USING SIMULATED SENSOR IMAGES FOR OBJECT RECOGNITION OF UNIVERSAL GOODS FOR AUTOMATIC UNLOADING OF CONTAINERS

Bernd Scholz-Reiter[(a)], Hendrik Thamer[(b)]

[(a), (b)] BIBA – Bremer Institut für Produktion und Logistik GmbH at the University of Bremen

[(a)]bsr@biba.uni-bremen.de, [(b)]tha@biba.uni-bremen.de

ABSTRACT
Unloading of standard sea containers is a time- and cost-expensive process step within global supply chains. Until now, existing solutions for automatic unloading are limited to goods of cubical shape. The main challenge in developing a robotic system for universal logistic goods is object recognition of goods that differ in size, shape and orientation. A common approach for object recognition systems is the comparison of sensor data to a model database. This paper presents a simulation system of sensor images for universal logistic goods. The objective of simulating sensor images is generating a model database and create training data for classification. Additionally, sensor images of complete packaging scenarios can be simulated for evaluating the object recognition algorithm with ideal sensor data that does not contain measurement noise or other inaccuracies. The simulation system is explained in detail by modelling of cubical objects.

Keywords: technical logistics , container unloading, sensor simulation, object recognition

1. INTRODUCTION

The dynamical development of global flow of goods in complex worldwide logistics networks creates challenging requirements for logistic service providers (Aden 2008). As for instance, the pressure of providing efficient logistic processes causes a request of technical systems for automation within the supply chain. The transported goods are generally packed in standardized packaging and loaded in carriers as containers, swap bodies and Uniform Load Devices (Echelmeyer 2008). Loading and unloading processes of containers are still a non-automated process in logistic markets and generates bottlenecks in efficiency (Burwinkel 2009). This paper focuses on the automatic unloading process.

In Europe, about 64% of the imported goods are suitable for automatic unloading due to their size, shape and weight (Akbiyik 2009). Hence, the economic relevance of automatic unloading is very high. The main shapes of packaged goods can be summarized to cubical, cylindrical and sack-shaped (Akbiyik 2009). Concerning cubical goods, the successful market launch of the 'Parcel Robot' (Scholz-Reiter 2008) has shown the feasibility of automatic unloading of cubical goods. The crucial step to extend the unloading system to universal goods is, amongst gripping technologies, the development of a suitable object recognition system, that is able to detect and classify different goods inside a container. The automatic unloading process for universal goods refers to the bin-picking problem and is not completely solved today (Kirkegaard 2006).

For an accurate determination of position and orientation of the goods, the sensor system must deliver 3D information about the scene. Instead of reconstructing 3D information from 2D images, a suitable sensor technology for object recognition inside a poorly lighted container is Time-of-Flight (TOF) laser scanning (Uriarte 2010). The resulting images store the depth information instead of grey-scale values. The following object recognition process analyses the images for significant characteristics and compares them with characteristics of predefined models of determined object classes.

This paper focuses on the simulation of the sensor images for an object recognition system for universal goods. These simulated sensor images will be used in two different approaches for object recognition in future research work. Afterwards, the two approaches will be evaluated and the result will be compared. The first approach considers the generation of a model database that is used for comparison of the real sensor images with models of every predefined object class. The models are generated by simulating the sensor measurement principle of a TOF laser scanner by means of standard computer graphic techniques. For every object class, several images from different positions are simulated and are stored in the model database. Due to possible occlusions in the packaging scenario, even models of object parts are generated. The second approach will use the simulated sensor images as training data for a classification task. Classification is the assignment of an input set to a finite number of discrete categories (Bishop 2006). Considering an automatic unloading system, the input set are 3D images of logistic goods. The classifier that is trained with the simulated sensor data, assigns each image to a predefined object class.

Additionally, simulating sensor images of entire packaging scenarios is also possible. These images can be used for evaluating object recognition algorithms. In order to determine the theoretical performance of different object recognition algorithms, the sensor image of the packaging scenario can be simulated without influence of negative measurement effects.

2. STATE OF THE ART

This paper presents a simulation methodology for TOF sensors in order to use it in the object recognition process of universal logistic goods. Therefore, the state of the art contains the TOF sensor technology and related work that use simulation techniques for this measurement principle.

