

I.7 SI-VE: SISTEMA DE VISIÓN EXPERTO*

*Ma. Sandra Camacho V., G. Patricia Gómez R.,
Jesús M. Olivares C.*

Resumen

Este trabajo describe una herramienta para el reconocimiento de imágenes usando entre otros el método de rastreo mediante la técnica de primeros vecinos para obtener el número de forma. Los patrones identificados se ubican y se relacionan para describir la escena completa.

INTRODUCCIÓN

Desde el descubrimiento de los metales hasta la fabricación en serie de computadoras sofisticadas de hoy en día, el hombre ha sabido mejorar y modificar su medio ambiente desarrollando investigaciones para diversas aplicaciones. Es así como el hombre ha evolucionado hasta llegar a los avances de los descubrimientos que actualmente son aplicados en todas las ramas de la ciencia. Una de esas ciencias es la computación, en la que se ha venido escudriñando últimamente, llegando a hacer uso de todo lo que se tiene para integrar herramientas muy sofisticadas y poderosas, teniendo entre ellas a la Inteligencia Artificial (IA).

Si bien es cierto que la IA ha tenido gran auge, también es cierto que son muchas y muy variadas sus aplicaciones, tal es el caso de la robótica, los sistemas expertos y el reconocimiento de patrones.

Los autores realizaron este trabajo como parte del curso de Lenguajes de Nivel Alto en el 4º semestre de la Licenciatura en Ciencias de la Informática en la Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas del Instituto Politécnico Nacional (UPIICSA-IPN), concluyéndolo en julio de 1986 y reportado en agosto del mismo año. El Sistema de Visión Experto funciona en una computadora HP-3000 con sistema operativo MPE.

Se pretende pues, seguir buscando materia y explotando el que ya se tiene para llegar a formar herramientas más y más adecuadas a las necesidades planteadas por el hombre. Uno de los temas de mayor aplicación dentro de la IA es el de “Visión por Computadora”, el cual se enfoca entre otras aplicaciones, a la percepción visual de los robots para su uso industrial.

El problema de desarrollar sistemas de Visión por Computadora se puede dividir en los problemas de adquisición, tratamiento e interpretación de una imagen.

- 1) La *adquisición* utiliza dispositivos tales como cámaras de televisión, láser, ultrasonido, infrarrojos, etc. y de un convertidor analógico-digital para captar y digitalizar la imagen.
- 2) El *tratamiento* incluye la preparación y mejora de la información adquirida (filtrado, ecualización del histograma, reconstrucción, etc.) y la segmentación; es decir la partición de la imagen en zonas con propiedades similares.
- 3) La etapa de *interpretación* de una imagen es la más compleja, pues de ella depende el éxito de la interpretación de la escena. Consta de dos procesos principales: extracción de características o primitivas y el análisis y manipulación de las mismas.

Uno de los trabajos realizados en el campo de visión por computadora, que se enfoca a la tercera etapa (interpretación de la imagen) trabajando con figuras bidimensionales es el Sistema de Visión Experto “SI-VE” cuyo nombre se debe a que posee la capacidad de aprender nuevos patrones.

SI-VE es el sistema que se analiza en este artículo con el propósito de proporcionar los métodos desarrollados para resolver el problema de la interpretación, llamada también "Reconocimiento de Imágenes”.

I EL PROBLEMA DE VISIÓN POR COMPUTADORA

La robótica trabaja continuamente en lo que es la percepción visual por computadora y procesamiento de imágenes, por lo que han sido desarrolladas numerosas investigaciones al respecto.

Así pues, la visión artificial es un tema causante de varias controversias y estudios que implican el reconocimiento de imágenes así como la ubicación de tales imágenes o figuras que conforman una escena dada.

Aunque resulte familiar el término “Visión por Computadora”, no es igualmente familiar el método y los medios utilizados para cubrir los requisitos necesarios para ello. Esto no implica que sea algo imposible de realizar, pero sí lo bastante abstracto y complicado como para tratarse con la delicadeza y cuidado necesario al desarrollar estudios minuciosos que conlleven a una solución exacta y general que indique el camino a seguir para cualquiera que sea su aplicación por muy específica que ésta sea.

