

SEGMENTACION DE IMAGEN PARA VISION DE COMPUTADORES

BERNARDO BARRAZA SOTOMAYOR*
Dpto de Electrónica
Universidad de Tarapacá-Chile
CLÉSIO LUIS TOZZI**
FEC - UNICAMP-BRASIL

RESUMEN

Este trabajo describe la fase de segmentación en el proceso de visión de computadores. La segmentación en el proceso de visión de computadores, consiste en la separación de los diferentes objetos que componen una escena. Se abordan las técnicas de segmentación por umbral ('thresholding') y la segmentación por detección de bordes. Se describen los métodos para el almacenamiento de las diferentes figuras obtenidas en la fase de segmentación, dando a conocer algoritmos aplicados a figuras representadas por su región y a figuras representadas por su frontera. La fase de segmentación finaliza con el almacenamiento ya descrito anteriormente.

INTRODUCCIÓN

El proceso de visión de computadores puede ser considerado formado por las siguientes fases:

- Pre-procesamiento
- Segmentación
- Extracción de atributos
- Clasificación

La fase de pre-procesamiento, consiste en una adecuación de la escena física a ser analizada en una forma tal que sea posible su procesamiento por el computador. En esta fase tenemos que existe una discretización, cuantización, filtraje, realzamiento, etc, lo cual permite que la escena física sea presentada en forma digital al computador.

Lo importante del proceso de visión de computador, es clasificar o identificar los diferentes objetos de la escena, lo cual es posible, considerando que cada objeto o figura posee características o atributos que lo diferencian de otros.

Con la finalidad de extraer características o atributos de cada objeto de la escena, es necesario hacer una separación de cada uno de los objetos que componen la escena que está siendo analizada. Esta fase de separación recibe el nombre de 'segmentación de imagen'.

SEGMENTACIÓN DE IMAGEN

Definiremos segmentación como la 'división de una imagen en objetos aislados', donde el objeto es un conjunto de 'pixels' (picture elements), que poseen niveles de gris similar, y cuyo conjunto tienen algún significado para el usuario.

*Casilla 270, Arica-CHILE

**Caixa Postal 1170, Campinas-S.P, BRASIL

La figura 1.1. muestra gráficamente como es entendida la segmentación según lo descrito anteriormente. Por tanto, la segmentación consistirá en la determinación para cada objeto de una frontera, una región conectada a esta frontera y un rótulo que identifica a este objeto.

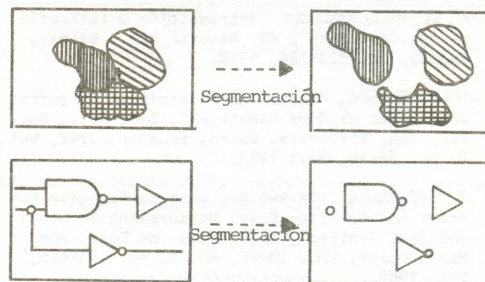


Fig. 1. Segmentación de una imagen.

El proceso de segmentación de una imagen está basado en los principios básicos de discontinuidad y similitud del nivel de gris, siendo las siguientes las técnicas utilizadas:

- Discontinuidad : Segmentación por detección de bordes
- Similitud : Segmentación por umbral o 'thresholding'

1.1. SEGMENTACIÓN POR DETECCIÓN DE BORDES.

La segmentación por detección de bordes se basa en el hecho que un borde es definido por una variación abrupta en el nivel de gris, o sea, existen niveles de gris relativamente consistente en cada una de las dos regiones y una variación sensible del nivel de gris en el límite de las esas regiones.

La figura 2 ilustra las variaciones del nivel de gris en los bordes de un objeto.

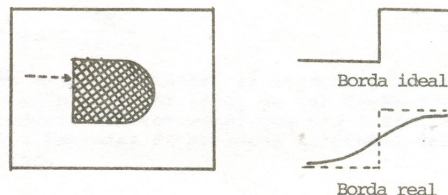


Fig. 2. Escena conteniendo objeto A con fondo B y los gráficos para niveles de gris.

