

V.1 ALGORITMO EVOLUTIVO PARA TRATAMIENTO DE IMÁGENES

*Angel Cesar Morales Rubio**

INTRODUCCIÓN

El presente trabajo es el resultado de seis años de investigación, que como se muestra, abarcó una amplia gama de especialidades o áreas de diversa índole, debido al interés particular por hallar un nuevo método de Reconocimiento de Patrones, que implementado en un Sistema Evolutivo para Tratamiento de Imágenes, lleve a éste a basar su funcionamiento de análisis en las características intrínsecas que éstas presentan.

Esta nueva alternativa responde a la necesidad de sintetizar una imagen de manera global con cierta naturalidad, teniendo un enfoque de paralelismo en la consecución de este objetivo.

Los frutos de esa investigación se encuentran contenidos en el presente documento estructurados en tres partes que son: Antecedentes, Análisis y Desarrollo.

I ANTECEDENTES

Reconocimiento de Patrones es el proceso por el cual sintetizamos heurísticamente la estructura de un objeto, encontrando así la firma o patrón característico de éste.

Pueden encontrarse algoritmos para reconocimiento de patrones que sin duda alguna cumplen con el objetivo de encontrar dicha firma en los objetos sobre los cuales actúan, éstos se apegan comúnmente a los siguientes lineamientos:

* Angel Cesar Morales Rubio elaboró este trabajo como parte de su tesis de Licenciatura en Ciencias de la Informática en la UPIICSA del IPN en 1992

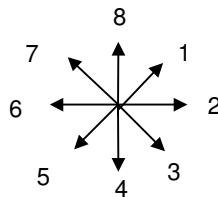
- 1) Se basan en refinamientos de métodos predecesores a ellos pero que siguen teniendo como origen una misma idea.
- 2) No captan de manera global el “significado” o “patrón” de una imagen.
- 3) En general el método de solución no utiliza paralelismo.

La mayoría de estos algoritmos tienen la peculiaridad de representar las imágenes del mundo exterior en un espacio de dos dimensiones, es decir, utilizan una matriz finita como modelo de representación. Esto es perfectamente lógico puesto que las figuras que intentamos reconocer “viven” en el mundo real aparentemente en un espacio de dos dimensiones. Por ello, una gran variedad de algoritmos (sin importar las diferencias en cuanto a su implementación) utilizan esta misma mesa de trabajo.

A continuación se describen las características más importantes de algunos algoritmos para Reconocimiento de Patrones, así como un resumen de sus principales ventajas y desventajas.

1 NÚMERO DE FORMA

Este procedimiento asume que dos puntos consecutivos de una imagen sólo pueden estar conectados entre sí en ocho direcciones, en similitud a los puntos cardinales:



Este método inicia tomando algún punto de la figura y marcándolo como inicial (figura 1), y si escogemos un recorrido en el sentido de las manecillas del reloj, podemos decir que existe una conexión de tipo 2 entre el punto de INICIO y su vecino derecho, siguiendo este procedimiento todo el objeto esta representado por

la oración 2222222244444666666666688888 o simplificando:
(9)2(5)4(9)6(5)8



número de forma: 2222222244444666666666688888

Figura 1. Una imagen y su número de forma

Así, si aplicamos el proceso en sentido inverso, (a partir de la secuencia de números) estaremos en la posibilidad de generar la imagen original; haciendo un análisis de este procedimiento encontramos dos problemas:

- 1) La sinterización de una *imagen sólida* requiere de la aplicación de algunos procedimientos adicionales.
- 2) Este método es altamente *sensible al factor de escala*.

2 ESQUELETOS DE UNA REGIÓN

Dado un objeto se divide este en lo que corresponde:

- 1) *Al borde*.
- 2) *Al interior*.

entonces, para cada punto en el interior se encuentra su vecino más cercano en el borde, si tiene más de un vecino se toma éste como el punto medio y pasa a formar parte del esqueleto del objeto, en caso contrario este punto se desecha y se toma el siguiente (figura 2). Resulta prácticamente imposible realizar una implementación directa de éste en un equipo de cómputo, ya que como se puede apreciar se tiene que estudiar cada punto individual y compararse con varios bordes en el peor de los casos, incrementándose en forma dramática el cálculo de distancias, además de existir objetos que presentan dificultades, como es el caso de círculos, y figuras abiertas.