A medida que se profundice en el tema con investigaciones, estudios y experimentos comprobados y puestos en práctica, se proporcionan, a su vez, nuevas herramientas útiles de las cuales partir para llegar a la construcción, de lo que se tiene como máxima pretensión en el área de “Robots Inteligentes” capaces de cubrir cualquier tarea que se les encomiende.

La extracción de características es la primera y más importante etapa de SI-VE, de la cual depende la eficiencia del sistema, la cual pretende llegar a la representación estructural única y exacta de cada figura.

El análisis de la representación estructural elabora un modelo base en forma de red, el cual contiene información tal como distancias y ángulos. Dicho modelo se encuentra listo para ser comparado con los patrones.

La descripción de la escena, que es la última etapa del sistema, establece la relación de todos los elementos que componen la

escena con respecto a su universo y la relación existente entre ellas mismas dentro de tal universo.

El crear un método exacto, rápido y confiable para la identificación y descripción de escenas es el objetivo principal que motivó la realización del SIstema de Visión Experto.

II SISTEMA DE VISIÓN EXPERTO PARA EL RECONOCIMIENTO DE ESCENAS

La obtención de características de los contornos de las figuras que conforman una imagen, es la pauta que sirve de partida hacia el reconocimiento de escenas compuestas de una o más figuras.

Una vez conocida la problemática involucrada en el área, se tiene que los pasos o puntos principales a cubrir para llegar a una efectiva obtención de resultados en reconocimiento de imágenes es la siguiente:

1) *Aceptación de la escena*: Se captura la imagen que contiene la información a procesar a través de una matriz de $M \times N$ caracteres en la que se vacía dicha información. La imagen puede estar formada por una o más figuras de tipo lineal, geométricas, compuestas o en forma de estrella. La digitalización de la escena, que será la forma de capturarla puede ser por pantalla o por editor.

2) *Detección del número de figuras que conforman la escena*: Como resultado de un primer seguimiento o búsqueda general de la escena captada, se encuentra el número de figuras que conforman la imagen, distinguiendo cada una de ellas de acuerdo al carácter de identificación asignado para su reconocimiento.

3) *Rastreo de figuras detectadas*: Una vez encontradas las figuras integrantes de la escena, se procede a hacer el rastreo por figura consistente en ir siguiendo punto por punto el contorno de la figura por medio del método de primeros vecinos, el cual se realiza siguiendo las localidades de la

matriz que han sido, ocupadas por la figura, partiendo del punto inicial.

4) *Análisis de figuras*: A partir del rastreo efectuado para cada figura, se obtiene el modelo correspondiente para cotejar con lo almacenado como patrón.

5) *Descripción de la escena*: La descripción de la escena es la parte culminante del proceso, la cual consiste en determinar la relación y posición que guarda cada figura componente de la escena.

Cada uno de los pasos mencionados involucran el uso de varias técnicas y métodos de Inteligencia Artificial y Reconocimiento de Patrones.

En general el proceso seguido para llegar al objetivo final es el que se observa en la figura 1.