La detección de bordes, aprovechando lo descrito anteriormente, es realizada con el uso de operadores de diferencias, que consideran los niveles de gris en el punto (x,y) y sus vecinos. Estos operadores se pueden clasificar en operadores de Gradiente, Laplacianos y Máscaras de convolución en general.

1.1.1. DETECCIÓN DE BORDES, USANDO EL MÉTODO DEL GRADIENTE.

Utiliza el operador derivativo de dos dimensiones. Sea $f(x,y)$, el nivel de gris en el punto (x,y) , entonces las derivadas parciales $\partial f/\partial x$ e $\partial f/\partial y$, miden las variaciones de los niveles de gris en las direcciones horizontal y vertical respectivamente. El operador que define el gradiente es:

$$G = ((\partial f/\partial x)^2 + (\partial f/\partial y)^2)^{1/2} \quad (1)$$

Para imágenes digitales, en vez del uso de la primera derivada es usada la primera diferencia.

$$\Delta_{xy} = |(f(x,y) - f(x+1,y)) + (f(x,y) - f(x,y+1))| \quad (2)$$

Existen otros operadores de este tipo y los más conocidos son los operadores de diferencias centrales, operadores de Robert, etc.

1.1.2. DETECCIÓN DE BORDES, USANDO EL MÉTODO DEL LAPLACIANO O SEGUNDAS DIFERENCIAS.

Una alternativa al método 1.1., es el uso del Laplaciano. Este método, tiene como característica principal una mayor sensibilidad en la detección de líneas, cantos y puntos aislados, siendo el Laplaciano definido por:

$$\nabla^2 f = \partial^2 f/\partial x^2 + \partial^2 f/\partial y^2 \quad (3)$$

Para imágenes digitales se usa:

$$\nabla^2 f(x,y) = f(x+1,y) + f(x,y+1) + f(x-1,y) + f(x,y-1) - 4f(x,y) \quad (4)$$

1.1.3. DETECCIÓN DE BORDES, POR MEDIO DE MÁSCARAS DE CONVOLUCIÓN.

Las máscaras de convolución son operadores de gradiente para la detección de bordes, en las cuales son considerados los pixels vecinos tanto en las direcciones horizontal y vertical como en las direcciones diagonales (± 45 y ± 135).

Su principio se basa en el desplazamiento de una máscara sobre el área de la escena por medio de un barrido, computándose para cada pixel (x,y) el valor de la área que se encuentra ubicada debajo de la máscara, obteniéndose una medida del gradiente en el punto y una medida de la dirección del borde.

Una de las máscaras mas usadas, es atribuida a Sobel y los valores a usar son los siguientes:

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

a	b	c
d	e	f
g	h	i

El valor del gradiente $g(e)$ en el punto $e=(x,y)$ en forma general, está definido por la siguiente ecuación:

$$g(e) = \left| \begin{aligned} &((f(g)+2f(h)+f(i)) - (f(a)+2f(b)+f(c))) \\ &+ ((f(c)+2f(f)+f(i)) - (f(a)+2f(d)+f(g))) \end{aligned} \right| \quad (5)$$

Para una imagen binaria, el máximo valor de este gradiente es 6 y corresponde a un canto o borde diagonal. Para un borde horizontal o vertical el valor del gradiente es 4.

Existen otros algoritmos que utilizan varias máscaras para la determinación de este gradiente. El valor máximo de una de estas máscaras determina el gradiente final en ese punto y la dirección del borde. Las máscaras posibles a usar son conocidas como de: Kirsch, Prewitt, 3-nivel y 5-nivel.

1.2. SEGMENTACIÓN POR UMBRAL ('THRESHOLDING').

Es uno de los métodos más usado en el proceso de segmentación. Consiste en escoger un nivel de gris T de modo que todos los pixels de nivel gris encima de este valor T, son considerados de los objetos (nivel 1) y los pixels con valores menores que T, son considerados del fondo de la escena. El método es efectivo si existe una diferencia en los niveles de gris de los objetos a analizar.