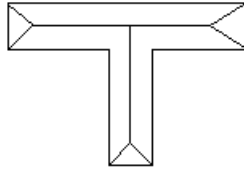


Figura 2. Esqueleto de una región

3 CUATRICES

El método toma como origen una figura en una matriz comúnmente cuadrada, y la divide en cuatro regiones iguales, como si fuera un plano cartesiano (figura 3), de ahí analiza cada cuadrante y verifica si hay información en él, si la hay divide este cuadrante a su vez en cuatro partes como sucedió originalmente, y aplica el mismo procedimiento en forma recursiva hasta encontrar un cuadrante que ya no se puede dividir, y después de haber analizado todos los cuadrantes de todos los niveles se obtiene una estructura en árbol que representa el objeto original:

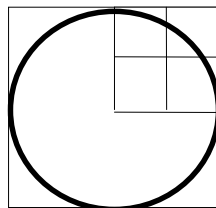


Figura 3. Una imagen con cuatrics

1) *La implementación* claramente puede realizarse en un ambiente de múltiples procesadores puesto que se puede asignar un procesador a cada uno de los cuadrantes que se vayan encontrando, existiendo la limitante teórica de crear un equipo de cómputo con un número ilimitado de procesadores,

lo que claramente marca una deficiencia en la funcionalidad del método, respecto al número de cuadrantes a procesar.

2) *El patrón o firma de la imagen* parece estar compuesta por una estructura de árboles, siendo la raíz el primer nivel o imagen original consiguiendo de esta manera un esqueleto, como en el método anterior, pero con una representación diferente.

II ANÁLISIS

Con el fin de encontrar un método de Reconocimiento de Imágenes que sintetice a éstas, se conjugaron tres métodos naturales de modelación de la realidad, estos métodos son: Visión Humana, Redes Neuronales y Fractales.

1 VISIÓN HUMANA.

Definitivamente el ser Humano parece ser una de las criaturas más perfeccionadas en este planeta en cuanto a su anatomía se refiere. Desde un punto de vista funcional, nuestros ojos dividen el campo visual en cuatro regiones simétricas. Esto es gracias a la distribución de las conexiones entre las fibras ópticas y la corteza visual del cerebro.

De lo anterior podemos concluir el siguiente hecho relevante en cuanto al sistema de captación de imágenes del ser Humano:

El mundo exterior a los ojos se encuentra dividido en cinco zonas, cuatro de las cuales tienen una correspondencia directa con la estructura anatómica de los ojos y la quinta más bien parece tener una función integradora de la realidad.

De esto obtenemos como hipótesis que el concepto de Cuatrices tiene semejanza a la estructura funcional del sentido de la vista humano.

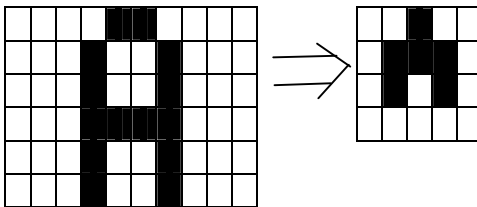
2 REDES NEURONALES FRACTALES

Explicar las ideas que sustentan cada uno de estos conceptos, esta fuera del alcance del presente documento, sin embargo el uso de estos será únicamente desde un punto de vista intuitivo.

III DESARROLLO DE LA SOLUCIÓN

Si dibujamos un objeto en una rejilla finita y definimos equivalencias de nivel, como se muestra en la figura 4 podemos realizar abstracciones continuas de la imagen original, hasta llegar a un solo elemento que contendría de manera sintetizada y global toda la información del objeto original.

Con la ayuda de la figura 4 podemos conceptualizar fácilmente que, el objeto original está disperso en la estructura de la pirámide que se forma al poner los diferentes niveles de abstracción, uno sobre otro. Entonces ya que toda la figura está dispuesta en un espacio de tres dimensiones, concluimos que para analizar imágenes en 2D necesitamos herramientas que actúen en 3D, resulta interesante pensar en el problema que se plantea al reconocer imágenes en 3D originalmente.