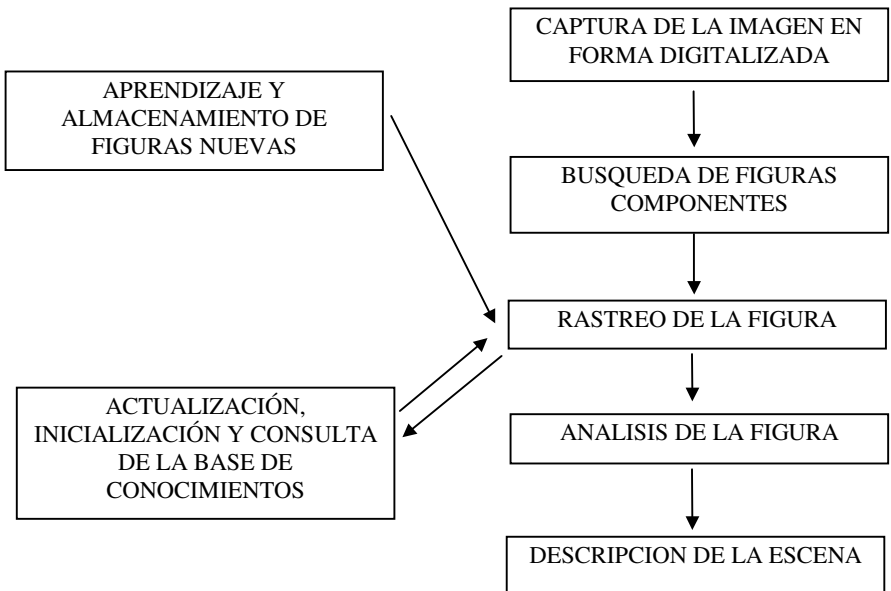


Figura 1. Etapas del método del sistema de visión por computadora

El aprendizaje es utilizado cuando la figura analizada no se encuentra en la base de conocimientos, en cuyo caso se puede o no adoptarla como patrón de acuerdo a la decisión del usuario.

La etapa de actualización del sistema permite dar de baja, modificar o consultar los patrones.

Una particularidad del sistema es que en principio sólo puede distinguir a las figuras dependiendo de la siguiente clasificación:

- 1) *Punto.*
- 2) *Líneas.*
- 3) *Figuras en forma de estrella.*
- 4) *Figuras compuestas.*

Al analizar figuras que no reconoce, realiza un aprendizaje guiado por la persona que maneja el sistema.

La arquitectura modular del SI-VE se muestra en la figura 2.

III MÉTODO DE RASTREO

Consiste en seguir una figura punto por punto, utilizando el método de primeros vecinos, adaptado especialmente para detectar vértices, puntos con bifurcaciones y puntos finales. Se llaman vecinos de un punto a todos los puntos que lo rodean. Estos vecinos tienen una jerarquía que determina la dirección a seguir. El valor dado a cada dirección se puede observar en la figura 3.

En cada avance se almacena la dirección respectiva para formar lo que se conoce como número de forma, que representa a la figura, con el cual, posteriormente, se analiza la inclinación de las rectas de la figura en análisis.

Cuando hay un cambio de dirección se toma el punto de análisis como fin de recta, e inicio de una nueva en el caso de que haya sólo dos vecinos. Cuando hay tres o cuatro vecinos se realiza un análisis para cada uno de ellos, que determine la

pendiente y con ello el punto que realmente deba ser tomado como inicio de esta nueva recta, almacenándolo para su posterior seguimiento. Véanse las figuras 4a y 4b.

