COMPUTATION E INFORMATICA

Recibido 02 May. 2024 *ReCIBE, Año 13 No. 1, May. 2024*

Aceptado 16 May. 2024

Segmentación rápida de imágenes con múltiples características

Fast Image Segmentation with Multiple Features

Sergio Gomez Vega[1]
checoskleyxd@gmail.com[*]

Alberto Luque Chang[1]

Hector Joaquin Escobar Cuevas[1]

Fernando Vega Parra[1]

[1]Universidad de Guadalajara

[*]Autor para correspondencia

Resumen:
La segmentación de características múltiples es superior a los enfoques unidimensionales en escalas de grises. El algoritmo del cambio medio (CM) se utiliza comúnmente para esta tarea. A pesar de sus interesantes resultados, el CM sigue siendo computacionalmente prohibitivo para escenarios de segmentación en los que el mapa de funciones está formado por características multidimensionales. El enfoque propuesto considera un mapa de características bidimensional que incluye el valor en escala de grises y la varianza local de cada píxel de la imagen. Para reducir el coste computacional, se modifica el algoritmo de CM clásico para que opere sobre un número menor de puntos. En tales condiciones, se diferencian dos conjuntos de elementos: datos implicados (el conjunto reducido de datos considerados en la operación CM) y datos no implicados (el resto de datos disponibles). A diferencia del CM clásico, que emplea funciones gaussianas, en nuestro enfoque el proceso de estimación del mapa de características se lleva a cabo utilizando un enfoque más preciso, como la función kernel de Epanechnikov. Una vez obtenidos los resultados del CM, se generalizan para incluir los datos no utilizados. Así, cada característica no utilizada se asigna al mismo clúster de datos utilizados más cercano. Por último, los grupos con menos características se fusionan con otros grupos vecinos. El método de segmentación propuesto se ha comparado con otros algoritmos del estado de la técnica usando de la base de datos de Berkeley. Los resultados experimentales confirman que el esquema propuesto produce imágenes segmentadas con un 50% más de calidad de percepción visual.

Palabras clave: Segmentación de imagen, Algoritmo de cambio medio, Estimador de densidad del kernel (EDK)

Summary:

Multifeature segmentation is superior to unidimensional approaches in grayscale images. The Mean Shift (MS) algorithm is commonly used for this task. Despite its interesting results, MS remains computationally prohibitive for segmentation scenarios where the feature map consists of multidimensional features. The proposed approach considers a two-dimensional feature map that includes the grayscale value and the local variance of each pixel in the image. To reduce computational costs, the classic MS algorithm is modified to operate on a smaller number of points. Under these conditions, two sets of elements are differentiated: involved data (the reduced set of data considered in the MS operation) and uninvolved data (the remaining available data). Unlike classic MS, which employs Gaussian functions, our approach uses a more precise method for the feature map estimation process, specifically the Epanechnikov kernel function. Once the MS results are obtained, they are generalized to include the unused data. Each unused feature is assigned to the same cluster as the nearest involved data. Finally, clusters with fewer features are merged with neighboring clusters. The proposed segmentation method has been compared with other state-of-the-art algorithms using the Berkeley database. Experimental results confirm that the proposed scheme produces segmented images with 50% higher visual perception quality.

Keywords: Image segmentation, Mean Shift algorithm, Kernel Density Estimator (KDE)

1 Introducción

Una de las principales etapas del procesamiento de imágenes es la segmentación de las mismas, que separa una imagen en distintas sub-áreas que no se solapan en función de un determinado criterio de similitud. Los resultados finales de distintos algoritmos de reconocimiento de objetos, recuperación basada en el contenido y clasificación semántica están totalmente condicionados por la eficacia del método de segmentación de imágenes empleado. Sin embargo, segmentar una sección o un elemento de una imagen complicada representa un proceso difícil. En las tres últimas décadas, se han introducido varios métodos potentes de segmentación de imágenes. Entre ellos se incluyen esquemas como los basados en histogramas [1,2], los basados en gráficos [3-5], los basados en la detección de contornos [6], los basados en umbrales [7,8], los difusos [9,10], los basados en campos aleatorios de Markkov [11], los basados en texturas [12], los basados en el análisis de componentes principales [13]y los basados en agrupaciones [14,15]. De estos enfoques, las técnicas de segmentación basadas en umbrales e histogramas se consideran las más populares en diversos campos de aplicación.

Los esquemas de segmentación basados en umbrales pueden clasificarse en métodos de segmentación de dos niveles y de varios niveles en función del número de objetos o regiones que contenga la imagen. En general, las imágenes con más de dos objetos o zonas requieren umbrales multinivel para su correcta segmentación. En función de la información que procesan, los métodos de segmentación de imágenes basados en umbrales también podrían dividirse en seis clases. también podrían clasificarse en seis clases [10]. Esta clasificación incluye métodos basados en clustering, métodos basados en histogramas, métodos basados en histogramas, métodos basados en histogramas, métodos basados en atributos de objetos, métodos basados en entropía, métodos locales y métodos espaciales.

Por otro lado, los esquemas basados en histogramas utilizan la información contenida en un histograma unidimensional, como sus curvaturas, valles y picos, para determinar el número de elementos presentes en la imagen. Sin embargo, dado que los histogramas unidimensionales no tienen en cuenta la asociación espacial entre píxeles, estos enfoques mantienen un rendimiento insatisfactorio. Para mejorar los resultados producidos por los histogramas unidimensionales, se han propuesto técnicas de histograma bidimensional para la segmentación de imágenes [16]. En estos enfoques, los valores de gris originales considerados en los histogramas unidimensionales se combinan con información sobre la posición de los píxeles. combinados con información sobre la posición de los píxeles [16,17] para producir histogramas bidimensionales (posición de grises). Según estos métodos, el conjunto de elementos diagonales del histograma bidimensional implica información esencial para clasificar objetos de fondo, mientras que los elementos no diagonales contienen información de imagen sobre píxeles de borde y datos de ruido. Normalmente, con las técnicas basadas en histogramas bidimensionales, los detalles finos como puntos, bordes y líneas no se pueden segmentar correctamente, ya que los histogramas bidimensionales suavizan los datos de la imagen [18] calculando la información media en escala de grises alrededor de los píxeles objetivo. Desde su introducción, varios algoritmos nuevos han tenido en cuenta el concepto de histograma bidimensional para producir resultados de segmentación interesantes [19,20]. Un ejemplo típico es el enfoque propuesto por Xueguang y Shu-hong [21]en 2012. En este método, se introduce un nuevo histograma bidimensional para segmentar objetos en imágenes genéricas. El histograma integra valores de píxeles en escala de grises y su respectiva información de gradiente. Aunque el algoritmo presenta resultados interesantes, nuevos experimentos [22]han demostrado que, en varios escenarios, este enfoque ofrece resultados inferiores a los esquemas de posición de grises. inferiores a los esquemas de posición de grises.

Técnicas metaheurísticas [20,23] También se han utilizado en combinación con histogramas bidimensionales como esquemas de segmentación [19,20,23-25]. Los métodos metaheurísticos son técnicas computacionales inspiradas en fenómenos naturales o sociales, que se utilizan habitualmente para resolver formulaciones de optimización [26] caracterizadas por su alta complejidad. Las técnicas de segmentación que integran métodos metaheurísticos e histogramas bidimensionales formulan la información proporcionada por el histograma como un problema de optimización.

Por tanto, mediante el uso de una función objetivo, se utiliza una técnica metaheurística particular para determinar el número de elementos presentes en la imagen, que minimizan/maximizan su valor. Se han propuesto varios enfoques metaheurísticos para la segmentación bajo el esquema de histograma bidimensional. Algunos ejemplos incluyen enfoques como la optimización del enjambre de partículas (PSO).[27-29], Evolución Diferencial (DE)[23], Optimización de colonias de hormigas (OCA)[30], Algoritmos genéticos (AG)[31], Colonia de abejas artificiales (CAA)[32], Recocido simulado (SA) [33], Algoritmo de enjambre de peces artificiales (AFS)[34], Optimización de enjambre de golondrinas (SSO) [35]Optimización similar al electromagnético (EMO) [36]y algoritmo de búsqueda gravitacional [37]. La mayoría de estas técnicas de segmentación tienen un problema crítico: el número de clases debe ser previamente fijado antes de su ejecución [36,37]. Este hecho limita seriamente su aplicación cuando se desconoce la naturaleza y características de la imagen. Los métodos de segmentación basados en técnicas metaheurísticas presentan buenos resultados en términos de optimización de características que relacionan el problema de segmentación con una función de costos, pero estos resultados no siempre se reflejan en la calidad visual [37]. Otra característica adversa de estos enfoques es su alto coste computacional. Todos los enfoques de segmentación basados en histogramas bidimensionales consideran el histograma como un mapa de características para relacionar dos características de píxeles diferentes. De hecho, los histogramas representan el método más sencillo para producir mapas de densidad. Para su generación, el espacio de datos se divide en clases disjuntas o contenedores, mientras que la densidad se estima calculando cuántos puntos de datos caen en cada contenedor. A pesar de su simplicidad, los histogramas presentan graves inconvenientes como baja precisión, discontinuidad de densidad y alto costo computacional [38]. Otro método útil para producir mapas de características es el método del estimador de densidad del kernel (KDE). KDE mantiene muchas características interesantes como regularidad asintótica, fundamentos matemáticos bien conocidos, propiedades de continuidad y diferenciabilidad [39,40]. A diferencia de los histogramas, los enfoques de KDE producen mapas de características con mayor precisión y propiedades suaves. Bajo el esquema KDE, la densidad en cada punto se calcula mediante el uso de funciones simétricas (funciones kernel) que calculan la acumulación ponderada de elementos vecinos. Varias funciones del kernel [38,41]Se han propuesto para su adopción en KDE métodos como las funciones gaussiana, triangular, uniforme, triweight, biweight y Epanechnikov. Aunque la función gaussiana es la más utilizada debido a su popularidad, el núcleo de Epanechnikov [42] presenta la mejor precisión de estimación cuando la cantidad de datos para calcular la densidad del mapa de características es muy limitada [42].

El esquema KDE también se ha empleado con fines de agrupación y clasificación. El ejemplo más representativo de su uso es el desplazamiento medio (CM) [43]algoritmo. El funcionamiento de CM considera dos pasos [44]. En la primera etapa, se realiza la estimación del mapa de características. Esta operación se logra mediante el uso del enfoque KDE. Luego, CM busca las posiciones que corresponden a los máximos locales del mapa de características de densidad mediante el uso de un método de gradiente. Estas posiciones, llamadas atractores, representan los prototipos de cúmulos de cada clase. Los resultados precisos y sólidos de CM han motivado su aplicación en varios contextos, como el filtrado de imágenes [45]y segmentación de imágenes [46]. En el proceso de segmentación de imágenes, CM detecta secciones unidimensionales homogéneas combinando el grupo de píxeles que corresponde a puntos suficientemente cercanos. puntos del tractor. En la literatura se han introducido varios algoritmos de segmentación basados en CM. Algunos ejemplos incluyen el enfoque propuesto por Tao et al [47]que incorpora un algoritmo de árbol con CM para la segmentación eficiente de barcos en imágenes infrarrojas. En [48], Parque y col. introdujo una técnica de segmentación basada en la fusión de un modelo de mezcla gaussiana con CM.

