

Clasificación

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

La segmentación se define como la partición de una imagen en regiones, en este laboratorio se creó una función de segmentación, se realizó la segmentación de dos imágenes en diferentes espacios de color y por diferentes métodos de segmentación. Se evaluó la función con el índice de Jaccard, dando como resultado que HSV como mejor espacio de color para la segmentación y K-means como mejor método de segmentación.

1.. Introducción

La segmentación de imágenes se define como la partición de una imagen en regiones constituyentes no solapadas, las cuales son homogéneas con respecto a alguna característica como una intensidad o textura. [1]

Se utilizó 4 métodos diferentes para realizar la segmentación a continuación se explicarán cada uno de ellos :

- **K-means** : el algoritmo de K-means se basa en el análisis de varianza, como resultado agrupa el conjunto de datos en un número predefinido de clases.
Se asume que el número de clusters K se conoce, cada cluster es representado por su centro de masa, El objetivo es minimizar las distancias cuadradas entre elementos y centroides. El algoritmo comienza con un conjunto aleatorio de k centroides, a cada pixel se le asigna centroides, Cada pixel corresponde al centroide cuya distancia sea la mínima Se detiene hasta que converge al criterio de parada.[2]
- **Mezcla de gaussianas** El objetivo de este método es representar cada grupo como distribuciones gaussianas o distribuciones paramétricas. Para esto se calcula la probabilidad de pertenencia a cada cluster, y se alcanza normalmente una clasificación asignando cada observación al cluster más probable. Es una generalización de K-means. La distancia de un punto en el cluster está dado por la métrica de Mahalanobis. Su utilidad radica

en que es una forma de determinar la similitud entre dos variables aleatorias multidimensionales. Se diferencia de la distancia euclídea en que tiene en cuenta la correlación entre las variables aleatorias.[3]

- **Segmentación jerárquica**: Este método proporciona un conjunto de segmentación de una imagen a diferentes niveles de detalle. Para esto define la distancia D entre grupos basándose en la medida de similaridad entre elementos, esta jerarquía puede ser presentada por un dendrograma. Inicialmente, cada punto es cluster, luego se agrupan los puntos mas cercanos, hasta obtener un único cluster.[2]
- **Watersheds**: 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. Este método calcula las líneas divisorias de aguas que separan cuencas geográficas, Esto se realiza mediante la inundación de cada uno de los valles (mínimo regional de la superficie topográfica), el agua procedente de la inundación de los valles forman lagos, cuando dos lagos se encuentran se construyen una represa para evitar que el agua de los dos lagos se mezclen. Cuando el conjunto está completamente sumergido, el conjunto de represas son las líneas divisoras.

2.. Métodos y materiales

2.1.. Materiales

La base de datos que se utilizó en este laboratorio, fue un subconjunto pequeño, aleatoriamente seleccionado, de la base de datos de segmentación BSDS500 esta cuenta con las anotaciones incluidas. Se utilizaron 2 imágenes de esta base de datos.

2.2.. Métodos

Método de segmentación El objetivo de este laboratorio era crear un función de segmentación, la sintaxis de la función se muestra a continuación:

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

Esta cuenta con varios hiperparametros:

- *rgbImage*: Inicialmente se pide que la imagen a segmentar este en el espacio de color RGB

Antes de segmentar la imagen por los distintos métodos de clustering fue necesario realizar un preprocesamiento de la imagen esta dependía de el espacio de color que se de como parámetro:

- *featureSpace* Se puede escoger entre 'rgb', 'lab', 'hsv', 'rgb+xy', 'lab+xy' o 'hsv+xy'. En cada espacio de color se le asignó a cada pixel las intensidades correspondientes a cada canal y además, se le asignó su ubicación en cordenadas X y Y. Se realizó una normalización de los datos en el espacio RGB, Hsv y Lab, interpolando el dato mínimo de cada canal a cero y el dato máximo a 1.

Únicamente se reescaló la imagen a la mitad de su tamaño en el método de segmentación jerárquica para poder ser procesada.

Ahora bien, teniendo las características de la imagen por pixel se realizó la segmentación segun el método de clustering elegido.

clusteringMethod : 'k-means', 'gmm', 'hierarchical' o 'watershed'.

