

(diapositiva) Sr. Consejero, Rector, autoridades académicas, civiles y militares, compañeras y compañeros, estudiantes

Quiero comenzar justificando el título que he elegido para esta lección inaugural. Desde que se me propuso ser quien pronunciara esta lección tuve dudas sobre la temática que debía elegir. Mi duda estribaba en si debía optar por desarrollar un tema científico concreto directamente relacionado con mi actividad investigadora y docente en los últimos 25 años, lo que podría resultar tedioso para un público heterogéneo, o elegir un tema más generalista para hacer esta lección inaugural más atractiva. Dado que la Secretaria General de la Universidad me indicó en su momento que tratara de limitar mi intervención a 30 minutos, he elegido la primera opción dado que mi experiencia en esta temática creo que me puede permitir resumir un siglo en media hora dando una idea general sobre una Ciencia que, frente a otras, es todavía muy joven. Con el objetivo de ser lo menos tedioso posible, trataré en mi exposición de obviar cuestiones técnicas y hacer un desarrollo con una orientación más atractiva, como si intentara que tras mis palabras muchos de ustedes quisieran dedicarse a esta Ciencia.

Dentro de 5 años se cumplirá exactamente un siglo desde la aparición de esta Ciencia a la que aún seguimos refiriéndonos como Visión por Computador. A diferencia de otras Ciencias, con mucha más tradición, conocemos el año en el que todo comenzó.

(diapositiva) En el año 1920 un grupo de personas con intereses periodísticos se plantearon el reto de hacer posible que una fotografía de un evento deportivo celebrado en la ciudad de Nueva York pudiera ser impresa en un periódico de la ciudad de Londres a las pocas horas de haberse tomado la citada fotografía. El objetivo fue conseguido mediante la codificación de la imagen en datos numéricos y la transmisión de ese conjunto de datos codificados, aprovechando el cable submarino que ya existía entre ambos continentes. Los datos eran decodificados en la ciudad de destino logrando de nuevo reproducir la fotografía mediante un instrumento muy parecido a lo que muchos años después denominamos una impresora de margarita. (diapositiva) Entender cómo se llevo a cabo esa codificación de datos y la decodificación de los mismos en el destino es muy simple. Si usted considera que sólo existe el blanco y el negro basta que asocie al negro el 0 y al blanco el 1 y que transmita en una secuencia de 0 y 1 los negros y blancos que hay en una fila de la imagen, produciendo un cierto retardo para transmitir la siguiente fila de la imagen. Prácticamente al mismo tiempo de este hecho alguien se dio cuenta de las enormes posibilidades que abría esa experiencia, porque simplemente modificando esos datos, (diapositiva) por ejemplo cambiando ceros por unos y unos por ceros, se podía obtener la imagen inversa de la original. De esta manera surge el concepto de “procesamiento de una imagen”. El caso es que rápidamente se formula el concepto de “Fundamento de una imagen digital” que hoy subsiste. (diapositiva) Una imagen es, para el ordenador, un conjunto de datos numéricos dispuestos en filas y columnas: en términos matemáticos una matriz. Por ejemplo, si usted asocia el 0 al color “negro absoluto” y el 255 al color “blanco absoluto”, usted dispone del intervalo [0,255] para expresar numéricamente el “color gris más oscuro o más claro” de cada punto de la imagen. De esa forma usted tiene una matriz numérica asociada a esa imagen, y si usted quiere “procesar esa imagen” (modificarla o extraer información de la misma) sólo puede usar esa matriz numérica.

Dando un salto de unos 30 años desde el origen de todo, comentarles que el primer campo aplicativo de la Visión por Computador fue la medicina. (diapositiva) Por aquellos años los instrumentos que permitían hacer radiografías eran ciertamente muy primitivos y por tanto proporcionaban imágenes de baja calidad, tal y como ven ustedes a la izquierda de la diapositiva. Los esfuerzos se centraban en mejorar la calidad de esas imágenes para conseguir resultados como los que ven ustedes a la derecha de la diapositiva. Como curiosidad indicarles que conseguir ese resultado que ustedes ven llevó bastantes horas de cómputo de un main frame. Un main frame sería el equivalente ahora a un supercomputador. A lo largo de varias décadas todos los esfuerzos de la visión por computador se centraron en la mejora de la calidad de las imágenes. Ello fue así porque existen multitud de actividades humanas en las que la imagen es una fuente muy importante de información. Y como consecuencia de este hecho la Visión por Computador se convierte en un área transversal de investigación debido a que los avances en muchos campos del saber se ven beneficiados de los avances en Visión por Computador. No debemos olvidar que ustedes pueden ver pero con su vista "no pueden medir cosas", y la Visión por Computador "puede ver y medir" (diapositiva) A título de ejemplo, pueden ver ustedes en la diapositiva dos ejemplos de mejora de calidad de una imagen en los que se trabajó durante aquellos años: por un lado como hacer desaparecer la falta de nitidez de una imagen y por otro como eliminar esas imperfecciones de una imagen, nosotros las llamamos ruido, para conseguir imágenes de una enorme calidad.

