

## Visto, memorizado, asimilado

Robots de aprendizaje autónomo resuelven tareas por medio de una cámara 3D Ensenso



Uno de los métodos de aprendizaje tradicionales consiste en comprobar qué sucede cuando nos comportamos de una manera u otra. El éxito o el fracaso de cada intento determina si se incorpora ese comportamiento y, por consiguiente, si se aprende o no. Este principio encuentra aplicación en la robótica. El Robot Learning Group (ROLE) del Departamento de Robótica y Automatización Inteligente de Procesos del Instituto Tecnológico de Karlsruhe (KIT) sigue diferentes líneas de investigación en el área del aprendizaje automático. Sus miembros investigan cómo los robots aprenden a resolver tareas de forma autónoma por medio del principio de prueba/error. Estos métodos se utilizan sobre todo para que los robots aprendan a manipular objetos, por ejemplo en los sistemas de bin picking. Una cámara 3D Ensenso N10 integrada directamente en la "cabeza" del robot proporciona las imágenes necesarias para que se desarrollen estos procesos.

La selección y agarre de objetos desordenados es una tarea clave en el ámbito de la automatización industrial. Sin embargo, las aplicaciones actuales de bin picking suelen estar adaptadas a la pieza para la que fueron diseñadas, por lo que no presentan la flexibilidad requerida. Los proyectos de investigación del Robot Learning Group proponen soluciones para que los robots que aprenden de forma autónoma sepan extraer objetos que no conocían de un contenedor. Para aprender esta tarea, el robot empieza por probar a coger un objeto de forma aleatoria, como haría una persona. Una red neuronal relaciona las imágenes 3D registradas con los intentos que han acabado en éxito o fracaso. Cada imagen se guarda con el resultado de agarre correspondiente, el cual se determina a través de un sensor de fuerza integrado en el mecanismo de agarre.

A partir de los datos registrados, la inteligencia artificial reconoce los puntos apropiados para agarrar los objetos y se "entrena" a sí misma. Como es habitual en los métodos modernos de "reinforcement learning"\*, resulta indispensable realizar una gran cantidad de intentos y acumular una ingente cantidad de datos. No obstante, los investigadores del KIT han podido reducir de forma considerable el número de intentos y, por tanto, el tiempo necesario para aprender.

### El agarre apropiado reduce el tiempo de entrenamiento



A diferencia de lo que ocurre en los métodos de agarre analíticos (o basados en modelos), el robot ROLE no precisa conocer de antemano las características de un objeto para poder reconocerlo. En este sentido desempeña un papel fundamental la frecuencia con la que el sistema ha podido agarrar un objeto a partir de imágenes similares. Los intentos de agarre del robot son decisivos para poder aprender más rápido. Una red neuronal permite predecir los resultados de las acciones de agarre por medio de los datos disponibles.

"Para que un sistema funcione bien necesitamos que se lleven a cabo unos 20.000 intentos, lo que equivale a unas 80 horas de entrenamiento del robot", explica Lars Berscheid, científico del KIT y miembro del Robot Learning Group. Estas cifras son valores aproximados y dependen de muchos factores, como el índice de éxito de los intentos aleatorios, que a su vez se ve condicionado, entre otros, por la geometría del objeto. Al igual que en los sistemas de aprendizaje autónomo, la cantidad de datos disponibles es el factor que limita las capacidades del sistema.

#### *Aprender practicando*

"Por ello, un objetivo importante de nuestra investigación es reducir el número de intentos necesarios. La pregunta principal para nosotros, los científicos, es por tanto la siguiente: ¿Cuántos intentos deben realizarse para obtener la máxima cantidad de datos en el menor tiempo posible y así acortar el tiempo de entrenamiento?", añade Berscheid.

Aquí es cuando interviene el llamado "aprendizaje por transferencia", por medio del cual pueden utilizarse los datos de una red neuronal entrenada para el reconocimiento de un objeto desconocido. Cuanto mayor sea la cantidad y la variedad de objetos para entrenar, mejor será la respuesta del sistema ante objetos no conocidos. A la larga, esto permitiría eliminar el entrenamiento específico de objetos para aplicaciones. El propósito a largo plazo es crear un sistema de control que pueda agarrar objetos desconocidos de forma autónoma, flexible y extremadamente fiable.