K-means: Se utilizó la función de matlab k-means, como parámetro de entrada recibe una matrix con las características de la imagen y el número de clusters.

gmm: Se utilizó la función de matlab fitgmdist, Esta función crea un modelo de mezclas de gaussianas con los datos y el número de clusters. Luego se somete la imagen al modelo y se obtienen los indices por medio de la funcion cluster.

hierarchical: Inicialmente, para este método fue necesario escalar las imágenes para que no generara un error al computar la funcion linkage, sin embargo, se buscó en la literatura de esta función y se encontró que con el parámetro savememory , el cual construye los cluster sin computar la matriz de distancia, se reescalo la imagen y se utilizó este parámetro para mejorar el tiempo de computación de este método. Se utilizó como distancia, la distancia euclidiana la cual es la única aplicable cuando se utiliza el parámetro savememory y como métrica de distancia ward.

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.

El hiperparámetro mas importante fué *numberOfClusters*;, se escogió uno distinto para cada imagen. Para esto se buscó en las anotaciones de cada imagen el menor cluster y este fue utlizado como parámetro, se escogió el menor k con el objetivo de evaluar nuestro algoritmo de una forma simple.

Método de evaluación

En los problemas de segmentación se utiliza como metodología de evaluación el índice de Jaccard, este mide el grado de similitud entre las anotaciones y el resultado de la segmentación. Otra estrategia de segmentación es computar una curva precisión y cobertura en los bordes detectados de los objetos.

Para este laboratorio se utilizó el índice de Jaccard como metodología de evaluación, Debido a que nuestra estrategia para determinar el rendimiento de la segmentación fue comparar con las anatociones y el resultado de la segmentacion, el objeto más grande de la imagen. Se evaluó por espacio de color utilizando como método watershed, y se utilizó la imagen que presentará más color y se evaluó el método de clustering tomando como espacio de color fijo HSV, se escogió la imagen de la iglesia como imagen a segmentar.

Se realizó los siguientes pasos para cada una de las imágenes.

1. Se tomó el objeto de mayor tamaño de la imagen, y se elige la imagen la imagen ground truths que contenga el menor número de clusters.Figura 1 se considera como mayor obtejo de la imagen el hombre con la mascara.
2. En la imagen segmentada se tomó todos los clusters que estan dentro de este objeto de mayor tamaño, Para indentificar las etiquetas asociados a este cluster se realizó manualmente: con ayuda del data cursor se visualizó las etiquetas y ubicación espacial pertenecientes a esos cluster.Por ejemplo, de la la Figura 2 se tomó como cluster asociados a la imagen colores asociados a las etiquetas: 2 morados distintas tonalidades,rosado,aguamariana, gris y azul.

3. Se binarizó la imagen siendo 1 los clusters pertenecientes a el objeto de mayor tamaño y 0 lo que no pertenece a esos clusters.
4. Se realizó el mismo procedimiento con la imagen de ground truth.
5. Se determinó el índice de jaccard de la imagen. Una de las deficiencias de este método es que no está automatizado es decir es necesario de la percepción humana para escoger los clusters que están asociados al objeto de la imagen.

3.. Resultados

A continuación se muestra los resultados para los espacios de colores y el método de evaluación dando como resultado que el método Jerárquico es el que tiene mayor tiempo de procesamiento.

Cuadro 1. Índice de Jaccard para cada espacio de color en segmentación Watershed.

Espacio de color	Jaccard
RGB	0.41
RGB+XY	0.53
HSV	0.82
HSV+XY	0.74
LAB	0.45
LAB+XY	0.50

Cuadro 2. Índice de Jaccard y tiempo de procesamiento para cada metodología de segmentación en el espacio de color HSV.

Método de segmentación	Jaccard	T. de procesamiento
K-means	0.94	2.46
GMM	0.93	8.50
Jerárquico	0.43	67.24
Watershed	0.52	0.47



Figura 1. Imagen a segmentar en prueba de espacios de color.

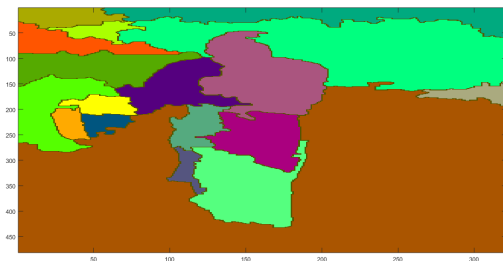


Figura 2. Segmentación de Figura 1 por HSV y Watershed.



Figura 3. Imagen a segmentar en prueba de métodos de segmentación.

