

Segmentación de imágenes de la base de datos BSDS500 mediante el uso de K-means y GMM

Cristian Alejandro Vergara Perico
Universidad de los Andes, Bogotá, Colombia.

Resumen—En este proyecto se desarrollarán dos métodos de segmentación basados en los métodos de clustering, k-means y mezcla de gaussianas. Estos métodos fueron ajustados mediante la determinación y modificación del espacio de características que definirán la agrupación y el número de subgrupos establecidos para dividir la imagen. Con el fin de evaluar su desempeño los métodos fueron empleados para segmentar imágenes de la base de datos BSDS5000 y fueron evaluados contra las notaciones de los seres humanos. Finalmente los resultados obtenidos, demostrarán que son métodos que permiten la segmentación sin embargo comparados con los métodos actuales y las segmentaciones humanas muestran deficiencias notorias.

Keywords—Segmentación, k-means, gmm,

I. INTRODUCCIÓN

EN el transcurso de los últimos 50 años la visión computacional ha tenido un gran avance, esto ha fomentado el desarrollo de algoritmos computacionales que permitan acercarse a la detección de objetos percibida por el sentido de visión humano.

Uno de los clásicos problemas de este campo en el cual se han enfocado diversos investigadores es la segmentación de imágenes. Algunos artículos que en los cuales los autores han desarrollado algoritmos de segmentación se pueden encontrar en [1], [2], [3], [4] entre otros.

Con el fin de comparar el comportamiento de los diversos algoritmos, se han recolectado bases de datos que contienen diversas imágenes segmentadas por humanos. La reunión de estas segmentaciones se toman como el la referencia contra la cual se comparan los algoritmos de la visión computacional. Este es el caso de la base de datos BSDS500 (Berkeley Segmentation Data Set and Benchmarks 500) descrita en [1].

En este laboratorio se desarrollaron dos algoritmos de segmentación basados en los métodos de clustering k-means y mezcla de gaussianas. Estos algoritmos se implementaron en las 200 imágenes de prueba de la base de datos de Berkeley (BSDS500) [1]. Finalmente los resultados se compararon con otros métodos de segmentación del estado del arte, estos resultados fueron graficados en un diagrama de precisión-cobertura.

En este artículo primero se describirá el problema de la segmentación desde el punto de vista de la visión humana (sección II). Luego en la sección III se describirán 4 algoritmos de agrupación que fueron evaluados para el desarrollo del algoritmo de segmentación. En seguida se encuentra la sección

IV donde se encuentra la calibración del algoritmo, selección de métodos y parámetros. Los resultados y su respectivo análisis se encuentran en la sección V. Finalmente se presentan las conclusiones del laboratorio en la sección VI.

II. DESCRIPCIÓN DEL PROBLEMA

El problema consiste en dividir una imagen en regiones que puedan definir objetos. Este problema es resuelto por la visión humana mediante la división de regiones basada en características como la forma, el color, la textura, ubicación entre otros.

Usualmente los objetos se pueden caracterizar por regiones que aglomeran toda la información adyacente en un área. La posición en la que se encuentra una de las características descritas previamente es relevante para la definición de objetos puesto que normalmente estos se encuentran definidos en una región unida del espacio cartesiano de posición.

La figura 1 muestra este concepto en segmentaciones realizadas por los seres humanos de la base de datos BSDS500. En esta figura es evidente que las regiones marcadas por el individuo diferencian objetos definidos por características de color y textura que se encuentran encerradas en áreas cerradas en el espacio de posición.

III. METODOS DE CLUSTERING

En esta sección se describirán los métodos empleados para la agrupación.



Figura 1: Anotaciones por los seres humanos. Izquierda imágenes reales, Derecha Anotaciones humanas

¹ Este trabajo fue realizado para el curso de maestría Visión Artificial dictado el primer semestre de 2015 en la Universidad de los Andes, Colombia

III-A. K-means

Este método de clasificación no supervisada recibe como parámetro de entrada el numero de regiones o clusters en los cuales se quiere subdividir la imagen. El algoritmo ubica aleatoriamente los cantroides desde los cuales se la mide distancia a los puntos del espacio de la característica a evaluar. Luego de realizar varias iteraciones el método busca minimizar la suma de las distancias entre puntos y el centroide. Ver referencia [5].

III-B. GMM

Al igual que K-means este método de clasificación no supervisada recibe como parámetro de entrada el numero de grupos en los cuales se subdividirá el espacio. El algoritmo asume que los puntos del espacio se encuentran distribuidos normalmente dentro los grupos. Este algoritmo realiza iteraciones para ajustar los datos al numero de distribuciones gaussianas definido previamente, el objetivo principal ajustar los datos a un modelo de mezcla de gaussianos aumentando la correlación entre este y los datos del espacio evaluado. Ver referencia [5].