Un operador de umbral puede ser definido por la siguiente ecuación:

$$T = f(x,y, N(x,y), f(x,y)) \quad (6)$$

donde:

x,y cordenadas del pixel
 $f(x,y)$ es el nivel de gris en el punto x,y
 $N(x,y)$ característica o propiedad del punto (x,y)

Para la obtención de una imagen binaria $g(x,y)$, a partir de una imagen original, se aplica el siguiente criterio:

$$g(x,y) = \begin{cases} 1 & \text{(objeto) si } f(x,y) > \text{valor de umbral} \\ 0 & \text{(fondo) caso contrario} \end{cases}$$

El valor de T o umbral es escogido por medio de un histograma que representa la distribución de los niveles de gris en la escena.

Si T depende:

- solo de $f(x,y)$, el umbral es llamado global
- de $f(x,y)$ y $N(x,y)$, el umbral es llamado local
- de las cordenadas x,y , de $N(x,y)$ y $f(x,y)$, el umbral es llamado dinámico.

1.3. ADELGAZAMIENTO ('THINNING').

Una técnica no comúnmente considerada como tal para el proceso de segmentación por algunos autores, pero con la cual se obtienen resultados excelentes en ciertas aplicaciones, es la llamada adelgazamiento de una imagen. El algoritmo de adelgazamiento usado, es aplicado a aquellas figuras que están representadas en la escena por un contorno de varios pixels de ancho. Al aplicar el procedimiento de adelgazamiento se obtiene finalmente la figura representada por un contorno de solamente un pixel de ancho.

El algoritmo considera que la figura está formada por pixels de contorno y pixels finales. Pixels de contorno son todos aquellos que pertenecen a la figura en un comienzo y pixels finales son los resultantes obtenidos después de aplicar el proceso de adelgazamiento. Las condiciones que un pixel debe cumplir para ser considerado de contorno o final están definidos en la referencia 5 y son determinados a través de máscaras. El algoritmo considera que si un pixel cumple la condición de pixel de contorno pero no la de un pixel final este debe ser borrado de la escena. El proceso es iterativo hasta obtenerse solamente pixels finales y estos deberán tener una conectividad entre ellos.

La figura 3 muestra graficamente el proceso de segmentación por detección de bordes y adelgazamiento. A figura 3.2 muestra la necesidad de aplicar la técnica de adelgazamiento para detectar el contorno o borde de una figura.

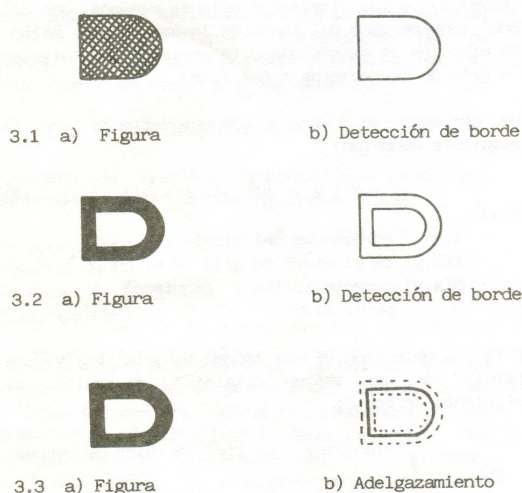


Fig 3. Proceso de segmentación por detección de bordes y adelgazamiento.

2. ALMACENAMIENTO.

El proceso de segmentación es finalizado con el almacenamiento de cada figura contenida en la escena en archivos separados. Este almacenamiento es necesario ya que, todavía los diferentes objetos continúan perteneciendo a la escena y para la clasificación de cada objeto deberemos medir características de cada uno de ellos en forma individual, por tanto, el proceso de segmentación debe finalizar con una separación física de figuras lo cual es conseguido con los algoritmos a discutir.