Tal vez llegado este momento esté usted pensando en que esto de trabajar con imágenes manipulando matrices numéricas para conseguir imágenes con los resultados que han visto debe ser algo muy difícil. Bien, no es tan difícil. (diapositiva) Como un ejemplo, vean en la diapositiva cómo se consigue el efecto de quitar ruido de una imagen. En la parte superior derecha tienen ustedes la operación en la realidad. Un punto de ruido en una imagen es sólo un punto cuya cuantificación numérica en la matriz es muy diferente a la de los puntos que la rodean. Por tanto basta que usted sustituya la cuantificación de ese punto por el resultado de aplicar la media de los valores que rodean a ese punto, incluyendo al propio punto, y esa nueva cuantificación será ya muy parecida a la del resto. El resultado desde un punto de vista visual es que milagrosamente ha desaparecido esa irregularidad porque, aunque sigue existiendo, su visión ya no es capaz de apreciarla. Naturalmente, cuando formalizamos científicamente esta operación no lo hacemos en estos términos. Decimos que hemos aplicado una operación matemática denominada "convolución con un núcleo gaussiano". La expresión la tienen ustedes a la izquierda de la diapositiva. Pero eso lo hacemos sólo porque en términos científicos ese es el lenguaje internacional en Visión por Computador. La realidad, cuando hablamos en términos prácticos, es tan sencilla como la que les he comentado en primer lugar.

Durante los últimos 20 años del siglo pasado hemos sido capaces de conseguir grandes avances en Visión por Computador. Y no sólo en la temática de mejorar la calidad de las imágenes sino también en procesos automáticos, sin intervención humana, que consiguen "extraer información muy relevante" de imágenes para el propósito que sea. A título de ejemplo (diapositiva) pueden ustedes ver en la diapositiva cómo es posible localizar automáticamente un objeto de interés, en este caso un cepillo, ignorando el resto de objetos o (diapositiva) cómo es posible extraer detalles de información sobre determinadas estructuras óseas presentes en una radiografía. Lo más significativo de estos dos ejemplos es

que el tiempo de cómputo actual necesario para ello, en un ordenador doméstico, es inferior a 3 segundos.

Situándonos ya en el comienzo de este siglo, indicarles que se produce un hito muy importante en la Visión por Computador. Dado que ya tenemos ordenadores con suficiente capacidad de cómputo en ese momento comienza el proceso de producir avances, no sólo usando una imagen sino varias a la vez. Al disponer de al menos dos imágenes de una misma escena obtenidas con dos cámaras más o menos paralelas tenemos la información necesaria para simular “exactamente” la visión humana. Y a partir de ahí cada cámara es un ojo y el ordenador el cerebro. Es entonces cuando la Visión por Computador evoluciona a un nuevo concepto que hoy llamamos “Visión Artificial” (diapositiva).

Tal vez la mejor forma de entender este nuevo concepto sea a través del símil que tienen en la diapositiva. La Visión Artificial es sólo una especialidad de una Ciencia, la Inteligencia Artificial, que abarca muchas más cuestiones al igual que la oftalmología lo sería para la Medicina. La única restricción para la Visión Artificial, con respecto a la Inteligencia Artificial, es que sólo podemos usar información visual para realizar cualquier aplicación.