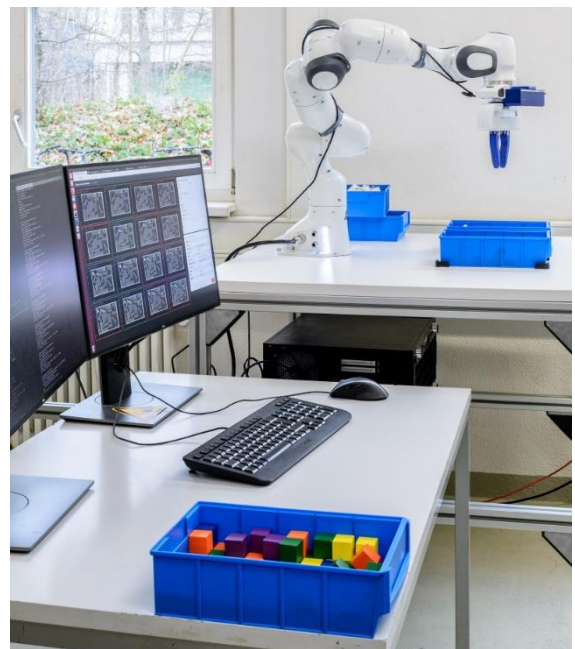
### Aprender sin un modelo definido

Esta es la diferencia determinante con respecto a las soluciones actuales de bin picking. El sistema de investigación del grupo ROLE trabaja sin un modelo de objeto aprendido que enseña al robot de antemano qué pieza tiene que agarrar. Por tanto, el robot es capaz de reconocer objetos desconocidos sin apenas limitaciones en cuanto a forma y características. Tampoco se requieren datos acerca de las propiedades de los materiales y las superficies, que el robot aprende de modo implícito.

*\*En el ámbito de la automatización, "aprendizaje por refuerzo" significa adquirir de forma autónoma una estrategia con la ayuda de recompensas.*

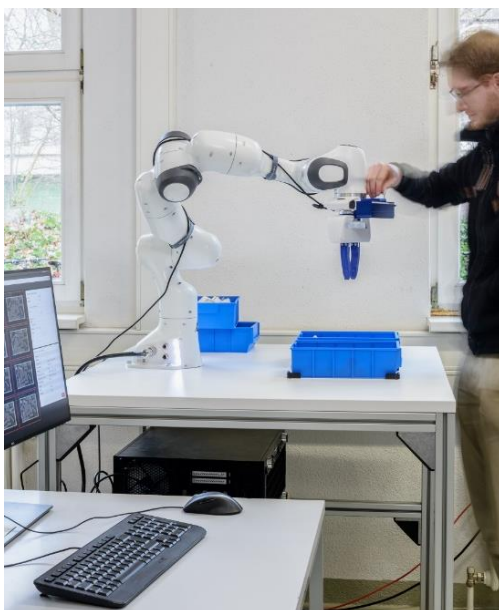
Esta es una de las grandes ventajas del enfoque sin modelo previo, ya que no se precisa ni la representación 3D de un objeto ni el modelado matemático del proceso de agarre. Por tanto, podría aportar grandes dosis de flexibilidad a las industrias y reducir considerablemente el trabajo de programación. Permitiría automatizar un gran número de nuevas aplicaciones en ámbitos que van desde la logística interna hasta la robótica de servicios. Aparte del agarre de objetos, también posibilita otros tipos de manipulación, como por ejemplo su desplazamiento. El robot aprende a desplazar objetos por sí mismo de tal modo que mejora su agarre en el siguiente paso, tal como sucede en la práctica de "Jenga de Hasbro", el popular juego de destreza consistente en retirar bloques de madera de una pila. Esto le permite vaciar por completo una caja sin ayuda de otros dispositivos (por ejemplo, placas de agitación).

