ORIGINAL PAPER

# Automated detection of euro pallet loads by interpreting PMD camera depth images

F. Weichert · S. Skibinski · J. Stenzel ·
C. Prasse · A. Kamagaew · B. Rudak ·
M. ten Hompel

**Abstract** In this study, a novel approach for the detection of parcel loading positions on a pallet is presented. This approach was realized as a substantial change in comparison with traditional system design of contour detection in de-palletizing processes. Complex 3D-vision systems, costly laser scanners or throughput decreasing local sensor solutions integrated in grippers are substituted by a low-cost photonic mixing device (PMD) camera. By combining PMD technology and a predetermined model of loading situations, stored during assembling the pallet, this approach can compensate for the drawbacks of each respective system. An essential part of the approach are computer-graphics methods specific to the given problem to both detect the deviation between the nominal and the actual loading position and if necessary an automated correction of the packaging scheme. From an economic point of view, this approach can decrease the costs of mandatory contour checking in automated de-palletizing processes.

**Keywords** Computer vision · De-palletizing · Contour check · Automated handling · PMD technology · 3D-Imaging

F. Weichert (✉) · B. Rudak
Department of Computer Science VII, Technical University
Dortmund, Dortmund, Germany
e-mail: frank.weichert@tu-dortmund.de

S. Skibinski · J. Stenzel · C. Prasse · A. Kamagaew
Fraunhofer Institute for Material Flow and Logistics,
Dortmund, Germany

M. ten Hompel
Chair for Materials Handling and Warehousing,
Technical University Dortmund, Dortmund, Germany

## 1 Introduction

Driven by the internationalization of supply chains and global competition, new trade corridors, privatisation in emerging countries [50], and typical processes in logistics became significantly more complex and dynamic during the last years. Among other things, shorter product life cycles, mass customization and stricter quality requirements [16, 54] demand for an increasing efficiency in this field of industry. If properly dimensioned and highly utilized, automated systems could be a solution. To keep up in the global competition, especially in high-wage countries, an increasing automation in facility logistics is unavoidable. But the use of modern technologies is not only limited to developed countries. For example, german companies producing intralogistic equipment registered in 2010 the highest rates of growth in export trade for the so-called BRIC-states (Brazil, Russia, India and China) [59]. Concerning some industrial sectors like food, brewery, chemistry, plastic and wood, automated palletizing is established as a quasi-standard for production subsequent packaging of goods [17]. Highly productive systems for specific palletizing tasks are available on the market (cf. [32, 33]).

Unstacking of palletized goods (de-palletizing) is usually done manually. High demands for flexibility (e.g., different size and shape of goods) are just one reason against fully automation. Formerly mentioned high-performance solutions are often too cost-intensive to make them economically justifiable. Generally, required information for (an automatic) de-palletizing is acquired by sensor units like laser scanners or expensive computer vision systems (cf. Sect. 1.1)

Accurate and stable handling of goods is no longer technical challenging but the detection of an exact gripping position. Especially, after removing typically used load

securing measures (e.g., wrapped film, etc.), transportation of the pallet could cause translational and/or rotatory displacements of the load. Hence, known nominal packing positions are no longer valid, and this gap has to be detected individually [61]. Humanization of labor should, beside the need to make de-palletizing more efficient, be a main objective. Adverse working conditions and high physical stress because of manually carrying heavy stocks (e.g., cold store) can be reduced or avoided by using automated handling solutions [17, 27]. However, the advent of new technologies may alleviate these challenges; especially, *Computer Vision using Photonic Mixing Device (PMD) technology* and *Embedded Systems* are particularly promising.

In this paper, a new approach is proposed to detect the loading situation on euro pallets for automated de-palletizing of packages. In the first instance, packages are automatically stored on a loading device (e.g., euro pallet) by means of a handling unit (e.g., consisting of an articulated robot and a gripper). In a further process step, single unit loads will be detached (de-palletized) by a similar device at the final destination or for order picking. In all cases, the outline of the loading unit has to be checked before processing de-palletizing to prevent collisions. To perform this task, the actual position of at least every object in the loading units top layer has to be detected. Therefore, a new challenging method of contour check, which can provide a reliable and fast gripping position to a handling unit, is highly desirable—especially if it is economically competitive to known solutions. An iterative two-stage algorithm adapted to the special problem of de-palletizing euro pallet loads will be presented in this paper. It matches the object geometry directly with the object representing point cloud and limits the possible transformations to three degrees of freedom. The new approach combines high cost efficiency, high robustness of the detection process, high time efficiency inter alia by reducing the search domain into three dimensions (translation about $x$-, $y$-axis and rotation about $z$-axis) and high flexibility in terms of supporting objects with variable dimensions and different material properties.

The task outlined above is structured in five sections. Following this introduction containing a general motivation for this work, state-of-the-art is presented for automated de-palletizing and detection of loading conditions and vision technology in Sect. 1.1. Next, an introduction to the TOF (time-of-flight) and PMD (photonic mixing device) technology is given (Sect. 1.2). Following these introductory comments, an overview on various deviation scenarios and the main challenges from the logistics point of view is given providing the basis for the following considerations (Sect. 2) The core of this paper—a novel automated estimation of the euro pallet load using the PMD

technology—is represented by a pipeline in Sect. 3. Section 4 deals with the proof of concept, including a description of the used test setup and several results. In Sect. 5, a qualitative comparison of existing commercial solutions and the new approach is presented. Finally, Sect. 6 summarizes the main statements of the paper and gives an outlook on future work.

## 1.1 State-of-the-art

In view of the amount of articles in the field of (automated) de-palletizing with a focus on logistic environments, the following presentation only takes contributions into account which have a direct impact on the current study.

### 1.1.1 Cases of automated de-palletizing

Two procedures of de-palletizing can be distinguished in practice. If the position and orientation of the loaded packages are available—usually stored in a central database—they are transmitted to the processing point (handling device). To prevent collision checking, the outline of the loading obtained by 3D-scanners or gripper mounted sensors is mandatory. A vision system has to detect the position and orientation of the packages, if the necessary set of data does not exist. This leads either to expensive solutions or to the extension of handling and processing times.