(diapositiva) Todos sabemos que en cualquier campo del saber el paradigma de la investigación, básica y aplicada, sigue el esquema que tienen ustedes en la diapositiva. Todos los investigadores hemos tenido y tenemos ideas que pensamos son utopías. Para mí una utopía es algo que estoy convencido no veré resuelto, por nadie, durante mi vida. Sin embargo producimos investigación básica parcial sobre esa utopía para hacer posible que en un futuro el reto pudiera ser conseguido. Cuando la investigación básica comienza a producir un número significativo de resultados parciales sobre ese problema, yo cambio mi estado y empiezo a pensar que “yo no voy a ser capaz de resolver la utopía, pero tal vez ya alguien de mi generación lo consiga y yo pueda verlo”. Es en ese momento en el que la utopía, para mí, se convierte a un nuevo estado: algo imposible de momento. Porque imposible ya es sólo para mí. Cuando alguien usa toda la investigación básica ya conocida, la combina adecuadamente y resuelve el problema ha convertido lo imposible en posible. Finalmente, el haber resuelto lo que fue una utopía generalmente conduce a formular una nueva utopía, algo así como “y ahora por qué no conseguir algo mucho más difícil aún”, con lo que se reinicia el ciclo.

(diapositiva) Bien, la Visión Artificial tiene formulada su propia utopía. La tienen ustedes en la diapositiva. Se trata de afirmar que “cualquier decisión inteligente que toma un ser humano a partir de la información visual que recibe es reproducible y mejorable por una máquina”. Naturalmente esto hoy es una utopía, pero trabajamos haciendo investigación básica y aplicada para convertir la utopía en, como mínimo, imposible. Y lo hacemos con un fin bueno: construir máquinas que ayuden a tomar decisiones en problemas complejos, nunca para sustituir a nadie porque en los Sistemas que construimos siempre la decisión final será de las personas.

(diapositiva) Citar los miles de aplicaciones de Visión Artificial que se han desarrollado en la última década es imposible para mí en este corto período de tiempo. La enorme expansión que ha tenido lugar se debe fundamentalmente a que los costes de material para hacer Visión Artificial se han reducido drásticamente y a que el número de investigadores dedicados a esta temática ha crecido exponencialmente. Con un ordenador doméstico y un par

de cámaras muy baratas usted puede desarrollar aplicaciones de Visión Artificial que sean muy atractivas. Les hablo por tanto de una cantidad que no debe exceder los 1000 Euros. Otra cuestión son “los costes de adquisición del conocimiento que debe poseer quien desarrolle estas aplicaciones”, pero evaluar ese coste es muy complicado porque ello conllevaría responder a la cuestión de cuánto vale “el saber decirle a una máquina que manipule unas cuantas matrices numéricas para producir la misma decisión de un ser humano y que eso funcione”.

(diapositiva) En función de la limitación de tiempo, y para que ustedes por si mismos puedan ver cuál es el presente de la Visión Artificial, y su carácter transversal como área de investigación, he seleccionado unos cuantos ejemplos, relativamente recientes, desarrollados en la Universidad de Córdoba.

(diapositiva) En primer lugar un sistema orientado a resolver un problema de salud. Se trata de que el especialista pueda tener datos objetivos sobre la evolución de la pérdida de movilidad de un paciente que padece una enfermedad que cursa con pérdida de movilidad. El objetivo del especialista era comprobar si la terapia indicada al enfermo conseguía detener o mejorar esa pérdida de movilidad y ello sin necesidad de tener que prescribir un número excesivo de radiografías al paciente. **(diapositiva)** En la diapositiva tienen ustedes la pantalla principal de una aplicación que permite grabar en video, desde 4 cámaras distintas, los movimientos realizados por el paciente a indicación del especialista. Pueden apreciar también una serie de puntos blancos adheridos a las articulaciones de las que se quiere averiguar el grado de movilidad. Tras la grabación de los movimientos, automática, el ordenador, también de forma automática, proporciona toda la información sobre la movilidad de una determinada articulación o de un subconjunto de ellas. **(diapositiva)** Esa información la ofrece de forma gráfica, como ven en la diapositiva, o en forma de texto indicándole el estadio de su enfermedad con respecto a individuos sanos o pacientes similares con diferentes grados de movilidad. El ordenador también recuerda su propia curva de movilidad en cada momento por lo que al repetir la prueba tantas veces como se quiera el especialista puede tener datos objetivos sobre su evolución y en consecuencia mantener la terapia o cambiarla. **(diapositiva para video)** También el especialista puede asegurarse de que los datos mostrados son los correctos puesto que puede ver lo que “el ordenador ha medido” para asegurarse de que no hubo fallos. Para ello el ordenador le muestra una reproducción del movimiento que han realizado las articulaciones durante el ejercicio realizado y que han sido el objeto de la medida. Y todo ello en una prueba completamente inocua porque usted no es radiado, sólo es grabado en sus movimientos.

