GROUP **LASER SCANNING**

Pure mathematics: Automatic object recognition

For humans, a lively street is like a very busy picture: cars, signs, trees, people and buildings, often one in front of the other. The brain is able to classify and pinpoint these objects effortlessly. But when an attempt is made to automate this process, it becomes clear how truly complex it is. As part of a research project, Fraunhofer IPM developed algorithms for the automatic identification and classification of objects on the road, with artificial neural networks forming the basis for pattern recognition.

Today, infrastructure is surveyed using high-performance cameras or laser scanners, which deliver high-resolution images and very accurate, georeferenced measurement data. Modern laser scanners record several million measurement points per second, providing a detailed map of the surrounding environment in the form of a 3D point cloud. The results are generally evaluated manually, which requires viewing extensive volumes of data (point clouds and image data). The automation of at least part of this time consuming process is the goal of a research project commissioned by Lehmann + Partner GmbH that Fraunhofer IPM is conducting for the Federal Highway Research Institute (BASt). The project involves developing algorithms to automatically identify, classify and locate elements of road infrastructure in 3D measurement data.

In order to evaluate road scenes, scientists are turning to complex learning algorithms based on the concept of deep learning using artificial neural networks (ANNs). This approach has been shown to be superior to traditional methods of object recognition. While the latter use feature sets provided by the developer, ANNs learn to recognize the relevant features on the basis of training data. In ANNs, the information provided passes through a large number of interconnected artificial neurons, where it is processed and transmitted to other neurons. ANNs learn the output patterns which correspond to specific input patterns with the help of manually annotated training data. On the basis of this »experience«, new types of input data can then be analyzed in real time. ANNs have proven to be very robust when confronted with variations on characteristic colors, edges and shapes.

Data fusion: Scanner data and camera images are merged to create the data pool

The more detailed the information in the data set, the more successfully objects can be recognized and classified. Camera and scanner data collected by a survey vehicle outfitted with laser scanners developed by Fraunhofer IPM and operated by project partner Lehmann + Partner form the basis of the project. Merged scanner and camera data serve as a suitable starting point for automatic object recognition. In one approach, georeferenced scanner data points are initially transferred to a grid format containing depth information and are then linked with RGB camera data. This pixel-based RGB-D(epth) data set contains a corresponding depth image for each RGB camera image, which makes it the ideal input format for ANNs. Scientists think that the depth values will help the network separate overlapping objects and generally make classifying and locating objects a more robust process.

DEEP LEARNING As a method of machine learning, deep learning is a subfield of artificial intelligence which relies on smart algorithms. For example, training data sets are used to identify objects in a picture. Deep learning is based on artificial neural networks and has been shown to be more robust when confronted with the varying forms and obscured, damaged or faded objects that are typical of street scenes than traditional methods of object recognition.

Using semantic segmentation to identify 3D georeferenced objects

The architecture of the network, in other words the number of network layers and the type of hierarchical links, is adapted to each specific task. The network is trained using a training data set. To this end, images are first semantically segmented manually and each pixel is attributed to a specific object class. Once a network has been trained with this data, it can be expanded to include additional object classes at any time with a new training data set. Using the pixel coordinates of the objects identified in the image data, the segmentation can be back-projected into the point cloud. To make this possible, the camera and laser scanner must be accurately aligned and the appropriate calibration parameters must be determined in a corresponding one-time process. Segmenting the point cloud enables georeferenced objects to be identified in 3D.

The topic of automatic object recognition is of interest to anyone faced with the task of evaluating large volumes of data. In the future, the challenges of surveying infrastructure will pale in comparison with the requirements of the automotive industry, which relies on this technology for self-driving vehicles. In autonomous driving, moving objects must be recognized in real time. This is where the neural networks of the human brain remain superior – for now.

1 *Recognizing, classifying and locating objects on the roads is time consuming. Pattern recognition using artificial neural networks automates this process.*

2 *Camera image and evaluated measurement image: Colored segmentation masks mark the forms of the identified objects.*

Prof Dr Alexander Reiterer, P +49 761 8857 - 183, alexander.reiterer@ipm.fraunhofer.de