In general, the existing solutions for contour detection can be sorted as shown in Fig. 1. Category A (cf. Fig. 1, left side) includes all solutions employing a known pattern model of the pallet load. The majority of systems employed in industrial de-palletizing applications so far do not contain any sensors (cf. Fig. 1, A1), for example, preprogrammed gantry robots process bulk de-palletizing task in a strictly controlled environment very efficient, but they fail in adverse environments where the pallet's position is not well defined [28]. Moreover, the grippers could be equipped with specific sensors (cf. Fig. 2a, b) to detect the pallet's outline (cf. Fig. 1, A2), due to iterative approximation of the detection process, this leads to high sampling intervals [44]. Further alternatives to the named solutions are simple systems employing light barriers (laser sensor) respectively light curtains, but this solution has no relevance in the industrial practice. An efficient alternative for contour detection (cf. Fig. 1, A3) is the combination of packaging pattern models and 3D-sensors (e.g., time-of-flight sensors), which will be presented in this paper and discussed in paragraph "Sensor-based automated de-palletizing" in more details. On the right part of the systematic B shown in Fig. 1, different ways to detect the outline of loading goods are represented, without having any a priori knowledge concerning the packing
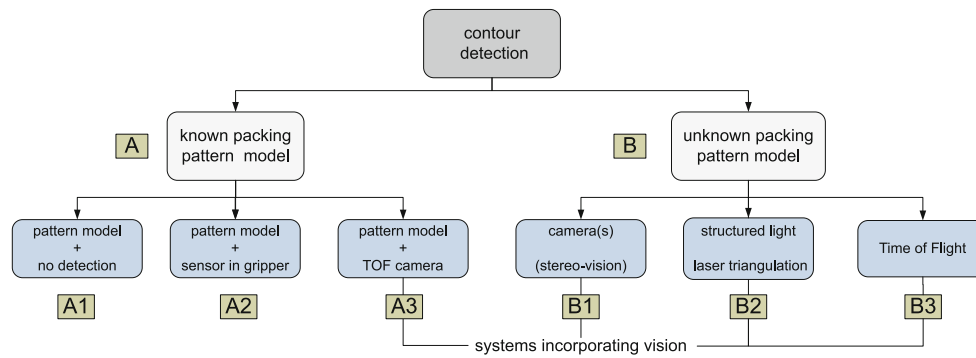
**Fig. 1** The figure shows different ways to detect the outline of the single loading goods. Distinction of cases could be done as follows: *A* are methods requiring a specific knowledge of the nominal pattern model; *B* describes solutions where the position is not necessary to determine the (gripping) position of the packages

**Fig. 2** Different technical solution for de-palletizing
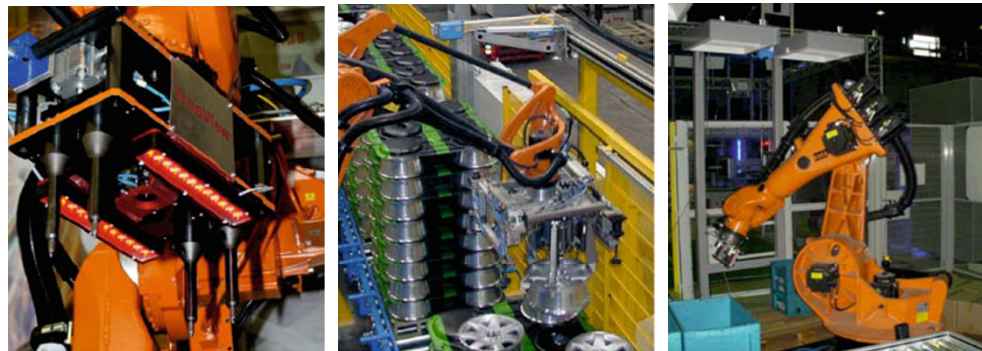


**(a)** Gripper equipped with specific sensors for the detection of the gripping position [1]

**(b)** Gripper-Camera-System for pick-up aluminum rims [29]

**(c)** De-palletizer with laser scanner and camera at Fraunhofer-Institute for Material Flow and Logistics

position. In these cases, the shape, position and orientation of loading goods have to be acquired by using complex sensor systems.

### 1.1.2 Storing and management of palletizing data

Basic requirement of using a known pattern model for detection—as proposed in this paper—is the availability of packaging data at the local handling device (de-palletizer). For that reason, the data from building the unit load (palletizing) have to be linked to the pallet. Beside a rather pragmatic approach of storing the packing position in a central database and marking the pallet with a number, in [46] detailed concepts using radio frequency identification (RFID) are introduced. Additionally, a data storage management is presented in order to assure efficient access times for reading and writing the tags.

### 1.1.3 Radio frequent localization

Besides mentioned vision-based solutions, locating the objects on the pallet by RF technologies would solve almost all problems. Main requirement is an adequate accuracy concerning distance and orientation of the palletized goods. Several approaches for the localization of active and passive RFID tags have been proposed in the past. For example, a system called LANDMARC uses the received signal strength (RSS) [39]. Ubisense, developed at the University of Cambridge, is another commercial ultra wide band (UWB)-based localization system with active tags. Localization is performed using both xTime Difference Of Arrival (TDOA) and Angle of Arrival (AOA) measurement [34]. In [4], the localization of passive tags relies on the TDOA. Mojix, an US-american company designed a real-time localization system called STAR, which is able to establish standard passive (UHF) RFID Tags [37]. Choi et al. [12, 13] present a passive tag localization based on an improved k-Nearest Neighbor algorithm (KGNN). Using minimal interrogation power and multilateration, Almaaitah et al. [2] introduce two methods [adaptive power with antenna array (APAA) and adaptive power multilateration (APM)] for three-dimensional localization of passive tags. In [43], a concept of zone-based and exact localization methods for passive tags

is described. An improvement of accuracy and speed is achieved by a combined cascaded coarse-to-fine approach. Following a semi-passive tag (Sensatag board) approach, Athalye et al. [3] describe a method of sensing proximity and standard backscatter modulation. Sample et al. [52] present an approach based on both RF and vision using a wireless identification and sensing platform (WISP). The tags are equipped with an integrated led light and detected by two sensors—a camera and a RFID antenna. Finally, they are able to control a gripping process performed by an industrial robot. Besides the localization of RFID tags, Bouzakis et al. [10] present an approach for detection based on RFID transponders. The method acquires the exact position of two RFID transponders on each parcel in order to detect the parcel's position and is adapted for detecting and handling with high-frequency accumulating and non-palletized parcels.

Summing up todays approaches of RF localization systems (see compare Table 1), an adequate accuracy (especially orientation) for the mentioned application cannot be guaranteed in the most cases (1–9). Furthermore, the monetary cost of localization, using solutions 1, 4, 9, and 10 (see Table 1), exceeds the actual cost of the object being tracked and hence is economically unfeasible.

### 1.1.4 Sensor-based automated de-palletizing

In view of falling prices systems incorporating vision (cf. Fig. 1, A3, B) is an increasing market segment compared to non-vision-based solutions [14]. Vision systems with additional assisting sensors like structured light [28, 45, 41], multiple laser spots [61] or laser triangulation sensors [22, 8] (cf. Fig. 1, B2) are available on the market [6, 18]. Time-of-flight sensors just like 2.5D laser scanner to acquire a point cloud [7, 40, 26] are widely spread in the industry [40, 60] (cf. Fig. 1, B3). These systems are

characterized by their high flexibility and robustness (cf. Fig. 2c). Usually designed for complex bin-picking tasks computing and cycle times are disproportionately high for unloading tasks. Additionally, further approaches for automated detection exist, especially regarding the bin-picking problem which can be seen as a part of the automated de-palletizing problem. Zhang et al. [61] describe an approach for robotic de-palletizing using uncalibrated vision and 3D laser-assisted image analysis. To compensate the absent calibration, the camera space manipulation method (CSM) is used which relies on a local calibration method. The authors do not consider the runtime which is a main task in the industrial process.