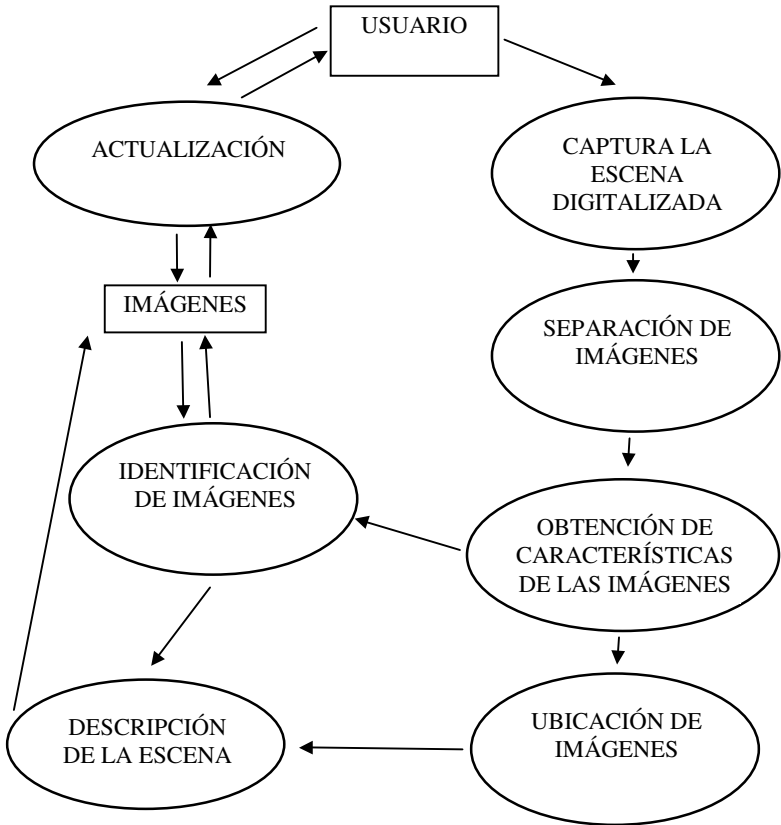


Figura 2. Arquitectura del SI-VE

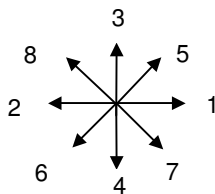
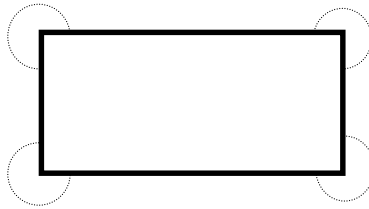
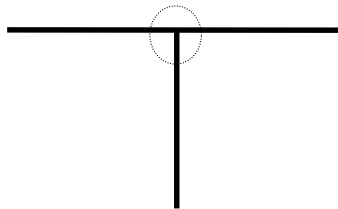


Figura 3. Direcciones de primeros vecinos



a. Dos vecinos para el punto en análisis



b. Tres vecinos para el punto en análisis

Figura 4. Cambios de dirección

En el caso de que un punto cuente con más de cuatro vecinos, se analizará el siguiente punto hasta llegar a aquel donde convergen las rectas, dado que cada vecino representa el inicio de una recta. Tal es el caso de las figuras 5a y 5b.

IV IDENTIFICACIÓN DE IMÁGENES

El proceso de identificación se realiza una vez que fueron extraídos los datos de las figuras, los cuales son:

- 1) *Los puntos que son terminales de una recta, vértices o intersección entre rectas.*
- 2) *Las rectas que tiene la figura.*

Lo anterior se almacena en una estructura que evita la repetición de puntos y rectas (base de datos), con los datos anteriores se genera la estructura en forma de red de la figura en

análisis la cual se empleará para ubicarla como patrón. Además de la estructura se obtiene la longitud de sus segmentos y los ángulos que contiene.

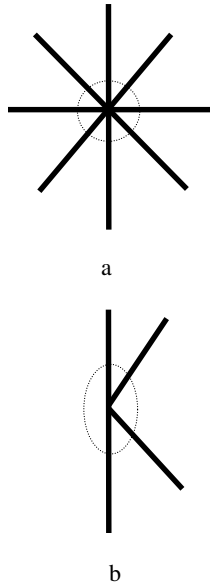


Figura 5. Cambios de dirección

Hecho lo anterior, se procede a ubicar a la figura según la familia topológica a la que más se parezca, en una de cinco posibilidades, que son:

- 1) *Punto*: Se trata de figuras que constan sólo de un punto.
- 2) *Cerrada*: Se especifica de esta forma a las figuras que terminan en el mismo punto en el que inician (figura 6a).
- 3) *Lineal*: Son aquellas figuras que tienen forma de línea y nunca llega a cerrarse (figura 6b).
- 4) *Estrellada*: Las figuras de esta familia se caracterizan por tener un punto central y líneas que salen de ella (figura 6c).
- 5) *Compuesta*: Son las figuras que tienen una combinación de las anteriores (figura 6d).