A pesar de sus interesantes resultados, CM mantiene un grave inconveniente: su coste computacional [42,49,50]. Durante la operación de CM, el mapa de características se calcula en cada punto, considerando la información de todos los datos disponibles. Asimismo, el proceso de asignación de conglomerados se lleva a cabo mediante la iteración de un método de gradiente para cada punto del mapa de características. Bajo esta limitación, el uso de CM es prohibitivo en escenarios de segmentación donde el mapa de características consta de características multidimensionales como las representadas por histogramas bidimensionales. Esta es la razón principal porque la mayoría de los algoritmos de segmentación de CM combinan datos

unidimensionales (valores en escala de grises) con otras técnicas computacionales que extraen información de clasificación sin utilizar características de píxeles adicionales (es decir, más dimensiones) [42]. El uso de más dimensiones en el mapa de características aumenta significativamente el costo computacional del método CM. Por lo tanto, el desafío es encontrar un esquema computacionalmente eficiente para manejar mapas de características multidimensionales. Una técnica común para reducir el costo computacional es centrar la operación del algoritmo en una pequeña cantidad representativa de elementos de todo el conjunto de datos. Los resultados parciales obtenidos luego se generalizan para incluir información que no está involucrada en el conjunto de datos de enfoque inicial. Luego, los resultados parciales obtenidos deben generalizarse para incluir información no involucrada. Esta metodología se ha aplicado con éxito en enfoques como Random Hough Transform (RHT) [51] y Consenso de muestra aleatoria (RANSAC) [52], para nombrar unos pocos.

En este artículo, se presenta un nuevo algoritmo de segmentación competitivo para imágenes en escala de grises. El enfoque propuesto considera un mapa de características bidimensional que incluye el valor de escala de grises y la varianza local para cada píxel de la imagen. Para reducir el costo computacional, el algoritmo de cambio medio (CM) se opera utilizando un conjunto representativo de elementos que contienen una cantidad muy limitada de puntos de todos los datos disponibles. En tales condiciones, se diferencian dos conjuntos de elementos: datos involucrados (el conjunto de datos reducido considerado en la operación CM) y datos no involucrados (el resto de datos disponibles). A diferencia de los enfoques tradicionales de CM que emplean funciones gaussianas, en nuestro enfoque, el proceso para estimar los valores de densidad en el mapa de características utiliza la eficiente función del núcleo de Epanechnikov. Una vez obtenidos los resultados de CM, se generalizan para incluir los datos no involucrados. Por lo tanto, cada elemento no utilizado se asigna al mismo grupo de datos utilizados más cercanos. Finalmente, los grupos con la menor cantidad de elementos se fusionan con otros grupos vecinos. El método de segmentación propuesto se ha comparado con otros algoritmos de última generación considerando el número total de imágenes del conjunto de datos de Berkeley. Los resultados experimentales confirman que el esquema propuesto presenta un mejor desempeño en términos de calidad, consistencia y precisión.

El resto del artículo está organizado de la siguiente manera: En la Sección 2, se da una breve explicación del algoritmo de cambio medio; en la Sección 3, se ilustran las principales características del problema ICE; en la Sección 4, se presentan los elementos más importantes sobre la CMA; En la Sección 5, se presenta el enfoque propuesto; La sección 6 exhibe los resultados experimentales; finalmente, en la Sección 7, se extraen las conclusiones.

2 Algoritmo de cambio medio

cambio medio [43]es un algoritmo que se ha utilizado comúnmente para la segmentación de múltiples características. Durante la operación CM, a cada característica $\boldsymbol{x}$ presente en el espacio de características, se le asigna un grupo $\boldsymbol{C_i}$ cuya posición prototipo $\boldsymbol{x_*}$ Corresponde a los máximos locales del mapa de densidad. Por lo tanto, a partir del punto $\boldsymbol{x}$, se calcula una nueva u bicación en la dirección del mayor aumento del mapa de densidad hasta un máximo local $\boldsymbol{x_*}$ ha sido alcanzado. La operación del desplazamiento medio considera dos etapas. Primero, se realiza el mapa de densidad del espacio de características. Esta operación se logra mediante el uso de un enfoque KDE. Luego, se lleva a cabo un paso de optimización basado en el ascenso del gradiente para encontrar los máximos locales (atractores de densidad) sobre el mapa de densidad [53].

2.1 Mapa de densidad

Un método útil para estimar el mapa de características de densidad $f(x)$ forman un conjunto de datos es el KDE no paramétrico. Bajo el esquema KDE, la densidad en cada punto se calcula mediante el uso de funciones del núcleo que calculan la acumulación ponderada de elementos vecinos. Una función del núcleo $\boldsymbol{k}$ se define como un modelo de kernel simétrico y no negativo $\boldsymbol{k}$, tal que:

$$K(x) \geq 0; K(-x) = K(x); \int K(x)dx = 1 \qquad (1)$$

Suponiendo un conjunto de $\boldsymbol{n}$ características $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots \boldsymbol{x_n}\}$, la función de densidad estimada $\boldsymbol{f(x)}$ se puede calcular en términos de $\boldsymbol{K}$ de la siguiente manera:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \qquad (2)$$

Dónde $\boldsymbol{h}$ representa la influencia del área de $\boldsymbol{K}$.

Aunque se han propuesto varias funciones del núcleo para los métodos de KDE, la mayoría de los esquemas de KDE consideran la función gaussiana. $\boldsymbol{K_G(x)}$ como modelo de núcleo. El núcleo gaussiano o normal se define en Ec.(3) mientras que su representación gráfica es como se muestra en la figura 1.

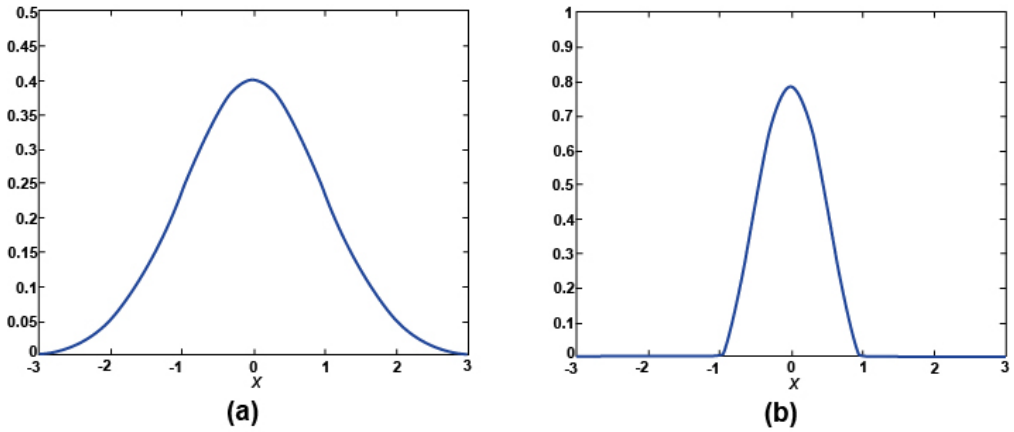


Figura 1. Diferentes modelos de kernel(s); kernel gaussiano $K_G(x)$ y (b) núcleo de Epanechnikov $\boldsymbol{K_E(x)}$

$$K_G(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2h}} \qquad (3)$$

Otra interesante alternativa del KDE es la función Epanechnikox $\boldsymbol{K_E(x)}$ (como se aprecia en la Figura 1(b)) [49], la cual está definida por el siguiente modelo:

$$K_E(x) = \begin{cases} \frac{3}{4}(1 - x^2), if\ |x| \leq 1 \\ 0, de\ otra\ manera \end{cases} \qquad (4)$$

Una forma común de evaluar la precisión de la estimación de un enfoque KDE es el análisis del error cuadrático medio (ECM) producido entre el mapa de características de densidad real $\boldsymbol{f}(\boldsymbol{x})$ y su representación estimada $\hat{\boldsymbol{f}}(\boldsymbol{X})$ de modo que

$$MSE\left(\hat{f}(x)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{f}(x_i) - f(x_i)\right)^2 \qquad (5)$$

En tales circunstancias, el núcleo de Epanechnikov $\boldsymbol{K_E}(\boldsymbol{x})$ presenta la mejor precisión de estimación en términos de CME criterio desde su expresión $\boldsymbol{K_E}(\boldsymbol{x})$, según varios estudios [38,41], representa la solución funcional que minimiza la CME formulación de Ec. (5). Este comportamiento se mantiene incluso cuando la cantidad de datos para calcular la densidad del mapa de características es muy limitada [41].

Para mostrar las capacidades de modelado del núcleo de $\boldsymbol{K_E}(\boldsymbol{x})$. Se considera el siguiente ejemplo. Suponga la función de densidad de probabilidad (PDF) producida por una mezcla gaussiana bidimensional $\boldsymbol{M}(\boldsymbol{X_1}, \boldsymbol{X_2})$. La mezcla $\boldsymbol{M}(\boldsymbol{X_1}, \boldsymbol{X_2})$ combina tres funciones gaussianas $(\boldsymbol{j} = \boldsymbol{1}, \boldsymbol{2}, \boldsymbol{3})$ según el modelo $\boldsymbol{N_j}(\mu_j, \boldsymbol{\Sigma_j})$, dónde μ corresponde a su punto central mientras $\sum$ simboliza su covarianza. La Figura 2 (a) presenta una mezcla gaussiana definida como $(\boldsymbol{X_1}, \boldsymbol{X_2}) = \left(\frac{2}{3}\right)\boldsymbol{N_1} + \left(\frac{1}{6}\right)\boldsymbol{N_2} + \left(\frac{1}{6}\right)\boldsymbol{N_3}$. De este PDF, se muestra un conjunto reducido de 120 puntos de datos (verFigura 2(b)). Teniendo en cuenta este espacio de características, los núcleos $\boldsymbol{K_G}(\boldsymbol{x})$ y $\boldsymbol{K_E}(\boldsymbol{x})$ se han utilizado para estimar el mapa de densidad $\hat{\boldsymbol{f}}(\boldsymbol{x_1}, \boldsymbol{x_2})$ obtención de Figura 2(c) y 2(d), respectivamente. De la Figura 2(c) y 2(d), está claro que el núcleo de Epanechnikov $\boldsymbol{K_E}(\boldsymbol{x})$ alcanza un mejor resultado de estimación que el kernel normal $\boldsymbol{K_G}(\boldsymbol{x})$. Es evidente que el núcleo gaussiano produce varios máximos locales falsos en$\hat{\boldsymbol{f}}(\boldsymbol{x_1}, \boldsymbol{x_2})$ como consecuencia de su incapacidad para modelar adecuadamente el mapa de densidad cuando el conjunto de datos disponibles se reduce drásticamente.

