

Segmentación

7 de junio de 2002

1 Segmentación

Los algoritmos para segmentar imágenes monocromáticas generalmente se basan en una de las dos propiedades básicas de los valores del nivel de gris: discontinuidad y similitud. En la primera categoría, la aproximación es particionar la imagen basándose en cambios abruptos en el nivel de gris. Las principales áreas de interés en esta categoría son de detección de puntos aislados, de líneas y de bordes en una imagen. La segunda categoría se basa en segmentar por niveles de umbral, en regiones de crecimiento y en corte y pegado de regiones. La técnica por un nivel de umbral global se trata en la subsección 1.1. La técnica basada en la detección de bordes se estudia en la subsección 3.3.2. Los otros esquemas para segmentación pueden revisarse en [1, 2, 3].

1.1 Segmentación por nivel de umbral global

Suponiendo que el histograma mostrado en la fig. 1(a) corresponde a una imagen, $f(x, y)$, compuesta de objetos brillantes sobre un fondo oscuro, de tal manera que los píxeles del objeto y el fondo tienen niveles de gris agrupados en dos modos dominantes. Un camino obvio para extraer los objetos del fondo es seleccionar un umbral T que separe estos modos. Entonces cualquier punto (x, y) para el cual $f(x, y) > T$ es llamado un punto objeto; de otra manera, el punto es llamado un punto del fondo. La misma aproximación básica clasificada un punto (x, y) como perteneciente a un tipo de objeto $T_1 < f(x, y) \leq T_2$, la otra clase de objeto si $f(x, y) > T_2$ y el fondo si $f(x, y) \leq T_1$, este es llamado multiniveles de umbral.

La técnica más simple de segmentación por nivel de umbral es particionar el histograma de la imagen usando un umbral simple T . La imagen particionada $g(x, y)$ esta definida como:

$$g(x, y) = \begin{cases} 1, & \text{si } f(x, y) > T \\ 0, & \text{si } f(x, y) \leq T \end{cases} \quad (1)$$

Entonces los píxeles etiquetados con 1 (o cualquier otro nivel de intensidad conveniente) corresponden a los objetos, mientras que los píxeles etiquetados con 0 corresponden al fondo.

En la práctica, el tipo de umbral global descrito por la ec. 1 es útil en ambientes altamente controlados. Una de las áreas donde es posible aplicarlo es en aplicaciones de inspección industrial, donde el control de la iluminación usualmente es posible y, por lo tanto, se puedan obtener histogramas bimodales.

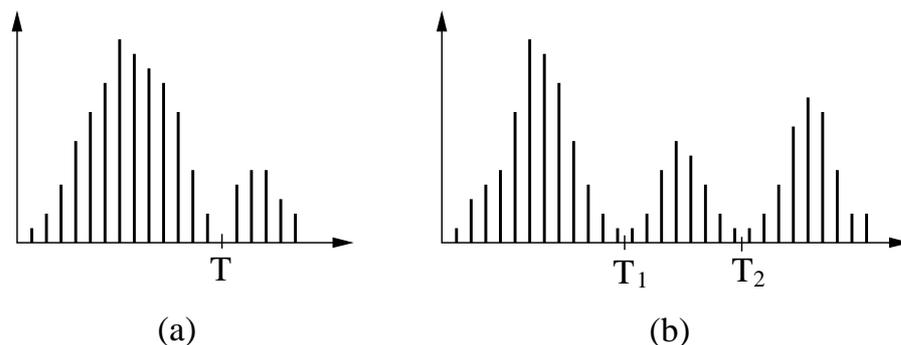


Figura 1: Histogramas de niveles de gris que pueden ser particionados por (a) en un umbral simple y (b) en umbrales múltiples.

El tipo de multiniveles de umbral es menos realizable. La razón estriba en la dificultad de establecer umbrales múltiples que aislen efectivamente regiones de interés, especialmente cuando el número de modos del histograma es grande. Típicamente, los problemas de esta naturaleza, si se van a manejar por varios niveles de umbral, se resuelven mejor usando un solo nivel de umbral variable.

