

La Ciencia, desde Morelos para el mundo

De la Visión Natural a la Visión Artificial: Segunda Parte

Por L. Enrique Sucar, INAOE,
Puebla
Miembro de la Academia de
Ciencias de Morelos

Al surgir las primeras computadoras hace más de 50 años, algunos científicos se preguntaron si esas máquinas podrían realizar tareas que asociamos con la inteligencia humana como resolver problemas, entender el lenguaje o analizar imágenes. Se empezaron entonces a desarrollar programas de computadora que buscaban emular la inteligencia humana, naciendo una nueva disciplina que se conoce como *inteligencia artificial*. En los primeros años se tuvieron avances aparentemente muy rápidos, con programas de computadora que juegan ajedrez, demuestran teoremas matemáticos o diagnostican ciertas enfermedades.

Ante tan impresionantes avances, cuenta la leyenda que un profesor encargó a un estudiante que resolviera el problema de visión en un verano; ¡aún no termina su proyecto! A través del tiempo los investigadores en inteligencia artificial nos hemos percatado de que las capacidades aparentemente fáciles para nosotros, como el ver, entender el lenguaje, planear nuestras actividades, etc., han resultado las más difíciles para las computadoras.

La visión artificial comienza, en forma análoga a la humana, con una imagen que es capturada a través de una cámara y guardada en la memoria de la computadora. Para la computadora la imagen es como un arreglo o matriz de puntos, llamados píxeles. Cada píxel tiene asociados un valor numérico, por ejemplo entre 0 y 255. En el caso de una imagen en blanco y negro, cada punto representa la intensidad o brillantez de la imagen. Las imágenes a color se representan con 3 arreglos de puntos, uno para el rojo, otro para el verde y otro para el azul; su combinación representa los diferentes colores. Una vez que se tienen las imágenes en la computadora, inicia el procesamiento de las mismas mediante programas de software, para intentar analizarlas y reconocer los objetos presentes en la imagen.

El análisis se va realizando por etapas, de cierto modo en forma similar a la visión biológica. Mediante



Esta imagen ilustra el proceso de segmentación de imágenes en la computadora. Podemos observar cómo se han separado el cielo, la piscina, y algunas partes de la casa (señalados con líneas blancas).

la estimación de las diferencias entre píxeles vecinos en la imagen, se obtienen los llamados bordes que indican donde termina un objeto y empieza otro. De esta forma se pueden delimitar los contornos de los diferentes objetos, o sus partes, en la imagen. También se puede hacer esto uniéndose píxeles que tienen características similares en color o intensidad. Una vez separadas estas regiones, lo que se conoce como *segmentación*, se obtienen características distintivas de las mismas, como su color, su forma o su textura. La textura se refiere a propiedades estadísticas de los objetos que los hacen no uniformes, como una alfombra, la madera o un bosque visto desde lejos. En base a dichas características se pueden distinguir algunos objetos o zonas de la imagen como el cielo, una mesa, o un tigre. Sin embargo, en general es muy difícil poder distinguir algunos objetos, ya sea porque varían en su apariencia visual (como las diversas clases de flores) o porque se ven distintos desde diferentes puntos de vista (como una persona vista de frente, de perfil o por atrás) o porque son muy complejos y están a su vez formados por otros objetos (como una cara humana). Por ello, los sistemas de visión

actuales se enfocan a reconocer un pequeño grupo de objetos, como puede ser el reconocer automóviles, partes industriales o personas. Pero no se ha podido desarrollar un sistema que tenga la capacidad de reconocer muchos objetos, con propiedades muy diferentes, como lo hacemos las personas.