(diapositiva) Todo el secreto de lo que ustedes acaban de ver estriba en el hecho que ustedes ven en la diapositiva. El material que ven brillar en la muñeca de la niña brilla mucho tanto si hay mucha luz como si hay poca. Es un material muy reflectante y lo mejor es que su coste es ridículo. Es el mismo material con el que se fabrican esas pulseras que muchos hemos comprado en alguna ocasión a nuestros hijos en la feria de Córdoba porque brilla mucho en la oscuridad. El ordenador lo que hace es localizar automáticamente y seguir el movimiento de esas bolas adheridas. El resto son cálculos matemáticos y presentación de datos.

(diapositiva) En segundo lugar, una aplicación para paliar la dependencia de las personas mayores. Cada vez vivimos más y eso conlleva un necesario incremento de fondos para atender la dependencia cuando seamos mayores. Por otro lado todos querríamos seguir viviendo en nuestro hogar de toda la vida, porque es nuestro nido, y ser atendidos allí. Bien aquí tienen ustedes una solución parcial basada en Visión Artificial.

(video 1) Por un lado es necesaria la seguridad sobre que un ordenador puede identificar a una persona entre varias, porque en ese hogar pueden entrar varias personas, para poder seguir la actividad del que interese. En el video tienen ustedes la prueba de que eso es perfectamente posible. Observen como incluso tratando de confundir al ordenador al cruzarse varias personas, el ordenador nunca se confunde en la identificación de cada persona. En la parte izquierda tienen ustedes el video real, sobre el que el ordenador identifica a cada persona mediante un prisma rectangular de diferente color y en la parte derecha lo que el ordenador nos está diciendo que ve, etiquetando también a cada persona con un color diferente.

(video 2) Por otro lado es necesario probar que una vez identificada una persona se pueden reconocer varias actividades diferentes de esa persona y no hacer nada hasta que se produzca una actividad que implique generar cualquier tipo de alarma para solicitar ayuda. En el video tienen ustedes la prueba de que esto también es posible. Se inicia con la localización de una persona que entra en una habitación y se siguen sus movimientos indicando en modo texto en la parte superior derecha en color azul actividades básicas que está realizando esa persona: **caminar, agacharse o sentarse**. Pueden ustedes ver como ese mensaje cambia cada vez que la acción correspondiente cambia. El ordenador no hace nada, sólo indicar cuál es la acción que se produce en cada momento, porque está entrenado para saber que esas actividades son normales en esa persona. Sin embargo observen lo que ocurre cuando la persona simula una caída. En la parte superior derecha se puede observar un mensaje que parpadea. Esto ocurre porque el ordenador sabe que esa actividad no es normal y está generando una alarma remota para solicitar ayuda. Observen también que si esa persona abandona esa posición y recupera una actividad normal el ordenador lo sabe, deja de generar la alarma y vuelve sólo a indicar la acción que está ocurriendo. Y todo ello en la más absoluta intimidad porque es sólo el ordenador quien le está viendo. En realidad es como si fuera usted, con 40 años menos y conservando todas sus facultades, el que se observa a sí mismo para cuidarse. Fíjense por último en los puntos negros que siguen apareciendo en la parte de la derecha cuando esa persona desaparece de la escena y fíjense en qué ocurre con esos puntos cuando la persona vuelve a aparecer.

(diapositiva) Todo el secreto de lo que ustedes han visto se basa en recordar que “los mosquitos detectan a un ser humano por el cambio de temperatura que se produce al tratarse de un cuerpo caliente que entra en la escena y a partir de ahí siguen ese cambio de temperatura”. Si cambian ustedes mosquitos por puntos negros en la imagen y el cambio de temperatura por el cambio en la imagen cuando algo entra en ella y se mueve pues ya saben por qué todo funciona así de bien. Naturalmente esta no puede ni debe ser la explicación científica que tiene que darse de una manera formal. Pero esa necesaria formalidad es solo aparato matemático e informático. Lo importante es la idea.

(diapositiva) En tercer lugar una aplicación destinada al almacenamiento de residuos radiactivos en condiciones de estricta seguridad para la población. Se trata de cómo la Visión Artificial puede permitir determinar con mucha exactitud la superficie total de un objeto 3D, que puede tener cualquier forma e incluso huecos dentro de la misma. Medir la superficie es lo que se necesita para poder determinar el índice de radiactividad de un objeto, y de acuerdo al valor de ese índice se podrá determinar el recipiente en el que debe ser almacenado ese objeto con el fin de que no exista radiación al exterior. Es importante resaltar que las medidas obtenidas deben ser de una alta precisión debido a las restricciones que impone el Consejo de Seguridad Nuclear.