1.2 Segmentación por selección del umbral basándose en las características de los bordes

Una aproximación para mejorar la forma del histograma es considerar solo aquellos pixeles que están situados sobre o cerca de los bordes entre los objetos y el fondo. Así se hace que el histograma sea menos dependiente del tamaño de los objetos y el fondo. Esto es, si se tiene una imagen con un objeto pequeño, el histograma dará un pico muy grande que correspondería al fondo, pero si se usa solo los límites entre los objetos, el histograma dará, aproximadamente, dos picos de la misma altura. Esto se puede lograr con los operadores gradiente y Laplaciano. La detección de bordes es uno de los métodos más comunes para la detección de discontinuidades en el nivel de gris. Primero se revisará la teoría en que se basa esta técnica y luego se describirá el algoritmo usado.

1.3 Formulación básica

Un borde es el límite entre dos regiones con propiedades relativamente distintas en el nivel de gris.

Básicamente, la idea fundamental de la mayoría de las técnicas para detectar bordes es la computación de un operador local derivativo. La figura 2 ilustra este concepto. La fig. 2(a) muestra la imagen de una banda brillante sobre un fondo oscuro, el perfil de los niveles de gris a lo largo de una línea horizontal, y la primera y segunda derivadas del perfil. Nótese que el perfil de un borde (la transición de oscuro a brillante) está modelado de manera suave, mas que de manera abrupta. Este modelo refleja el hecho de que los bordes en las imágenes digitales son generalmente un poco borrosas como resultado del muestreo.

La fig. 2(b) muestra que la primera derivada del perfil del nivel de gris es positiva en el principal borde de una transición, negativa en el borde siguiente, y, como es de esperarse, cero

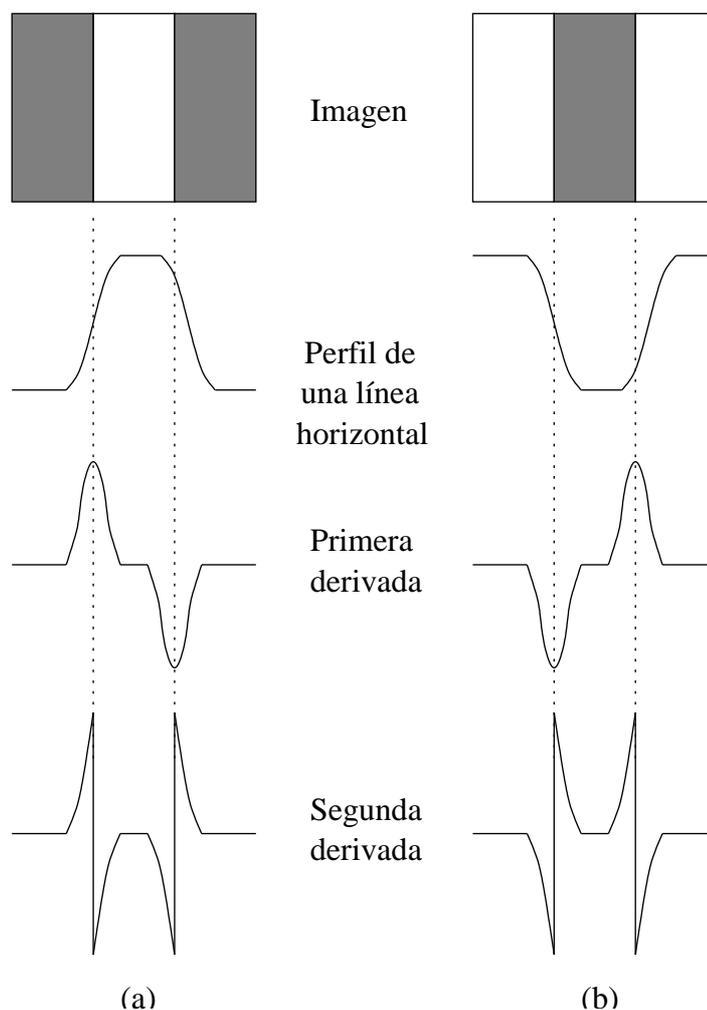


Figura 2: Detección de bordes por los operadores derivativos: (a) Una banda brillante sobre un fondo oscuro, (b) Una banda oscura sobre un fondo brillante. Nótese que la segunda derivada tiene un cruce por cero en la localización de cada borde.