2.1 Atractores de densidad

La idea principal detrás de la estimación de un mapa de densidad. $\hat{\boldsymbol{f}}(\boldsymbol{x_1}, \boldsymbol{x_2})$para un conjunto de datos dado es encontrar los puntos característicos $\boldsymbol{x_i^*}$ que representan los prototipos de clúster que resumen mejor la distribución de los puntos de datos modelados de $\hat{\boldsymbol{f}}(\boldsymbol{x})$. Estos puntos corresponden a los máximos locales presentes en el mapa de densidad estimado $\hat{\boldsymbol{f}}(\boldsymbol{x})$. Estos puntos se pueden encontrar aplicando un método de optimización basado en gradientes. Considerando un conjunto de datos $\boldsymbol{x}$ de $\boldsymbol{d}$ elementos $(\boldsymbol{x} = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x_d}\})$, en este esquema, un punto aleatorio $\boldsymbol{x_j}$ primero se selecciona $(\boldsymbol{j} \in \boldsymbol{1}, \ldots, \boldsymbol{d})$. Luego, una nueva ubicación $\boldsymbol{x_i^{t+1}}$ se calcula en la dirección del gradiente de densidad $\boldsymbol{\nabla}\hat{\boldsymbol{f}}(\boldsymbol{x_j})$ hasta el máximo local $\boldsymbol{x_i^*}$ ha sido alcanzado. Esta regla de actualización se formula de la siguiente manera:

$$x_j^{t+1} = x^t + \delta \cdot \nabla\hat{f}\left(x_i^t\right) \qquad (6)$$

Dónde $\boldsymbol{t}$ denota el número de iteración actual, mientras que δ representa el tamaño del paso. Bajo esta formulación, el gradiente de densidad en cualquier punto $\boldsymbol{x}$ se calcula calculando la derivada de la PDF de la siguiente manera:

$$\nabla\hat{f}(x) = \frac{\partial}{\partial x}\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n}\frac{\partial}{\partial x}K\left(\frac{x - x_i}{h}\right) \qquad (7)$$

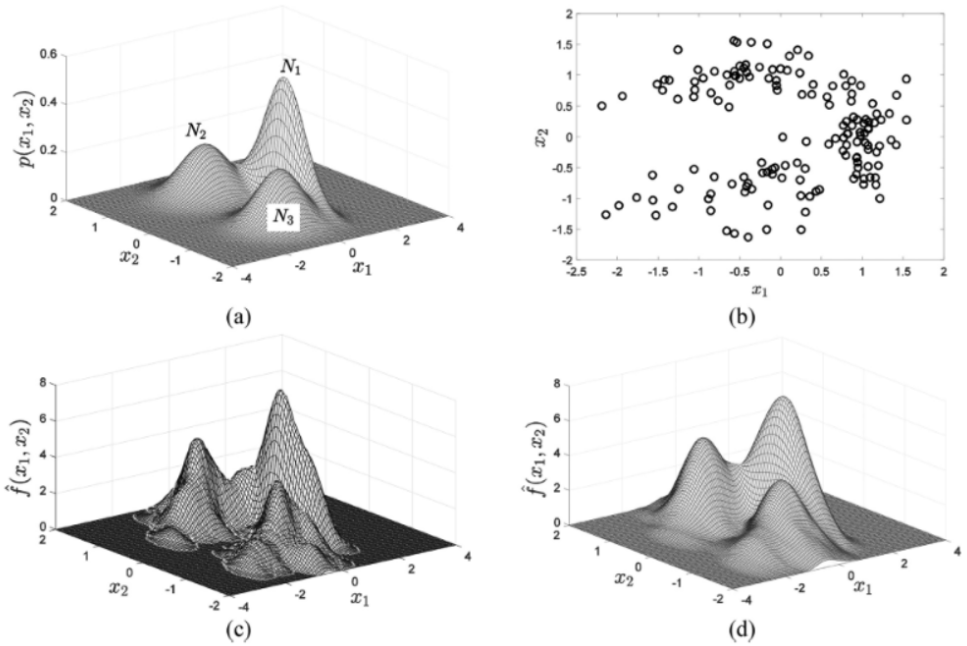


Figura 2. Capacidades de modelado del kernel gaussiano. $\boldsymbol{K_G(x)}$ y Epanechnikov $\boldsymbol{K_E(x)}$ cuando se reduce el conjunto de datos disponibles. (a) PDF bidimensional, (b) 120 características muestreadas del PDF, (c) estimación del mapa de densidad $\boldsymbol{\hat{f}(x_1, x_2)}$ considerando el núcleo kernel gaussiano y (d) $\boldsymbol{K_G(x)}$ y Epanechnikov $\boldsymbol{K_E(x)}$.

Dónde $\boldsymbol{n}$ representa el número de dimensiones de $\boldsymbol{x} = (\boldsymbol{x_1}, \ldots, \boldsymbol{x_n})$. Una vez aplicado el método del gradiente a todos de los $\boldsymbol{d}$ elementos del conjunto de datos $\boldsymbol{x}$, varios puntos de datos serán atraídos por el mismo máximo local $\boldsymbol{x_i^*}$. Estos puntos se clasificarán luego como una parte del grupo $\boldsymbol{C_i}$. Bajo estas condiciones, como resultado final, obtenemos una lista de $\boldsymbol{C}$ atractores de densidad $\{\boldsymbol{x_1^*}, \ldots, \boldsymbol{x_c^*}\}$ dónde cada atractor $\boldsymbol{x_i^*}$ debe mantener un conjunto de elementos $X_i = \{\boldsymbol{x_a}, \ldots, \boldsymbol{x_g}\}$ que corresponde a todos los datos atraídos por $\boldsymbol{x_i^*}(\boldsymbol{x_a}, \boldsymbol{x_c} \in \boldsymbol{C_i})$. La figura 3 presenta una ilustración del proceso de gradiente para encontrar los máximos locales. En la imagen se muestra la información de contorno del mapa de densidad $\boldsymbol{\hat{f}(x_1, x_2)}$. En tales condiciones, comenzando en puntos $\boldsymbol{x_p}$,$\boldsymbol{x_q}$ y $\boldsymbol{x_r}$ se calculan nuevas ubicaciones en la dirección del mayor aumento del mapa de densidad$\boldsymbol{\hat{f}(x_1, x_2)}$. hasta los máximos locales, $\boldsymbol{x_1^*}$, $\boldsymbol{x_2^*}$ **y** $\boldsymbol{x_3^*}$ respectivamente, ha sido alcanzado.

3 Método propuesto

En este enfoque, llamado CAM-SEG, la segmentación es producida mediante el análisis del espacio de características bidimensionales, los cuales incluyen el valor de escala de grises y la varianza local para cada pixel en la imagen. El proceso agrupa los pixeles de acuerdo con su intensidad dentro del mapa de objetos bidimensional. Esto se logra mediante el algoritmo CM modificado. Para reducir el costo computacional, el algoritmo CM es adaptado para contener elementos representativos, que tienen un número limitado de pixeles de las características limitando así, la información disponible, por lo tanto, dos grupos de elementos son identificados: características usadas (la reducción del conjunto de datos) y características no usadas (el resto de la información). En este enfoque, el proceso para estimar la densidad de los valores en el mapa de características es adoptado de la función de Epanechnikov para su kernel. Una vez se obtiene el CM resultante, se generaliza para incluir las características no integradas. A continuación, a cada elemento no utilizado se le asigna el grupo más cercano y para aumentar la eficiencia grupos con poca cantidad de datos se fusionan con otros grupos.

3.1 Definición de características

Si se asume que $\boldsymbol{I}(\boldsymbol{x},\boldsymbol{y})$ corresponden a niveles de escala de grises de 0 a L-1 para cada pixel en el espacio coordenado $(\boldsymbol{x},\boldsymbol{y})$ en cualquier imagen de dimensión **M*N.** $\boldsymbol{V}(\boldsymbol{x},\boldsymbol{y})$ representa la varianza del valor del pixel en las coordenadas $(\boldsymbol{x},\boldsymbol{y})$, generando una ventana cuadrada de dimensión **n*n**. Por lo tanto, el objeto $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ de un pixel localizado en $(\boldsymbol{x},\boldsymbol{y})$ se simboliza bidimensionalmente como el siguiente vector:

$$\boldsymbol{F}(x,y) = [\boldsymbol{I}(x,y)\boldsymbol{V}(x,y)] \qquad (8)$$

En $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ ambos elementos $\boldsymbol{I}(\boldsymbol{x},\boldsymbol{y})$ y $\boldsymbol{V}(\boldsymbol{x},\boldsymbol{y})$ son normalizados de 0 a 1.

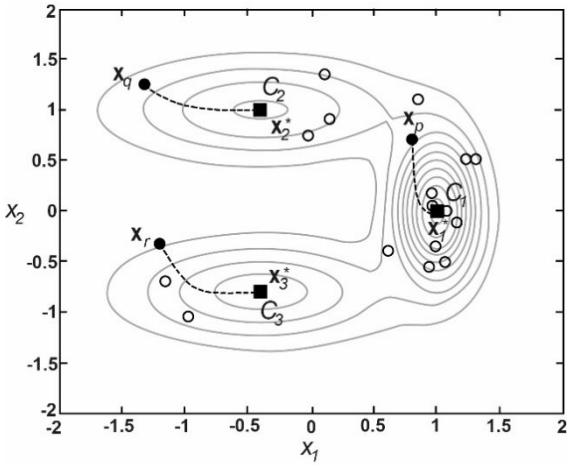


Figura 3 Ilustración del proceso de gradiente para encontrar los máximos locales.

3.2 Reducción del conjunto de datos

Para reducir el conjunto de datos, CM usa elementos representativos que contienen información limitada de toda la disponible $F(x,y)$. Bajo las condiciones anteriores, todo el conjunto de datos es dividido en dos grupos: elementos usados U y elementos no usados $\widehat{U}$. El grupo U son los elementos considerados para el procesamiento, mientras que $\widehat{U}$ representa el resto de los elementos, con las definiciones anteriores, se puede concluir que $\widehat{U}$ junto con U forman el conjunto F.

El número de elementos presentes en U es un aspecto determinante en la segmentación propuesta. Para ver el impacto de este parámetro, se a conducido un experimento de sensibilidad, en el experimento, se sabe que el tamaño de U es un porcentaje de toda la información disponible, por lo que se varía desde 1% hasta 100% mientras los otros parámetros permanecen constantes.

Los elementos U son seleccionados de manera aleatoria de la información disponible F. Por lo tanto, para minimizar los efectos estocásticos, se ejecutan 30 corridas independientes para cada tamaño de U. Los resultados sugieren que después de aumentar el porcentaje arriba del 5% el costo computacional crece exponencialmente mientras que mejora marginalmente el resultado final, se puede concluir que el mejor porcentaje para el tamaño de U es de 5%.

Al reflexionar sobre lo anterior se puede concluir que de toda la información disponible **M*N** del conjunto de datos $\boldsymbol{F}$ solo el 5% se usara para producir el conjunto $\boldsymbol{U}$. Los elementos de $\boldsymbol{U}$ serán seleccionados aleatoriamente de $\boldsymbol{F}$. Por lo tanto, el conjunto de datos $\boldsymbol{U}$ incluye vectores bidimensionales tal que $\boldsymbol{0} = \boldsymbol{int}(\boldsymbol{0.05} \cdot \boldsymbol{M} \cdot \boldsymbol{N})$, donde $\boldsymbol{int}(\cdot)$ produce un valor entero. Por otro lado, el grupo de elementos en $\widehat{\boldsymbol{U}} = \{\widehat{\boldsymbol{u}}_{\boldsymbol{i}}, \widehat{\boldsymbol{u}}_{\boldsymbol{j}}, \cdots, \widehat{\boldsymbol{u}}_{\boldsymbol{k}}\}$es integrado por elementos donde $\widehat{\boldsymbol{U}} = \{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k} \in \boldsymbol{F} | \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k} \notin \boldsymbol{U}\}$.

