

# LA ESTRUCTURA DE DATOS EN EL PROCESAMIENTO DE IMAGENES

ING. JORGE G. ERAZO, MSECE

**Abstracto:** El presente artículo hace una descripción de las estructuras de datos usadas en el procesamiento de imágenes y que están clasificadas dentro de las llamadas estructuras jerárquicas. Hace, además, una presentación de una de las técnicas de segmentación de imágenes: Descomposición Regular.

La demanda causante de procesos para el tratamiento de datos de imágenes, implica la aparición de nuevas aplicaciones. Muchas de éstas aplicaciones necesitan consideraciones de tiempo real y, con más importancia reconfiguraciones dinámicas de los sistemas debido al número de variables en la solución de los problemas de nuestros días (13).

Los avances tecnológicos que representa VLSI (very large scale integration) ha hecho posible contar con circuitería cada vez más pequeña y rápida, al tiempo de permitir poner en práctica la construcción de arquitecturas especiales de multiprocesamiento para la solución de problemas específicos. El fundamento utilizado para el diseño de los nuevos CHIPS es aquél de explícitamente implementar en hardware una estructura de datos requerida para la solución de una clase específica de problemas.

Muchas estructuras de datos han sido ya implementadas, pero de todas ellas, las estructuras jerárquicas son las que muestran ser las más convenientes para el procesamiento de imágenes principalmente por las siguientes razones.

- 1.- Da un conocimiento jerárquico de los objetos en la escena.
- 2.- Es representativo de la forma en que los humanos percibimos una escena.
- 3.- Conserva las relaciones geométricas entre pixels, lo cual facilita la identificación de características geométricas de los objetos en una escena.
- 4.- Facilita la implementación de técnicas de transmisión progresiva.

Extremadamente importante para el procesamiento de imágenes es el estudio de las estructuras de datos usados para la representación de imágenes y su contenido. Esto se debe a que el conocimiento de estas estructuras ayudará en el entendimiento de la naturaleza paralela - serie de la visión.

Una estructura de datos es la relación de elementos de datos. El escogitamiento de una estructura de datos apropiada puede simplificar enormemente la solución de un problema dado.

La estructura de datos para una imagen puede dividirse en:

- Estructura de datos no jerárquicas: Entre otras tenemos:

- \* Matriz de dos dimensiones

- \* De lazo
- \* Grafos
- \* Listas
- \* Bases relacionales

- Estructuras de datos jerárquicas: Son:

- \* Pirámides
- \* Descomposición regular
- \* Árboles de 4 elementos
- \* Conos de procesamiento

En este trabajo, presentaré una breve descripción de las estructuras de datos para imágenes que están clasificadas dentro del grupo de estructuras jerárquicas.

## a) Pirámides matriciales

Una pirámide matricial es una secuencia de arreglos de la forma:

$$\{M(n), M(n-1), \dots, M(i), M(i-1) \dots M(0)\}$$

donde:

- M(n) representa la imagen original
- M(0) representa un pixel
- M(i-1) representa una versión de M(i) a la mitad de la resolución de M(i).

Una pirámide permite la selección de posibles versiones de una imagen a diferentes grados de resolución. La escala de la resolución puede usar las potencias de dos, o puede ser más fina si se consideran las potencias de 2 y 3. En este caso, la pirámide puede tener arreglos cuyos tamaños pueden ser: 128, 108, 96, 81, 72, ... 6, 4, 3, 2, 1 si damos el límite "menor o igual" a 128.

Una de las ventajas de la pirámide matricial es que el algoritmo usado en el procesamiento de la imagen tiene acceso a ésta a diferentes niveles de resolución.

- b) Pirámide árboles es un conjunto de nodos-p de la forma:

$$P = \{(k, i, j)\}$$

donde:

k representa niveles de la imagen; por lo tanto:

$$0 \leq k \leq n$$

i, j son las coordenadas del arreglo original. Considerando una resolución de potencias de dos:

$$0 \leq i \leq 2^{**} (n-1); \text{ y} \\ 0 \leq j \leq 2^{**} (n-1).$$

La estructura pirámide T tiene una función F llamada padre que es definida para todos los P, excepto para P(0,0,0) que representa la raíz de la pirámide.

$$F(k, i, j) = (k-1, [i/2], [j/2])$$

donde:

$[i/2]$  ( $[j/2]$ ) representa el entero más alto posible que sea menor o igual a  $i/2$  ( $j/2$ ).