Los métodos discutidos corresponden a los siguientes tipos de figuras obtenidas en la segmentación:

- 1.- Almacenamiento por trazo, aplicado a figuras obtenidas en la segmentación por umbral.
- 2.- Almacenamiento por la regla de la cadena, aplicado a figuras obtenidas en la segmentación por detección de bordes y adelgazamiento

2.1. ALMACENAMIENTO POR TRAZO.

En este método es realizado un barrido a través de líneas horizontales; para cada línea es almacenado el valor de las coordenadas del pixel de comienzo y final de cada trazo. En la figura 4, está representado graficamente este proceso para dos objetos.

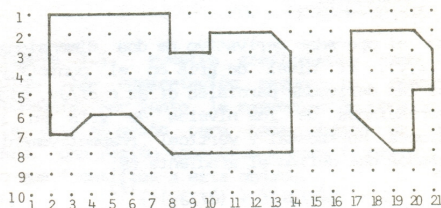
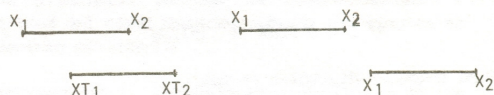


Fig. 4. Proceso de almacenamiento por trazo.

De la figura se observa que:

- en la línea 1, es generado el archivo A
- en la línea 2, son generados los archivos B y C
- en la línea 3, se verifica que los archivos A y B corresponden a una misma figura
- al término del proceso se tendrán sólo dos archivos que corresponden AUB y C

Las condiciones o criterio que se debe cumplir para que el trazo de la línea analizada, tenga una relación con el trazo de la línea anterior son:



Sea:

X_1, X_2 coordenadas X, del trazo de la línea i
 XT_1, XT_2 coordenadas X del trazo de la línea i+1

Si : $(XT_2+1 \leq X_1)$.OR. $(XT_1 -1 \geq X_2)$

Entonces, los trazos no pertenecen a la misma figura

El almacenamiento por trazo, es obtenido a través de los siguientes pasos:

- 1.- Haga un barrido de la línea i(i=1). Determine los valores de las coordenadas de los trazos y almacene los pares de coordenadas de la línea i, en K archivos.
- 2.- En la línea i+1, determine un trazo existente.
- 3.- Aplique el criterio de verificación de relación entre trazos, para el trazo obtenido en 2, verificando si pertenecen a figuras cuyos archivos fueron creados anteriormente.
- 4.- Si pertenece, almacene el trazo en el archivo respectivo. Vuelva al paso 2.
- 5.- Si no pertenece, cree un nuevo archivo y almacene en el mismo las coordenadas correspondientes. Vuelva al paso 2.
- 6.- Haga una comparación de los archivos creados para las posibles figuras y determine se existen elementos comunes. En caso positivo haga el ordenamiento de las figuras finales.

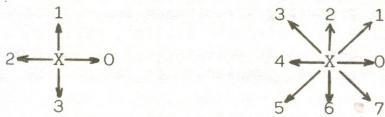
2.2. ALMACENAMIENTO POR LA REGLA DE LA CADENA.
('BOUNDARY CHAIN CODE').

La regla de la cadena es una forma de representación para figuras descritas por su contorno y que resultan normalmente de la segmentación por detección de bordes y adelgazamiento.

La regla de la cadena se basa en el hecho de que muchos objetos pueden ser definidos por su contorno, no siendo necesario el almacenamiento de los puntos interiores de la figura. Así, el código de la cadena es una representación compacta del contorno cerrado de una figura, siendo esta cadena una secuencia ordenada de elementos $a_1, a_2, a_3, \dots, a_n$, juntamente con las coordenadas del pixel de inicio de esta cadena.

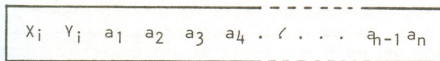
Existen dos convenciones para representar una cadena. Una considera que un pixel central es asociado o tiene conexiones con los pixels vecinos en las direcciones horizontales y verticales, y es denominado de 4-vecinos. A otra convención considera que el pixel central esta asociado con los vecinos horizontales, verticales y diagonales, y recibe el nombre de 8-vecinos. La justificativa para el uso de las dos representaciones viene del hecho que para algunas aplicaciones el contorno es definido sólo por pixels en las direcciones horizontal y vertical, siendo este el caso de figuras obtenidas por detección de bordes. En el caso de figuras obtenidas por adelgazamiento, muchas veces para algunos pixels de la figura no existe conectividad en las direcciones horizontales y verticales, pero si en las direcciones diagonales, siendo este el caso para el uso de la convención de 8-vecinos.