Besides laser scanner with long acquiring times because of mechanical movement (pan) *Photonic Mixing Devices* (PMD) are an innovative approach to Time-Of-Flight (TOF) sensors. Today PMD technology is used for classification of single objects [24, 30] providing volume and coordinates (e.g., for packages [55]) or bin picking [51, 36, 26]. Compared to the presented new approach, these solutions are adversely affected by the low resolution of PMD cameras. They have to compensate it with long acquiring times, for example, caused by the need of multiple pictures. Even though nominal position data of the pallet load are not required, complete knowledge of the object is mandatory (e.g., image or CAD model) for all mentioned approaches, respectively primitive segmentation methods are used [41]. An integrated bin-picking concept based on a hybrid state machine using a PMD camera and a lightweight robot is presented by Fuchs et al. [21] which is commercially available. The underlying architecture includes a human–robot interaction on the one hand to compensate detection faults by the integrated computer vision system and on the other hand to serve human safety. The computer vision system is hierarchically composed and consists of hough transformation on the coarse level followed by particle filtering and clustering and finally a local search with iterative closest point (ICP) algorithm. Because of the light weight robot, the solution is inappropriate for handling with high dimensions and heavy loads as euro pallet loads. While the focus of this solution lies in the closely human–robot interaction, robots for automatic de-palletizing euro pallets are due to human safety isolated by a gate from the environment. A hybrid state machine as described by Fuchs is due to human safety inappropriate in relation with heavy weight robots. Hence, the commercial available solution presented Fuchs tend special bin-picking scenarios with small dimension and light weight objects. Another approach for industrial bin picking can be found in [11] which is based on a modified RANSAC algorithm [19] in order to match the CAD model with the laser scan data. Ghodabi et al. [23] describe a system for detection and classification of moving objects

**Table 1** Comparison of accuracy of RF localization systems

| No. | Approach | Tag | Accuracy [distance error (cm)/ rotation error] |
|---|---|---|---|
| 1 | LANDMARC (RSS) [39] | Active | 100/n.a. |
| 2 | Ubisence (AOA) [34] | Active | 15/n.a. |
| 3 | TDOA [4] | Passive (SAW) | 20/n.a. |
| 4 | Mojix [37] | Passive | 100/n.a. |
| 5 | KGNN [12, 13] | Passive | 9.5/n.a. |
| 6 | APM (multi reader) [2] | Passive | 32/n.a. |
| 7 | APPA (single reader) [2] | Passive | 48/n.a. |
| 8 | Cascaded coarse-to-fine [43] | Passive | 37/n.a. |
| 9 | Sensatag [3] | Semi-passive | 14/n.a. |
| 10 | WISP [52] | Passive and led | 0.1/n.a. |

based on support vector machines (SVM) using 3D data of a PMD camera. To train the SVM classifier, two data sets are derived from each image set, one based on heuristic features and the other based on the principal component analysis (PCA). In [53], the authors in contrast to this paper mainly focused on a supervised classification (by an artificial neural network) of universal goods, which limits the flexibility for multiple environments and industrial applications.

In contrast to the presented approaches solving the bin-picking problem, the approach presented in this paper additionally detects and verifies the complete (visible) euro pallet load and visualizes it afterward. This allows an early intervention of the user in case of unacceptable deviations according to the reference positions of the euro pallet load respectively in case of the absence of particular parcels.

## 1.2 TOF and PMD technology

Common 3D-sensors often use time-of-flight measurement principles for environment detection. In this area, photonic mixing device sensors have appeared as a cost-efficient solution for a broad variety of applications in automation engineering. In the following, a short introduction into the TOF and PMD technology is given. For more details see, for example, Kolb et al. [30]. A PMD sensor measures the distance to an object by emitting modulated infrared light followed by determining the phase shift between the emitted and reflected light. Thus, the relation between the measured phase shift $\mathbf{\Phi}$ and the modulation frequency $\omega$ results in the object distance $d$:

$$d = \frac{c \cdot \mathbf{\Phi}}{2 \cdot \omega}. \tag{1}$$

The output of the sensor is a two-dimensional depth map where typical resolutions reach from $16 \times 1$ to $200 \times 200$ depth measurement points. In addition to the distance measurements, the sensor also delivers a gray scale image with the same resolution.

Figure 3 visualizes the unambiguity range of a PMD camera using a modulation frequency of $f_{mod} = 20$ MHz. Objects that are beyond the range of 7.5 m will be seen at a distance of $d$ mod 7.5 m. This constraint of the PMD camera does not have an impact in the test setup as all objects are less than 7.5 m away from the sensor. The PMD
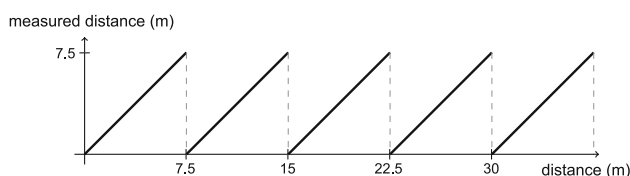
sensor used in this paper has a resolution of $64 \times 50$ depth pixels and a field of view of 30° (horizontal) and 40° (vertical). The accuracy of the sensor mainly depends on the used modulation frequency and the reflectivity of the observed objects. A higher modulation frequency results in a higher accuracy of the sensor measurements (cf. [47]) as objects with high reflectivity (e.g., white paper) also do. The measurement errors are classified into systematic errors like the integration-time-related error, the wiggling error or the temperature-related error and non-systematic errors like signal noise and multiple light reception [20]. While the non-systematic errors are hardly to compensate, the focus lies on the correction of the systematic errors by identifying the error function. The measurement principle also leads to an unambiguity range of the PMD sensor that is dependent on the modulation frequency.

## 2 Logistics environment

In this section, possible loading scenarios, the general approach to detect gaps between the nominal and actual loading situations and the challenges to solve the defined task while achieving all logistical demands are described.

### 2.1 Detection of loading conditions on a pallet

Due to intra-company transportation and handling of pallets after removing securing measures like wrapped film, stretch hood etc., translational and/or rotatory displacements of the load are most probable. Possible variations of the loading situation are visualized in Fig. 4. Figure 4a shows the ideal state: the actual load is identical with the nominal one. From the load shown in Fig. 4b, a parcel (front right) has been removed (e.g., intervention by a human), and a parcel has been manually added at the top of the load. Figure 4c shows a possible deviation from the nominal state caused by an abrupt slow down (e.g., braking maneuvers of fork lift). In Fig. 4d, some packages have been randomly shifted and rotated. The proposed approach for contour detection contains process steps shown in Fig. 5. A flow process chart visualizes the automated load detection. At first, a depth image of the actual load is acquired by a PMD camera. Further the nominal model is being compared with certain actual load represented by the acquired depth image. If both, the nominal model and the acquired depth image match and deviations are within tolerance, for example, automated de-palletizing can proceed. If deviations are too huge or unidentified objects are present, an operator can be notified. After adjusting the load, the automated de-palletizing could continue.