2.1. TOF Laser Scanning

In order to identify universal goods inside a container under unfavourable lightning conditions, sensor systems which provide depth information by TOF laser scanning are suitable for image acquisition (Uriarte 2010). These sensors measure the distance between the object and the sensor by sending a light ray from a light source to the object to be measured. The light is emitted and partially reflected from an object to the sensor, which detects it. The sensor measures the time between sending and receiving the ray and by knowing the speed of light the distance to the object point can be calculated. The acquired data is usually delivered as a set of points with (x,y,z) coordinates for each point. The LMS-200 from Sick is a frequently used ToF scanner, with a rotating mirror. It has an angle of 180° and a angular resolution of 0,25° and a measuring range from 0.1 to 30 m. The lateral resolution of this scanner is about 22 mm when measuring objects at 5 m. Figure 1 illustrates the measuring principle of the sensor.

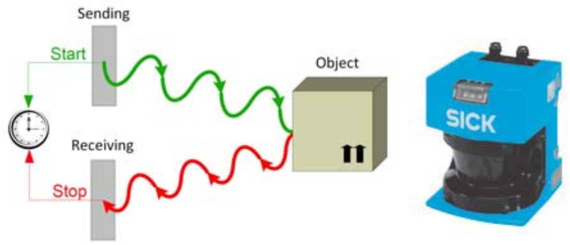

Figure 1: TOF Scanning (Uriarte 2010)

Since the LMS 200 works line-based, the camera is mounted on a pitch device in order to scan the whole packaging scene.

2.2. TOF Sensor Simulation

Simulation of TOF sensors can avoid the development of cost expensive prototypes. Thereby, the design of sensor hardware and application development can be realized by simulation. Common algorithms that are tested and evaluated by simulated are in the area of sensor calibration and sensor data processing (Keller 2007).

In order to verify a calibration tool, *Meissner et al.* have generated a realistic 3D simulation of an urban intersection and simulate several 4-layer laser scanner (Meissner 2010). A simulation system for system analysis and algorithm development is given by *Kukko et al.* (Kukko 2007). The purpose of this work is providing a tool for analyzing systematic properties of scanning systems and algorithm development.

Simulating TOF sensor images of complex 3D scenes is a time expensive processing step. In order to perform the simulation in real time, *Keller* makes use of the programmability of modern Graphic Processing Units (GPUs) (Keller 2009). Their simulation approach is motivated in particular by physics of Photonic Mixing Device (PMD) sensors, which are a specific type of TOF sensor. Additionally, they consider typical measurement inaccuracies like deviation errors and resolution artefacts. An approach in which simulating TOF images are used as training data for classification is presented by *Shotton et al.* (Shotton 2011). They use simulated TOF images of humans of many shapes and sizes in highly varied poses for human pose recognition. They also use GPU programming techniques for recognition in real-time.

Considering the field of automatic unloading of containers, none of the existing approaches for cubical goods use simulated sensor images whether for evaluation of algorithms, generating a model database or classifier training.

3. SYSTEM ARCHITECTURE

This section describes the complete architecture of an object recognition system for universal logistic goods. The system identifies the shape of logistic goods and determines their position and orientation (pose). The object recognition process and the use of simulated sensor images is explained in detail.

3.1. Architecture

The system architecture covers the complete logistic process of automatic unloading of containers. For handling of logistic goods a mechanical robotic system is used. Figure 2 illustrates the system architecture. The process is starting with a packaging scenario that is made up with logistic goods with shapes from all predefined object classes. In the first step, a 3D image from the scenario is acquired by a TOF sensor. The sensor delivers a point cloud from which a 3D range image can be generated. A range image contains distance information instead of grey scale values. Afterwards, this image is analysed by object recognition

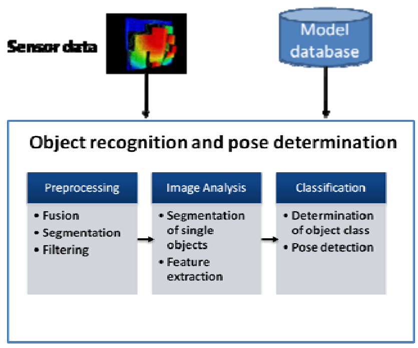

Figure 2: System Architecture

techniques that are described in the following subsection. The result of object recognition are position and possible gripping points of logistic objects that are suitable for automatic unloading by the mechanical robot system. Unloading another logistic good requires a new image acquisition process, because gripping by the mechanical robot may have influences on the complete packaging scenario.