La figura 5 muestra la representación de las convenciones mencionadas, para un pixel central X, y la figura 6 una representación de una cadena.



a) 4-vecinos directos (0-1-2-3) b) 8-vecinos indirectos (1-3-5-7)

Fig. 5. Representación de la convención usada.



- x_i, y_i , coordenadas de inicio.
- $a_1, a_2, a_3, \dots, a_n$, convención que representa el contorno
- existe conectividad entre a_1 y a_n

Fig. 6. Representación de una cadena

La regla de la cadena permite una compresión de los datos, siendo necesario almacenar en un sólo archivo y sin necesidad de ordenamiento, las coordenadas del pixel de inicio de la cadena y la secuencia de los elementos de la cadena.

La figura 7 muestra una aplicación de almacenamiento por la regla de la cadena para un objeto.

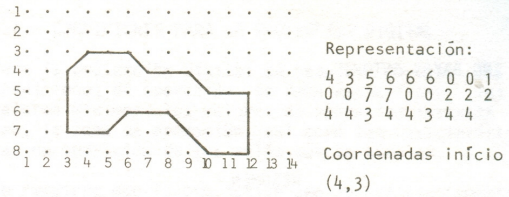


Fig. 7. Aplicación de la regla de la cadena.

3. CONCLUSIONES.

Lo expuesto anteriormente corresponde a una parte de un trabajo desarrollado en la Universidad Estadual de Campinas-FEC, sobre Reconocimiento de Padrões para Aplicaciones en CAD/CAM

Fueron desarrollados en la fase de segmentación los algoritmos y programas en un computador DEC-10, simulando para las técnicas discutidas, entrada de datos de diferentes escenas, con lo cual se consigue abarcar diferentes situaciones que podrían presentarse en un proceso real.

Con la incorporación de la fase de almacenamiento dentro de la segmentación, se consigue que los objetos sean físicamente separados, finalidad de la segmentación. Además el uso de almacenamiento por figuras individuales permite el uso de menor cantidad de memoria secundaria, problema que siempre es importante en procesamiento de imágenes.

En el trabajo no son discutidas las condiciones de mejoramiento de la imagen de entrada, considerando que ya ha existido una etapa previa de adecuación y que el ambiente de trabajo es controlado como en el caso de aplicaciones de Robótica en procesos industriales, de fabricación.

Las fases que siguen a la de segmentación son las de extracción de atributos y clasificación, con las que se concluye el proceso de identificación de las figuras de una escena.

4. REFERENCIAS.

- 1.- R.C.Gonzalez and R.Safabakhsh, "Computer Vision Techniques for Industrial Applications and Robot Control," Computer, Dec,1982,17-32.
- 2.- A.Rosenfeld and A.C.Kak, Digital Picture Processing, Academic Press, N.Y.,1976.
- 3.- I.E.Abdou and W.K.Pratt, "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors," Procc. IEEE, Vol.67, No5, May1979
- 4.- C.P.Robinson, "Edge Detection by Compass Gradient Mask," CGIP.,6,429-501(1977)
- 5.- J.S.Weszka, "A Survey of Threshold Selection Techniques," CGIP.,7,259-265(1978)
- 6.- R.Stefanelli and R.Rosenfeld, "Some Parallel Thinning Algorithms for Digital Pictures," JACM,18,255-264(1971)
- 7.- K.Castleman, Digital Image Processing, Prentice-Hall, N.J.,1979
- 8.- B.Barraza.Sotomayor, "Reconhecimento de Padrões para Aplicações em CAD/CAM," Tese de Mestrado, UNICAMP-FEC, 1984.