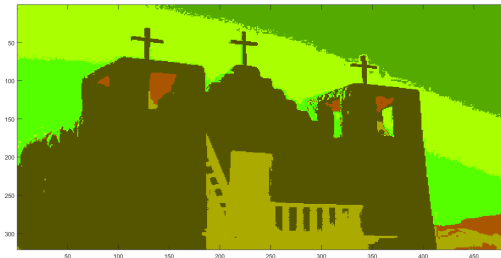


Figura 4. Segmentación de Figura 3 por HSV y K-Means.

4.. Discusión y Conclusiones

Usar otros espacios de color conlleva a un entendimiento más intuitivo de las cualidades de cada píxel. En el caso de HSV, valores como saturación de color dan una idea de qué tan puro es el color que se está manejando en la imagen y la cercanía que existe entre los tonos. Por otra parte, en el caso de Lab, se busca manejar distancias equitativas entre colores, de manera que se puedan realizar operaciones entre estos usando la norma euclideana. Para la Figura 1 se tiene que el mejor espacio de color es HSV, esto se debe a que la cantidad de bordes detectados por el método de Canny es menor que en el espacio de color Lab, de esta manera, se obtiene un mínimo impuesto que abarca toda la región de la camisa del personaje central, en la cual los pliegues de ésta no crean bordes que separen las regiones en pequeños componentes conexos, tal y como ocurre en el caso de Lab. El mejor método de segmentación es relativo a lo que se desea en el problema, al observar la Figura 2 se tiene que Watershed es una buena aproximación para imágenes con bordes altamente denotados al obtener índices de Jaccard de gran magnitud en cualquier espacio de color, como se muestra en el Cuadro 1. Pero en imágenes con

bordes transitorios y suaves, resulta ser una mejor opción el método de K-Means al priorizar información de color antes que componentes conexos en la imagen, como se muestra en la Figura 4. Sin embargo, el método de Watershed, posee un tiempo de procesamiento mucho menor en comparación a los demás métodos como se muestra en el Cuadro 2.

Las principales limitaciones presentes en el método de Watershed, consisten en la implementación del método de Canny como detección de bordes para obtener componentes conexos en la imagen y de esta forma, usar los de mayor tamaño como mínimos locales. Esto se nota al observar la segmentación obtenida en la Figura 2, donde los zapatos de la persona presente no fueron segmentados debido a que el componente conexo existente en dicho lugar no era lo suficientemente grande para ser escogido como mínimo local.

Por otra parte, la no convergencia de los métodos de k-means y GMM conlleva a una demora en procesamiento, la cual es dependiente de la cantidad de clusters que se deseen obtener.

Por último, el método más ineficiente es la segmentación jerárquica, donde toma demasiado tiempo de procesamiento y espacio en memoria para calcular las distancias entre cada uno de los píxeles al momento de formar el dendrograma.

El espacio de color con mayor discriminación para la Figura 1 es Lab, siendo L el canal que mejor discrimina la información de contornos y texturas en la imagen.

Una de las fallas que presentan todos los métodos en general, es ignorar la información de texturas presentes en un área local debido a que no existe una librería de textones que comprenda todas las posibles texturas en la base de datos.

Otra falla es ignorar la información espacial local en la imagen para todos los métodos exceptuando Watershed. A pesar de que se incluyen las variables de espacio como cualidad en la imagen, la mayoría de métodos separan regiones homogéneas en varias secciones si éstas son amplias y abarcan gran parte de la imagen.

La estrategia de evaluación puede ser mejorada si se obtienen anotaciones en las cuales únicamente se tiene un objeto en la imagen, de esta forma se evita el error humano al agrupar los grupos de etiquetas que conforman un objeto completo dentro de la imagen.

Para mejorar el método, se requeriría que éste no imponga mínimos en la imagen dependiendo del tamaño de componentes conexos que encuentre en ella, ya que pueden haber objetos de tamaño pequeño, separados por bordes que no son detectados por el método de Canny y en consecuencia, forman un objeto de gran tamaño. Se puede mejorar el método y evitar la imposición de mínimos al emplear superpíxeles en la imagen, ya que estos reducirán todo ruido dada la textura de la localidad.

Referencias

- [1] M. González and V. Ballarin, “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] M. T. E. Portillo and J. A. S. P. Mendoza, “P. ch. mahalano-bis y las aplicaciones de su distancia estadística,” *CULCyT*, no. 27, 2015.