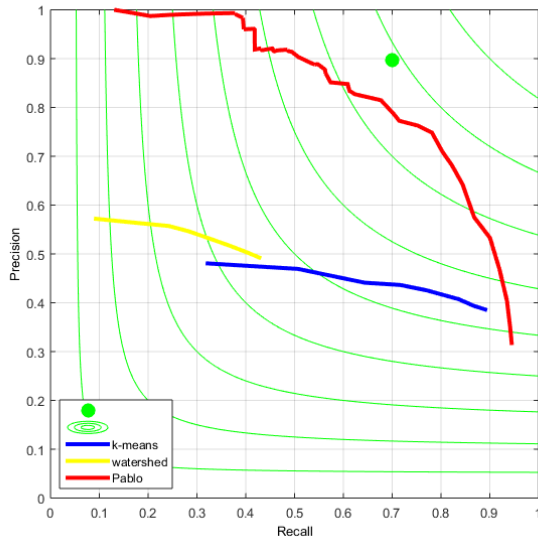


Figura 1. Precisión-cobertura para imágenes de entrenamiento.

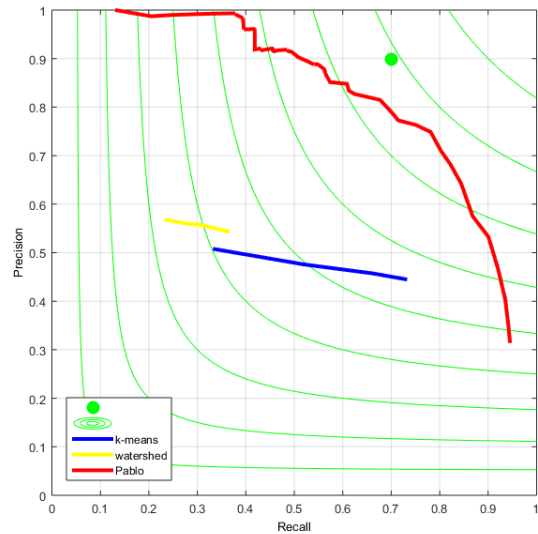


Figura 2. Precisión-cobertura para imágenes de prueba.

Como se observa en el Cuadro 1, en el método de k-means, los clusters escogidos como óptimos dado su puntaje fueron 2, 4, 6 y 8. Mientras que en el Cuadro 2, en el método de watershed, los clusters escogidos fueron 6, 8, 10 y 15.

Cuadro 1. Valores de cobertura para número de clusters en método de k-means.

Número de clusters	Umbral de cobertura
2	0.388
4	0.386
6	0.376
8	0.357
10	0.333
15	0.293
20	0.258
25	0.233

Cuadro 2. Valores de cobertura para número de clusters en método de watershed.

Número de clusters	Umbral de cobertura
2	0.384
4	0.423
6	0.436
8	0.434
10	0.434
15	0.430
20	0.426
25	0.423

A continuación, en el Cuadro 3, se observan los puntajes globales para cada uno de los métodos empleados, usando la métrica de la F-medida en bordes y regiones de segmentación, respecto al punto óptimo de la anotación en humanos.

Cuadro 3. Comparación entre métodos.

Tipo de clusterizado	F-medida bordes global	F-medida regiones
K-means	0.55	0.57
Watershed	0.46	0.52
Pablo (UCM)	0.76	0.84

4.. Discusión y Conclusiones

El método más eficiente es K-means, ya que este supera en todas las medidas al método de watershed como se observa en el Cuadro 3. Este funciona mejor ya que tiene en cuenta saturación e información espacial, mientras que watershed realiza un promedio de todas las variables de entrada al funcionar únicamente en un espacio 2D.

Asumiendo que existe un rango de K óptimo general, escoger un valor por debajo de éste conllevaría a un underfitting en la imagen, juntando regiones separadas por un borde suave, mientras que al aumentar K ocurre un overfitting, donde cada detalle mínimo de cambio de iluminación u otra variación en un mismo objeto, lo dividiría en dos segmentaciones distintas.

No existe un valor K óptimo en la base de datos y si se desea

establecer un rango óptimo, éste no sería el mismo para todas las imágenes debido a la complejidad y alta variabilidad de estas. Es decir, lo que puede ser un valor óptimo para una imagen, puede producir un underfitting o un overfitting en otra imagen. Esto puede ejemplificarse en la Figura 3, donde múltiples detalles de tonos, iluminación y bordes se encuentran en el avión, asumiendo un K ideal igual a dos donde se espera obtener la silueta del avión, sin embargo al aumentarlo un K , el cual es óptimo para otra imagen, podrían surgir detalles irrelevantes dentro de éste, lo que resultaría difuminando la silueta del avión.