**Fig. 3** Unambiguity range of a PMDcamera
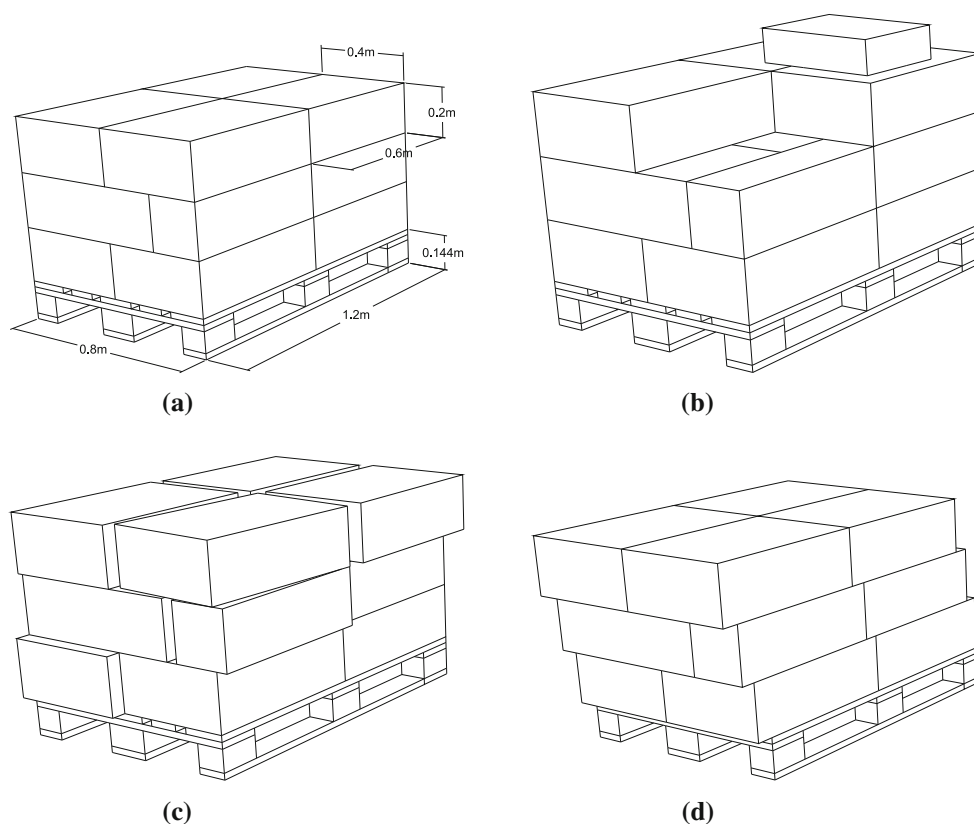
Fig. 4 Deviations between different nominal and actual loads
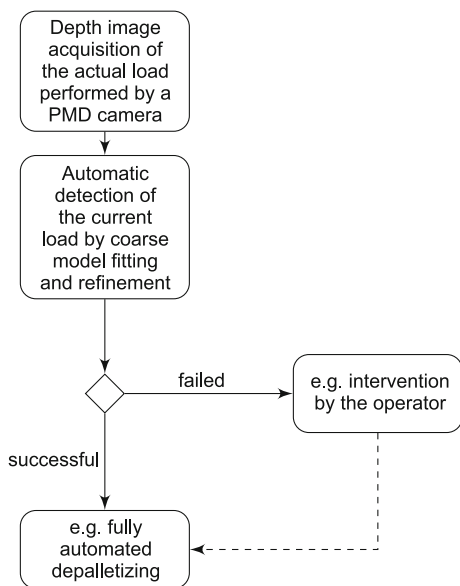


Fig. 5 Flow process chart of the automated load detection

## 2.2 Gap analysis

In order to summarize the main advantages and drawbacks of the existing approaches, Table 2 shows a survey of the mentioned solutions. Analyzing the specific characteristics

of the mentioned approaches, most of the solutions promise an adequate accuracy. Drawbacks are either high cycle time or high cost for extensive equipment. Hence, the new approach should stand out due to low cost, high performance and adequate accuracy. This leads to the challenges in a logistics environment.

## 2.3 Challenges

Depending on the specific handling system, especially the gripper, accomplishable accuracy of position detection is the essential factor of success. Only by ensuring a sufficiently low tolerance between real contour and processed camera data, significant enhancements toward existing applications could be achieved. In case of necessary additional sensors to compensate inadequate accuracy, considerably advantages regarding the price are no longer likely. Equally reduction of handling speed caused by additional sampling in the close-up range to prevent collision will increase the cycle time of the whole process. Beyond accuracy of detection, further requirements concerning robustness and flexibility have to be mentioned. In this context, robustness can be defined as repeatability and stability of the implemented algorithms. Flexibility mainly refers to different surfaces of the loaded goods (bright,

**Table 2** Main advantages and drawbacks of existing methods for load detection in the field of de-palletizing

| Category cf. Fig. 1 | Method cf. State-o- the-art | Pro Advantages | Contra Drawbacks | Accuracy Position (mm)/Orientation (deg) |
|---|---|---|---|---|
| A1 | No detection | Very simple and cheap | Inflexible, instable against changing positioning conditions | Providing an accurate adjustment of the pile <10/<2 |
| A2 | Border detection using laser sensors [44] | Simple and cheap | Slow (12s per cycle) | n.a. |
| B1 | Stereo cameras [1] | Simple and cheap (one camera), high accuracy | Sensitivity to lighting conditions, quite slow (3s for detection) | ±1/n.a. |
| B1 + B2 | Uncalibrated vision and 3D laser-assisted image analysis [61] | Reliable gap insertion (gripper), flexible to perform less complex 2,5D task | Extensive equipment: min. 2 cameras, single laser pointer and multiple laser pointer | n.a. |
| B3 | 2D range imagery [28] | No influence of illumination and appearance, minimal equipment (one laser scanner) | Manual detection of orientation | <20/<5 |
| B2 + B2 | Model based range images [8] | High accuracy, scalability | High computational cost, very complex object model, sensors with high resolution | ±1/±5 |
| B1 + B2 | 2D camera and laser scanner [9] | Flexible, performs complex 3D task | Long cycle times (movement of scanner), expensive (equipment) | n.a. |

dark, matt, glossy, etc.) as well as changing environmental influences (light, temperature etc.). To differentiate from established solutions, achieving the defined requirements on the highest level is mandatory. Briefly consolidated the main challenges are:

– high cost efficiency
– high robustness
– high time efficiency
– high flexibility

In view of these challenges, the automated load detection pipeline was designated, which is represented in the following section.

# 3 Automated load detection pipeline

In order to perform an automatized de-palletizing of euro pallets, it is crucial to determine the actual positions of all parcels. As described in Sect. 2.1, the actual and nominal parcel positions often differ from each other, so that a refinement of the given model is a must. The novel automated estimation of the euro pallet load can be presented by a pipeline—the whole process is visualized in Fig. 6. The schematic representation of the automated load detection pipeline can be subdivided into three groups of tasks. The first one is the acquisition of the PMD depth data and preprocessing as well as of the loading state facilitated by a central database to store the RFID tag data. Accordingly, the main part of the pipeline, the model fitting, follows. This step can be again subdivided into two tasks, coarse fitting of the model **T** into the PMD point cloud

**P** and fine fitting of the individual models $t_i \in T$ (representing single parcels). The result of the previously outlined model fitting is the status of the euro pallet load. Optionally, an additional step for verifying the fitting results can be applied in order to prevent passing of uncorrectable load information to the following logistic process chain. In the following sections, each of these steps is explained in detail.