3.3 operación CM

Ya reducido el conjunto a $\boldsymbol{U}$, el CM se puede ejecutar. En la operación **CM**, se deben configurar dos parámetros importantes: El núcleo $\boldsymbol{h}$ y el tamaño de paso $\boldsymbol{\delta}$. Para que el CM procese la información de las características bidimensionales de $\boldsymbol{F}$, ambos elementos vectores $\boldsymbol{h}$ y $\boldsymbol{\delta}$ deben ser bidimensionales con valores para cada dimensión. Estos factores son inicializados automáticamente atreves del análisis del conjunto de datos $\boldsymbol{U}$. El proceso usado para esta configuración fue reportado en la regla de Scott [54]. Bajo esta regla el valor de ambos factores $\boldsymbol{h}$ y $\boldsymbol{\delta}$ están estrechamente relacionados con la desviación estándar $\boldsymbol{\sigma}$ de los datos $\boldsymbol{U}$. Cada dato $\boldsymbol{u_d} = [\boldsymbol{u_d^1}, \boldsymbol{u_d^2}]$ donde representa $\boldsymbol{u_d^1}$ la información en escala de grises de $\boldsymbol{u_i}$ mientras que $\boldsymbol{u_d^2}$ corresponde a su varianza. Por lo tanto, la desviación estándar para cada dimensión $\boldsymbol{\sigma_1}$ y $\boldsymbol{\sigma_2}$ son computadas con las siguientes formulas:

$$\sigma_1 = \sqrt{\frac{1}{o}\sum_{d=1}^{o}(u_d^1 - \bar{u}^1)^2} \tag{9}$$

$$\sigma_2 = \sqrt{\frac{1}{o}\sum_{d=1}^{o}(u_d^2 - \bar{u}^2)^2} \tag{10}$$

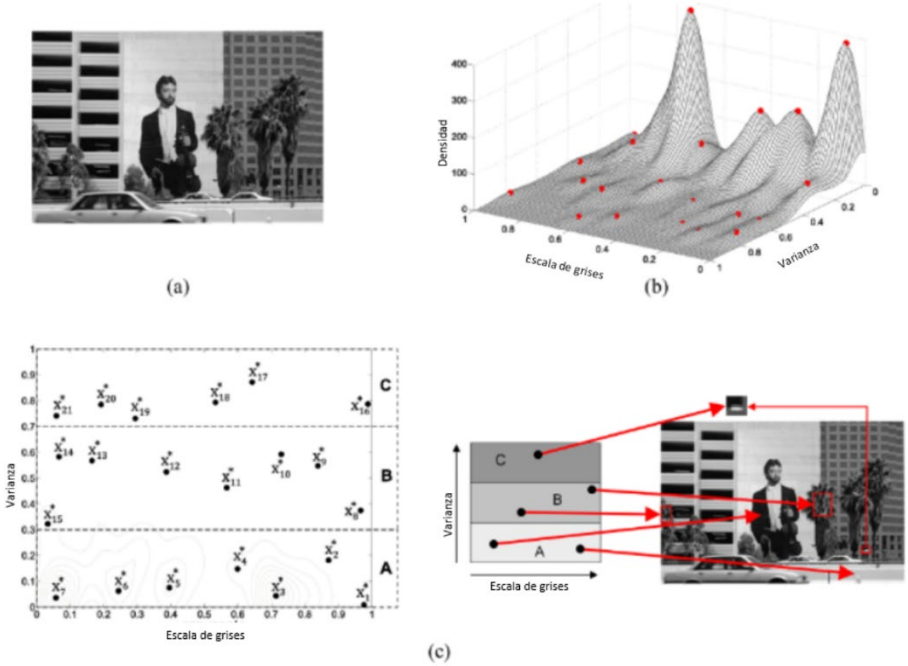
Donde $\bar{\boldsymbol{u}}^{\boldsymbol{1}}$ y $\bar{\boldsymbol{u}}^{\boldsymbol{2}}$representan el valor medio tal que $\bar{\boldsymbol{u}}^{\boldsymbol{s}} = (\boldsymbol{1}/\boldsymbol{o})\sum_{\boldsymbol{d}=\boldsymbol{1}}^{\boldsymbol{o}} \boldsymbol{u_d^s}$. Atreves de una exhaustiva experimentación, la regla de Scott [54] establece que los factores los factores están relacionados de la siguiente manera:

$$h_s = 3.5 \cdot \sigma_s \cdot o^{-\frac{1}{3}}, \qquad \delta_s = \sqrt{0.5} \cdot \sigma_s \tag{11}$$

Donde $\boldsymbol{s} \in [\boldsymbol{1}, \boldsymbol{2}]$.

3.4 Análisis de los resultados

La operación CM se genera con dos pasos. Primero, la densidad del mapa es producida, después se detecta el máximo local con la información del mapa de densidad. Por lo tanto, cuando se le



aplica este procesoal conjunto de datos $\boldsymbol{U}$, la función CM produce un mapa de densidad, un conjunto $\boldsymbol{c}$ de puntos $\{\boldsymbol{x_1^*}, \ldots, \boldsymbol{x_c^*}\}$ atractores y una lista de elementos $\boldsymbol{x_j} = (\boldsymbol{u_a}, \ldots, \boldsymbol{u_g})$ para cada atractor $\boldsymbol{x_j^*}$donde enlista los puntos atractores como $\boldsymbol{u_a, u_g} \in \boldsymbol{C_i}$.

Figura 4 Análisis del efecto de los puntos de atracción. (a) Imagen original, (b) mapa de densidad producido con el conjunto de datos reducido, (c) división del mapa de densidad y sus características de correspondencia en la imagen.

En la figura 4 se puede observar una imagen y como esta se segmentará. Esta imagen tiene un tamaño de 214 X 320. De esta imagen, se computa un conjunto de características $\boldsymbol{F}$ integradas por valores de escala de grises y varianza de cada pixel. Entonces, un conjunto de 3,424 elementos $(\boldsymbol{o = int(0.05 \cdot 214 \cdot 320)})$ son aleatoriamente seleccionados de $\boldsymbol{F}$ para producir el conjunto reducido $\boldsymbol{U}$. para usar los puntos $(u_1, \ldots, u_o)$ de U y posteriormente ejecutar el algoritmo CM. Como resultado, CM produce la información mostrada en la ilustración 2 (b). La ilustración 2 también representa el mapa de densidad, así como sus puntos de atracción dibujados en rojo. Acorde con la información, el mapa de densidad puede dividirse en 3 diferentes regiones: área A, B, C ilustración 2 (c).

Los puntos de atracción localizados en el área A son caracterizados por tener una varianza pequeña, bajo estas condiciones, se pueden representar en regiones homogéneas de escala de grises en una imagen. Por otro lado, el área B representa objetos texturizados o detalles de la imagen tal como bordes. Finalmente, el área C corresponde en general a ruido no deseado. El uso del espacio bidimensional permite que el método propuesto obtenga un mayor desempeño en los resultados al detectar mejor los detalles e incorporarlos al resultado final.

3.5 Generalización de los resultados del CM

Una vez obtenidos los grupos los resultados del CM se pueden reducir a un conjunto de datos $\boldsymbol{U}$, los resultados parciales deben de generalizar la información no incluida $\widehat{\boldsymbol{U}}$. En este enfoque, para cada uno de los elementos $\widehat{\boldsymbol{u}}_{\boldsymbol{k}} \in \widehat{\boldsymbol{U}}$, es asignado al mismo grupo $\boldsymbol{C}_{\boldsymbol{i}}$ o en su defecto al grupo mas cercano $\boldsymbol{u}_{\boldsymbol{a}} \in \boldsymbol{X}$, asumiendo que $\boldsymbol{u}_{\boldsymbol{a}}$ pertenece a:

$$u_a = \arg\min d\left(\hat{u}_k, u_j\right), \qquad \hat{u}_k \in \hat{U}, \quad u_a \in X_i \wedge u_a \in U \tag{12}$$

Donde $\boldsymbol{d}\left(\widehat{\boldsymbol{u}}_{\boldsymbol{k}}, \boldsymbol{u}_{\boldsymbol{j}}\right)$ representa la distancia euclidiana entre $\widehat{\boldsymbol{u}}_{\boldsymbol{k}}$ y $\boldsymbol{u}_{\boldsymbol{j}}$.

Una vez asignado los elementos no usados $\widehat{\boldsymbol{U}}$ a su respectivo grupo $\boldsymbol{C}_{\boldsymbol{i}}$ ($\boldsymbol{i} \in \boldsymbol{1}, \ldots, \boldsymbol{c}$), todas las listas de los grupos incrementaran sus números de elementos al incluir los elementos no usados $\widehat{\boldsymbol{U}}$. Por lo tanto, la nueva lista de grupos $\boldsymbol{X}_{\boldsymbol{i}}^{\boldsymbol{new}}$ son actualizadas como $\boldsymbol{X}_{\boldsymbol{i}}^{\boldsymbol{new}} = \boldsymbol{X}_{\boldsymbol{i}} \cup \left\{\widehat{\boldsymbol{u}}_{\boldsymbol{k}} \in \widehat{\boldsymbol{U}} \wedge \widehat{\boldsymbol{u}}_{\boldsymbol{k}} \in \boldsymbol{C}_{\boldsymbol{i}}\right\}$.

3.6 Generalización de los resultados del CM

con la operación CM muchos grupos son producidos. Sin embargo, no todos son visualmente predominantes. Por esta razón, es necesario combinar aquellos que son irrelevantes en terminaos del número de elementos contenidos. La idea principal es integrar grupos con poca población con otros grupos que presentan características similares para preservar la información visual. Por lo tanto, la idea principal es mantener los grupos con suficiente concentración mientras que los demás son fusionados.

El problema de determinar el número de elementos que caracterizan un clúster predominante es similar a detectar el punto de quiebre en ingeniería de sistemas. La ubicación del punto de quiebre representa el "punto de decisión correcto" en el cual el valor relativo de una variable ya no es significativo en términos de su contribución final. La detección mediante el punto de quiebre se aplica porque requiere un costo computacional muy reducido. En esta investigación, se experimentó con varios enfoques para discriminar los grupos producidos según su importancia. Algunos de ellos incluyen el análisis de componentes principales, la teoría del diseño de experimentos, etc. Sin embargo, estos métodos aumentan el costo computacional sin una mejora significativa en comparación con el punto rodilla. El enfoque principal de CAM-SEG es reducir el costo computacional, por eso el método de punto rodilla para separar grupos relevantes de grupos irrelevantes, es la mejor opción. Bajo este esquema, un conjunto de c pares es producidos. Cada par $\boldsymbol{V}_{\boldsymbol{l}} = (\boldsymbol{v}_{\boldsymbol{1}}^{\boldsymbol{x}}, \boldsymbol{v}_{\boldsymbol{2}}^{\boldsymbol{x}})$ es asociado a la siguiente información:

$$l = g(|X_k^{New}|); \quad v_l^x = \frac{g(|X_k^{New}|)}{c}; \quad v_l^x = \frac{|X_k^{New}|}{\sum_{i=1}^{c}|X_k^{New}|}, \ k \in 1, \ldots, c; \tag{13}$$

Donde $\left|\boldsymbol{X_k^{New}}\right|$ representa el número de elementos contenidos en el grupo $\boldsymbol{c_k}$. $\boldsymbol{g}\left(\left|\boldsymbol{X_k^{New}}\right|\right)$ es una función que renquea $\left|\boldsymbol{X_k^{New}}\right|$ dependiendo de su numero de elementos. Por lo tanto, el número uno $\boldsymbol{g}\left(\left|\boldsymbol{X_k^{New}}\right|\right) = \boldsymbol{1}$ es producido cuando el grupo $\boldsymbol{c_b}$ tienen el máximo número de elementos. Por otro lado, el índice más grande $\boldsymbol{g}\left(\left|\boldsymbol{X_k^{New}}\right|\right) = \boldsymbol{c}$ es asignado al grupo más pequeño $\boldsymbol{c_s}$.