en las áreas que tienen el nivel de gris constante. La segunda derivada es positiva para la parte de la transición asociada con el lado oscuro del borde, negativa para la parte de la transición asociada con el lado brillante del borde, y cero en las áreas con nivel de gris constante. Por lo tanto, la magnitud de la primera derivada puede usarse para detectar la presencia de un borde en una imagen, y el signo de la segunda derivada puede usarse para determinar si un pixel está situado sobre el lado oscuro o brillante de un borde. Nótese que la segunda derivada tiene un cruce por cero en el punto medio de la transición del nivel de gris.

Aquí la discusión se limitó a un perfil horizontal en una dimensión, similares argumentos se aplican a un borde en cualquier orientación en una imagen. Simplemente se define un perfil perpendicular a la dirección del borde en cualquier punto deseado y se interpretan los resultados como en la discusión anterior. La primera derivada en cualquier punto de una imagen es obtenida usando la magnitud del gradiente en aquel punto. La segunda derivada es obtenida de manera similar usando el Laplaciano.

1.4 Operadores de gradiente

El gradiente es el método más común de diferenciación en aplicaciones de procesamiento de imágenes. Para una función $f(x, y)$, el gradiente de f en las coordenadas (x, y) está definido como el vector

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix} \quad (2)$$

Es bien conocido en análisis de vectores que el vector gradiente apunta en la dirección del cambio máximo de f en (x, y) . En la detección de bordes una cantidad importante es la magnitud de este vector, referido generalmente como gradiente y denotado por f , donde:

$$\nabla f = \text{mag}(\nabla \mathbf{f}) = [G_x^2 + G_y^2]^{1/2} \quad (3)$$

Esta cantidad es igual a la máxima velocidad de incremento de $f(x, y)$ por unidad de distancia en la dirección de f . En la práctica es común aproximar el gradiente con los valores absolutos:

$$\nabla f = |G_x| + |G_y| \quad (4)$$

ó como:

$$\nabla f = \max(G_x, G_y) \quad (5)$$

las cuales son mucho más simples de realizar, particularmente con hardware dedicado.

La dirección del vector gradiente también es una cantidad importante. Si (x, y) representa el ángulo de dirección del vector (x, y) , entonces del análisis de vectores

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y}{G_x} \right) \quad (6)$$

donde el ángulo es medido con respecto al eje de la x .

Las derivadas de la ec. 2 pueden implementarse en forma digital de varias maneras. Una es usando los operadores Sobel mostrados en la fig. 3 que tienen la ventaja de proveer además un efecto de suavizamiento. Y debido a que la derivada enfatiza el ruido, el efecto de suavizamiento es una característica particularmente atractiva de los operadores Sobel.

| | | | | | | | | |
|-------|-------|-------|-----|---|---|-----|----|----|
| z_1 | z_2 | z_3 | -1 | 0 | 1 | -1 | -2 | -1 |
| z_4 | z_5 | z_6 | -2 | 0 | 2 | 0 | 0 | 0 |
| z_7 | z_8 | z_9 | -1 | 0 | 1 | 1 | 2 | 1 |
| (a) | | | (b) | | | (c) | | |

Figura 3: (a) Una región de una imagen de tamaño 3×3 píxeles. (b) Máscara usada para computar G_x en el punto central de la región (a). (c) Máscara para computar G_y en el mismo punto. Estas máscaras son referidas frecuentemente como operadores Sobel.

Entonces, la ec. 2 queda usando las máscaras de la fig. 3 como:

$$G_x = (Z_3 + 2Z_6 + Z_9) - (Z_1 + 2Z_4 + Z_7) \quad (7)$$

$$G_y = (Z_7 + 2Z_8 + Z_9) - (Z_1 + 2Z_2 + Z_3) \quad (8)$$

1.5 El Laplaciano

El Laplaciano de una función en 2-D $f(x, y)$ es una derivada de segundo orden definida como

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (9)$$

Como en el caso del gradiente, ec. 9 puede implementarse de varias maneras en forma digital. Para una región de 3×3 , la forma más frecuentemente encontrada es la que se muestra en la fig. 4.