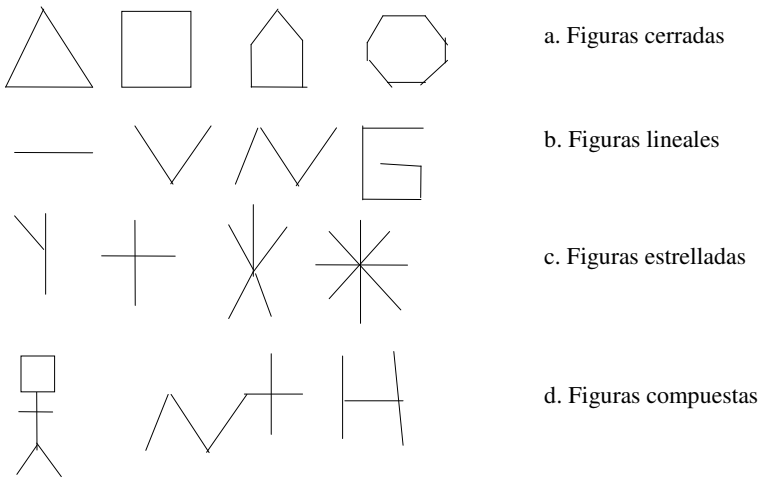


Figura 6. Familias topológicas

Una vez que la figura queda reconocida en alguna de las familias topológicas, se busca alguna estructura a la cual se parezca. Para esto se emplea un isomorfismo de redes, si la encuentra procede a tomarla como patrón de semejanza y continua el análisis como se describe posteriormente; de otra forma, si no existe alguna estructura a la cual se parezca la figura en análisis, el sistema preguntará si la adopta como patrón o no; si llega a hacerlo almacena los datos que ya tiene y le pide un nombre al usuario para etiquetar la estructura en cuestión. El siguiente nivel de análisis es por figura específica donde el patrón analizado además de pertenecer a la misma familia topológica y tener la misma estructura, posee ángulos y medidas proporcionales al patrón, en cuyo caso se procederá a realizar un aprendizaje, pero ahora por figura específica. *El resultado es el nombre de la figura según el nivel de análisis al cual se haya llegado.*

V DESCRIPCIÓN DE LA ESCENA

El conocimiento de la figura en sí, muchas veces no es de utilidad por sí solo, pues el uso de la información se da cuando se establecen las relaciones entre estos conocimientos. Hablando de figuras, supongamos una escena cuyos componentes sean letras, es de gran importancia establecer la relación o agrupación de las mismas para conocer el significado de la escena en su totalidad. En el caso de la industria, la posición de las herramientas en la banda y la separación de cada una determinan el movimiento del robot.

Dependiendo del uso o aplicación de la información que se desee obtener, será el planteamiento de la descripción. En el caso del SI-VE la aplicación general utiliza una forma bastante sencilla aunque no muy exacta para describir la escena lo cual se debe a la consideración y libertad de poder modificar en la medida que se requiera dependiendo de su uso. En ella se establecen dos tipos de relaciones. Primeramente cada figura de acuerdo a la posición que mantiene dentro de la escena; y el segundo la relación que guarda cada elemento con los demás. Llamemos a la primera "ubicación" y a la segunda "relación".

1) *La ubicación:* Se realiza tomando las características extraídas de la figura y con ellas se sobrepone un cuadrilátero que abarca toda la extensión de la figura, dicho cuadrilátero se ubica en la escena según el lugar que ocupe dentro de la misma. Por ejemplo: arriba, abajo, arriba a la derecha, izquierda, etc.

2) *La relación:* Se efectúa después de la ubicación, consiste en determinar que posición ocupa una figura con respecto de otra, comparando todos los puntos que pertenecen a la figura y buscando, entre todas las posibilidades la que sea la más adecuada. Las posibilidades son: arriba, dentro, fuera, junto, encima, etc.

CONCLUSIONES

Cuando alguien habla de casos como el hecho de ver, oír, hablar, entender, etc. se piensa que son funciones obvias y el realizar un programa o sistema que lleve a cabo esto, no tiene mayor problema. Avocándonos al problema particular visto en el desarrollo del presente trabajo, podemos apreciar que no es algo trivial pues involucra una extensa investigación para poder comprender cómo es que los seres vivos ven. Para que un sistema artificial pueda, aunque sea de una forma muy restringida, simular este proceso. Las soluciones planteadas en los estudios hasta hoy realizados no son únicas y definitivas; la solución más adecuada sólo puede obtenerse si se relacionan todas las herramientas y métodos desarrollados en la Inteligencia Artificial.

El beneficio que aporta un dispositivo de visión artificial sería amplio, ya que se utilizaría en robots, para realizar actividades que para el hombre son muy peligrosas, rutinarias o pesadas; ayudará a invidentes o bien en ampliar el campo de visión de tal manera que le permitan ver en lugares donde no puede llegar, como es el caso del océano o del espacio.