Cada nodo-p, por ejemplo 1, tiene cuatro nodos hijos, digamos 1'. Los nodos hijos

responden a la ecuación:

$$F(1') = 1$$

Para construir una pirámide T partiendo de un arreglo, se aplica una regla. Si por ejemplo, la regla escogida es la del promedio, entonces:

$$V(k, i, j) = \begin{cases} \frac{1}{4} \sum V(k+1, 2i+x, 2j+y) & \text{para } k < n \\ x = 0, 1 \\ y = 0, 1 \\ A(i, j) & \text{para } k = 1 \end{cases}$$

donde:

$V(k, i, j)$  es una función que mapea a un valor de rango tal como:

$$R = \{0, 1, 2, \dots, 255\}$$

$A(i, j)$  es el arreglo original

Otras reglas que pueden usarse son:

- Regla de muestreo
- Regla de nodos

### c) Descomposición Regular

El método es el siguiente:

- 1.- La imagen es arreglada en una matriz  $N \times N$ .
- 2.- Se define un nodo raíz
- 3.- La imagen se divide en 4 o 9 partes.
- 4.- Se define un nodo por cada subimagen.
- 5.- Cada nodo de subimagen es conectado a la raíz.

6.- El proceso continúa hasta que se alcanza el nivel deseado de resolución de imagen.

La ventaja de este método es que la imagen es dividida en matrices pequeñas y, por lo tanto, la imagen puede ser almacenada por grupos. Si la imagen no entra en la memoria, los grupos pueden ser almacenados en cualquier medio de almacenamiento y llamados cuando son requeridos en el proceso.

En la descomposición regular, el factor de división que limita el número de ramas-hijo que cada nodo puede tener, es variable.

#### a) Árboles de 4 elementos.

Topológicamente es idéntico a la descomposición regular pero siempre usa su factor de división de 2, de manera que cada nodo puede tener 4 hijos. Es posible hacer una reducción cuando los 4 hijos de un mismo nodo padre tienen el mismo valor. En este caso, los hijos son eliminados ya que el valor del padre los representa a todos.

Una aplicación que utiliza la estructura de árboles de 4 elementos es la de gráficos computarizados debido a la facilidad de manipulación en la computación como rotaciones, intersecciones, superposiciones, etc.

El problema de la determinación de cuales subimágenes son adyacentes a otras, ha llevado a métodos que usan una estructura modificada de árboles de 4 elementos llamada "enlazada".

Para el procesamiento de imágenes en un computador, es necesario considerar dos áreas principales:

- Segmentación: Que se refiere a la reducción de los datos de la imagen mediante la partición de la imagen de acuerdo con la relación escogida.
- Reconocimiento: Referida a la descripción de los objetos que pueden ser identificados dentro de la imagen.

La segmentación puede ser obtenida por dos métodos:

- \* Restringiendo el dominio del problema a su limitado número de variables de entrada.
- \* Usando técnicas de reducción.

Una de las técnicas usadas es la ya mencionada como estructura de datos: Descomposición Regular.

El método consiste en la subdivisión sucesiva de la imagen en búsqueda de áreas que son simples de mostrar gráficamente.

Áreas simples (de un mismo color, por ejemplo) requieren poco tratamiento; mientras áreas complejas (como una foto de un pueblo), son tratadas como subproblemas - pueden ser divididas hasta que son regulares o se llega al tiempo límite de ejecución.

La descomposición regular tiene dos objetivos:

- 1.- Dividir una imagen digitalizada en porciones que son:
  - informativas, Y
  - no informativas.
- 2.- Eliminar los elementos no informativos.

Si el algoritmo no puede decidir si una porción de la imagen es informativa o no, esa porción es dividida. La información es discriminada basándose en la intensidad, número de bordes, etc, o puede involucrar cálculos complejos.

La importancia de cada cuadrante es medida con relación a su cuadrante padre. La importancia relativa de un subcuadrante "a" dentro de un cuadrante A está definida como:

$$d(a,A) = \frac{\text{parámetro discriminante en subcuadrante a}}{\text{parámetro discriminante en un cuadrante A}}$$

Si los límites de informativo y no informativo de los cuadrantes son conocidos entonces se tiene un método de discriminación:

- $V2 \leq d(a,A) \leq 1.0$  INFORMATIVO
- $V1 \leq d(a,A) < V2$  INSEGURO
- $0.0 \leq d(a,A) < V1$  NO INFORMATIVO

donde:

- V1 = límite menor de la variable de discriminación.
- V2 = límite mayor de la variable de discriminación

La representación de una imagen dividida y subdividida se logra con mejores resultados si se utiliza la estructura de árboles de 4 elementos. Para entender esta estructura, es necesario introducir algunos conceptos:

