Seen, stored, learned

Self-learning robots solve tasks with the help of an Ensenso 3D camera

Trying out different behaviours is one of the classic learning methods. Success or failure decides which behaviour is adopted. This principle can be transferred to the world of robots. At the Institute for Intelligent Process Automation and Robotics of the Karlsruhe Institute of Technology (KIT), the Robot Learning Group (ROLE) focuses on various aspects of machine learning. The scientists are investigating how robots can learn to solve tasks by trying them out independently. These methods are used in particular for learning object manipulation, for example for grasping objects in a typical bin picking scenario. An Ensenso N10 3D camera directly at the "head" of the robot provides the required image data.

The gripping of randomly lying objects is a central task, especially in industrial automation. However, current Bin Picking solutions are often inflexible and strongly adapted to the workpiece to be gripped. The research projects of the Robot Learning Group promise a remedy, e.g. with robots that independently learn to pick up previously unknown objects from a container. In order to learn such a task, the robot first begins with random gripping attempts, as a human would do. A neural net connects the 3D images taken with the successful or unsuccessful gripping attempts. For each image, the gripping result, which was determined by a force sensor in the gripper, is stored. The AI (artificial intelligence) uses the stored data to identify meaningful gripping points for the objects and thus "trains" itself. As is usual with modern methods of reinforcement learning*, large amounts of data and many gripping attempts are essential for this. However, the researchers at KIT were able to significantly reduce the number of the latter and thus also shorten the time required for learning.

*) *Strengthening learning in the machine area, in which a strategy is learned independently supported by rewards*

The right grip reduces training time

How it was grasped so it is learned

In contrast to analytical (or model-based) gripping methods, the ROLE robot does not need to have the features required for recognition described in advance. However, it plays an important role in how often the system has been able to successfully capture an object with "similar" images. The grip that the robot tries out is critical for faster learning success. With the help of a neural network, gripping results can be predicted using existing knowledge.

"For a well-functioning system, we currently need about 20,000 gripping experiments, which corresponds to about 80 hours of training time on the robot," explains Lars Berscheid, researcher at KIT and part of the Robot Learning Group. These figures are approximate values and depend on many factors, such as the gripping rate of random grips, which in turn is influenced, among other things, by the component geometry. As it is common with learning systems, the amount of data available is the limiting factor for the system's capabilities.

"Therefore, an essential task of our research is to reduce the number of necessary gripping attempts. The central scientific question here is therefore: Which grips have to be tried in order to gain as much information as possible as quickly as possible and thus shorten the training time," adds Berscheid.

At this point, transfer learning is also used. The knowledge of an already trained neural network can be used for the recognition of previously unknown objects. The larger the number and range of training objects of the system, the better it can generalize to unknown objects. This could permanently eliminate the need for specific training of objects for applications. The long-term goal is a control system that can independently and flexibly grasp arbitrary and unknown things with industrial reliability.

Learning without a given model

This is exactly the crucial difference to today's Bin Picking solutions. The research system of the ROLE group works without a "taught" model of the workpiece to be gripped and thus also for unknown objects. In principle, there are no restrictions with regard to their form and nature. Knowledge of material and surface properties is also not necessary and is implicitly learned. This is the great advantage of the model-free approach, for which neither the 3D shape of an object nor the mathematical modelling of the gripping process are necessary. It could be used flexibly in industry and with less programming effort. The automation of many new applications would be possible - from intralogistics to service robotics. At the same time, other types of object manipulation, such as moving, can be realized in addition to the gripping itself.

The robot learns to move objects independently in such a way that they can be grasped better in the next step. This allows a box to be completely emptied without any further equipment (e.g. shaking plates).

The training of the robot works completely without human intervention. In practice, this very automation of the learning process is one of the greatest challenges. Only when the system meets all requirements, e.g. a given cycle time, can it be used productively in production and, of course, learn further.

Here, too, there are time advantages over the usual procedure for bin picking applications. On the one hand, the ROLE system is very fast with only 20 milliseconds when calculating the next grip. On the other hand, manual programming is reduced when the system is commissioned. Bin picking could thus gain considerably in flexibility.

The system of the ROLE group also works for unknown objects.