3.2. Object Recognition

The object recognition system consists of three parts. First, the image is preprocessed. In this step, inaccuracies that are caused by the image acquisition process are reduced by applying filter algorithms. Additionally, an initial segmentation step is performed to distinguish between container body and container content. The distinction is necessary for a collision detection during the unloading process. For object recognition the distinction is not mandatory. However, due to a reduced image size the computation time of the image analysis step reduces. In the case of using more sensor that scan the same packaging scenario from different point of views, a fusion step is necessary. This is realized by performing a registration between the sensor images. Image registration techniques define a transformation that maps the first point set onto the second one (Forsyth 2002).

After preprocessing, an image analysis step identifies characteristic features of logistic objects. This requires another segmentation step in order to identify regions that represent single logistic objects. Then, feature extraction is performed. A feature is a characteristic attribute in the image that can describe a specific object. Usual features of objects are geometric characteristics like corner points, surfaces, patches and related areas.

In the last step the detected object is classified according to a related object class. Therefore, the detected features of a segmented region are compared to features of predefined model class. The geometric models of logistic objects from all predefined shape classes are stored in a model database.

The simulated sensor images will be used in the object recognition process in two different ways. The first approach uses the simulated sensor images for generation of the model database. This means for real automatic unloading processes, the sensor data will be analyzed, and single regions that possibly contain a logistic object are compared to the model database by matching operations. The second approach uses a different classification concept. Here, a learning algorithm is trained with simulated sensor images of single logistic objects. After the training phase, the classifier should be able to distinguish between the predefined object classes. For processing in real-time both approaches will be implemented by using GPU programming techniques.

4. SENSOR SIMULATION

This section describes the simulation of the TOF measurement principle by reference to the laser scanner LMS 200 from the company SICK. The objectives of sensor simulation are creating 3D images for every predefined object class in various scanning directions. Additionally, sensor images of individual parts of objects classes are generated. This is useful in case when objects are partially occluded by other objects in the scene.

One benefit of simulating sensor images is the generation of sensor images without scattering effects or measurement noise. Thereby, the object recognition system can be tested under ideal conditions and the theoretical performance can be evaluated. The simulated sensor data is generated by a simulation software that is implemented in MATLAB. The user is able to create different packaging scenarios with objects from predefined object classes. Afterwards, the sensor parameters like scan resolution and sensor position can be set. Figure 3 illustrates the scene generator of the simulation software.

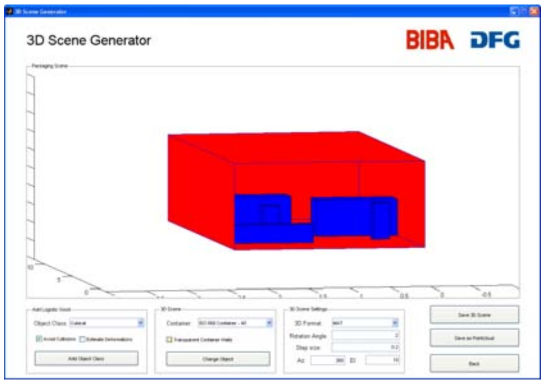

Figure 3: Screenshot of the 3D Scene Generator

In real scanning processes, the light ray from the sensor unit is emitted sequentially over the whole scene. The sensor has a scan angle of 180° degrees and the angular resolution can be chosen from 0.25°, 0.5°, 1°. The light ray of the sensor is modelled by a line g like equation 1 with a position vector $\vec{p}$ and a direction vector $\vec{v}$.

$$g : \vec{x} = \vec{p} + \lambda \vec{v} \quad (1)$$

For each ray the intersection between the ray and the objects in the scene is computed. Therefore, the objects must be described geometrically. Actually, the simulation platform is able to simulate packaging scenarios with cubical goods. Implementing geometric models for the other objects will be one following step of the research work.