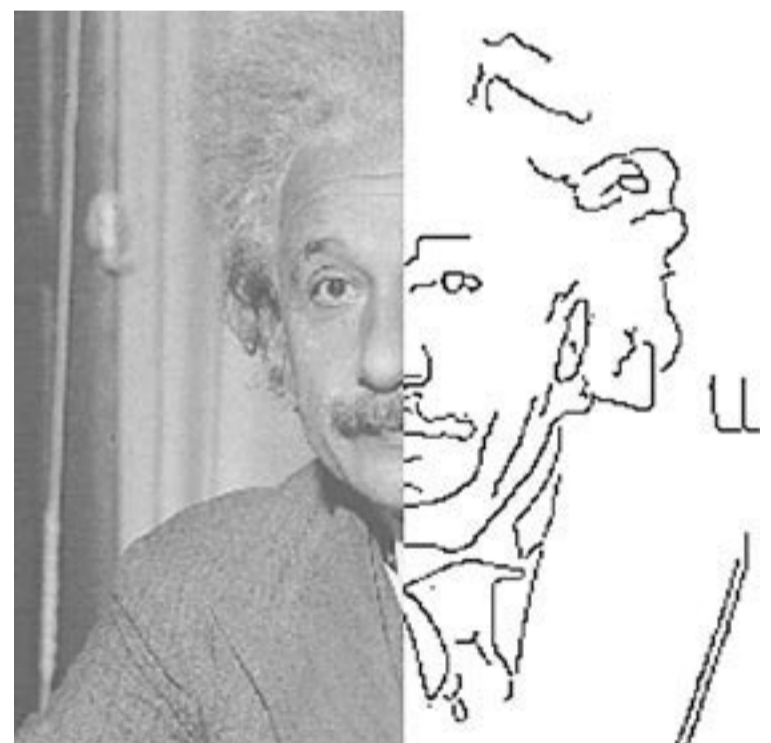
Otro aspecto en el cual ha tenido avances significativos la visión artificial es en recuperar la tercera dimensión a partir de imágenes. Una imagen tiene sólo 2 dimensiones, ¡pero el mundo es tridimensional! Si queremos que un robot pueda tomar un objeto, o estimar la distancia a la pared para no chocar con ella, queremos estimar la tercera dimensión. Existen varias estrategias para hacer esto. La más directa, en forma análoga a nosotros, es utilizar dos cámaras, en lo que se conoce como *visión estereoscópica*. Para ello la computadora analiza las imágenes obtenidas por dos cámaras, separadas cierta distancia. Al encontrar los puntos correspondientes entre una imagen y la otra, y medir su diferencia en posición (*disparidad*), se puede calcular la distancia a dicho punto en el mundo mediante relaciones geométricas. Sin embargo, determinar los puntos correspon-

dientes entre el par de imágenes no es una tarea fácil. Otras estrategias para recuperar la profundidad se ba-

san en los cambios de luminosidad debidos a la forma del objeto (forma de sombreado), a las distorsiones de las texturas por su orientación (forma de textura), o en usar video; es decir, varias imágenes consecutivas moviendo la cámara o el objeto (forma de movimiento).

Los avances en la visión artificial han permitido desarrollar diversas aplicaciones prácticas de sistemas de visión por computadora. Por ejemplo, hay sistemas en plantas industriales que ayudan a detectar defectos de fabricación en ciertas partes o productos. También hay aplicaciones en medicina, con sistemas que ayudan a los médicos a analizar imágenes de rayos-X o tomografías para encontrar tumores, por ejemplo. Se ha utilizado la visión artificial en robots lo que les permite no chocar con las paredes y localizarse en un mapa del entorno. También existen programas de visión que pueden distinguir huellas o caras de personas para identificarlas en sistemas de seguridad.

A pesar de estos avances aún estamos lejos de un sistema de visión artificial que se tenga las capacidades de la visión natural humana, en cuanto a su generalidad y flexibilidad. Posiblemente el llegar a comprender mejor como es que funciona el sistema visual en nuestro cerebro, nos ayude a desarrollar computadoras y robots con mejores capacidades visuales, ¡y nuestro legendario amigo pueda concluir su proyecto de verano!



Esta fotografía de Einstein muestra el proceso de detección de bordes en una imagen. Del lado izquierdo vemos la imagen original, del lado derecho los contornos obtenidos mediante la detección de bordes por la computadora.