## 3.1 Palletizing model

As proposed in Sect. 2, the nominal model and the initial orientation by considering the attenuation of the response signal are obtained from RFID tags attached to each euro pallet. Furthermore, the new approach is currently designed and optimized for cuboid-shaped parcels that can be of different dimensions and proportions arranged on multiple layers on a euro pallet. The defined model **T** will be now initialized and later fitted into the point cloud as following:

$$\mathbf{T} = (T_i), \quad t_i = \left(t_i^z, a_i^x, a_i^y, b_i^x, b_i^y, c_i^x, c_i^y, d_i^x, d_i^y, h_i, i_i, f_i, p_i\right), \quad 1 \le i \le n. \tag{2}$$

The model **T** consists of $n$ tuples $t_i$. Each tuple $t_i$ describes one cuboid (e.g., one parcel). Figure 7 visualizes the model description. The variables $a_i^x$ till $d_i^y$ depict the coordinates of each of the corners of the cuboid and $t_i^z > 0$ stands for the $z$ coordinate of the upper plane of the cuboid, while $h_i$ is describing the distance between the upper and the bottom plane. The variable $i_i$ defines if the specified cuboid needs individual refinement in the fine fitting state of the pipeline (see Sect. 3.5). Further $f_i \in \{0, 1\}$ denotes detected cuboids
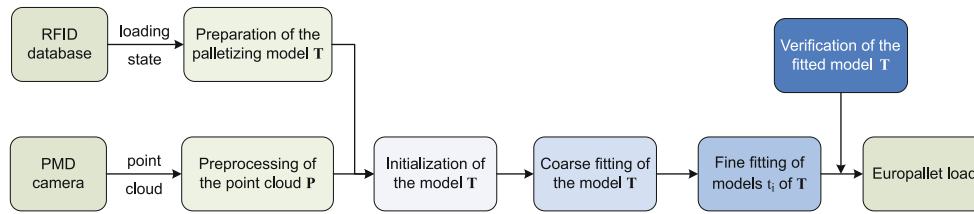
**Fig. 6** Schematic representation of the automated load detection pipeline. The steps of the detection pipeline can be subdivided into three groups of tasks. The first one is the acquisition of the PMD depth data and preprocessing, and the second one does the model fitting for providing the euro pallet load. Optionally, a third step of verification of the fitting results is applied
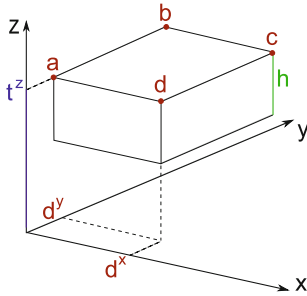


**Fig. 7** Visualization of the model coordinates of a particular cuboid

and $p_i$ holds the supporting point count of cuboid $i$. It should be mentioned that by adapting the model, different shaped parcels are supported.

### 3.2 Preprocessing of PMD depth data

As described in Sect. 1.2, the evaluated PMD cameras are suitable to deliver depth information and intensity information in various formats. The approach presented in this paper focuses, because of robustness concerns, on the depth imagery only and does not require any intensity information. Further the depth information is transformed into world coordinates. So that each acquired point constitutes of a tuple $X_{ij}$ of Cartesian $x_{ij}$, $y_{ij}$ and $z_{ij}$ coordinates. In order to be able to dismiss points, the formerly introduced tuple is extended by an additional boolean variable $d_{ij} \in \{0, 1\}$. The whole point cloud data structure $\mathbf{P}$ can be noted as following:

$$\mathbf{P} = (X_{ij}), \quad X_{ij} = (x_{ij}, y_{ij}, z_{ij}, d_{ij}), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n. \tag{3}$$

The resolution $m \times n$ of the matrix $\mathbf{P}$ is given by the resolution of the PMD sensor (cf. Sect. 1.2)

Previously to the actual algorithm, a preprocessing step is essential in order to exclude erroneous pixels caused by the acquisition process from further model fitting pipeline. The point here is to distinguish between two characteristics of artifacts, which are responsible for invalid depth tuples:

– artifact type 1: under- or overexposed pixels
– artifact type 2: outliers.

All sorts of invalid depth tuples have to be removed from the point cloud in order to ensure robustness of the detection following. Commonly under- or overexposed sensor pixels (artifact type 1) are caused by inappropriate integration times, but can also have their origin in surfaces of high reflectance, for example, parcel tape. The sensor center has shown itself as very sensitive to reflecting surfaces. Overexposed sensor pixels are marked by the PMD camera with $z_{ij} = -1$ and underexposed sensor pixels with $z_{ij} = -2$:

$$z_{ij} \in \mathbf{P} = \begin{cases} -1, & \text{if overexposed} \\ -2, & \text{if underexposed.} \\ z \geq 0, & \text{else} \end{cases} \tag{4}$$

In contrast to over- or underexposed pixels, outliers (artifact type 2) are not detected automatically by the PMD camera, so that the outlier detection has to be done by the proposed pipeline (cf. Fig. 8). Outliers can be caused like overexposured pixels by surfaces of a high reflectance, but also by shadowing which comes along with the use of an active light source for depth measurement. Both types of outliers can be eliminated by the use of a significance bound [15]. In order to better differentiate between the two formerly introduced types of outliers, the computation of the absolute value is abandoned, so that the formulation of the outlier detector can be noted as following:

$$\text{out}(z_{ij}) = \begin{cases} 1, & \text{if } \frac{z_{ij} - \overline{Z}}{\sigma(Z)} \geq \delta_{\max} \\ 1, & \text{if } \frac{z_{ij} - \overline{Z}}{\sigma(Z)} \leq \delta_{\min} \\ 0, & \text{else} \end{cases}. \tag{5}$$

If $\text{out}(z_{ij}) = 1$ then $z_{ij}$ is supposed to be an outliner and can be removed. The two constants $\delta_{\max}$ and $\delta_{\min}$ define the threshold from which on values are supposed to be outliers. $\overline{Z}$ stands for the arithmetic average and $\sigma(Z)$ for the standard deviation. So the point cloud $\mathbf{P}$ without over- and underexposured pixels and without outliers that constitutes the input of the next pipeline step can be formulated like this:
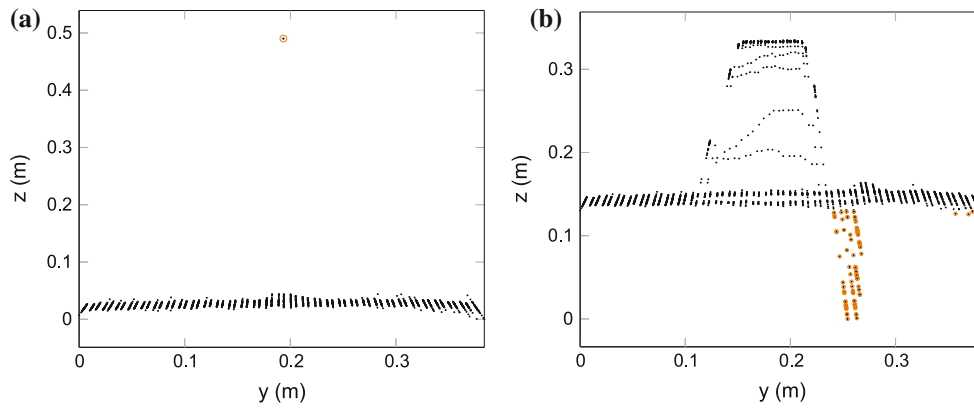
**Fig. 8** Occurrences of different invalid depth tuples (*circle* highlighted in *orange*): **a** suspiciously high $z$ value, which is caused by a surface of high reflectance and **b** suspiciously low $z$ values, which are caused by shadowing—further all objects of the scene have been shifted upwards along the $z$ axis. The axis dimensions are given in meters