4.1. Modelling of Cubical Objects

A cubical object consists of four vertices and linking edges. The vertices defines the corner points of the cubical objects. The cubical object is limited by six faces that are described by equation 2. A face f in coordinate form is described by a normal vector $\vec{n}$ and a real number b.

$$f : \vec{n}\vec{x} = b \tag{2}$$

Within the simulation platform, all kinds of cubical logistic objects can be defined and placed inside the standard container. Collisions of cubical objects inside the container are prevented by a collision check algorithm. For simplification reasons, deformations of the objects are not considered.

4.2. Generation of the point cloud

The TOF sensor output is a point cloud with (x,y,z) coordinates. In order to simulate the point cloud, the intersection of each ray with cubical logistic objects is computed. These intersection points represent the point cloud of a virtual TOF sensor. As the packaging scenario is made up with cubical objects, the intersection of each ray with every cubical object has to be determined. For a cubical object, that implies a possible intersection computation with six faces. In the case of a packaging scenario with many logistic objects and a high scanning resolution the calculation time increases significantly.

Therefore, visible surface determination algorithms from the field of computer graphic techniques are applied. A suitable algorithm is back face elimination. A back face is an oriented face with respect to a vector v (mostly the view direction of the camera) if the angle between its normal vector n and v is between 0 and 90 degrees. Expressed mathematically, the dot product between n and v must be greater or equal zero. Figure 4 illustrates the principle of back face culling for a cubical object. (Agoston 2005)

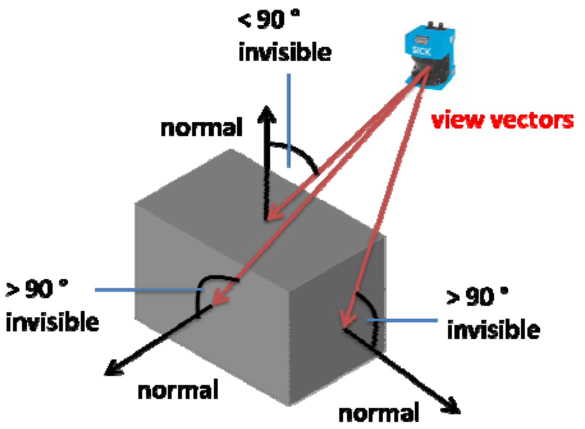

Figure 4: Back face culling for a cubical object

Result of the back face culling operation is a reduced set of visible surfaces for which the intersection with the ray must be computed. However, back face culling is a local operation, therefore it could be possible that a surface that is defined as visible can be occluded by another logistic object. This issue is addressed in the following subsection.

Computing intersection points of a ray and all visible surfaces of a logistic object requires methods from the field of analytical geometry. The intersection between a line and a face has to be computed. Therefore, equation 1 is inserted in equation 2 which results in equation 3.

$$\vec{n}(\vec{p} + \lambda\vec{v}) = b \tag{3}$$

Equation 3 is solved for lambda. This value is inserted in equation 1 and thereby the possible intersection point is computed. As the face is in principle infinite, the intersection point is checked whether it is within the boarder of the face of the cubical object. After the checking the intersection with every object of the packaging scenario, each ray has a set of intersection points. The number of elements of the set reaches from zero till the number of all faces of logistic objects in the packaging scenario. Because of the possible long computation time, the intersection computations are implemented in C-Code instead of using the script language of MATLAB.

4.3. Sensor Image Generation

The generation of a TOF sensor image requires a processing step that identifies from the set of intersection points the closest point to the position of the sensor for each simulated ray. For this purpose, a standard computer graphic technique is used again. A z-buffer algorithm is a two-dimensional array that saves current depth information for each pixel (Agoston 2005). Here, the number of columns of the z-buffer is equal to the number of scanning points per line of the simulated sensor. The number of rows depends to the settings of the rotating unit of the sensor. The z-Buffer can store a value for each ray that is simulated. In order to generate the simulated sensor image, the virtual packaging is scanned line by line. The distance of the scan points depends on the angular resolution of the simulated sensor. If an intersection between a ray and a face of a logistic object is detected, the distance from sensor to intersection points is computed and stored in the z-buffer. When another intersection point of the same ray with another face is detected, the distance is computed again and compared with the one in the z-buffer. In the case of a closer intersection point, the distance stored in the z-buffer is overwritten with the new distance value. Thereby, the closest intersection point for every ray is determined and the TOF sensor image that contains depth information is simulated.