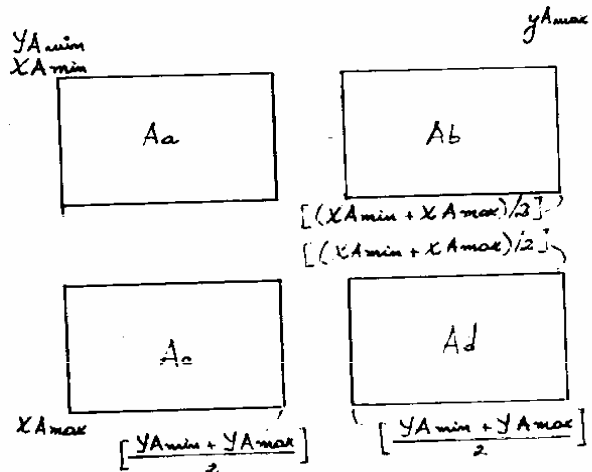
- CUADRANTE: Es la raíz del árbol. Pictóricamente es un nodo.
- CUADRANTE TERMINAL: Es un cuadrante que no tiene subcuadrantes
- IMAGEN REDUCIDA: Es un set de cuadrantes terminales.

Cada nodo en un árbol tiene propiedades que pueden ser expresadas como:

$$C = (\text{propiedad}, X_{min}, X_{max}, Y_{min}, Y_{max})$$

donde "propiedad" es cualquiera de las características de una imagen como intensidad, tamaño, etc.

Una vez que las propiedades del padre son conocidas éstas se computan para sus 4 hijos dividiendo cada lado del cuadrante por la mitad y asignando nuevos vértices a los subcuadrantes mediante el uso de las funciones FLOOR y CEILING. Esto permite tener cuadrantes aislados como se muestra en la figura



Si las propiedades del cuadrante son:

$$CA = [IA, XA_{min}, XA_{max}, YA_{min}, YA_{max}]$$

entonces, las propiedades de los 4 hijos son:

$$CAa = \left\langle \sum_{j=YA_{min}}^{YA_{max}} \sum_{i=XA_{min}}^{XA_{max}} I(i,j), XA_{min}, \left[ \frac{XA_{min} + XA_{max}}{2} \right], XA_{max}, YA_{min}, \left[ \frac{YA_{min} + YA_{max}}{2} \right] \right\rangle$$

$$CAb = \left\langle \sum_{j=YA_{min}}^{YA_{max}} \sum_{i=XA_{min}}^{XA_{max}} I(i,j), \left[ \frac{XA_{min} + XA_{max}}{3} \right], \left[ \frac{YA_{min} + YA_{max}}{2} \right], YA_{max} \right\rangle$$

$$CAc = \left\langle \sum_{j=YA_{min}}^{YA_{max}} \sum_{i=XA_{min}}^{XA_{max}} I(i,j), \left[ \frac{XA_{min} + XA_{max}}{2} \right], XA_{max}, YA_{min}, \left[ \frac{YA_{min} + YA_{max}}{2} \right] \right\rangle$$

$$CAD = \left\langle \sum_{j=YA_{min}}^{YA_{max}} \sum_{i=XA_{min}}^{XA_{max}} I(i,j), \left[ \frac{XA_{min} + XA_{max}}{2} \right], XA_{max}, \left[ \frac{YA_{min} + YA_{max}}{2} \right], YA_{max} \right\rangle$$

En el método hay dos consideraciones importantes:

- 1.- La división de cuadrantes es arbitraria
- 2.- Las funciones de discriminación usan solo los parámetros especificados.

Estas consideraciones traen consigo dos problemas:

- 1.- La división arbitraria de la imagen puede causar pérdida de información, haciendo aparecer áreas informativas como no informativas.
- 2.- El uso de la función de discriminación puede dejar una imagen reducida con información incompleta.

Para evitar estos problemas, es importante buscar áreas vecinas toda vez que estas pueden tener información adicional.

Estos cuadrantes vecinos se añaden a la estructura árbol desde que el padre real sea **NO INFORMATIVO**. En este caso, es necesario tener una función de discriminación que pueda ser el valor promedio de los cuadrantes informativos e inseguros o la culminación de pixels que están en ciertos bordes de los cuadrantes.

Una vez que el árbol ha sido armado, la descripción de objetos o áreas conectadas de la imagen puede ser hecha en términos de una subestructura específica. Por ejemplo, un objeto puede ser representado por varios cuadrantes terminales. En este caso es necesario decidir si los dos cuadrantes terminales pertenecen al último objeto. Este proceso se conoce como **PODA** y consiste en agrupar cuadrantes terminales que están geográficamente conectados.