El robot se entrena sin intervención humana. En la práctica esta automatización del proceso de aprendizaje es uno de los mayores desafíos. Una vez que el sistema sea capaz de satisfacer todos los requisitos, como por ejemplo un tiempo de ciclo predefinido, se podrá utilizar en la producción, donde tendrá la posibilidad de seguir aprendiendo. Esto permite ganar tiempo frente al procedimiento habitual de las aplicaciones de bin picking. Por un lado, el sistema ROLE destaca por su velocidad, ya que tarda solo 20 milisegundos en calcular el siguiente agarre. Por otro, se reducen los trabajos de programación manual en la puesta en marcha del sistema. El bin picking podría ganar de forma significativa en flexibilidad.



El sistema del grupo ROLE también funciona con objetos desconocidos

### Datos gráficos 3D como punto de partida



Una cámara 3D Ensenso es la encargada de proporcionar la base visual para el agarre del robot. La cámara mira desde arriba el contenedor que contiene objetos de uno o varios tipos ordenados de manera aleatoria. El sistema de procesamiento de imágenes proyecta una textura de alto contraste sobre el contenido de la caja y genera una nube de puntos 3D de las superficies visibles desde arriba como base para el cálculo de la imagen de profundidad en matices de gris. Este paso ya se encuentra implementado en el Ensenso SDK. A continuación, la imagen de profundidad se escala con una resolución de tan solo 12.000 píxeles y se utiliza como introducción de algoritmos de IA. La red neuronal se encarga entonces de analizar las imágenes y de los pasos sucesivos para extraer el siguiente objeto de la caja. La cámara se halla integrada directamente en la "cabeza" del robot para poder realizar diversos experimentos de manera flexible.

Los investigadores del grupo ROLE utilizan la biblioteca NxLib de Ensenso SDK para capturar imágenes de profundidad y para procesarlas por medio de OpenCV y TensorFlow

"Nos hemos decidido por una cámara Ensenso N10, porque permite una distancia mínima de tan solo 30 centímetros aproximadamente con respecto al objeto y, en general, dispone de una amplia gama de distancias. Satisface todas nuestras exigencias como cámara estéreo activa en el espectro infrarrojo que también registra escenas en movimiento", argumenta Berscheid respecto a la elección de este modelo de cámara.

La robusta y compacta carcasa de aluminio de la cámara Ensenso N10, con conectores GPIO atornillables para disparador y flash y conexión USB 2.0, está equipada con dos sensores CMOS monocromo (global shutter, 752 x 480 píxeles) y un proyector (infrarrojo en el espectro invisible, 850 nm). La cámara 3D se entrega precalibrada y con una interfaz MVTEC HALCON, además de una API (C++, C#, .NET) orientada a objetos. Con unas distancias focales de 3,6 a 16 mm, resulta apropiada para distancias de hasta 2000 mm e incluso para el registro en 3D de objetos en movimiento. Los investigadores del grupo ROLE utilizan la biblioteca NxLib de Ensenso SDK para capturar imágenes de profundidad y para procesarlas por medio de OpenCV y TensorFlow.

## Perspectiva

Aunque los procesos desarrollados en el KIT son extremadamente prometedores, los investigadores todavía no han alcanzado su objetivo. "Si bien el bin picking funciona con fiabilidad con objetos como tornillos, hace falta seguir investigando en el caso de operaciones de agarre de objetos más complejos cuyos datos desconoce previamente el robot. Estos sistemas todavía no están lo suficientemente maduros para salir al mercado. Los métodos que estamos desarrollando pueden aplicarse para numerosas tareas de forma sencilla y flexible", indica Lars Berscheid.

En el campo de la investigación robótica los sistemas de aprendizaje autónomo siguen ocupando un lugar central. A partir de proyectos prácticos los científicos han podido identificar la creciente demanda de flexibilidad en los procesos de producción, lo que en el área de la robótica implica un mayor uso de sensores y, por consiguiente, unas necesidades incrementadas de procesamiento de imágenes.

El KIT se centra en dos líneas principales de investigación. En primer lugar, explora cómo se pueden mejorar y acelerar los métodos básicos de aprendizaje. Resultan prometedoras algunas técnicas como el entrenamiento con ayuda de un gemelo digital o por medio de la simulación virtual y su posterior aplicación a un robot de verdad, así como la transferencia de datos entre diferentes tareas. En segundo lugar, los científicos investigan qué nuevas aplicaciones podrían automatizarse mejor —o incluso por primera vez— con robots de aprendizaje autónomo. En este sentido se abren interesantes posibilidades en la manipulación de textiles (coger y doblar toallas y prendas de ropa, por ejemplo), el desmontaje de piezas industriales como electromotores para su reciclaje, el lacado de objetos no conocidos a partir de los datos de una cámara o la manipulación de líquidos o granulados. Estas tareas se aprenden durante la simulación y luego se transfieren a los robots reales.