III-C. Jerárquico

Este métodos se encuentra en la categoría de clasificación supervisada. Se basa en la subdivisión de la imagen en regiones en muchos subgrupos. EN el primer nivel cada pixel conforma una subdivisión. Conforme se avanza en los nivel de la jerárquia los subgrupos se van uniendo dependiendo de la distancia simple entre estos, esta union forma subgrupos de mayor tamaño. Despues de avanzar por todos los niveles de la jerárquia se forma un único grupo el cual contiene todos los subgrupos de niveles mas bajos. A partir de este análisis es posible definir el número de subgrupos deseado con el fin de segmentar la imagen en un determinado número de regiones. Ver referencia [5].

III-D. Watersheds

Al igual que el método jerárquico watersheds es un algoritmo de clasificación supervisada. Este método considera las representación de imágenes como terrenos topográficos en los cuales se producen inundaciones desde los puntos mínimos. La generación de contornos se logra a partir de la unión de pequeños lagos conforme aumenta el nivel de agua. En su estado basico el algoritmo considera todos los mínimos locales como precursores de subgrupos generando una sobresegmentación de la imagen, con el fin de evitar el fenomeno descrito anteriormente se define un nivel desde el cual se escogen los mínimos. [5].

IV. SELECCIÓN DE PARÁMETROS

En esta sección se encuentra la calibración del metodo desarrollado.

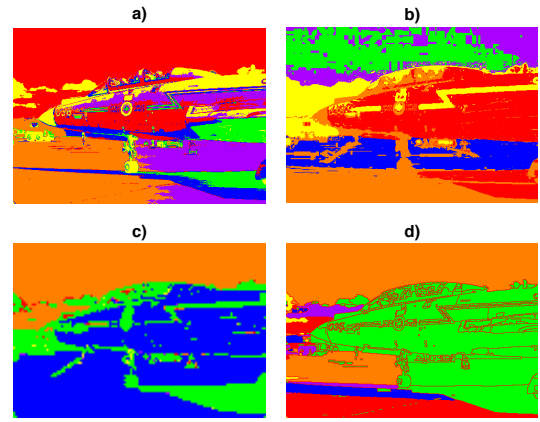


Figura 2: Segmentaciones realizadas por los 4 metodos de clustering. a) Corresponde al metodo de k-means con 9 grupos, b) al metodo GMM con 9 grupos, c) al metodo jerárquico con un nivel de distancia de 100 y b) al metodo de watershed con un ajuste de minimos en 40

IV-A. Espacio de características a evaluar

El espacio de color empleado para realizar las segmentaciones es el espacio LAB dado a que este brinda información tanto de colores como de luminosidad en uno de sus tres canales. Este canal de luminosidad añade una característica que permite segmentar de mejor manera los objetos.

A este espacio se le añadieron las coordenadas espaciales de los pixeles a evaluar, con el fin de forzar la agrupación de pixeles que se encuentren en posiciones adyacentes. Con lo anterior la dimensión final del espacio de características a evaluar fue de 5. Las 2 ultimas dimensiones correspondientes a la ubicación espacial se ponderaron con el fin de no sesgar la segmentación por la posición de los pixeles. La información de cada uno de los canales se encuentra resumida en la tabla I.

Canal	Mínimo	Máximo
L	0	100
a	-128	127
b	-128	127
x	0	20
y	0	20

Cuadro I: Valores máximos y mínimos de los canales del espacio utilizado para realizar la segmentación

IV-B. Selección métodos de agrupación

Al implementar los métodos descritos en la sección III se obtuvieron las imágenes de la figura 2

De estos se seleccionaron los métodos supervisados mezcla de gaussianas (gmm) y k-means debido al tiempo de procesamiento y los resultados de segmentaciones obtenidas, visibles en la fitura 2.

IV-C. Numero de clusters

Idealmente se debería escoger un rango amplio del numero de clústeres con el fin lograr una mejor evaluación del algoritmo de segmentación, sin embargo al emplear un número elevado de clústeres el tiempo de procesamiento aumenta. De esta manera el número de clústeres empleado para evaluar el algoritmo de segmentación fue: [3,6,9,12,15]

V. RESULTADOS

En la figura 3 se encuentra la comparación de las segmentaciones realizadas por los seres humanos y los método de k-means y mezcla de gaussianas. Se puede apreciar que el peso dado a los canales de posición fue demasiado alto, lo cual genero sesgo demasiado alto por la característica de posición.

El metodo de mezcla de gaussianas segmento de mejor manera las imagenes, separando características relevantes que fueron diferenciadas por las segmentaciones humanas, sin embargo al tener en cuenta la posición, se separaron objetos que pertenecían al mismo grupo. En las regiones del cielo y el agua los algoritmos generan divisiones que no son marcadas por los humanos.

En la figura 4 se presentan los resultados de ambos algoritmos en un diagrama de precisión y cobertura, estos indices muestran el desempeño de los metodos generados y son comparados con el metodo desarrollado en [1], los resultados se encuentran resumidos en el cuadro II.