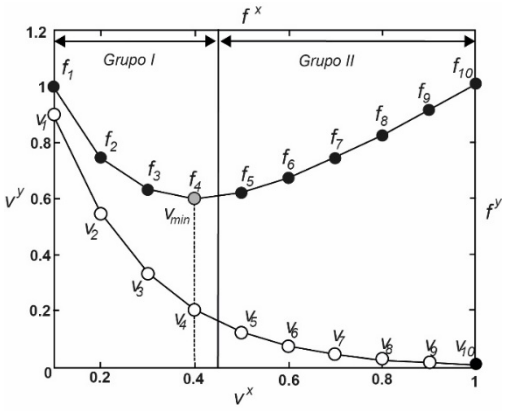


Figura 5. Detección del punto de inflexión para dividir los grupos en relevantes e irrelevantes.

La Figura 5 representa la información contenida por el conjunto de pares. Considerando la distribución $\boldsymbol{v_l}$ $(\boldsymbol{l} \in \boldsymbol{1}, \ldots, \boldsymbol{c})$, se genera una función de costo $\boldsymbol{f_l} = (\boldsymbol{f_l^x}, \boldsymbol{f_l^y})$ de $\boldsymbol{v_l}$, que relaciona el incremento $\boldsymbol{v_l^y}$respecto a $\boldsymbol{v_l^x}$. Esta función es calculada como:

$$f_l^x = v_l^x; \quad f_l^y = v_l^y + v_l^x; \tag{14}$$

Una característica importante de la función $\boldsymbol{f_l^y}$es que posee solo un mínimo global $\boldsymbol{v_{min}}$. Este valor corresponde al punto rodilla, el cual se define como:

$$v_{min} = \arg min\ f_l^y; \quad 0 \leq l \leq c; \tag{15}$$

En este método el punto $\boldsymbol{v_{min}}$ divide todos los grupos obtenidos por CM en los dos diferentes grupos (relevante e irrelevantes) plasmados en la Figura 5. Los grupos relevantes representan los pixeles predominantes en la segmentación de la imagen. Los grupos irrelevantes serán combinados con sus vecinos debido a su población reducida. Por lo tanto, el conjunto de grupos $\boldsymbol{c_a}, \ldots, \boldsymbol{c_d}$ es relevante, por consecuente $\boldsymbol{c_a}, \ldots, \boldsymbol{c_d}$ ∈Grupo I. Por otro lado, el conjunto de grupos $\boldsymbol{c_q}, \ldots, \boldsymbol{c_t}$ no es relevante, por consiguiente $\boldsymbol{c_q}, \ldots, \boldsymbol{c_t}$ ∈Grupo II. Grupos irrelevantes se combinan mientras que los relevantes prevalecen. En la combinación, los pixeles de cada grupo $\boldsymbol{c_w}$; $\boldsymbol{w} \in \boldsymbol{Grupo\ II}$ son integrados con los grupos $\boldsymbol{c_z}$; $\boldsymbol{z} \in \boldsymbol{Grupo\ I}$, donde $\boldsymbol{c_z}$ corresponde al *Grupo I* y la distancia entre los grupos $\left|\left|\boldsymbol{X_w^*} - \boldsymbol{X_z^*}\right|\right|$es la mínima posible, por lo tanto, estos dos grupos se fusionan. La operación final consiste en cada grupo $\boldsymbol{X_q^*}$ incluye 2 dimensiones una para la escale de grises $\boldsymbol{X_{q1}^*}$ y otra para la varianza $\boldsymbol{X_{q2}^*}$, todos los elementos de c_q son asignados a la misma escala de grises dada por $\boldsymbol{X_{q1}^*}$.

Los grupos irrelevantes contienen tan pocos elementos que son irreconocibles en la imagen a simple vista. En general, la mayoría de los grupos irrelevantes están ubicados cerca de un grupo representativo como consecuencia de diferentes factores de la imagen, como la iluminación, el ruido, etc. Para construir un grupo representativo a partir de varios grupos irrelevantes, es necesario fusionar grupos irrelevantes con características muy diferentes, lo que produce resultados inexactos en la imagen segmentada final.

3.7 Procedimiento computacional

El método propuesto fue diseñado en base a un procesamiento secuencial. El procedimiento se puede resumir en forma de seudo código en Algoritmo 1. Este esquema considera como entrada

una imagen con tamaño M x N. como primer paso, es computar la varianza local para cada pixel $\boldsymbol{V(x,y)}$ generada usando una ventana de 3 x 3. Se continúa, construyendo el espacio bidimensional de características agrupando las dos características (línea 3) escala de grises y varianza $\boldsymbol{F(x,y) = [I(x,y), V(x,y)]}$. Procediendo con el algoritmo, se selecciona el grupo reducido $\boldsymbol{U}$ (línea 5) a partir de la población $\boldsymbol{F}$, al hacer esto, se diferencian los dos conjuntos $\boldsymbol{U}$ y $\boldsymbol{\hat{U}}$, al diferenciar los dos conjuntos se puede procesar la desviación estándar $\boldsymbol{\sigma_1}$ y $\boldsymbol{\sigma_2}$ de los grupos relevantes $\boldsymbol{U}$ (línea 6). Considerando ambas desviaciones estándar, se calculan los parámetros $\boldsymbol{h_1, h_2, \delta_1}$ y $\boldsymbol{\delta_2}$ (línea 7) para ejecutar el algoritmo (línea 8) y obtener el conjunto $\boldsymbol{c}$ de puntos atractores $\{\boldsymbol{x_1^*, \dots, x_c^*}\}$ y sus respectivas listas $\boldsymbol{X_i = (u_a, \dots, u_g)}$ para cada punto de atracción $\boldsymbol{x_i^*}$ donde se enlista los puntos de datos atraídos por el ($\boldsymbol{u_a, u_g \in C_i}$). Ya generala la agrupación del algoritmo, los resultados parciales deben ser generalizados para incluir la información irrelevante $\boldsymbol{\hat{U}}$ (líneas 9-13). En esta operación, a cada elemento no utilizado $\boldsymbol{u_k \in \hat{U}}$ se le asigna el mismo clúster $\boldsymbol{C_i}$ del dato utilizado más cercano $\boldsymbol{u_a}$. Luego, se produce una función de costo $\boldsymbol{f}$, que se relaciona con el incremento del número de elementos con respecto a su importancia (línea 14). Con la información de $\boldsymbol{f}$, se detecta el punto de quiebre $\boldsymbol{v_{min}}$ (línea 15). Una vez calculado $\boldsymbol{v_{min}}$, se clasifica el número total de clústeres (línea 16) en clústeres relevantes (Grupo I) y clústeres irrelevantes (Grupo II). Posteriormente, en el proceso de combinación (líneas 17-20), los elementos de píxeles de cada clúster $\boldsymbol{C_j}$ ($\boldsymbol{j \in Grupo\ II}$) se integran con los elementos del clúster $\boldsymbol{C_a}$ ($\boldsymbol{a \in Grupo\ I}$), donde $\boldsymbol{C_a}$ corresponde al clúster del Grupo I de manera que la distancia entre sus prototipos de clúster $\boldsymbol{\left|\left|X_j^* - X_a^*\right|\right|}$ sea lo más mínima posible. Por lo tanto, los elementos de $\boldsymbol{C_a}$ también incluirán los elementos de $\boldsymbol{C_j}$ ($\boldsymbol{C_a = C_a \cap C_j}$). Finalmente, la salida del algoritmo es el clúster acumulado final del Grupo I.

Algoritmo 1

Seudo-código para el CAM-SEG propuesto

1 $input: I(x,y) de\ tamaño\ MxN$
2 $V(x,y) \leftarrow CalcularVarianza(I(x,y))$
3 $F(x,y) \leftarrow Construirlascaracterísticasdelespacio(I(x,y), V(x,y);$
4 $o \leftarrow int(0.05MxN)$
5 $[U, \hat{U}] \leftarrow Selecciódecaracterísticasparareducciondedataset(o)$
6 $[\sigma_1, \sigma_2] \leftarrow Calculardesviaciónestandard(U)$
7 $[h_1, h_2, \delta_1, \delta_2] \leftarrow ObtenerParametrosCM(\sigma_1, \sigma_2)$
8 $[(x_1^*, \dots, x_c^*), \{X_1, \dots, X_c\}] \leftarrow correrCM(U, h_1, h_2, \delta_1, \delta_2)$
9 $para\ cada\ \hat{u}_k \in \hat{U}$
10 $u_a = argmind(\hat{u}_k, u_j), u_a \in X_i \wedge u_a \in U$
11 $si\ U_a \in X_i entonces\ \hat{u}_k \in C_i$
12 $X_i^{New} \leftarrow X_i \cup \{\hat{u}_k\}$
13 $fin\ del\ para$
14 $f \leftarrow obtenerfunciónobjetivo(X_1^{New}, \dots, X_2^{New})$
15 $v_{min} \leftarrow obtenerpuntoderodilla(f)$
16 $[Grupo\ I, Grupo\ II] \leftarrow Divide\ Grupos(v_{min}, X_1^{New}, \dots, X_c^{New})$
17 $Para\ cada\ C_i \in Group\ II$
18 $$C_a = argmin\ d(x_i^*, x_w^*), c \in Grupo\ I$$
19 $$C_a \leftarrow C_a \cup C_j$$
20 $fin\ del\ para$
21 $Salida: \{C_q, C_r, \dots, C_t\} \in Grupo\ I$

4 Resultados experimentales

Con el propósito de ilustrar los resultados obtenidos con este algoritmo se utilizó un set de 300 imágenes para comparar con los resultados de otros algoritmos obtenido de Berkeley Segmentation Dataset and Benchmark (BSDS300), este data set fue seleccionado con el propósito de tener mayor compatibilidad con otros métodos. Además de contar con las imágenes, el dataset tiene un dataset con las imágenes llamadas ground truth, las cuales son imágenes segmentadas por humanos expertos en estos temas.

En los experimentos el algoritmo CAM-SEG fue comparado con los algoritmos Fast and Robust Fuzzy C-Means Clustering (FRFCM), 2D histogram and exponential Kbest gravitational search (EKGSA) y Electromagnetics-like Optimization segmentation (EMO),. En la comparación de estos algoritmos, se utilizó la mejor configuración para cada algoritmo según se especifica en su referencia, todos los experimentos fueron procesados en MATLAB en una PC con un Rysen 9 12-core a 3.7 GHz y una memoria RAM de 32 GB.

Las métricas utilizadas para evaluar el procesamiento de cada algoritmo en este análisis fueron: Structural similarity index measurement (SSIM), Feature Similarity index (FSIM), Root mean squared error (RCME), Peak Signal to Noise Ratio (PSNR), Normalized Cross-Correlation (NK), Average Difference (AD), Structural Content (SC), Maximum Difference (MD), Normalized Absolute Error (NAE), Visual Perception (VP) and Computational Cost (CC). Las primeras 9 métricas son para evaluar la calidad de la segmentación en la imagen resultante I procesada de cada algoritmo, contrastada con la imagen R la cual será su respectiva imagen segmentada por humanos o también llamada ground truth antes del procesamiento.

