



Los altavoces inteligentes han conseguido que podamos hablar con la tecnología, que, a partir de ahora, ya dialoga con nosotros. El siguiente paso es dotar de ojos a la inteligencia artificial, para que pueda conocer nuestras necesidades y darles respuesta, sin necesidad de que meditemos palabra.

El año 2018 supuso la llegada de los altavoces inteligentes a los hogares españoles. Esto no quiere decir que anteriormente no tuviésemos contacto con la inteligencia artificial. Muy al contrario, los sistemas inteligentes llevan ya mucho tiempo entre nosotros, aunque a veces no nos demos cuenta. Cuando nuestra plataforma de vídeo bajo demanda nos sugiere títulos que se ajustan a nuestros gustos cinematográficos, es porque un algoritmo ha ido aprendiendo sobre nosotros en función de las selecciones que realizamos día tras día. Lo mismo ocurre cuando al escribir un texto en el teléfono, en el corrector aparece el topónimo del lugar donde solemos pasar las vacaciones, con solo escribir las dos primeras letras. Son innumerables las situaciones de

nuestra vida diaria en las que, conscientes o no, estamos tratando con máquinas inteligentes.

Quizá esta relación con la inteligencia artificial se ha hecho más patente al estar asociada a los altavoces o asistentes virtuales. Y, relacionado con lo anterior, también influye que la forma de relacionarnos con ella está cada vez más próxima a cómo nos comunicamos los humanos, es decir, con la voz y el lenguaje natural. Nuestra forma de hacernos entender por los ordenadores ha ido simplificándose y haciéndose más intuitiva década tras década.

De los complejos lenguajes de programación a las aplicaciones que funcionaban con comandos estándar; de estas, a las interfaces gráficas como Windows; el hipertexto permitió navegar a través de mares de recursos de información y, ya en este siglo, los teléfonos inteligentes -verdaderas computadoras en miniatura- incorporan las pantallas táctiles, para poder operar distintos programas con solo utilizar el dedo.

Ahora, la nueva forma de comunicarnos con las máquinas es nuestra voz, pues gracias a las tecnologías de procesamiento del lenguaje natural pueden entendernos, tal y como lo haría una persona, y contestarnos. El paradigma de la voz son los asistentes inteligentes, como Amazon Echo, Google Home, HomePod de Apple o Movistar Home, que comenzaron a comercializarse el pasado año en nuestro país.

Pero todo indica que pronto asistiremos a un nuevo salto en este campo, y que los sistemas inteligentes domésticos, además de la voz, utilizarán la visión, utilizando cámaras para comprender y reaccionar con su entorno. La inteligencia artificial abre los ojos.

Un mercado en alza

Las máquinas inteligentes están por todas partes a nuestro alrededor, pero el fenómeno más destacable en este campo en los últimos años es el éxito que están cosechando los altavoces inteligentes en todo el mundo. Se calcula que a finales de 2018 había en torno a 100 millones de estos asistentes en el planeta.

El año pasado Google vendió más de 5,4 millones de su dispositivo y Amazon en torno a 4 millones de su Echo, solamente en los seis primeros meses. Para 2024 se prevé un mercado mundial valorado en 30 mil millones de dólares.

Centrándonos en países, Estados Unidos experimenta la mayor penetración de asistentes, aunque China va ganando terreno y se acerca a su volumen. En el mercado americano, Amazon Echo tiene presencia en el 10% de los hogares, seguido de Google Home (4%) y de Cortana de Microsoft (2%).

Por su parte, la consultora Gartner predijo que, en el año 2020, el 50% de todas las búsquedas en internet se realizarán a través de la voz, y que el 30% se llevarán a cabo a través de un dispositivo sin pantalla. Teniendo en cuenta que solo quedan seis meses para que llegue la fecha indicada, parece un vaticinio exagerado, si bien nos da una idea del futuro brillante que se augura para los interfaces de voz.

El próximo paso de la inteligencia artificial doméstica podría incluir la combinación de voz y gestos para comunicarnos con los sistemas, lo que implica que las máquinas tendrán que poder vernos.

La visión artificial

Hablamos de visión artificial (*computer vision*) para hacer referencia al proceso mediante el cual una máquina captura, procesa y analiza imágenes tomadas del mundo real con el objeto de extraer de ellas información contextual sobre el entorno físico.

Se trata de una rama de la inteligencia artificial que pretende replicar las funciones del ser humano de ver, describir y comprender. Técnicamente, la vista ha sido lo más fácil de replicar, gracias a los espectaculares avances que han tenido lugar dentro de la fotografía y del terreno audiovisual en general.

Enseñar a las máquinas describir lo que ven resulta mucho más complejo. Los ordenadores han conseguido detectar los bordes de los objetos retratados e incluso interpretar la perspectiva y el movimiento. Todo esto solo ha resultado posible gracias a las mejoras recientes en la capacidad de computación paralela, una disciplina informática que se basa en ejecutar simultáneamente numerosos cálculos o procesos que son parte de un problema más grande a resolver.