Algunos de los retos clave son seguir mejorando el índice de éxito de los intentos de agarre así como la solidez del sistema. "Con el sistema ROLE se alcanzan en principio tasas de éxito superiores al 95 %", señala Berscheid. Esto significa que de 100 intentos de agarre como máximo 5 no tienen éxito. La pregunta que se plantea es la siguiente: ¿Pueden reducirse aún más los intentos que acaban en fracaso gracias a un aprendizaje más largo?

Otra cuestión compleja es cómo reacciona el sistema cuando faltan datos 3D en la imagen de profundidad. En la actualidad, el equipo ROLE se limita a agarrar objetos solo en vertical y desde arriba. Pero ¿cómo puede utilizar el sistema los seis grados de libertad? También se investigan nuevos enfoques para superar otros de los retos que suele plantear el bin picking sin modelo, sobre todo en las etapas sucesivas del proceso, como el almacenamiento o la manipulación posterior de los objetos agarrados.

A los investigadores del KIT aún les queda trabajo, pero las soluciones y los resultados obtenidos hasta la fecha muestran el inmenso potencial que alberga el aprendizaje automático para su aplicación en el ámbito industrial. El procesamiento de imágenes 3D, intrínsecamente unido a esta especialidad, proporciona datos básicos para guiar las "manos del robot" hacia el agarre perfecto. Visto, memorizado, asimilado: este es el futuro.

## Cliente

El grupo ROLE (Robot Learning Group) del Departamento de Robótica y Automatización Inteligente de Procesos del Instituto Tecnológico de Karlsruhe sigue diferentes líneas de investigación en el ámbito del aprendizaje automático aplicado a robots. Esta área abarca el aprendizaje por refuerzo



- de estrategias para trayectorias de movimiento ([T. Kröger](#))
- para la manipulación de objetos por parte de robots y máquinas herramienta ([T. Kröger](#))
- de tareas complejas secuenciales o compuestas por trayectorias de movimiento sucesivas ([P. Meißner](#))

De todas estas aplicaciones el grupo se interesa sobre todo por la transferencia Sim-To-Real.

<https://www.ipr.kit.edu/3234.php>

## Cámara

### Enseno N10 – Visión artificial 3D rápida y precisa

- Dotada de una interfaz USB 2.0 para una aplicación universal y flexible
- Carcasa de aluminio robusta y compacta
- Sensores CMOS global shutter y proyector de patrones integrados
- Hasta 30 imágenes por segundo con la máxima resolución y 64 niveles de disparidad
- Diseñada para distancias focales de hasta 2000 mm (N10) y campos visuales variables
- Generación de una única nube de puntos 3D a partir de todas las cámaras del modo multicámara
- Generación directa de las nubes de puntos 3D desde varias perspectivas
- Sistema "Projected Texture Stereo Vision" para la captura de superficies sin textura
- Captura de objetos fijos y en movimiento
- Paquete de software incluido con controlador y API para Windows y Linux
- El paquete de software soporta tanto los modelos USB como los modelos GigE
- Programas de ejemplo con código fuente para HALCON, C, C++, C#
- Precalibrada y por consiguiente fácil de configurar
- Función integrada para la calibración de mano-ojo en un robot mediante placa de calibración
- Integración mediante software de las cámaras industriales uEye, por ejemplo para registrar información cromática adicional o códigos de barras
- Subsampling y binning para tasas de transferencia de datos y frecuencias de imagen flexibles



## Autor

Sabine Terrasi

IDS Imaging Development Systems GmbH  
Dimbacher Str. 6-8  
74182 Obersulm  
Alemania

T: +49 7134 96196-0

E: [s.terrasi@ids-imaging.de](mailto:s.terrasi@ids-imaging.de)

W: [www.ids-imaging.es](http://www.ids-imaging.es)

© 2019 IDS Imaging Development Systems GmbH