(video) Tienen ustedes en el vídeo la forma en que la Visión Artificial puede resolver el problema. Un conjunto de 5 cámaras primero se autocalibran de forma independiente y automática. Básicamente cada una de ellas observa donde están ubicadas las demás e intenta referir su posición y las del resto a un sistema de coordenadas común. Lo único necesario es colocar “alguna cosa” que pueda ser observada por todas las cámaras. Puede ser un simple tablero de ajedrez o cualquier otro patrón con formas concretas. El proceso de autocalibración finaliza cuando las 5 cámaras se han puesto de acuerdo en la posición que ocupa cada una con respecto a un sistema de coordenadas global. Posteriormente, cada cámara, también de forma independiente, observa todos los elementos que hay en la escena. Es necesario que cada una analice cada uno de los objetos que está viendo para después poder distinguir el objeto que se quiere medir y diferenciarlo del resto de objetos en la escena. Estos dos procesos son los que ocurren mientras ustedes no aprecian nada. A partir de ahí, en el proceso de medir un trozo de roca, observarán ustedes que es prácticamente instantáneo, se establece una especie de tertulia entre las cámaras en el que se transmiten lo que cada una de ellas ha visto desde su posición para entre todas discutir primero cuales tienen razón y cuáles no en cada instante y efectuar después una especie de reconstrucción del objeto pero midiendo de forma simultánea al proceso de reconstrucción. El ordenador juega el papel de moderador en esa tertulia y es quien toma la decisión final. Tras ofrecer la medida requerida no existiría ninguna garantía del grado de fiabilidad de la misma. Por ello el sistema permite mostrar una visualización de lo que ha medido, desde cualquier punto de vista, para que usted compare con el objeto real y pueda validar la precisión de la medida obtenida.

(diapositiva) En esta aplicación, el secreto de lo que acaban de ver se debe a dos observaciones distintas. La primera consiste en recordar el hecho de que exactamente igual que dividiendo una escena 2D en cuadrados de área conocida y contando los cuadrados en el interior usted tiene una medida aproximada del área de la superficie, y la precisión puede ser aumentada haciendo que los cuadrados sean más pequeños, para un objeto 3D basta dividir la escena en cubos de volumen conocido y contar el número de cubos del interior del objeto para tener un volumen aproximado del mismo. Y a partir del volumen y de la forma, que se conoce a partir del proceso de reconstrucción, se puede obtener la superficie. (diapositiva) La segunda observación, tal vez la más interesante para que esta aplicación funcione, se corresponde con la dinámica del juego de dados denominado “el mentiroso”, en el que se trata de adivinar si el anterior jugador me miente sobre la información que me proporciona, lo que puede depender a su vez de que a él le hayan mentido antes y deba continuar con la mentira porque no quiere ser excluido del juego. Si ustedes han practicado este juego alguna vez sabrán que la atención de cada jugador no debe ser sólo hacia el que le precede sino a todos porque de esa

observación se puede intuir información añadida sobre el comportamiento que tendrá el jugador que me precede. Exactamente eso es lo que sucede en la tertulia entre las cámaras y el ordenador, cuando se está midiendo. Durante el proceso, en cada instante, no siempre una cámara puede estar dando la información correcta. Puede estar engañando porque no ve correctamente la escena desde su punto de vista o porque al no ver esté copiando y transmitiendo lo mismo que la que se ubica más próxima a ella. El ordenador que sigue nuestro modelo, es un moderador con gran experiencia porque la ha adquirido repitiendo el proceso muchas veces, gracias a su velocidad de cómputo, antes de dar la medida final. Y esas repeticiones las usa para asignar un grado de creencia de lo que dice cada cámara en cada instante. Una vez que cree conocer muy bien a cada cámara sabe a qué conjunto de cámaras debe hacerles caso durante cada parte del proceso y a cuales no para no cometer errores. Realmente el sistema fallaría estrepitosamente si durante el proceso una mayoría de cámaras mintieran sistemáticamente, porque nuestro modelo en el ordenador no está preparado para que todas las cámaras mientan. Deben existir al menos dos cámaras en cada instante diciendo la verdad, para poder tener la profundidad, y además las que dicen la verdad deben estar alejadas entre ellas porque de lo contrario existirían partes ocultas del objeto y no sería posible medirlo en su totalidad. (diapositiva) Afortunadamente en Visión Artificial al menos dos cámaras, en cada instante, no mientan. Y buena prueba de ello lo tienen ustedes en la diapositiva. En ella pueden ustedes ver imágenes de los resultados de las mediciones de dos grandes objetos llevadas a cabo en la Central Nuclear ubicada en Zorita, lugar donde está instalado el sistema y que se está empleando durante todo el proceso de desmantelamiento de la Central Nuclear. En la parte superior tienen ustedes los objetos reales y en la inferior los resultados visuales de lo que el Sistema ha medido realmente. Observen la similitud entre ambas. En la pieza de la derecha, ubicada dentro de un edificio de la Central, pueden tener ustedes una referencia visual del tamaño real de la misma.