Metodo	F
Humano	0.79
gPb-owt-ucm	0.71
GMM	0.54
kmeans	0.49

Cuadro II: Valores de F para diferentes metodos de segmentación

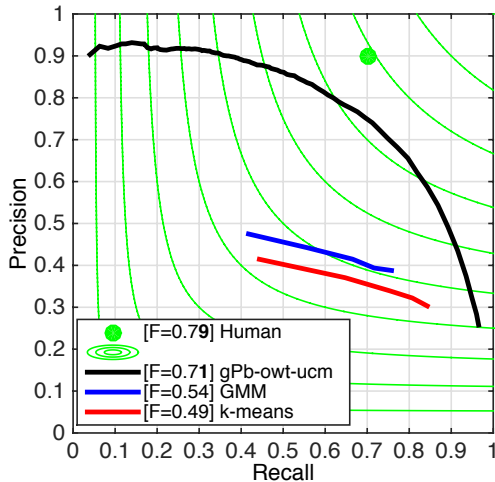


Figura 4: Evaluación de los metodos en diagrama precisión-cobertura

La diferencia entre los metodos generados y las anotaciones humanas es bastante significativa. Sin embargo la diferencia entre los 2 metodos empleados es de tan solo 0.05 porciento, ademas el tiempo empleado por el metodo de mezcla de gaussianas es mucho mayor que el tiempo empleado por el metodo de k-means, lo cual describe el ultimo como un metodo mas apropiado si la variable de tiempo es requerida.

Estos metodos de segmentación generarón líneas cortas en el diagrama de precisión vs cobertura. El aumento de número de grupos establecidos para el metodo k-means y gmm puede aumentar la cobertura y precisión de ambos metodos, sin embargo no se garantiza que que ambos indices aumenten simultaneamente.

VI. CONCLUSIONES

Los algoritmos empleados para la segmentación de las figuras son medianamente eficaces en comparación de las segmentaciones humanas o el metodo descrito en [1]. Sin embargo su simpleza induce un corto tiempo en el procesamiento.

En las segmentaciones realizadas por los seres humanos, se puede observar que estos segmentaron las áreas con misma textura como una misma región, por ejemplo la segmentación de cebras en la figura 3. Los metodos implementados se encuentran limitados en el reconocimiento de estas características.

Por otro lado al incluir las características de posición, el metodo es incapaz de reconocer un objeto que se encuentre ocluido por otro, dado a que este generaría regiones separadas pertenecientes al mismo.

Se recomienda disminuir el peso de los canales de posición del espacio a evaluar, con el fin de lograr mejores resultados al no sesgar demasiado las segmentaciones por la posición. De igual manera se aconseja ampliar el rango de número de clusters para de esta manera conseguir un mayor número de muestras y tener una evaluación más significativa

REFERENCIAS

- [1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011. [Online]. Available: <http://dx.doi.org/10.1109/TPAMI.2010.161>
- [2] M. Donoser, M. Urschler, M. Hirzer, and H. Bischof, "Saliency driven total variational segmentation," *ICCV*, no. *, pp. 810–813, 2009.
- [3] E. Sharon, M. Galun, D. Sharon, R. Basri, and A. Brandt, "Hierarchy and adaptivity in segmenting visual scenes," *Nature*, vol. 442, no. *, p. 810–813, 2006.
- [4] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. *, no. *, p. *, 2004.
- [5] P. Arvelaez, "Computer Vision. IBIO4490. Lecture 5, Clustering."

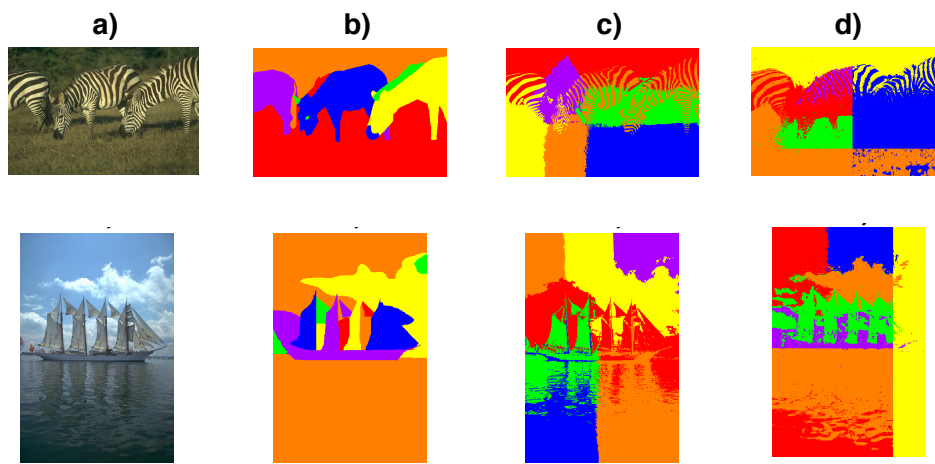


Figura 3: Segmentaciones obtenidas. Columna a) imagenes, b) segmentaciones humanas, c) segmentaciones k-means y d) segmentaciones gmm