$$\forall i \in \{1,\ldots,m\}$$
$$\times \left( \forall j \in \{1,\ldots,n\} \left( z_{ij} \in \mathbf{P} \wedge \left( \underbrace{z_{ij} \geq 0}_{\text{artifacttype1}} \wedge \underbrace{\text{out}(z_{ij}) = 0}_{\text{artifacttype2}} \vee d_{ij} = 1 \right) \right) \right).$$
(6)

The point cloud $\mathbf{Z}$ that has been pruned by all for deletion marked points will from now on be used instead of $\mathbf{P}$ and is defined as following:

$$\forall i \in \{1,\ldots,m\} \big( \forall j \in \{1,\ldots,n\} \big( d_{ij} = 0 \Rightarrow X_{ij} \in Z \big) \big).$$
(7)

The following step is crucial for reducing the search domain into the three dimensions, translation about $x$-, $y$-axis and rotation about $z$-axis which requires a planar alignment of the PMD camera to the ground plane. Hence, the camera tilt is corrected by calculating the ground plane using the random sample consensus (RANSAC) [19] method and transforming the point cloud such that the ground plane is planar to the $x$-$y$-plane of the reference coordinate system.

### 3.3 Initialization of the model

Initialization of the model is mainly done because of performance, but also of robustness considerations. Therefore, a suitable initial point for the model is needed. As a good initial point for the model the arithmetic average of all significant points has proven successful. The set of all significant points $S$ as all points above average is defined:

$$\forall X_{ij} \in Z \big( z_{ij} > \overline{Z} \Rightarrow X_{ij} \in S \big).$$
(8)

After determining the initial point, the whole model is translated according to the $x$- and $y$-axis so that the center of the model and the initial point correspond. In this pipeline stage, another performance improvement can be applied: cuboids which are according to their $t_i^z$ value above the upper bound of the point cloud $r_z$ are obviously not present and can therefore be directly discarded while considering some additional tolerance factor $\upsilon > 0$. Formally the discarding procedure can be noted as following:

$$\mathbf{T}' = \bigcup_{t_i \in \mathbf{T}} \big\{ \big( t_i \Rightarrow t_i^z \leq r_z \cdot (1 + \upsilon) \big) \big\}, \quad \mathbf{T} = \mathbf{T}'.$$
(9)

### 3.4 Coarse granularity model fitting

In the previous pipeline step, a proper initial position has been determined (cf. Fig. 10a). Now in the coarse model fitting step, the position of the whole model is supposed to be refined. This is done by us using an iterative approach based on key components of the ICP concept [5]. The basic idea of the ICP algorithm is to register two point clouds in a common coordinate system by an iterative approach. In each iteration step, the algorithm selects the closest points as correspondences and calculates the transformation (translation and rotation) for minimizing the deviation between the two point clouds. Many variants of the ICP algorithm have been proposed, that is, other strategies for point selection or parameters affecting the weighting. For a review of the state-of-the-art refer to the report of Rusinkiewicz and Levoy [49]. Despite the similarity to the ICP algorithm, the algorithm introduced in this work differs in a fundamental way from the general ICP approach. In general, the problem of estimating the point-to-plane correspondence is solved by an orthogonal projection of all points onto suitable planes. The detection of the euro pallet load utilizes only the point cloud $\mathbf{P}$ respectively $Z$ (the PMD depth data) in terms of Cartesian coordinates and a 3D palletizing model $\mathbf{T}$ but no second (additional) point cloud. While the general point-to-plane-ICP algorithm which requires the time-consuming calculation of

correspondences between the two (source and destination) point clouds, our approach is based on an efficient detection of supporting points for the model **T** in order to match directly the geometry with the point cloud.

Figure 9 visualizes parts of the following detailed description. In each iteration step of the coarse fitting approach, several transformations $\theta_i \in \Theta$ are computed (translations regarding the $x$ and $y$ axis at $\pm\tau$, rotations regarding the $z$ axis at $\pm\rho$), and afterward, the best fitting model is selected. The variable $\tau$ defines the upper bound for the sum of all translations and $\rho$ the upper bound for all rotations. Both $\tau$ and $\rho$ are multiplied after each fitting iteration by the corresponding factors $\tau_f$ and $\rho_f$. Additionally, the number of iterations can be limited by a constant $\mu_{max}$. For a closer consideration of the consequential trade-off between accuracy and speed refer to the evaluation (see Sect. 4) Transformations $\theta_i$ that are not within the point cloud's bounds are rejected. In order to prevent worsening the fitting by inappropriate transformations, the current model is also evaluated and compared with the models resulting from the current iteration step:

$$\Theta = \left\{ \theta_0 = \left( \theta_0^p = \text{point}(\mathbf{T}, Z), \theta_0^e = \text{euclid}(\mathbf{T}, Z), \theta_0^{\mathbf{T}} = \mathbf{T} \right) \right\}.$$

(10)

A transformation triple $\theta_i = (\theta_i^p, \theta_i^e, \theta_i^T)$ is compound of the number of points $Z$ that are covered by the model $\mathbf{T}$ ($\theta_i^p$), the overall costs of this coverage ($\theta_i^e$) and the model ($\theta_i^t$). The function that computes the point count can be defined as follows:

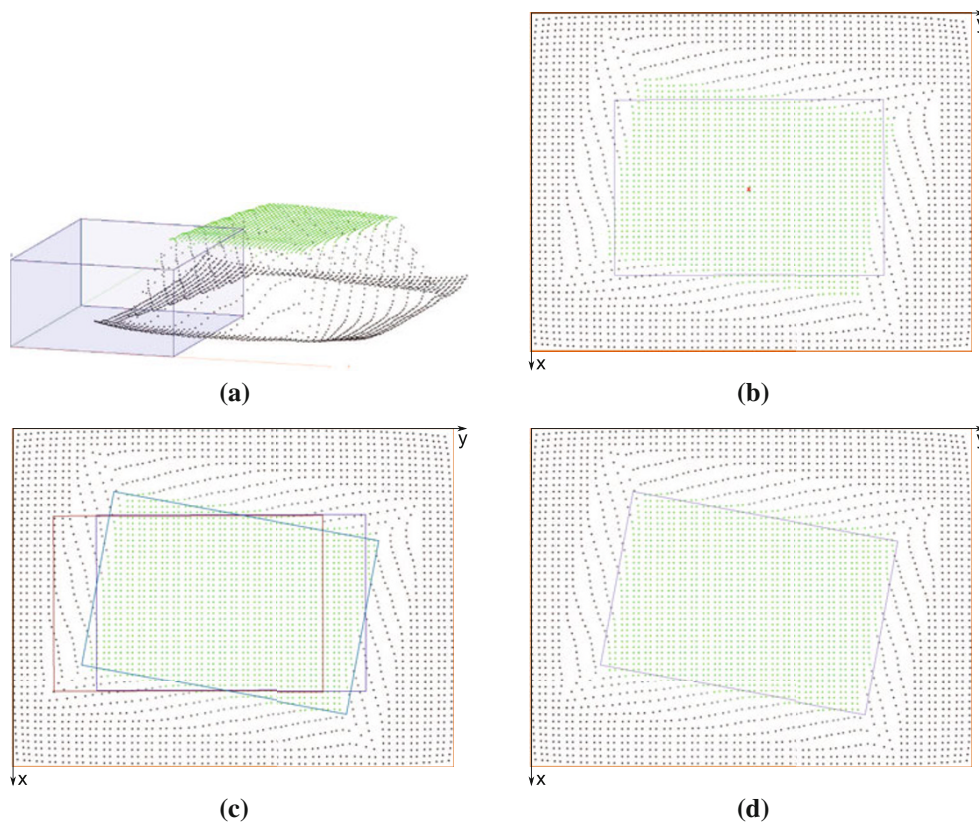$$\text{point}(\mathbf{T}, Z)$$
$$= \sum_{X_{ij} \in Z} \left( \begin{cases} 1, & \text{if} \quad \exists t_k \in \mathbf{T}\left( \text{inrect}_{\omega_k, ij} = 1 \wedge e_{k,ij}^z \leq \phi_c \right) \\ 0, & \text{else} \end{cases} \right).$$

(11)

It mainly consists of two components, the point in rectangle test and the calculation of the Euclidean distance:

$$\text{inrect}_{\omega_k, ij} = \left( \omega_k, x_{ij}, y_{ij} \right)$$

(12)

$$e_{k,ij}^z = \sqrt{\left( t_k^z - z_{ij} \right)^2}.$$

(13)

The first one is called $\text{inrect}_{\omega\_k, ij}$ assures that only points $x_{ij}$, $y_{ij}$ whose orthogonal projection lies interior the bounds



**Fig. 9** Particular steps of the detection algorithm. The relevant points with euclidean distance to the model according the $z$-coordinate serve as supporting point candidates to the model. The *orange rectangle* visualizes the bounds. **a** Model and point cloud before initialization (perspective view) **b** initialization of the model (orthogonal view) **c** an iteration step with the original and two (of six) modified templates, shifted and rotated (orthogonal view) **d** the resulting template maximizes the number of supporting points (relevant points which lie in the *rectangle*) (orthogonal view)

$\omega_k$ of at least one rectangle is counted—$\omega_k$ are the coordinates of the upper surface of a cuboid $k$. The second one, $e^{\tilde{z}}_{k,ij}$, asserts that only point-to-plane correspondences are counted which fulfill some upper bound threshold concerning the Euclidean distance. The cost function based on the Euclidean distance can be noted formally similar to the former points function. In practice, the values of both functions can be determined together by using a joined computation:

$$\text{euclid}(\mathbf{T}, Z)$$
$$= \sum_{X_{ij} \in Z} \left( \min \left( \left\{ e_k \mid t_k \in \mathbf{T} \Rightarrow (\text{inrect}_{\omega_k, ij} = 1) \wedge \left( e^{\tilde{z}}_{k,ij} \leq \phi_c \right) \right\} \right) \right).$$
(14)

In each iteration of the coarse model fitting step, a decision has to be made according to the selection of the best suitable model $\theta^{\mathbf{T}}_i \in \Theta$. This approach tries to maximize the number of points matched and to minimize the cost function. If multiple models are equal, one can be randomly picked afterward. Formally, the selection rule can be written as following:

$$\text{select}(\Theta) = \theta^{\mathbf{T}}_k \in \text{one}\left( \min_e \left( \max_p(\Theta) \right) \right).$$
(15)

### 3.5 Fine granularity model fitting

In the previous step, a fitting of the whole model into the formerly acquired point cloud was performed. Because misplacements of packages are very likely to be locally present, a individually refinement of each detected cuboid-shaped parcel $t_i \in T$ (cf. Fig. 10b) is needed. Induced by the 2.5D depth image and performance considerations the refinement procedure is performed always only for the top-most layer of parcels called R:

$$R = \left\{ r \mid (r = t_i \in \mathbf{T}) \Rightarrow \left( i_i = 1 \wedge f_i = 1 \wedge t^{\tilde{z}}_i \right) \right\}$$
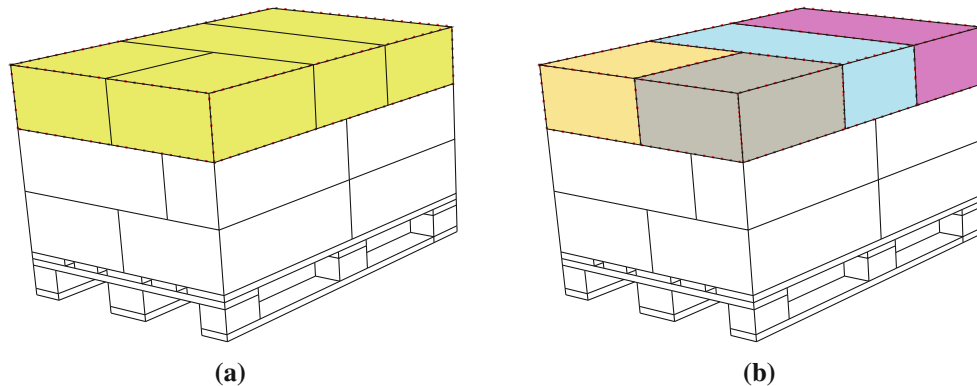(16)

with

$$t^{\tilde{z}}_i = \max \left\{ t^{\tilde{z}}_j \mid t_j \in \mathbf{T} \Rightarrow (i_j = 1) \wedge (f_j = 1) \right\}.$$
(17)

An initial parcel detection can be achieved by computing the area $A(t_i)$ of the upper parcel side, followed by a multiplication with the average number of points per unit of area. The number of expected points is now compared with the number of points actually covered. Additionally, a threshold $\phi_a$ is considered. Formally, the initial parcel detection can be written as following:

$$\forall t_i \in \mathbf{T} \left( f_i = \begin{cases} 1, & \text{if } p_i \geq A(t_i) \cdot \lambda \cdot \phi_a \\ 0, & \text{else} \end{cases} \right).$$
(18)