(diapositiva) En cuarto y último lugar sólo dos pequeños comentarios sobre la aplicación de la Visión Artificial para dotar de “visión real en tres dimensiones del interior del cuerpo” al cirujano que maneja el robot, que ha sido bautizado con el acrónimo de CORDOBA, mientras el mismo lo usa cómodamente sentado frente a los mandos del robot y usando unas gafas. Seguramente muchos de ustedes habrán tenido noticias sobre la construcción del robot y sus posibilidades por lo que el detalle que les comento se refiere única y exclusivamente a cual es la idea sobre la que se soporta toda la visión 3D del robot. (diapositiva) En la diapositiva, en la parte inferior a la izquierda tiene lo que usted vería sin gafas. La superposición incorrecta de las dos imágenes producidas por las dos microcámaras en el interior del cuerpo porque tienen un pequeño desplazamiento entre ellas. A la derecha tienen ustedes la imagen 3D correcta que ve el cirujano con las gafas. Realmente el modelo en el ordenador sólo se asegura, usando para transmitir información a su cerebro las gafas, de que su ojo izquierdo vea sólo la imagen de la cámara izquierda y que su ojo derecho vea únicamente la imagen de la cámara derecha. Y no hace nada más. Porque su cerebro está programado para producir la sensación visual de profundidad, la visión 3D, al recibir dos imágenes independientes con un pequeño desplazamiento entre ellas. Hay quien dice que esta idea “es engañar al cerebro” pero yo no participo de esa opinión. Mi opinión sobre esto es que su cerebro sabe hacer muchísimas cosas que usted no sospecha, y que lo único que hay que

hacer es darle al cerebro la información de una manera particular para aprovechar todo su potencial.

(diapositiva) Si ustedes quieren que el robot haga cosas más sofisticadas, por ejemplo impedir que el cirujano con las pinzas “pueda invadir zonas de riesgo preestablecidas” entonces el ordenador sí tiene que calcular numéricamente la tercera coordenada de cada punto para delimitar esa zona de riesgo. Pero es que hacer esto también es fácil. Tienen ustedes dos gráficos elementales que muestran la sencillez con la que si el ordenador dispone de dos imágenes diferentes de una misma escena y las coordenadas 2D de un mismo punto 3D en ambas escenas, el ordenador puede obtener cual es la tercera coordenada que le hace falta para saber todo de ese punto. Para ello, sólo hay que recordar el concepto elemental de “ semejanza de triángulos ” para, a partir del gráfico en la izquierda, obtener las relaciones que tienen en el gráfico de la derecha y obtener lo que desea. Ni siquiera el ordenador tiene que hacer eso para cada punto de la escena. Sin entrar en detalles técnicos comentar que sólo hace falta hacerlo para unos 11 puntos de la imagen y a partir de ahí resolver un sistema de ecuaciones que además es lineal. Y con eso ya tiene usted la profundidad completa de la escena. Naturalmente, como en los ejemplos anteriores, tras los desarrollos realizados en el robot existe un aparato matemático e informático de cierta relevancia, pero eso de nuevo es sólo para que el producto, el robot funcionando bajo las órdenes de un cirujano, sea más cómodo y atractivo. Las ideas originales de la visión 3D del robot son única y exclusivamente las que les he comentado.