Pero, sin duda, lo más difícil es conseguir que un sistema inteligente comprenda lo que ve. El ser humano desde la cuna asimila la información que recibe de los sentidos y aprende a conceptualizarla, pero lograr que haga lo mismo un algoritmo es uno de los mayores retos actuales de la ciencia computacional.

Para entrenar la visión de las máquinas se les suministran grandes cantidades de imágenes. Por ejemplo, para conseguir que sepan reconocer una manzana, tendrán que alimentarse de todas las fotos posibles de manzanas, para que puedan distinguir una de lo que no es. Al analizar la imagen de una manzana, compararán cada uno de los *pixels* de la misma con los de las imágenes de manzanas que han recibido en su entrenamiento.

Uno de los campos en el que la visión artificial es una necesidad fundamental es en el coche autónomo. El sistema inteligente de abordaje debe ser capaz de identificar los objetos del entorno -vehículos, peatones, señales- para poder circular de forma óptima.

Los altavoces tendrán ojos

A finales del siglo pasado se hablaba de un futuro caracterizado por la computación ubicua, es decir, una era en la que los ordenadores no fueran algo diferenciado, sino que estuvieran incorporados en todo tipo de objetos de nuestro entorno. Y, ciertamente, esa predicción -atribuida a Mark Weiser de Xerox- parece haberse materializado en el mundo en el que vivimos. Cada vez en mayor medida estamos rodeados de cosas que llevan incorporada capacidad de computación e incluso, en ocasiones, inteligencia artificial: teléfonos, televisores, automóviles, robots de limpieza doméstica...

Un fenómeno añadido es que la tecnología se hace transparente, o lo que es lo mismo, que, aunque está por todas partes, no se hace patente. Esto ha sido bautizado como *computación ambiental* (*ambient computing*). La principal consecuencia de ello es que cada vez necesitamos menos conocimientos técnicos para relacionarnos con la tecnología. El salto más reciente en este sentido ha sido la llegada de los asistentes domésticos virtuales, que cualquiera puede manejar con solo usar la voz.

Y, sin embargo, esto solo es el principio. Llegará el momento en que ni siquiera habrá que hablarle a la tecnología para que satisfaga nuestras necesidades, pues esta será capaz de percibir las por sí misma. Será el día en que los sistemas inteligentes podrán ver e interactuar con el entorno que perciban.

La visión artificial aplicada al entorno doméstico requerirá el despliegue de cámaras por el hogar para suministrar de "ojos" a los algoritmos.

Las aplicaciones de la visión artificial son ilimitadas. La seguridad en el hogar es quizá la más evidente, por ejemplo, un sistema de domótica que avise cuando un bebé se aproxime a una situación de riesgo, como unas escaleras, un enchufe o cuando esté jugando con un objeto afilado. También podría alertar de descuidos que pueden acabar en desastre, como un grifo que se ha dejado abierto o un guiso olvidado en el fuego. Por

supuesto, esto es igualmente útil de cara a proteger la vivienda ante intrusos.

Aunque parezca una aplicación más trivial, gracias al registro de nuestros movimientos a lo largo de la casa, la inteligencia artificial nos puede ayudar a encontrar objetos extraviados, como pueden ser unas llaves. Esta función puede resultar de especial relevancia para personas que sufren demencias o enfermedades neurodegenerativas.

Poder reconocer la presencia de cada miembro de la familia permitirá a los asistentes domóticos personalizar el entorno en cada caso: modificar la luz o la temperatura de la estancia, ofrecer una selección de canales o contenidos de televisión en función de las preferencias de cada uno...

Las experiencias pioneras

Sin duda, Amazon ha conseguido llevarle la delantera en este terreno a sus competidores al poner en marcha una experiencia de visión artificial con su asistente Alexa. La iniciativa se llama Echo Look y asesora a su propietario sobre cómo vestir, ofreciendo consejos y recomendando el atuendo más adecuado para cada ocasión. La empresa de Jeff Bezos ha comercializado en Estados Unidos este servicio por unos 200 dólares, que está integrado en su altavoz inteligente Alexa.

Por su parte, Samsung ha presentado, en el último Consumer Electronics Show celebrado en Las Vegas, Family Hub, un refrigerador que incorpora Bixby Vision (Bixby es la inteligencia artificial desarrollada por la compañía coreana). El sistema identifica todos los productos que contiene la nevera en cada momento e informa al dueño cuando es necesario reponer alguno.

Un último caso: la empresa Piccolo está trabajando en un asistente para el hogar con visión, que por medio de una cámara analiza nuestra geometría -mediante una técnica denominada *skeletal tracking*-, e introduce nuestra forma de movernos en un modelo que registra la posición del cuerpo. De esta forma, el sistema aprende nuestros gestos y es capaz de reaccionar ante ellos, por ejemplo, encender la televisión cuando apuntamos con el dedo hacia la pantalla.

No hay duda de que en un futuro cercano las máquinas ya no necesitarán escuchar nuestra voz para saber qué necesitamos en cada momento pues lo podrán ver con sus propios ojos.