The Structural similarity index measurement (SSIM) evalúa la similitud entre la imagen segmentada a su referencia (ground truth). Asumiendo que I es una imagen de las mismas dimensiones que su referencia, SSIM es calculado con la siguiente formula:

$$SSIM = \frac{(2u_I u_R + Q_1)(2\sigma_{IR} + Q_2)}{(u_I^2 + u_R^2 + Q_1)(\sigma_I^2 + \sigma_R^2 + Q_2)} \qquad (16)$$

Donde $\boldsymbol{Q_1}$ y $\boldsymbol{Q_2}$ simbolizan dos constantes positivas (típicamente 0.01). $\boldsymbol{u_I}$ y $\boldsymbol{u_R}$ corresponden al valor promedio de la imagen segmentada y la de referencia respectivamente. $\boldsymbol{\sigma_I}$ y $\boldsymbol{\sigma_R}$ son las variancias de la imagen segmentada y su respectivo ground truth respectivamente y $\boldsymbol{\sigma_{IR}}$ es la covarianza de entre la imagen procesada y el ground truth el valor SSIM varia entre 0 y 1 mas cerca de 1 representa una mejor similitud.

Feature Similarity index (FSIM) es la comparación basada en índices que contrasta las estructuras locales mediante la extracción de la fase de congruencia (PC) y el valor de los gradientes (GM). El mapa de similitud S_{pc} se calcula con la siguiente formula:

$$S_{pc}(\mathrm{I},\mathrm{R}) = \frac{2 * I * R + T_1}{I^2 * R^2 + T_1} \qquad (17)$$

Donde $\boldsymbol{T_1}$ representa un valor constante positivo (típicamente 0.01). El gradiente de magnitud es se calcula utilizando las mascaras convoluciones GI y GR obtenidas de I y R respectivamente. Así, el mapa de gradiente es calculado de la siguiente manera:

$$S_{GM}(\mathrm{GI},\mathrm{GR}) = \frac{2 * GI * GR + T_1}{GI^2 * GR^2 + T_1} \qquad (18)$$

Al calcular $\boldsymbol{S_{pc}}$ y $\boldsymbol{S_{GM}}$ es posible calcular el FSIM sustituyendo valores en la siguiente formula:

$$FSIM = \frac{\sum_{j=1}^{2} S_{pc} * S_{GM} * I_j}{GI^2 * GR^2 + T_1} \tag{19}$$

Un valor más alto de FSIM representa una mejor calidad en el procesamiento de la imagen.

El valor RCME también evalúa la similitud substrayendo los valores del ground truth en la procesada para después obtener la raíz del promedio del error total, en esta ocasión un valor menor en el error representa una mejor similitud y se calcula de la siguiente manera:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - R(i,j))^2}{MN}} \tag{20}$$

Donde I es la imagen procesada, R es la orignal, M y N son las dimensiones de las imágenes.

Peak Signal to Noise Ratio (PSNR) relaciona el valor máximo posible de un píxel de una imagen MAXI con su respectiva similitud en términos de la magnitud del PSNR. Un valor de PSNR más bajo implica un mejor rendimiento. Bajo estas condiciones, el PSNR se calcula de la siguiente manera:

$$PSNR = 20log_{10}\left(\frac{MAXI}{RMSE}\right) \tag{21}$$

Normalized Cross-Correlation (NK) evalúa la coincidencia entre la imagen segmentada y la imagen de referencia en términos de su correlación. Un valor alto de NK representa un mejor procesamiento y se calcula de la siguiente manera:

$$NK = \frac{\sum_{i=1}^{M}\sum_{i=1}^{N} I(i,j) \cdot R(i,j)}{\sum_{i=1}^{M}\sum_{i=1}^{N} R(i,j)} \tag{22}$$

The Average Difference (AD) evalúa la diferencia entre la imagen segmentada y la de referencia simplemente obteniendo el promedio de la diferencia entra la procesada menos su respectivo ground truth, el valor absoluto de lo anterior fija un valor que diferencia la imagen de referencia respecto a la procesada, por lo tanto, un valor más cerca del 0 representa una menor diferencia y mejor procesamiento, se obtiene de la siguiente manera:

$$AD = \left|\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j) - R(i,j)}{MN}\right| \tag{23}$$

The Structural Content (SC) evalúa similitud mediante la autocorrelación indexada. Un menor valor implica un mejor procesamiento. Se calcula con la siguiente formula:

$$SC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j)^2}{\sum_{i=1}^{M}\sum_{j=1}^{N} R(i,j)^2} \tag{24}$$

Maximun Difference (MD) mide la máxima disparidad entre una imagen procesada y la imagen de referencia. Un valor menor de MD representa una mejor segmentación, calculándose de la siguiente manera:

$$MD = MAX|I(i,j) - R(i,j)|;\ \forall i,j \in M \times N \tag{25}$$

The Normalized Absolute Error (NAE) evalúa la similitud entre las imágenes considerando el valor absoluto de los errores normalizados producido entre las imágenes, un valor menor implica un mejor resultado, calculándose de la siguiente manera:

$$NAE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|I(i,j) - R(i,j)|}{\sum_{i=1}^{M}\sum_{j=1}^{N}R(i,j)} \tag{26}$$

Visual Perception (VP) provee una manera cuantitativa de una imagen en términos de percepción humana. Aunque aun es tema de estudio, algunos estudios [63-64] indican que la percepción humana esta relacionado con las diferencias de los detalles percibidos por un espectador. VP considera las diferencias entre las principales características de los resultados de una imagen. Para evaluar VP de una imagen I, el modelo adoptado involucra la intensidad de los bordes de los pixeles, el numero de bordes existentes y la entropía completa de la imagen. Esta información se procesa de la siguiente manera:

$$VP = \log\big(log\big(E(I)\big)\big) \cdot \frac{NE(I)}{M \times N} \cdot H(I) \tag{27}$$

Donde $E(I)$ representa la suma de todas las intensidades de los bordes obtenidos por el filtro de sobel. $NE(I)$ corresponde a el numero de bordes obtenido del procesamiento. $H(I)$ simboliza la entropía de la imagen. Un valor más alto de VP reprecenta más información de la imagen y por tanto un mejor procesamiento. $H(I)$ es calculado de la siguiente manera:

$$H(I) = \sum_{i=0}^{L-1} P(q) \cdot \log\left(P(q)\right) \tag{28}$$

Donde $P(q)$ se refiere a la función de densidad de probabilidad en el nivel de intensidad $P\ (q)\ \in 0, \dots, l-1)$ de la imagen. Es evidente hasta cierto punto las imágenes son visualmente aceptables por el ojo humano, sin embargo, después de cierto umbral la imagen se vuelve muy compleja para su visualización, algo que no toma en cuenta esta fórmula.

Los métodos de segmentación considerados en las comparaciones son, en general, procesos complejos con varias operaciones aleatorias y subrutinas estocásticas. Por lo tanto, es impráctico realizar un análisis de complejidad desde un punto de vista determinista. Por esta razón, se evalúa el Costo Computacional (CC) considerando el esfuerzo computacional invertido para cada algoritmo. El CC se evalúa, asumiendo el esfuerzo computacional promediado en segundos de 30 ejecuciones independientes. Un valor de CC bajo significa un mejor rendimiento.

Tabla 1
Comparación de características del método propuesto CAM-SEG con otros métodos considerando la dataset de Berkleley.

Métricas	FRFCM	EKGSA	EMO	**CAM-SEG**
SSIM	0.6621	0.6087	0.5901	0.7981
FSIM	0.7821	0.7897	0.7697	0.8025
RMSE	50.5218	52.0279	54.0974	48.05110
PSNR	22.3987	23.1187	24.9874	20.4410
NK	0.9875	0.8821	0.7889	1.02174
AD	42.8714	43.1101	44.0124	42.1621
SC	1.3781	1.2871	1.2721	1.0874
MD	110.9871	111.1498	112.0074	110.2187
NAE	0.1674	0.1724	0.1821	0.1081
VP	0.2887	0.3887	0.4150	0.6825
CC	48.97	84.87	108.21	27.14

4.1 Comparación

Esta subsección muestra la comparación de rendimiento de los diferentes esquemas de segmentación cuando se utilizan en las 300 imágenes del conjunto de datos Berkeley. El objetivo es determinar, en términos generales, la eficiencia y efectividad de cada método. La Tabla 1 muestra los valores promedio de cada índice considerado sobre el conjunto de 300 imágenes de BSDS300. Para el funcionamiento de los métodos EKGSA y EMO, es necesario definir de antemano el número de clases en las cuales se segmentarán las imágenes. Por lo tanto, en las comparaciones, el número de clases se fija en cuatro niveles. Este número se ha decidido para mantener la compatibilidad con los resultados obtenidos para tales esquemas en sus respectivas referencias [36,37].

Inspeccionando la tabla 1, se observa que el método propuesto presenta mejores valores promedios entre los competidores. Esto da a entender que el algoritmo CAM-SEG conlleva la mejor semejanza entre los resultados segmentados y sus correspondientes referencias (ground truth). El algoritmo FRFCM presenta la segunda mejor segmentación mientras que el EKGSA y EMO producen los peores resultados. Por otro lado, la propuesta CAM-SEG ofrece el mejor valor en cuanto a percepción visual (VP), seguida de cerca por EKGSA y EMO, mientras que FRFCM presenta la peor evaluación de percepción visual. De acuerdo con estos valores, queda claro que la propuesta CAM-SEG mantiene más del 50% de la VP considerando su competidor más cercano (EMO). También es importante destacar que la propuesta CAM-SEG mantiene el menor esfuerzo computacional en comparación con otras técnicas de segmentación. Para el costo computacional, se puede decir que el algoritmo propuesto es cerca de dos veces más rápido que el FRFCM, aproximadamente tres veces más rápido que el EKGSA y cerca de cuatro veces más rápido que el EMO.