Los cuadrantes conectados en grupos usando el criterio de los "8 vecinos"; que consiste en, dado un punto  $P_0(i, j)$  tiene 8 vecinos que son:

$(i-1, j-1), (i-1, j), (i-1, j+1), (i, j-1), (i, j+1), (i+1, j-1), (i+1, j+1), (i+1, j)$

y están representados como:

P1	P2	P3
P4	P0	P5
P6	P7	P8

Una vez que el árbol esté disponible, las regiones conectadas se obtienen a través de la aplicación de las siguientes definiciones:

- Dos cuadrantes  $Q_1$  y  $Q_2$  están conectadas si existen dos puntos  $P_1 \in Q_1$  y  $P_2 \in Q_2$ .
- Si  $X = \{Q_1, Q_2, \dots, Q_n\}$  es un cuadrante extraído, una región  $R$  es un sub-set de  $X$  tal que

$R = Q_{i1}, Q_{i2}, \dots, Q_{im}$

donde:

$\{i_1, i_2, \dots, i_m\} \subseteq \{1, 2, \dots, n\}$

de forma que:

$0 < j \leq m$   
 $0 < k \leq m$

donde:

$j \neq k$  ( $Q_{ij}$  conectado a  $Q_{ik}$ )

y de forma que

$P \{X-R\}$  ( $P$  conectado a  $Q_{ij}$ )

Los cuadrantes están conectados en regiones mediante el chequeo de puntos de borde al rededor de cada cuadrante. Usando la información de los vertices en cada cuadrante, la aplicación de las definiciones enunciadas permite obtener regiones de imagen conectadas a partir de los cuadrantes terminales.

#### CONCLUSIONES

Hay una tendencia a hacer de la interrelación hombre - máquina lo más natural posible.

Esto incluye la habilidad de hablar directamente a la máquina o la de mostrarle imágenes. Para conseguir esto es necesario contar con un hardware especializado para determinar las características básicas de la palabra y de la imagen; y con el desarrollo de un software capaz de inducir entendimiento del lenguaje o de la imagen. En resumen, existe la necesidad de un eficiente almacenaje de información de imágenes, de su sistema de procesamiento y de transmisión.

El contenido de este trabajo ha presentado solo una parte del problema y su solución con la sola intención de motivar futuras investigaciones.

#### REFERENCIAS

- NIK. A. ALEXANDRIDIS.- Image Processing in Supersystems; Ohio University, 1984.
- KASUHIRO FUCHI.- Aiming for Knowledge Information Processing Systems. Japón/81.
- A.P. REEVES.- Parallel Computer Architectures for Image Processing; Proceedings of the International Conference on Parallel Processing; OSU, Agosto 1981.
- S. FERNBACH.- Parallelism in Computing; Proceedings of the International Conference on Parallel Processing; OSU, Agosto 1981.
- C.C. GOTLIEB & L.R. GOTLIEB.- Data types and Structures; Prentice Hall, 1978.
- STEVEN L. TANIMOTO.- Image Data Structures; University of Washington, 1980.
- CHARLES DYER.- A VLSI Pyramid Machine for Hierarchical Parallel Image Processing; University of Illinois at Chicago Circle, 1981.
- A. KLINGER & C.R. DYER.- Experiment on Picture Representation Using Regular Decomposition; University of California, Los Angeles, 1984.
- CHARLES R. DYER.- Pyramid Algorithms and Machines; University of Illinois at Chicago Circle, 1982.
- R.J. BARON & L.G. SHAPIRO.- Data Structures and their Implementation; Van Nostrand Reinhold Co., 1980.



#### BIOGRAFIA

ERAZO M. JORGE GUSTAVO. Nació en Quito el 13 de Julio de 1954. Obtuvo el título de Ingeniero en Electrónica y Telecomunicaciones en la Escuela Politécnica Nacional en 1979. En 1981 egresó de Ingeniería Industrial en la Escuela Politécnica Nacional. En Marzo de 1985 obtuvo el título de Master of Sciences en Ingeniería Electrónica en Ohio University, USA. Siendo becario FULBRIGHT. Ha realizado estudios de Doctorado en Ohio University desde Marzo de 1985 hasta Septiembre de 1986, fecha en que retornó al país. Sus áreas de interés son el Control Microprocesado de Procesos, el procesamiento de imágenes y la arquitectura de computadores. Actualmente desempeña el cargo de Gerente de la División Electrónica de QUIPUS INFORMATIKA.