

Discurso de aceptación

20 de junio de 2024

Takeo Kanade, galardonado en la categoría de *Tecnologías de la Información y la Comunicación* (XVI edición)

Es para mí un gran honor recibir el prestigioso Premio Fundación BBVA Fronteras del Conocimiento en la categoría de Tecnologías de la Información y la Comunicación.

La visión artificial o por ordenador es el campo que permite dotar de las funciones que realizan los «ojos» a ordenadores y robots. He tenido ocasión de presenciar el tremendo avance que se ha producido en este campo y de participar en él desde sus inicios, en 1970, hasta las indispensables y omnipresentes tecnologías actuales.

Los humanos somos excelentes máquinas de visión. Los primeros enfoques de la visión por ordenador fueron heurísticos y *ad hoc*, y se basaban sobre todo en la reflexión introspectiva de los investigadores. Podríamos caracterizarlos como enfoques del tipo «programemos lo que creo que estoy haciendo». Pero, sin embargo, como los datos visuales son resultado de procesos ópticos, geométricos y físicos, consideré que debíamos centrarnos más bien en los modelos matemáticos y científicos subyacentes. Con este enfoque más riguroso, mis alumnos y yo logramos crear varios algoritmos fundacionales de la visión artificial. Uno de ellos es un algoritmo conocido como el flujo óptico de Lucas-Kanade, que en definitiva es la base de casi todo lo que abarca la codificación de vídeo, como el seguimiento del movimiento y la compresión de vídeo. Otro ejemplo está en el algoritmo de factorización de Tomasi-Kanade, que resultó ser uno de los primeros algoritmos de funcionamiento para el problema de reconstruir fielmente la estructura de una escena tridimensional a partir de un conjunto de imágenes, algo que durante mucho tiempo había sido un reto para los investigadores.

A lo largo de mi trayectoria, he disfrutado mucho creando soluciones para multitud de problemas prácticos. El rostro humano ha sido uno de mis temas

favoritos. Mi tesis doctoral, realizada allá por 1973 en la Universidad de Kioto (Japón), es reconocida como el primer programa informático que localizaba automáticamente rasgos faciales como la nariz, los ojos y la boca en imágenes digitales utilizadas para el reconocimiento facial. Era capaz de procesar ochocientas imágenes: un número modesto para los estándares actuales, pero inusualmente grande en aquella época. El reconocimiento facial por ordenador es habitual hoy en día, desde el inicio de sesión en nuestros móviles hasta el control de pasaportes en los aeropuertos. A mediados de los años 90, demostré la fiabilidad de la detección facial con algoritmos de aprendizaje automático de redes neuronales, una tecnología hoy omnipresente en los teléfonos móviles. Más adelante, en la primera década de los años 2000, me pasé al reconocimiento automático de expresiones faciales. Nuestro equipo, en colaboración con investigadores del campo de la psicología, desarrolló métodos para reconocer microexpresiones faciales pormenorizadas junto con movimientos de la cabeza y el cuerpo, que se utilizan para enriquecer la interacción entre humanos y robots, así como para el diagnóstico médico y psiquiátrico.

La conducción autónoma es otro de mis temas favoritos. A mediados de los 80, fundé un proyecto en este campo en la Universidad Carnegie Mellon. Desarrollamos una serie de funciones para la conducción autónoma basada en la visión artificial, como seguimiento de carriles, detección de coches y peatones, planificación de trayectorias y estacionamiento paralelo. Estas funciones se instalaron en un vehículo controlado por ordenador y equipado con cámaras que recibió el nombre de NAVLAB (abreviatura de Laboratorio de Navegación). En 1995, el NAVLAB-5 de quinta generación logró «Cruzar América sin manos» por la autopista de Pittsburgh a San Diego, aproximadamente 5000 kilómetros con un 98,2 por ciento de control autónomo de la dirección. Fue un importante hito en la historia de la conducción sin conductor. Durante el trayecto, un policía del estado de Texas se acercó a nuestro vehículo, al parecer sospechando del cartel de «Nadie a bordo» que llevaba. Sin embargo, al descubrir que se trataba de un vehículo experimental de conducción autónoma, se ofreció a acompañarnos hasta la frontera del estado por seguridad.

La «realidad virtualizada» es el concepto que acuñé a mediados de los 90. Consiste en crear una representación virtual espacio-temporal completa de un acontecimiento dinámico del mundo real utilizando un gran número de cámaras, y que los usuarios puedan verlo libremente desde diversos puntos de vista. Para su estudio y desarrollo, construí una cúpula única en su especie cubierta por muchas cámaras: primero 50, luego hasta 400. En 2001, en la Edición XXXV de la *Super Bowl* de fútbol americano, se realizó una demostración del sistema *EyeVision* con 33 cámaras controladas por robótica, que retransmitía las repeticiones de las jugadas en 360 grados, al estilo de una película de *Matrix*, a más de cincuenta millones de espectadores de todo el mundo. Durante la emisión aparecí brevemente para explicar cómo funcionaba el sistema, lo que me valió el divertido título de «El único profesor jamás aparecido en la *Super Bowl*». Hoy el concepto se ha extendido, y vemos más medios parecidos para todo tipo de eventos televisados.

La visión por ordenador se utiliza hoy en día en todas partes, desde la vida cotidiana hasta la exploración espacial, pasando por las instalaciones fabriles y la medicina. Sin embargo, puede que no hayamos hecho más que arañar la superficie de su potencial. Es posible descubrir capacidades mucho mejores y más aplicaciones. Podríamos hacer visible incluso lo invisible. Con sensores avanzados, ordenadores de alto rendimiento y algoritmos de aprendizaje de inteligencia artificial, el campo de la visión por ordenador se encuentra en medio de una «tormenta perfecta» de lo más emocionante, en el mejor de los sentidos.

Por último, debo expresar la mayor gratitud a mis muchos estudiantes, colaboradores y colegas. Sin ellos no habría sido posible ninguno de los sistemas de visión artificial que he descrito. Doy las gracias a mi mujer, Yukiko, que me acompaña hoy aquí, y a mis dos hijos, Shunichi y Sayaka, que por desgracia fallecieron jóvenes. Ellos me han dado el amor y el apoyo que me permiten seguir investigando día y noche.