| | | |
|---|----|---|
| 0 | 1 | 0 |
| 1 | -4 | 1 |
| 0 | 1 | 0 |

Figura 4: Máscara de 3×3 para computar el Laplaciano

$\nabla^2 f$ de la ec. 9, y viendo las figs. 3(a) y 4, se aproxima como:

$$\nabla^2 f = (Z_2 + Z_4 + Z_6 + Z_8) - 4Z_5 \quad (10)$$

El requerimiento básico en la definición del Laplaciano digital es que los coeficientes asociados con el pixel central sea positivo y que coeficientes asociados con los pixeles externos sean negativos. Porque el Laplaciano es una derivada, la suma de los coeficientes tiene que ser cero. Por lo tanto la respuesta es cero siempre que el punto tratado y sus linderos tengan el mismo valor.

El Laplaciano responde a las transiciones en intensidad, y este es raramente usado por la detección de bordes por varias razones. Como es una derivada de segundo orden el Laplaciano típicamente es inaceptablemente sensible al ruido. Aún mas, el Laplaciano produce bordes dobles y es incapaz de detectar la dirección del borde. Por estas razones, el Laplaciano usualmente juega un papel secundario como detector para establecer cuando un pixel esta sobre el lado brillante u oscuro de una imagen.

1.6 Algoritmo

El gradiente y el Laplaciano pueden usarse para formar una imagen en tres niveles, como sigue:

$$s(x, y) = \begin{cases} 0, & \text{si } \nabla f < T \\ +, & \text{si } \nabla f \geq T \text{ y } \nabla^2 f \geq 0 \\ -, & \text{si } \nabla f \geq T \text{ y } \nabla^2 f < 0 \end{cases} \quad (11)$$

donde los símbolos O , $+$, $-$ representan cualesquiera de tres distintos niveles de gris, T es un umbral, y el gradiente y el Laplaciano son computadoras en cada punto (x, y) . Para un objeto oscuro sobre un fondo brillante el uso de la ec. 11 produce una imagen $s(x, y)$ en el cual todos los pixeles que no están sobre un borde (como los determinados por f al ser menores que T) son etiquetados a 0 , todos los pixeles en el lado oscuro de un borde son etiquetados a $+$ y todos los pixeles sobre el lado brillante son etiquetados a $-$. Los símbolos $+$ y $-$ se intercambian para un objeto brillante sobre un fondo oscuro.

La información obtenida por este procedimiento puede usarse para generar una imagen binaria segmentada, en la cual los 1s corresponden al objeto de interés y los 0s corresponden al fondo. La transición (a lo largo de una línea horizontal o vertical) de un fondo brillante a un objeto oscuro debe caracterizarse por la ocurrencia de un ' $-$ ' seguido de un ' $+$ ' en $s(x, y)$. El interior del objeto está compuesto de pixeles que están etiquetados con ' O ' ó ' $+$ '. Finalmente, la transición del objeto al fondo esta caracterizado por la ocurrencia de un ' $+$ ' seguido de un ' $-$ '. Así una línea horizontal o vertical que contiene una sección de un objeto tiene la siguiente estructura:

$$(\dots)(-, +)(O \text{ ó } +)(+, -)(\dots)$$

donde (\dots) representa cualquier combinación de $+$, $-$ y O . Los paréntesis más internos contienen puntos de objeto que son etiquetados a 1. Todos los demás pixeles de la misma línea son etiquetados a O , con la excepción de cualquier otra secuencia de $(O \text{ ó } +)$ limitados por $(-, +)$ y $(+, -)$.

Referencias

- [1] P.C. Chen and T. Pavlidis. Segmentation by texture using a co-occurrence matrix and a split merge algorithm. *Computer Graphics and Image Processing*, 10:172–182, 1979.
- [2] Theo Pavlidis and Y. Liow. Integrating region growing and edge detection. *IEEE Tran. on Patt. and Match. Intelligence*, 12(3), 1990.
- [3] T.R. Reed and J.M. Hans du Buf. A review of recent texture segmentation and feature extraction techniques. *Image Understanding*, 57(3):359–372, 1993.