4.4. Experiments

For evaluating the simulation methodology, complete packaging scenarios with cubical logistic objects are imported in the simulation platform and a TOF sensor image is simulated. For illustration two test scenarios are presented. The first one represents a simple packaging scenario with a few cubical logistic objects. Figure 5 illustrates the packaging scenario and the corresponding TOF sensor image.

The second scenario is generated within the context of (Scholz-Reiter 2009). It represents a preferably optimal loading solution of packaging plan by using wall building approaches. Through the two examples, the simulation platform is validated whether the simulation methodology is suitable to generate realistic simulated sensor images of virtual packaging scenarios. The simulation results of the second scenario are illustrated in figure 6.

a) b)

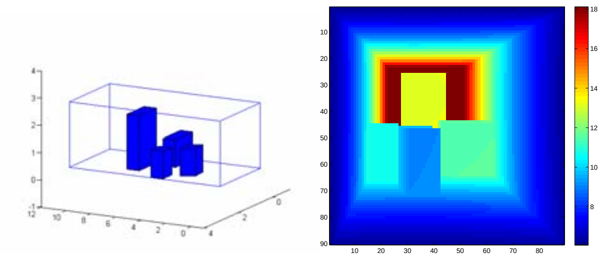

Figure 5: a) Packaging Scenario b) Sensor Image

The packaging scenario of experiment 1 has 4 logistic objects and 10 visible faces to be scanned. The computation time was 0.04 seconds. The sensor scanned 8011 points including container walls. Figure 5b visualizes the simulated sensor image that contains depth information. The depth is coded by colour, whereby reds represents more remote distance than blue.

a) b)

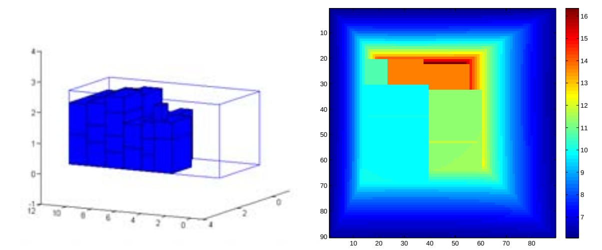

Figure 6: a) Packaging Scenario b) Sensor Image

Experiment 2 has 25 logistic objects and 51 visible faces to be scanned. The computation time was 0.19 seconds. Since the sensor setting was the same like experiment 1, the number of scanned points including container wall was equal to 8011 points. In this experiment, a great challenge for object recognition becomes visible. The two closest logistic goods that are stacked directly above each other on the left side of the container can be a great problem for an object recognition system. Based on the 3D sensor image, figure 6b, the object recognition system cannot distinguish whether it is a large object or several smaller objects. This issue must be solved by a suitable segmentation algorithm during the image analysis step.

5. CONCLUSION

Automatic unloading of universal logistic goods out of a container is a big technical challenge, because of undefined size and shape of goods. In order to identify the position and orientation of a good, a suitable object recognition system is necessary. Therefore, a new approach is presented that uses simulated sensor images. These images are generated by using standard computer graphic techniques. They will be used in two different ways. Firstly, they will be used for creating a model database for template matching. Secondly, they will be applied as training data for a classification procedure. These two aspects will be implemented and evaluated independently from each other.

The next step in the presented research work will be the integration of geometric models for cylindrical logistic objects and sacks. After integration, suitable intersection algorithms must be implemented in order to simulate the TOF measurement principle. Additionally, object recognition algorithms will be tested with ideal sensor images that are created by the simulation platform. Furthermore, a demonstrator platform will be constructed in future research activities. Then, realistic packaging scenarios with logistic objects can be generated and scanned from a TOF scanner. By recreating and scanning the same scenario in the simulation platform, a comparison of the real and virtual sensor image can be performed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support by DFG (German Research Foundation) for the EMOSES project.

REFERENCES

Aden, D., 2008. Seehäfen – Logistische Netzknoten der Globalisierung. *Proc. Das Beste der Logistik*, Berlin, Heidelberg (Germany), 391–403.

Agoston, M. K., 2005. Computer Graphics and Geometric Modeling – Implementation and Algorithms, USA.

Akbiyik, H., Kirchheim, A., Echelmeyer, W., 2009. *Latest Trends in the Container Market*. Bremer Value Report, vol. 3, Bremen (Germany).

