

Evaluación de detección de contornos en dos métodos de segmentación tradicionales

Juan Sebastián Cuéllar
Universidad de los Andes, Bogotá, Colombia
js.cuellar169@uniandes.edu.co

Abstract

Este artículo presenta una evaluación cuantitativa de detección de contornos en dos métodos de segmentación tradicionales (kmeans y modelos de mezcla de gaussianas) adoptando la metodología del set de datos de segmentaciones de Berkeley (BSDS500). Los resultados se comparan con métodos de segmentación más recientes en cuanto a su cobertura y precisión. Tras la evaluación se evidencia que los métodos tradicionales presentan un desempeño inferior en cuanto a su precisión. Se mencionan posibles alternativas para mejorar su rendimiento.

1. Introduction

En este artículo se describen y se evalúan dos métodos de clustering tradicionales para solucionar dos de los problemas fundamentales de la visión computacional: detección de contornos y segmentación de imágenes. Para evaluar el desempeño de los métodos se hará uso del data set de segmentación de Berkeley (BSDS 500) introducido en [7] Aunque se tiene evidencia de métodos con resultados mucho más satisfactorios[4][6][8][1] el objetivo de este trabajo es determinar una base cuantitativa en cuanto al desempeño de los métodos elegidos y proponer alternativas para mejorar su rendimiento.

Se desarrolló un algoritmo de segmentación para cuatro métodos de clustering tradicionales, kmeans, modelos de mezcla de gaussianas (GMM), watersheds y clustering jerárquico. El algoritmo se utilizó en un pequeño grupo de imágenes del set BSDS 500 empleando diferentes combinaciones de argumentos (método de clustering, espacio de representación y número de clústeres, altura de mínimos extendidos o umbral de inconsistencia, ver fig. 1). Tras definir los clústeres y observar las segmentaciones se tomaron en cuenta los métodos de clustering que mejor conformaban una representación visual de la separación de las regiones homogéneas de la imagen. Para el caso particular de este estudio se eligieron los métodos de kmeans y GMM en el

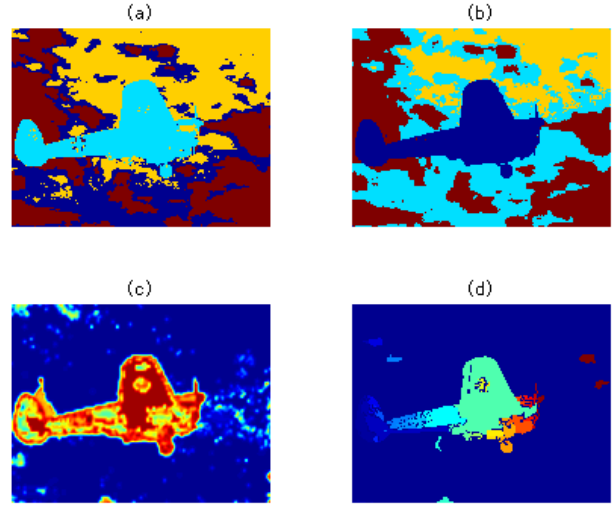


Figure 1. Segmentaciones sobre una imagen del set de entrenamiento en espacio RGB usando (a) kmeans (k=4) (b) GMM (k=4) (c) segmentación jerárquica (umbral de coeficiente de inconsistencia = 1) y (d) watersheds (altura de mínimos extendidos = 40) .

espacio de representación RGB. Los métodos elegidos se describen a continuación.

1.1. Clustering por kmeans

Este método busca particionar n observaciones en k clústeres dentro de los cuales cada observación pertenece al clúster con la media más cercana. Para encontrar los clústeres se emplean dos pasos iterativos hasta que la suma de los cuadrados de los errores dentro del grupo no puede reducirse más [3], matemáticamente se busca:

$$\operatorname{argmin} \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (1)$$

donde μ_i es la media de los puntos en S_i .

Los pasos iterativos son los siguientes: (1) Calcular la media de cada clúster. Los puntos pertenecientes a los

clúster iniciales pueden elegirse de forma aleatoria y (2) Calcular la distancia entre cada punto y cada clúster usando la distancia hacia cada media del clúster correspondiente. Se asigna cada punto al clúster que se encuentra a una menor distancia.

Los grupos obtenidos son tan compactos geoméricamente como sea posible alrededor de su clúster respectivo [3].