Equivalencias:

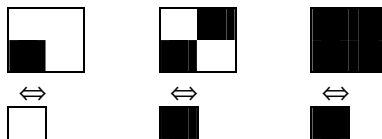


Figura 4. Una rejilla y equivalencias

Regresando a la visión humana recordemos que ésta divide el espacio en cinco regiones, si intentamos aplicar el concepto de pirámide dentro del contexto de las cinco "regiones humanas", tenemos como resultado una pirámide por cada región.

De lo anterior se concluye lo siguiente:

- 1) Se puede *asignar un procesador para el control de cada una de las pirámides de la periferia y otro para la pirámide central* actuando como integrador de las otras cuatro, así el tiempo de reconocimiento de un objeto se decrementaría.
- 2) Si se analizan las áreas que tienen en común las diferentes pirámides, podemos inferir que los puntos del objeto en esas regiones se encuentran definidos en dos lugares diferentes, lo que nos lleva a *redundancia en la información* y para poder quitarla sería necesario que los procesadores pudieran comunicarse entre sí, lo que implicaría que el modelo se estuviera moviendo con el propósito de automantenerse, y como el modelo es una imitación de la realidad posiblemente ésta también se estuviera moviendo en su estado original.
- 3) Es factible encontrar una solución en la que se obtengan *múltiples pirámides* controladas por varios procesadores, sin tener la limitante de cinco impuesta por la visión humana. Este enfoque requiere que el método de análisis de la imagen encuentre el número de procesadores dependiendo de la estructura interna del objeto de estudio.

Así, el Algoritmo Evolutivo para Tratamiento de Imágenes contempla las siguientes habilidades:

- 1) Utilizar como medio de análisis el *enfoque piramidal* sustentado en que para analizar imágenes en 2D debemos utilizar herramientas de 3D, de tal forma que encuentre una representación global del objeto original.
- 2) Debe encontrar la *firma característica* de la imagen que se le esta presentando, basándose en la estructura de ésta.
- 3) El proceso de *análisis debe ser evolutivo*, es decir, conforme avance en la elaboración de la síntesis de la imagen, vaya

guiando su proceder con base a la experiencia obtenida hasta ese momento.

Dadas las características anteriores, comenzaremos con la descripción del planteamiento propuesto que contempla los requisitos de operación detallados.

Teniendo como referencia la cordillera que se forma con las pirámides por cada región de la vista humana, cada una de las pirámides puede contener a su vez una gran variedad de "pirámides", como las llamaremos, de menor tamaño o grado. Tendremos pirámides de grado uno, dos, tres, etc., dependiendo de cuanta información contenga el objeto analizado.

También podemos plantear que teniendo la pirámide de tamaño 1 como fractal fundamental, podemos construir a partir de éste la pirámide total del objeto de estudio. Sin embargo describimos a continuación otro enfoque que puede ser más interesante.

Suponiendo que nos encontramos analizando una pantalla en la que se encuentra una letra A, acompañada de otros elementos que pudieran ser considerados como ruido, entonces dependiendo del grado de información a lo largo y ancho de la imagen podemos formar una especie de zona montañosa, que abarca toda la superficie original, donde los picos más altos revelan zonas con una alta densidad de información, logrando obtener por así decirlo, la montaña característica de la letra A, además del patrón del ruido que la rodea. Así de esta forma tendremos que no solo hemos analizado la A original, sino que además se ha captado de una manera global el significado de la escena donde se encuentra ésta.

El enfoque de solución empleado para alcanzar los objetivos propuestos, es tener como vía de solución una derivación de las redes neuronales.

Recordando los planteamientos descritos en antecedentes, las imágenes se trasladan a un espacio bidimensional sobre la cual se aplican los algoritmos que se deseen. El algoritmo de solución abarca el análisis de la imagen con varios colores, iniciando con la

lectura de ésta y transportándola a una matriz bidimensional, en la cual cada píxel de la pantalla se encuentra en un nivel de abstracción cero (0) o inicial.

Se establece que el comportamiento de cada píxel en la matriz depende de sus vecinos (red neuronal) tanto en el mismo nivel como en los niveles superiores, llegando a ocupar una posición aleatoria dentro de la "montaña" general de la imagen. Después, si listamos todos aquellos picos que forman parte de las crestas de las montañas que fueron encontradas tendremos como resultado la *"Firma Característica de toda la Imagen Analizada"*.