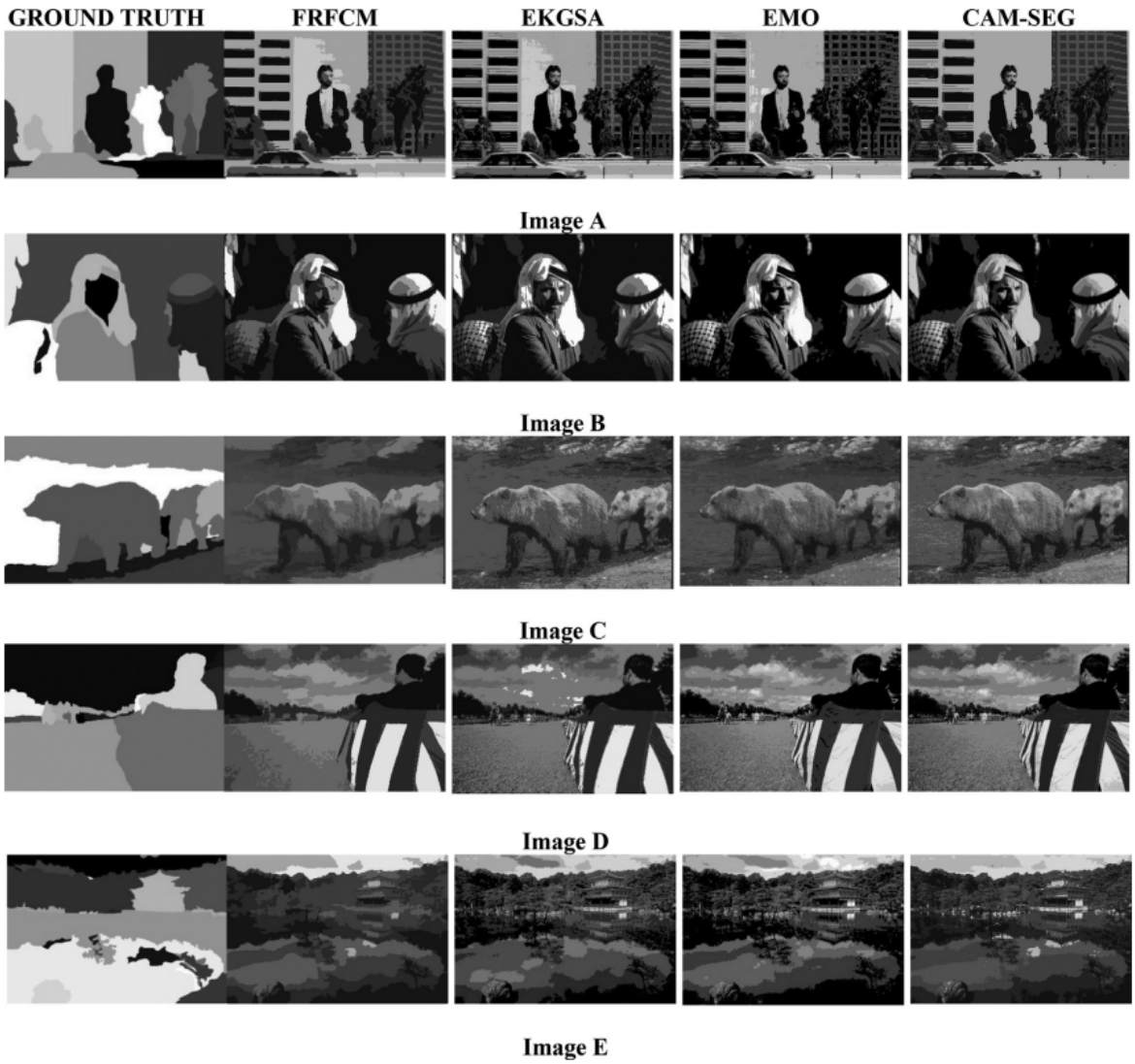


Figura 6. Resultado de la segmentación visual de un conjunto representativo de imágenes del conjunto de datos de Berkeley.

Con objeto de visualización de los resultados se utilizaron 6 imágenes para ilustrar los resultados obtenidos. Estas imágenes se obtuvieron de un conjunto de datos llamado BSDS300 generado por la universidad de Berkeley. Estas imágenes fueron seleccionadas por su complejidad al momento de generar la segmentación. Se puede observar que el método propuesto presenta una mejor percepción visual en comparación con los otros competidores. En todas las imágenes, detalles como textura o pequeños elementos son claramente mostrados por CAM-SEG, mientras que los otros algoritmos eliminan detalles interesantes. Este rendimiento notable es consecuencia de su mecanismo robusto para detectar detalles a partir del mapa de características generado por el algoritmo CM.

La tabla 2 presenta el resumen de los resultados de las imágenes seleccionadas. La tabla reporta las métricas anteriormente descritas.

Tabla 2
Métricas de características para el conjunto representativo de las imágenes presentadas en la Figura 6.

Imagen	Métricas	FRFCM	EKGSA	EMO	**CAM-SEG**
Imagen A	SSIM	0.7093	0.7030	0.6103	0.7630
	FSIM	0.7539	0.6950	0.7788	0.8364
	RMSE	24.9397	19.2821	32.2896	23.0181
	PSNR	20.1929	22.4277	17.9495	20.8893
	NK	0.8520	0.9327	0.8103	0.9746
	AD	31.1213	37.254	36.9640	19.7554
	SC	1.3461	1.1112	1.4582	1.0282
	MD	80	70	89	65
	NAE	0.2024	0.1391	0.2585	0.1207
	VP	0.2170	0.4954	0.4087	0.6414
	CC	42.14	80.12	100.12	26.21
Imagen B	SSIM	0.4531	0.4381	0.5901	0.8883
	FSIM	0.7703	0.8553	0.7697	0.8551
	RMSE	22.2047	22.7306	54.0974	12.8758
	PSNR	22.9871	20.9985	24.9874	25.9352
	NK	0.9505	0.8073	0.7889	0.9669
	AD	32.1436	39.5537	44.0124	25.7359
	SC	1.0733	1.4834	1.2721	1.0493
	MD	150	56	112.0074	111
	NAE	0.1654	0.3115	0.1821	0.1129
	VP	0.3010	0.4971	0.4150	0.7121
	CC	39.14	81.14	108.21	25.14
Imagen C	SSIM	0.6411	0.7652	0.3557	0.8030
	FSIM	0.6934	0.8473	0.8047	0.8667
	RMSE	13.2809	18.3313	24.5442	10.6946
	PSNR	25.6662	22.8669	20.3317	27.5474
	NK	0.9600	0.8521	0.8491	0.9921
	AD	41.9045	45.1565	35.9246	38.9610
	SC	1.0669	1.3569	1.3000	1.0045
	MD	41	45	59	74
	NAE	0.0941	0.1594	0.1339	0.0757
	VP	0.4107	0.6921	0.5472	0.8721
	CC	37.14	87.97	97.24	26.87
Imagen D	SSIM	0.6597	0.7702	0.7521	0.7776
	FSIM	0.7406	0.7733	0.8417	0.7985
	RMSE	17.7907	17.7575	19.0758	17.3683
	PSNR	23.1269	23.1413	22.5211	23.3356
	NK	0.9654	0.8931	0.8498	0.9734
	AD	41.8971	44.6256	46.2638	31.7830
	SC	1.0517	1.2406	1.3602	1.0354

	MD	154	36	44	135
	NAE	0.0972	0.1403	0.1710	0.0978
	VP	0.3802	0.5901	0.6014	0.8511
	CC	40.14	90.01	89.21	28.57
Imagen E	SSIM	0.6286	0.6875	0.7190	0.8015
	FSIM	0.6754	0.8233	0.7364	0.8242
	RMSE	17.8809	22.9722	20.0168	14.0111
	PSNR	23.0829	20.9067	22.1028	25.2013
	NK	0.9595	0.8251	0.8796	0.9832
	AD	41.1631	49.7261	46.9961	37.3525
	SC	1.0521	1.4242	1.2748	1.0144
	MD	121	54	46	50
	NAE	0.1424	0.2320	0.1526	0.1201
	VP	0.5109	0.8001	0.6587	0.9127
	CC	49.01	110.61	114.07	30.17

En la tabla 2, se puede observar la mejora basada en las métricas de segmentación mediante el método propuesto. En general EKGSA y EMO tiene un mal desempeño, Esto dado que se ejecutaron variar veces para producir resultados consistentes. En el caso del FRFCM, mantiene un buen procesamiento desde la perspectiva numérica. Sin embargo, la percepción visual respecto a las mejores métricas que obtiene no es congruentes, observándose valores en las métricas altos mientras que la imagen presenta una mala calidad de segmentación. A diferencia de FRFCM, la propuesta CAM-SEG obtiene los mejores valores de VP. Este comportamiento notable es consecuencia de su capacidad para identificar grupos de píxeles que representan detalles en las imágenes. Por otro lado, los métodos EKGSA y EMO alcanzan mejores valores de VP que FRFCM, pero peores que CAM-SEG. Estos resultados son producto de sus principios de funcionamiento.

Una inspección detallada de la Tabla 2 muestra que el método propuesto CAM-SEG presenta valores deficientes en algunos índices como MD o PSNR en algunas imágenes. El algoritmo propuesto presenta una notable capacidad para identificar regiones homogéneas que representan detalles finos en las imágenes. En tales circunstancias, dado que las imágenes de referencia discriminan principalmente información semántica (como un oso, una cara, una persona, etc.), en algunas regiones (secciones con detalles) podrían aparecer grandes diferencias entre la imagen segmentada de CAM-SEG y su respectiva imagen de referencia. Aunque estas diferencias producen resultados numéricos deficientes en algunos índices de rendimiento, aumentan la percepción visual (VP) de la imagen segmentada.

En la tabla 2, al observar el costo computacional de los algoritmos, está absolutamente claro que el método propuesto mantiene el valor mas bajo y por tanto un menor costo computacional. Aunque FRFCM compite con el algoritmo propuesto mejor que EKGSA y EMO, sus resultados son peores que los del método CAM-SEG, acorde con los resultados se puede concluir que el método propuesto presente el mejor balance entre calidad de segmentación y velocidad.

5. Conclusiones

En este artículo, se presenta un nuevo algoritmo de segmentación competitivo para imágenes en escala de grises. El enfoque propuesto basado en el algoritmo de desplazamiento medio considera un mapa de características bidimensional que incluye el valor de escala de grises y la varianza local para cada píxel de la imagen.

Para reducir el costo computacional, el algoritmo de cambio medio (MS) se modifica para operar con un número muy limitado de puntos de todos los datos disponibles. En tales condiciones, se diferencian dos conjuntos de elementos: elementos usados (el conjunto de datos reducido considerado en la operación MS) y elementos no utilizados (el resto de datos disponibles).

A diferencia del algoritmo MS clásico que emplea funciones gaussianas, en nuestro enfoque, el proceso para estimar los valores de densidad en el mapa de características utiliza la función del núcleo de Epanechnikov. Este núcleo presenta la mejor precisión de estimación cuando la cantidad de datos para calcular la densidad del mapa de características es muy limitada [39].

Una vez obtenidos los resultados del MS, se generalizan para incluir los datos no implicados. Por lo tanto, cada elemento no utilizado se asigna al mismo grupo de datos utilizados más cercanos. Finalmente, los clusters con la menor cantidad de elementos se fusionan con otros clusters vecinos.

En los experimentos, hemos aplicado el algoritmo CAM-SEG a las 300 imágenes del conjunto de datos de Berkeley, y sus resultados se comparan con los producidos por el algoritmo de agrupación de medias C difusa rápido y robusto (FRFCM) [53], el histograma 2D y el algoritmo de búsqueda gravitacional exponencial Kbest (EKGSA)[24]y la segmentación de optimización similar a la electromagnética (EMO) [54]. El primer método, FRFCM, es un enfoque de agrupación de última generación para fines de segmentación, mientras que EKGSA y EMO representan dos algoritmos de segmentación representativos basados en esquemas metaheurísticos. En nuestro análisis se han considerado once índices de desempeño: Medición del índice de similitud estructural (SSIM), índice de similitud de características (FSIM), raíz del error cuadrático medio (RMSE), Relación pico señal-ruido (PSNR), correlación cruzada normalizada (NK), diferencia promedio (ANUNCIO), Contenido Estructural (CAROLINA DEL SUR), diferencia máxima (Maryland), error absoluto normalizado (NAE), Percepción visual (vicepresidente) y costo computacional (CC). Los resultados experimentales confirman que el esquema propuesto produce imágenes segmentadas con una calidad de percepción visual un 50% mejor aproximadamente dos veces (≈1,8 − 2) más rápido que sus competidores.

Referencias

[1] K.S. Tan, N.A.M Isa, Color image segmentation using histogram thresholding– fuzzy c-means hybrid approach, Pattern Recognit. 44 (2011) 1–15.