1.2. Clustering por modelo de mezclas de gaussianas (GMM)

Utiliza una función de densidad de probabilidad paramétrica representada como una suma ponderada de M componentes de densidades gaussianas para explicar la distribución de los píxeles [5]. La probabilidad está dada por la siguiente ecuación

$$p(x|\lambda) = \sum_{i=1}^M w_i g(x|\mu_i, \Sigma_i) \quad (2)$$

donde $w_i, i = 1, \dots, M$, son las ponderaciones de las mezclas y $g(x|\mu_i, \Sigma_i)$ son las componentes de las densidades de las gaussianas de la forma

$$\frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu_i)' \Sigma_i^{-1} (x - \mu_i) \right\} \quad (3)$$

donde μ_i es el vector de medias y Σ_i es la matriz de covarianzas. Para definir los parámetros que determinan las densidades de probabilidad gaussiana (μ_i, Σ_i y w_i) se utiliza el algoritmo de esperanza – maximización. Este algoritmo es otro método iterativo que busca la máxima verosimilitud de los parámetros estimados hasta hallar un punto máximo de convergencia. Los pasos iterativos son

- Se estiman responsabilidades con ciertos parámetros dados, paso E.

$$z_{ik} = \frac{1}{z_i} \pi_k N(x|\mu_k, \Sigma_k) \quad (4)$$

sueto a

$$\sum_k z_{ik} = 1 \quad (5)$$

- Se calculan los parámetros que maximizan las responsabilidades, paso M.

$$\mu_k = \frac{1}{N_k} \sum_i z_{ik} x_i \quad (6)$$

$$\Sigma_k = \frac{1}{N_k} \sum_i z_{ik} (x_i - \mu_k)(x_i - \mu_k)^T \quad (7)$$

$$\pi_k = \frac{N_k}{N} \quad (8)$$

Los grupos obtenidos se agrupan en M clústeres y son incluidos de acuerdo a la densidad de probabilidad descrita por los parámetros encontrados [5]

2. Metodología

Tras una primera evaluación visual del algoritmo con algunas imágenes de entrenamiento de BSDS 500 se segmentaron las imágenes del directorio de prueba usando una función de segmentación dentro del software Matlab®. Cada imagen se segmentó con cinco números de clúster k diferentes ($k = 2, 3, 4, 5, 6$) usando cada método elegido (kmeans y GMM). Las cinco segmentaciones de cada imagen se guardaron en una celda en formato .mat y luego se movieron a un folder individual para emplear la función boundaryBench para Matlab de [7]. Para permitir una base de comparación de los algoritmos por kmeans y GMM se utiliza el método de segmentación globalized probability of boundary – oriented watershed transform – ultrametric contour map (gPb-OWT-UCM) de [7]. La función boundaryBench permite evaluar cada método usando el marco basado en fronteras de precisión – cobertura de [2] para determinar la calidad de detección de las fronteras y usa el set BSDS como "ground truth" para establecer esta medida de calidad. Dicha metodología de evaluación mide el desempeño del algoritmo en términos de la precisión, lo cual es la fracción de verdaderos positivos, y la cobertura, la fracción de píxeles de fronteras del "ground truth" detectados. Esto permite obtener una medida F que resume el resultado del método y que se puede reportar basada en los resultados de todo el set de imágenes para un escala óptima fija (ODS) y como un agregado de medidas F basado en los resultados del set de imágenes para la escala óptima en cada imagen (OIS), y adicionalmente la precisión promedio (AP) que en otras palabras es el área bajo la curva precisión – cobertura.

3. Results

Table 1. Referencia para detección de contornos sobre el BSDS500. Resultados de la evaluación de detección de contornos para tres métodos de segmentación. Se muestran las medidas F de todo el set de imágenes para un escala óptima fija (ODS), para la escala óptima en cada imagen (OIS) y el AP

Método	ODS	OIS	AP
Human	0.8	0.8	–
gPb-OWT-UCM	0.73	0.76	0.73
GMM	0.5	0.55	0.14
kmeans	0.45	0.48	0.09

En la tabla 1. se pueden observar los valores calculados de ODS, OIS y AP para los tres métodos evaluados. Tras observar la fig.2 y la tabla 1 es evidente que los resultados obtenidos al usar los métodos de kmeans y GMM no