Bishop, C., 2006. Pattern Recognition and Machine Learning. New York (USA).

Burwinkel, M., Echelmeyer, W., Wellbrock, E., 2009. Development of robotic systems for seaports by means of modular approaches. *Proceedings of the International Conference on Robotics and Applications*, Cambridge (Massachusetts, USA).

Echelmeyer, W., Kirchheim, A., Wellbrock, E., 2008. Robotics-Logistics: Challenges for Automation of Logistic Processes. *Proceedings of the IEEE International Conference on Automation and Logistics*, Qingdao (China), 2099-2103.

Forsyth, D.A., Ponce, J., 2002. Computer Vision – A modern approach. Prentice Hall, Upper Saddle River (USA).

Keller, M., Orthmann, J., Kolb, A., Peters, V., 2007. A Simulation Framework for Time-Of-Flight

Sensors. *International Symposium on Signals, Circuits and Systems*, Iasi (Romania).

Keller, M., Kolb, A., 2007. Real-time simulation of time-of-flight sensors. *Simulation Modelling Practice and Theory*, vol. 17, issue 5.

Kirkegaard, J., Moeslund, T.B., 2006. Bin-picking base on Harmonic Shape Context and Graph-Based Matching. *18th Conference on Pattern Recognition*, Hong Kong (China).

Kukko, A., Hyyppä, J., 2007. Laser Scanner Simulator for system analysis and algorithm development: A case with forest measurements. ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo (Finland), 234-240.

Meissner, D., Dietmayer, K., 2010. Simulation and Calibration of Infrastructure based Laser Scanner Networks at Intersections. *IEEE Intelligent Vehicles Symposium 2010*, San Diego (USA), 670-675.

Scholz-Reiter, B., Burwinkel, M., Isenberg, M.-A., Heger, J., Lampe, W., 2009. Automatische Konsolidierung von Seecontainern – Beschleunigung des Beladevorganges mittels dynamischer Beladungsplanung und automatischer Beladung. *Werkstattstechnik online vol. 99*, 166-171.

Scholz-Reiter, B., Echelmeyer, W., Wellbrock, E., 2008. Development of a Robot-Based System for Automated Unloading of Variable Packages out of Transport Units and Containers. *Proceedings of the IEEE International Conference on Automation and Logistics*, Qingdao (China), 2766-2770.

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A. 2011. Real-Time Human Pose Recognition in Parts from Single Depth Images. *IEEE Computer Vision and Pattern Recognition 2011*, in Press, Colorado Springs (USA).

Uriarte, C., Burwinkel, M., Gorldt, C., 2010. Trends der Objekterkennung von Robotiksystemen in Logistikfabriken. *15. Magdeburger Logistiktagung – Effiziente Infrastrukturen*. Magdeburg (Germany).

AUTHORS BIOGRAPHY

Bernd Scholz-Reiter (born in 1957) holds a degree in Industrial Engineering from the Technical University Berlin (TUB). Following his PhD (TUB, 1990), Bernd Scholz-Reiter served as post-doctoral fellow in the department for Manufacturing Research at IBM T.J. Watson Research Center in Yorktown Heights, U.S.A. From 1994 to 2000 he held the Chair for Industrial Information Systems at the newly founded Brandenburg Technical University at Cottbus, Germany. In 1998 he founded the Fraunhofer Application Center for Logistic Systems Planning and Information Systems in Cottbus, which he headed until 2000.

Since November 2000 Bernd Scholz-Reiter is full professor and holds the new Chair for Planning and Control of Production Systems at the University of Bremen, Germany. Since 2002 he also serves as Managing Director of the BIBA – Bremer Institut für Produktion und Logistik GmbH (BIBA) at the University of Bremen.

Since July 2007, Bernd Scholz-Reiter is Vice President of the German Research Foundation (DFG).

Hendrik Thamer was born in Bad Soden/Taunus, Germany in 1981. He graduated from the University of Oldenburg in 2009 with a diploma in computer science focusing on embedded systems and micro robotics.

Since 10/2009 he has been working as a research scientist at the University of Bremen with Prof. Scholz-Reiter. His research interests are in the field of 3D object recognition and image processing as well as in developing of image processing systems for quality management in micro production.