[2] H.-D. Cheng, X. Jiang, J. Wang, 2002. Color image segmentation based on homogram thresholding and region merging, Pattern Recognit 35 (2002) 373–393.

[3] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2002) 888–905.

[4] Y. Tian, J. Li, S. Yu, T. Huang, Learning complementary saliency priors for foreground object segmentation in complex scenes, Int. J. Comput. Vis. 111 (2015) 153–170.

[5] P.F. Felzenszwalb, D.P Huttenlocher, Efficient graph-based image segmentation, Int. J. Comput. Vis. 59 (2004) 167–181.

[6] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 898–916.

[7] X. Zhang, C. Xu, M. Li, X. Sun, Sparse and low-rank coupling image segmentation model via nonconvex regularization, Int. J. Pattern Recognit. Artif. Intell. 29 (2015) 1–22.

[8] A. Dirami, K. Hammouche, M. Diaf, P.. Siarry, Fast multilevel thresholding for image segmentation through a multiphase level set method, Signal Process. 93 (2013) 139–153.

[9] H. Zhang, J.E. Fritts, S.A Goldman, Image segmentation evaluation: a survey of unsupervised methods, Comput. Vis. Image Underst. 110 (2008) 260–280.

[10] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, J. Electron. Imaging 13 (2004) 146–168.

[11] M. Mignotte, A label field fusion Bayesian model and its penalized maximum rand estimator for image segmentation, IEEE Trans. Image Process. 19 (2010) 1610–1624.

[12] M. Krinidis, I. Pitas, Color texture segmentation based on the modal energy of deformable surfaces, IEEE Trans. Image Process. 18 (2009) 1613–1622.

[13] Y. Han, X.-C. Feng, G. Baciu, Variational and PCA based natural image segmentation, Pattern Recognit. 46 (2013) 1971–1984.

[14] Z. Yu, O.C. Au, R. Zou, W. Yu, J. Tian, An adaptive unsupervised approach toward pixel clustering and color image segmentation, Pattern Recognit. 43 (2010) 1889–1906.

[15] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, A.K. Nandi, Significantly fast and robust fuzzy C-Means clustering algorithm based on morphological reconstruction and membership filtering, IEEE Trans. Fuzzy Syst. 26 (5) (2018) 3027–3041.

[16] A.S. Abutaleb, Automatic thresholding of gray-level pictures using two-dimensional entropy. Comput. Vis. Graph. Image Process.. 47, (1089), 22–32.

[17] A. Brink, Thresholding of digital images using two-dimensional entropies, Pattern Recognit. 25 (1992) 803–808.

[18] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

[19] A.B Ishak, Choosing parameters for Rényi and Tsallis entropies within a two-dimensional multilevel image segmentation framework, Physica A 466 (2017) 521–536.

[20] X. Zhao, M. Turk, W. Li, K.-c Lien, G. Wang, A multilevel image thresholding segmentation algorithm based on two-dimensional K–L divergence and modified particle swarm optimization, Appl. Soft Comput. 48 (2016) 151–159.

[21] W. Xue-guang, C. Shu-hong, An improved image segmentation algorithm based on two-dimensional Otsu method, Inform. Sci. Lett. 1 (2012) 77–83.

[22] C. Sha, J. Hou, H. Cui, A robust 2D Otsu's thresholding method in image segmentation, J. Vis. Commun. Image Represent. 41 (2016) 339–351.

[23] S. Sarkar, S. Das, Multilevel image thresholding based on 2D histogram and maximum Tsallis entropy—a differential evolution approach, IEEE Trans. Image Process. 22 (2013) 4788–4797.

[24] A. Nakib, S. Roman, H. Oulhadj, P. Siarry, Fast brain MRI segmentation based on two-dimensional survival exponential entropy and particle swarm optimization, in: Proceedings of the International Conference on Engineering in Medicine and Biology Society, 2007.

[25] A. Nakib, H. Oulhadj, P. Siarry, Image thresholding based on pareto multiobjective optimization, Eng. Appl. Artif. Intell. 23 (2010) 313–320.

[26] X.-S. Yang, Nature-inspired Optimization AlgorithCM, Elsevier, 2014.

[27] Y.-g. Tang, D. Liu, X.-p. Guan, Fast image segmentation based on particle swarm optimization and two-dimension otsu method, Control Decis. 22 (2007) 202–205.

[28] X. Lei, A. Fu, Two-dimensional maximum entropy image segmentation method based on quantum-behaved particle swarm optimization algorithm, in: Proceedings of the International Conference on Natural Computation, 2008.

[29] C. Qi, Maximum entropy for image segmentation based on an adaptive particle swarm optimization, Appl. Math. 8 (2014) 3129–3135.

[30] X. Shen, Y. Zhang, F. Li, An improved two-dimensional entropic thresholding method based on ant colony genetic algorithm, in: Proceedings of the WRI Global Congress on Intelligent SysteCM, 2009.

[31] H. Cheng, Y. Chen, X. Jiang, Thresholding using two-dimensional histogram and fuzzy entropy principle, IEEE Trans. Image Process. 9 (2000) 732–735.

[32] S. Kumar, T.K. Sharma, M. Pant, A. Ray, Adaptive artificial bee colony for segmentation of CT lung images, in: Proceedings of the International Conference on Recent Advances and Future Trends in Information Technology, 2012.

[33] S. Fengjie, W. He, F. Jieqing, 2D Otsu segmentation algorithm based on simulated annealing genetic algorithm for iced-cable images, in: Proceedings of the International Forum on Information Technology and Applications, 2009.

[34] L. Xiao-Feng, L. Hui-Ying, Y. Ming, W. Tai-Ping, Infrared image segmentation based on AAFSA and 2D-renyi entropy threshold selection, in: Proceedings of the Joint International Conference on Artificial Intelligence and Computer Engineering and International Conference on Network and Communication Security, 2016.

[35] R. Panda, S. Agrawal, L. Samantaray, A. Abraham, An evolutionary gray gradient algorithm for multilevel thresholding of brain MR images using soft computing techniques, Appl. Soft Comput. 50 (2017) 94–108.

[36] D. Oliva, E. Cuevas, G. Pajares, D. Zaldivar, V. Osuna, A multilevel thresholding algorithm using electromagnetism optimization, Neurocomputing 139 (2014) 357–381.

[37] H. Mittal, M. Saraswat, An optimum multi-level image thresholding segmentation using non-local means 2D histogram and exponential Kbest gravitational search algorithm, Eng. Appl. Artif. Intel. 71 (2018) 226–235.

[38] M.P. Wand, M.C. Jones, Kernel Smoothing, Springer, 1995.

[39] J.E. Chacón, Data-driven choice of the smoothing parametrization for kernel density estimators, Can. J. Stat. 37 (2009) 249–265.

[40] K.S Duong, Kernel density estimation and Kernel discriminant analysis for multivariate data in R, J. Stat. Softw. 21 (2007) 1–16.

[41] A. Gramacki, Nonparametric Kernel Density Estimation and Its Computational Aspects, Springer, 2018.

[42] V.A. Epanechnikov, Non-parametric estimation of a multivariate probability density, Theory Probab. Appl. 14 (1969) 153–158.

[43] Y.Z. Cheng, Mean Shift, mode seeking, and clustering, IEEE Transact. Pattern Anal. Mach. Intel. 17 (8) (1995) 790–799.

[44] Y. Guo, A. S¸engür, Y. Akbulut, A. Shipley, An effective color image segmentation approach using neutrosophic adaptive mean shift clustering, Measurement 119 (2018) 28–40.

[45] G. Domingues, H. Bischof, R. Beichel, Fast 3D mean shift filter for CT images, in: Proceedings of the Scandinavian Conference on Image Analysis, Sweden, 2003, pp. 438–445.

[46] D. Comaniciu, P. Meer, Meanshift: a robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intel. 24 (5) (2002) 603–619.

[47] W.B. Tao, J. Liu, Unified mean shift segmentation and graph region merging algorithm for infrared ship target segmentation, Opt. Eng. 46 (2007) 12.

[48] J.H. Park, G.S. Lee, S.Y. Park, Color image segmentation using adaptive mean shift and statistical model-based methods, Comput. Math. Appl. 57 (2009) 970–980.

[49] Q. Li, J.S Racine, Nonparametric Econometrics: Theory and Practice, Princeton University Press, 2007 ISBN 978-0-691-12161-1.

[50] Y. Luo, K. Zhang, Y. Chai, Y. Xiong, Muiti-parameter-setting based on data original distribution for DENCLUE optimization, IEEE Access 6 (2018) 16704–16711.

[51] L. Xu, E. Oja, P. Kultanen, A new curve detection method: Randomized Hough transform (RHT), Pattern Recognit. Lett. 11 (5) (1990) 331–338.

[52] M.A. Fisher, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.

[53] I. Horová, J. Kolácek ˇ , J. Zelinka, Kernel Smoothing in Matlab, World Scientific, 2012.

[54] D.W. Scott, Scott's rule, Wires Comput. Stat. 2 (4) (2010) 497–502.

[55] https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/.

[56] S.K. Choy, T.C. Ng, C. Yu, Unsupervised fuzzy model-based image segmentation, Signal Process. 171 (2020) 107483.

[57] N. Dhanachandra, Y.J. Chanu, An image segmentation approach based on fuzzy c-means and dynamic particle swarm optimization algorithm, Multimed. Tools Appl. (2020), https://doi.org/10.1007/s11042-020-08699-8.

[58] C. Wu, Y. Chen, Adaptive entropy weighted picture fuzzy clustering algorithm with spatial information for image segmentation, Appl. Soft Comput. 86 (2020) 105888.

[59] A.A. Hernandez del Rio, E. Cuevas, D. Zaldivar, Multi-level image thresholding segmentation using 2d histogram non-local means and metaheuristics algorithCM, in: D. Oliva, S. Hinojosa (Eds.), Applications of Hybrid Metaheuristic AlgorithCM for Image Processing. Studies in Computational Intelligence, 890, Springer, 2020.

[60] B. Vinoth Kumar, S. Sabareeswaran, G. Madumitha, A decennary survey on artificial intelligence methods for image segmentation, in: R. Venkata Rao, J. Taler (Eds.), Advanced Engineering Optimization Through Intelligent Techniques. Advances in Intelligent SysteCM and Computing, 949, Springer, 2020.

[61] M. Chouksey, R.K. Jha, R. Sharma, A fast technique for image segmentation based on two Meta-heuristic algorithCM, Multimed. Tools Appl. (2020), https://doi.org/10.1007/s11042-019-08138-3.

[62] A. Draa, A. Bouaziz, An artificial bee colony algorithm for image contrast enhancement, Swarm Evol. Comput. 16 (2014) 69–84.

[63] A. Draa, A. Bouaziz, An artificial bee colony algorithm for image contrast enhancement, Swarm Evol. Comput. 16 (2014) 69–84.

[64] L.D.S. Coelho, J.G. Sauer, M. Rudek, Differential evolution optimization combined with chaotic sequences for image contrast enhancement, Chaos Sol. Fract. 42 (2009) 522–529.

[65] C. Munteaunu, A. Rosa, Gray-scale image enhancement as an automatic process driven by evolution, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 34 (2) (2004) 1992–1998.

[66] M. Braik, A. Sheta, A. Ayesh, Particle swarm optimisation enhancement approach for improving image quality, Int. J. Innov. Comput. Appl. 1 (2) (2007) 138–145.