Era de esperarse que los resultados en los dos métodos propuestos no superaran al de Pablo, dado que el método de Pablo propone un algoritmo que construye una jerarquía de regiones a partir de cualquier detector de contorno. Une la información de un detector de contornos gPb con un algoritmo de agrupamiento. El algoritmo de agrupamiento tiene en cuenta dos partes: una nueva OWT, que produce un conjunto de regiones iniciales a partir de la salida del detector de contorno, y la segunda parte forma un mapa de contorno ultra métrico. Éste define un gráfico inicial, donde los nodos son las regiones en Po , los enlaces unen regiones adyacentes y son ponderados por una medida de similitud entre regiones. Organiza por similitud y combina iterativamente las regiones más similares. Produciendo un árbol de regiones, donde las hojas son los elementos de Po , la raíz es el dominio de la imagen y las regiones son ordenadas por la relación de inclusión [3]. Mientras que, k -means intenta optimizar un objetivo global (varianza de los clusters) y logra un local óptimo.

Como se había comentado en el laboratorio pasado, el método de watershed posee una alta precisión en imágenes con bordes definidos y uniformidad dentro de las regiones de segmentación, dichas cualidades le permiten obtener una alta precisión, sin embargo, se pierde cobertura al no poder definir los bordes suaves. Ahora bien, comparando los resultados del método propuesto de evaluación en el anterior laboratorio con el de este laboratorio. Los resultados presentan una correspondencia con los vistos en el laboratorio anterior, siendo el mejor método de segmentación K -means. Sin embargo, los resultados no son tan altos como el anterior laboratorio. Esto puede deberse al número de imágenes evaluadas en este laboratorio las cuales fueron 200 imágenes, mientras que en el anterior se evaluó nuestros métodos con sólo una imagen. Esto genera, que los resultados obtenidos anteriormente no tengan validez puesto que no se tienen un resultado global del algoritmo.

Una de las limitaciones que pueden tener nuestros métodos en comparación al de Pablo es que ninguno de los dos métodos creados tiene en cuenta la textura en

la imagen, por ende, se pierde gran cantidad de información.

Como se observa en la Figura 3, un patrón que siguen las segmentaciones de K -means es que las regiones uniformes grandes se tienden a tratar como distintos clusters al usar información espacial, mientras que en una región cercana, a una misma saturación, pero distintos tonos, se tienden a agrupar, debido que la información fundamental no se le da el peso suficiente. Por otra parte, los patrones en la segmentación de watershed se definen en el contraste de los bordes, se observa el seguimiento fiel a los bordes en objetos como el carro amarillo a la izquierda de la imagen, o la parte superior del avión, mientras que en bordes poco definidos como la parte inferior del avión se tiende a mezclar éste con la pista de carretero.

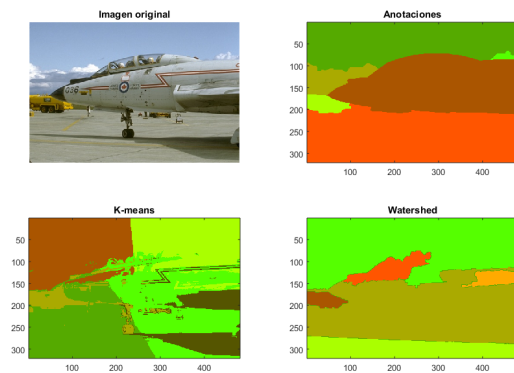


Figura 3. Imagen segmentada de muestra. En orden: Imagen original, anotaciones, segmentación por K -means y segmentación por Watershed.

Se podría mezclar las ventajas de ambos métodos empleados, usando la agrupación de colores en el método de K -means junto a la invarianza de watershed en regiones uniformes de área extensa en la imagen. A esto se le puede añadir una respuesta de la imagen a un diccionario de textones, de manera que la segmentación cubra variables de color, forma, textura y la no división de regiones uniformes extensas.

También se podría generar un algoritmo que siga la idea de watershed, pero pueda trabajar en un espacio N -dimensional.

Referencias

- [1] M. González and V. Ballarín, "Segmentación de imágenes utilizando la transformada watershed: obtención de marcadores mediante lógica difusa," *IEEE Latin America Transactions*, vol. 6, no. 2, pp. 223–228, 2008.

- [2] P. Arbeláez, “Grouping,” https://sicuaplus.uniandes.edu.co/bbcswebdav/pid-1777468-dt-content-rid-18350020_1/courses/201710_IBIO4680_01/06_representation_01.pdf, 2016.
- [3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “From contours to regions: An empirical evaluation,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 2294–2301, IEEE, 2009.