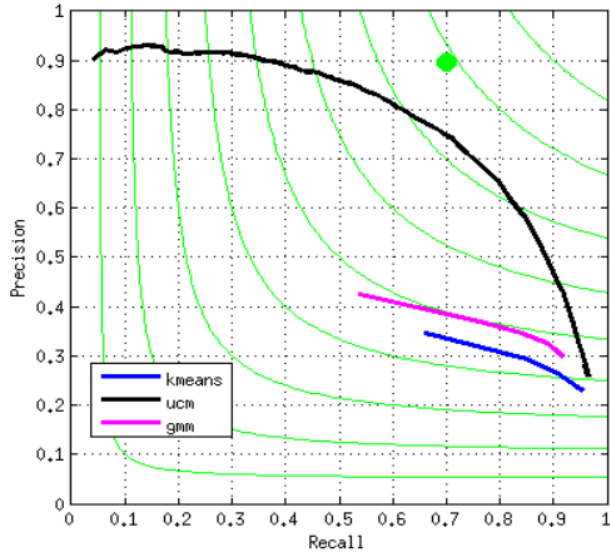


Figure 2. Curvas precisión - cobertura para los tres métodos evaluados. La diferencia en precisión de los métodos de kmeans y GMM en relación con gPb-OWT-UCM es evidente.

son satisfactorios puesto que su desempeño es mucho más reducido en comparación con el método UCM. Aunque el método de GMM es ligeramente mejor al de kmeans, los dos métodos alcanzan precisiones y AP muy reducidas.

4. Discusión

Los resultados obtenidos podrían explicarse a partir de los inconvenientes de los dos métodos respecto al número de clústeres finales puesto que dependen directamente de la persona que emplea el algoritmo, es decir, para iniciar la segmentación es necesario que el usuario indique un número de clústeres constante. Esto limita mucho la precisión del método debido a que en todo el set de imágenes BSDS 500 el número de los objetos o individualidades visuales no es constante y por lo tanto no es viable el uso de un número constante de clúster para la segmentación de todo el set. Debido a que los métodos elegidos aseguran el mismo número de segmentaciones que de clústeres de entrada, los algoritmos, en su afán de determinar dichos grupos, pueden definir fronteras entre objetos que no están definidas en la base "ground truth" con la frecuencia suficiente como para reducir la precisión significativamente. Los grupos conformados son susceptibles a "outliers" cuando se usa kmeans, esto implica que en presencia de algunos pixeles muy diferentes (mucha distancia con el resto de pixeles), las medias se van a calcular de tal forma que se puedan explicar dichos pixeles dentro de algún clúster. Esto implicará ocasionalmente que algunas medias se establezcan en posiciones poco adecuadas y en la agrupación incluirán pixeles más distanciados de lo necesario. Como consecuencia se ob-

tendrá una segmentación con fronteras poco precisas lo que posiblemente explicaría la razón por la cual presenta un rango de precisión menor en comparación con el método de GMM dentro de la evaluación realizada. Por otro lado, los clústeres encontrados por GMM son soluciones aproximadas porque dependen mucho de la inicialización de los parámetros de entrada y en ocasiones puede converger a un mínimo local, esto implicaría cierta varianza en la solución final y se esperarían curvas ligeramente diferentes cada vez que se calcularan las segmentaciones.

Para mejorar los métodos utilizados se podría desarrollar una metodología que determinara la cantidad de regiones homogéneas de la imagen para así utilizar un número de clústeres apropiado en cada una de ellas. Adicionalmente se podría buscar una alternativa para encontrar los centros de los clústeres de manera más precisa, esto considerando que los procesos iterativos no siempre convergen o establecen las medias en las posiciones más adecuadas y además consumen mucha memoria computacional.

References

- [1] D Comaniciu and P Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(5):603 – 619, May 2002.
- [2] D R Martin, C C Fowlkes, and Jitendra Malik. Learning to Detect Natural Image Boundaries Using Local Brightness, Color and Texture Cues. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5):530 – 549, May 2004.
- [3] David MacKay. Chapter 20. An Example Inference Task: Clustering. In *Information Theory, Inference and Learning Algorithms*, pages 284 – 292. Cambridge University Press, 2003.
- [4] E Sharon, M Galun, D Sharon, R Basri, and A Brandt. Hierarchy and Adaptivity in Segmenting Visual Scenes. *Nature*, 442:810 – 813, 2006.
- [5] G McLachlan. *Mixture Models*. Marcel Dekker, New York, 1988.
- [6] P F Felzenszwalb and D P Huttenlocher. Efficient Graph-Based Image Segmentation. *Int'l J. Computer Vision*, 59:167 – 181, 2004.
- [7] Pablo Arbeláez, Michael Maire, Charles Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE transaction on pattern analysis and machine intelligence*, 33(5):898 – 916, May 2011.
- [8] T Cour, F Benezit, and J Shi. Spectral Segmentation with Multiscale Graph Decomposition. *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, 2005.