(diapositiva) Llegado este momento creo que procede una reflexión sobre cuestiones comunes que han sido repetidas en los cuatro ejemplos que les he mostrado y que son extensibles a la mayoría de las aplicaciones de la Visión Artificial que existen. Recordarán ustedes cada una de las ideas originales que han soportado las aplicaciones que les he mostrado. Las tienen ustedes en la diapositiva. Todas ellas son elementales y nada tienen que ver ni con las matemáticas ni con la informática. También en Visión Artificial lo importante son las ideas originales. Y si son elementales mejor. Esas ideas se ocurren sólo si uno es muy observador de sus propios comportamientos y decisiones, de las de los demás y además recuerda uno todo lo que ha visto y le han enseñado a lo largo de su vida en cualquier temática por muy alejada que ésta pueda estar del objetivo que se persigue en Visión Artificial. Y ahí es donde se encuentra el caladero de las ideas que nos permiten crear cosas como las que ustedes han visto. El aparato matemático e informático necesario se puede aprender en cualquier Universidad del mundo, y en particular en la Universidad de Córdoba lo enseñamos. También podemos enseñarle a incrementar su carácter observador. Lo único que no podemos enseñarle es a ser observador porque esa característica suele ser innata. Ser observador consiste en fijarse mucho en los detalles sin hacer ningún esfuerzo especial. Algo así como "sin prestar ninguna atención a nada darse cuenta de todo".

Es muy difícil para mí predecir cuál será el futuro de la Visión Artificial a largo plazo. Sólo puedo decirles que sobre cosas que ahora formula uno como una utopía se pueden ver avances en muy pocos meses, por lo que sólo me atrevo a hablar de un plazo corto de unos cinco años. Y lo que es creo que podremos ver cosas que hoy sólo podemos soñar.

(diapositiva) Para ilustrar esta última afirmación imaginen ustedes que alguien les propusiera crear un sistema que usando sólo cámaras convencionales nos permitiera ver en tiempo real algunas cosas del interior de su cuerpo y cómo evolucionan esas cosas cuando usted se mueve. No valen radiografías convencionales, ni TAC, ni resonancias magnéticas. Sólo cámaras de bajo precio. Su respuesta lógica sería que eso es una utopía porque una cámara convencional no puede ver el interior del cuerpo. Mi respuesta hace relativamente poco tiempo coincidía con la suya. Pero ahora...vean ustedes este vídeo de resultados que acabamos de obtener sobre un caso particular dentro de la utopía formulada. Van a ver ustedes desde diferentes puntos de vista a una persona caminar en un centro de investigación muy lejano a Córdoba. Nosotros sólo cogemos ese video y aplicamos nuestro modelo que está en el ordenador. Este modelo tiene como objetivo que el ordenador localice y siga unas cuantas articulaciones importantes de nuestro aparato locomotor y aprovechamos para tratar de mostrar que el ordenador también sabe cuáles son nuestras extremidades, nuestro tronco y nuestra cabeza. Para mostrar que eso es posible el ordenador superpone en el video colores distintos a cada parte del cuerpo y también traza segmentos sobre lo que sería parte de nuestro esqueleto uniendo las articulaciones deseadas. (video) Pueden ustedes ver cómo se mueven las articulaciones en el interior del cuerpo mientras caminamos. Observen que la cosa empieza a funcionar relativamente bien y en tiempo real, es decir, mientras el individuo se mueve. Y ello sólo usando un ordenador doméstico y siete cámaras, y sin ser radiado en ningún momento. Tras ver esto, ¿qué responderían a la cuestión que hace unos minutos les formulé como una utopía? Yo diría que tal vez esto en vez de ser una utopía ya es sólo imposible.

Para no dejarles con demasiada intriga sobre lo que acaban de ver comentarles que tras esta aplicación, que es real, existe también una idea original que es elemental y que explica por qué esto funciona. En este caso, la idea original se encuentra en los primeros minutos de la película de Pedro Almodóvar titulada "La piel que habito". Para tratar de incrementar el carácter observador, sólo de quienes tengan curiosidad, les propongo el trabajo de adivinar cuál es la idea.

(diapositiva) Para finalizar indicarles que no había error en el título de mi charla porque recuerden les he hablado también de lo que creo pasará en los próximos cinco años, por tanto de un siglo, agradecerles a todos su enorme paciencia al escuchar esta breve charla, y expresar un agradecimiento especial a mis compañeras y compañeros de la Escuela Politécnica Superior de Córdoba por haber permitido que fuera yo, en su nombre, el que impartiera esta lección inaugural.

Muchas gracias a todos.