3D image data as a basis

The visual data for the robot is provided by an Ensenso 3D camera. It looks from above at the container, which is randomly filled with objects of one or more types. The image processing system projects a high-contrast texture onto the content of the box and generates a 3D point cloud of the surfaces visible from above as a basis for calculating the depth image in greyscale. This step is implemented directly in the Ensenso SDK. The depth image is then scaled to a resolution of only 12,000 pixels and used as input for the AI algorithms. The neural network then takes care of the image analysis and the logical steps for the next grip into the box.
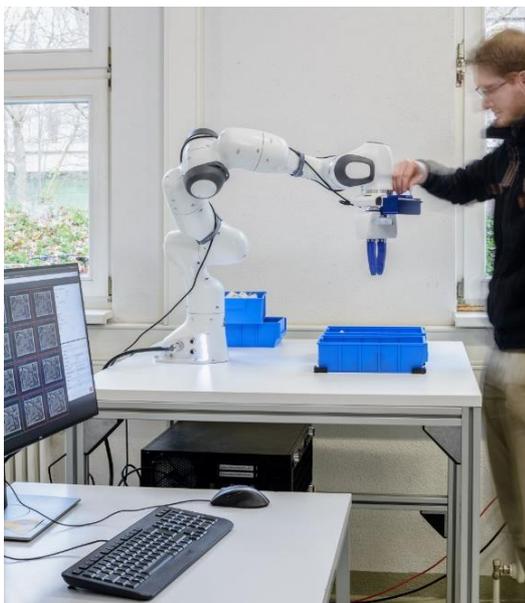
The camera is mounted directly on the "head" of the robot in order to be able to flexibly realize different experiments. "We decided on an Ensenso N10 camera because the model allows a minimum distance of only 30 cm to the object and has a large overall distance range. As an active stereo camera in the infrared range, which is also suitable for moving scenes, it meets all our requirements," Berscheid explained the choice of camera model.The robust, compact aluminium housing of the Ensenso N10 camera, with screwable GPIO connectors for trigger and flash and USB 2.0 connection, is equipped with two monochrome CMOS sensors (global shutter, 752 x 480 pixels) and an infrared wavelength pattern projector operating at 850 nm.

ROLE Group researchers use Ensenso SDK to capture depth images and process them with OpenCV and TensorFlow.

Pre-calibrated and supplied with an MVTec HALCON interface and object-oriented API (C++, C#/ .NET), the 3D camera has focal lengths from 3.6 to 16 mm and is suitable for working distances up to 2,000 mm and can even be used for 3D detection of moving objects.ROLE group researchers use the NxLib of the Ensenso SDK to capture depth images and process them with OpenCV and TensorFlow.

Outlook

The methods developed at KIT are trend-setting, but the researchers have not yet reached their goal. "While Bin Picking already works reliably with simple objects such as screws, there is still some research required to reach product maturity, especially for gripping more complex, unknown objects. However, the methods we develop are basic and flexible and can be used for different tasks," said Lars Berscheid. Self-learning systems remain a central topic within robotics research. From application-oriented projects, scientists are also realising the need for more flexibility in production, which in robotics often leads to the use of more sensors and thus also more image processing.

At KIT, research will continue to focus on two central topics: On the one hand, how basic methods of learning can be improved and accelerated. Techniques such as training with the help of a digital twin, simulation on the computer and subsequent transfer to a real robot, but also the transfer of knowledge between different tasks are promising. On the other hand, the scientists are researching which new applications could be automated better or even for the first time with learning robot systems. There are exciting possibilities here, for example, in the handling of textiles (gripping and folding towels and clothing), the dismantling of industrial parts such as electric motors for recycling, the painting of unknown objects based on camera data, or the handling of liquids or granular media. This work is learned in simulation and then transferred to the real robot.

Challenges are, for example, how a further increase of the gripping rates and the robustness of the system can be achieved. "In principle, we can achieve gripping rates of over 95% with the ROLE system," explains Berscheid. This means that a maximum of 5 out of 100 gripping attempts are unsuccessful. The question that now arises: Can the remaining failed attempts be further reduced by longer learning?

Another non-trivial problem is how the system deals with missing 3D data in the depth image. Currently, the ROLE team limits itself to grabbing objects vertically from above. But how can the system use all six degrees of freedom? New solutions are also sought for other typical challenges of model-free bin picking, in particular for subsequent process steps such as storage or further processing of the gripped objects.

KIT scientists still have a lot of research to do, but the approaches and previous results show the immense potential of machine learning for industrial use. 3D image processing is inevitably associated with this and provides important basic data for controlling the "robot hands" towards the perfect grip. Seen, saved, learned - that is the future.

Client

The research group "ROLE - Robot Learning Group" of the Institute for Intelligent Process Automation and Robotics of the Karlsruhe Institute of Technology (KIT) deals with different focal points in the areas of machine learning for robots. This includes Reinforcement Learning

- strategies for movement trajectories (T. Kröger)
- for object manipulation by robots and machine tools (T. Kröger)
- of complex tasks that are sequentially or concurrently composed of movement trajectories (P. Meißner)

The group is particularly interested in Sim-To-Real Transfer across all the above applications.
https://www.ipr.kit.edu/3234.php

Camera

Ensenso N10 - 3D vision, fast and precise

- With USB 2.0 interface – versatile and flexible
- Compact, robust aluminium housing
- Integrated global shutter CMOS sensors and pattern projector
- Up to 30 frames per second at full resolution and 64 disparity levels
- Designed for working distances of up to 2.000 mm (N10) and variable picture fields
- Output of a single 3D point cloud with data from all cameras used in multi-camera mode
- Live composition of the 3D point clouds from multiple viewing directions
- "Projected texture stereo vision" process for capturing surfaces without texture
- Capture of both stationary and moving objects
- Free software package with driver and API for Windows and Linux
- One software package supports USB and GigE models
- HALCON, C, C++ and C# sample programs with source code
- Pre-calibrated and therefore easy to set up
- Integrated function for robot hand-eye calibration with calibration plate
- Integration of uEye industrial cameras on the software side, for example, to capture additional colour information or barcodes
- Subsampling and binning for flexible data and frame rates

Author

Sabine Terrasi

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

T: +49 7134 96196-0
E: s.terrasi@ids-imaging.de
W: www.ids-imaging.com