In contrast to the coarse fitting step, transformations have to be computed for each $t_i \in R$ and all supported kinds of transformations $j$. The parameters $\lambda$ and $\phi_a$ define the expected number of points per unit area and factor for the permissible deviation respectively. All transformations have to fulfill two requirements in order to be appended to the set of possible transformations $\Theta$. At first, the model must stay interior the bounds of the point cloud (ensured by the "inrange" function). Second, the transformation of a singular parcel must avoid any kind of intersection with other detected parcels (ensured by the "intersects" function). Considering this, both assumptions $\Theta$ can be defined for one fine granularity iteration as following:

$$\Theta = \bigcup_{t_i \in R} \bigcup_{j \in \{0, \dots, 6\}} \theta$$
$$= \begin{cases} \{\theta_{ij}\}, & \text{if } \left( \text{inrange}(\theta^{\mathbf{T}}_{ij}, Z) = 1 \right) \wedge \left( \text{intersects}(\theta^{\mathbf{T}}_{ij}) = 0 \right) \\ \emptyset, & \text{else} \end{cases}.$$
(19)



**Fig. 10** Schematic representation of the comparison of two the steps of the model fitting: **a** global model **T** for each layer for the coarse granularity model fitting and **b** individual models (representing a single parcel) $t_i \in T$ for the fine granularity model fitting
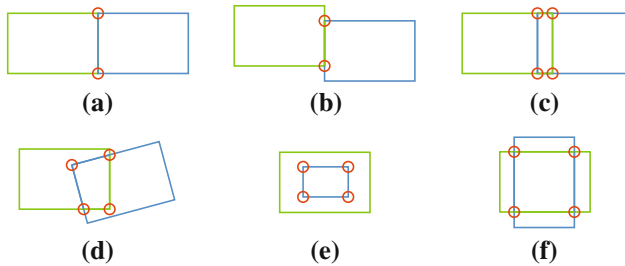
**Fig. 11** Schematic representation of possible collision situations: collisions **a**, **b** are allowed, in contrast to the collisions **c**, **d**, that are not allowed. The occurrence in (**e**) is a special but also a not allowed case—there is no intersection between the line segments. Finally, case (**f**) shows that it is not sufficient to check whether the points of a *rectangle* lie inside of another *rectangle*

### 3.6 Verification of the automatic detection

Finally, after the coarse and fine granularity fitting procedures, the results have to be verified in order to ensure robustness and prevent passing of maybe uncorrectable load information (cf. Fig. 11) to the following process chain. Because of the restriction to the top-most load layer, a corresponding subset of the point cloud called I is defined:

$$I = \left\{ x | (x = X_{ij} \in Z) \Rightarrow z_{ij} > \left(t_{max}^z - h_{max}^z\right) \cdot \left(1 + \phi_f\right) \right\}$$
(20)

The number of points $X_{ij} \in I$, that are mapped by the model **T**, are determined for the set of the significant points I. The final verification step can then be noted as following:

$$\text{verify}(I, \mathbf{T}) = \begin{cases} 0, & \text{if } \sum_{X_{ij} \in I} \text{match}(X_{ij}, \mathbf{T}) \geq \phi_v, \\ 1, & \text{else} \end{cases}$$
(21)

with

$$\text{match}(X_{ij}, \mathbf{T}) = \begin{cases} 1, & \text{if } \exists t_k \in \mathbf{T} \left( \text{inrect}_{\omega_k, ij} = 1 \wedge e_{k,ij}^z \leq \phi_f \right) \\ 0, & \text{else} \end{cases}.$$
(22)

If the number of unmatched points is above a threshold $\phi_v$, the verification function returns 0 (in the other case 1). By altering the threshold parameter, the tolerance of the verification step can be adjusted.

## 4 Evaluation

The evaluation of the automated load detection presented in this paper is divided into three parts. First, the test set-up is described. Second, the accuracy of automated detection is examined for a diversity of test setups followed by a runtime analysis of the detection algorithm.

### 4.1 Test set-up

The basic experimental setup is shown schematically in Fig. 12. A euro pallet with standardized dimensions 1,200 × 800 × 144 mm (length × width × height) serves as a test object regarding the evaluating measures. The test load of the euro pallet consists of parcels with different material properties as shown in Fig. 12c, white parcels with the dimensions 400 × 300 × 260 mm, and brown parcels with the dimensions 380 × 280 × 280 mm in three different variants, (1) without additional materials, (2) with plastic envelope, (3) with plastic envelope and label. Different scenarios with up to three layers were constructed resulting in a maximal total height of 924 mm. The sensor being used for capturing the scene is a PMD camera (type: O3D201, company: ifm electronic) which is mounted 3.1 m above the ground plane. This construction ensures parcel loads up to 1.2 m height lying completely in the angular view of the sensor. In the optimal case, the PMD sensor is parallel to the ground plane. In practice, the ground plane often is not entirely parallel. Therefore, the calculation of a correction plane, for example, using the RANSAC-algorithm is a remedy. Afterward, the image plane of the PMD sensor is parallel to the ground plane, and the center point of the sensor on ground plane serves as point of origin for measurements of the positions of the euro pallet and the particular parcels. The system configuration according to the evaluation of the run time consists of an Intel Q6700 3GHz (Quad-Core Processor) with 2GB DDR2 PC1066 (Main Memory).

### 4.2 Testing of the automated detection of euro pallet loads

The quality of the detection algorithm depends on two main aspects, namely the accuracy and the runtime of the load detection process. The following experimental results are based on the detection of one specified parcel on particular layers of the euro pallet. An important factor for the quality of the raw data of the PMD camera is the integration time. The measures were done with an automatic calibration of integration time that is supported by the used PMD camera.

#### 4.2.1 Detection quality

According to the evaluation of the accuracy, one particular parcel was picked out on a specified (first, second or third) visible layer. Based on a reference configuration, the selected parcel was displaced as shown in Fig. 13a, b. The test parcel was shifted toward the *x*- and *y*-axis by the specified values 5, 10, 15 cm and rotated around the *z*-axis by the specified values 5, 15, 30 deg. The deviation

**Fig. 12** **a** Real and **b** schematic view of the test setup used for image acquisition to evaluate the quality of estimation between the nominal and actual package positions. The PMD camera is mounted 3.1 m above the ground plane. This ensures that euro pallet loads up to 1.2 m height are fully visible to the camera (including 10 % of additional tolerance). **c** Parcels with different material types serve as test objects

**(a)** View of test setup
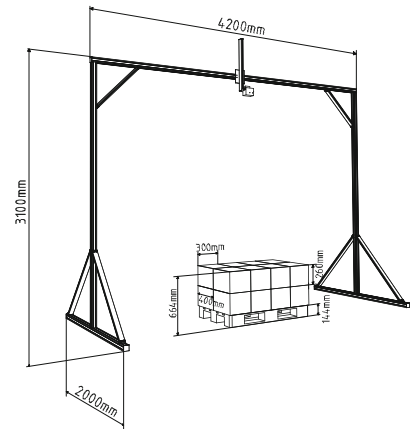
**(b)** Schematic setup

**(c)** Test objects

between the detected parcel position and the calculated reference position regarding to the fix position was recorded. The evaluation process was proceeded for each type of the parcel with special material properties as shown in Fig. 12c.
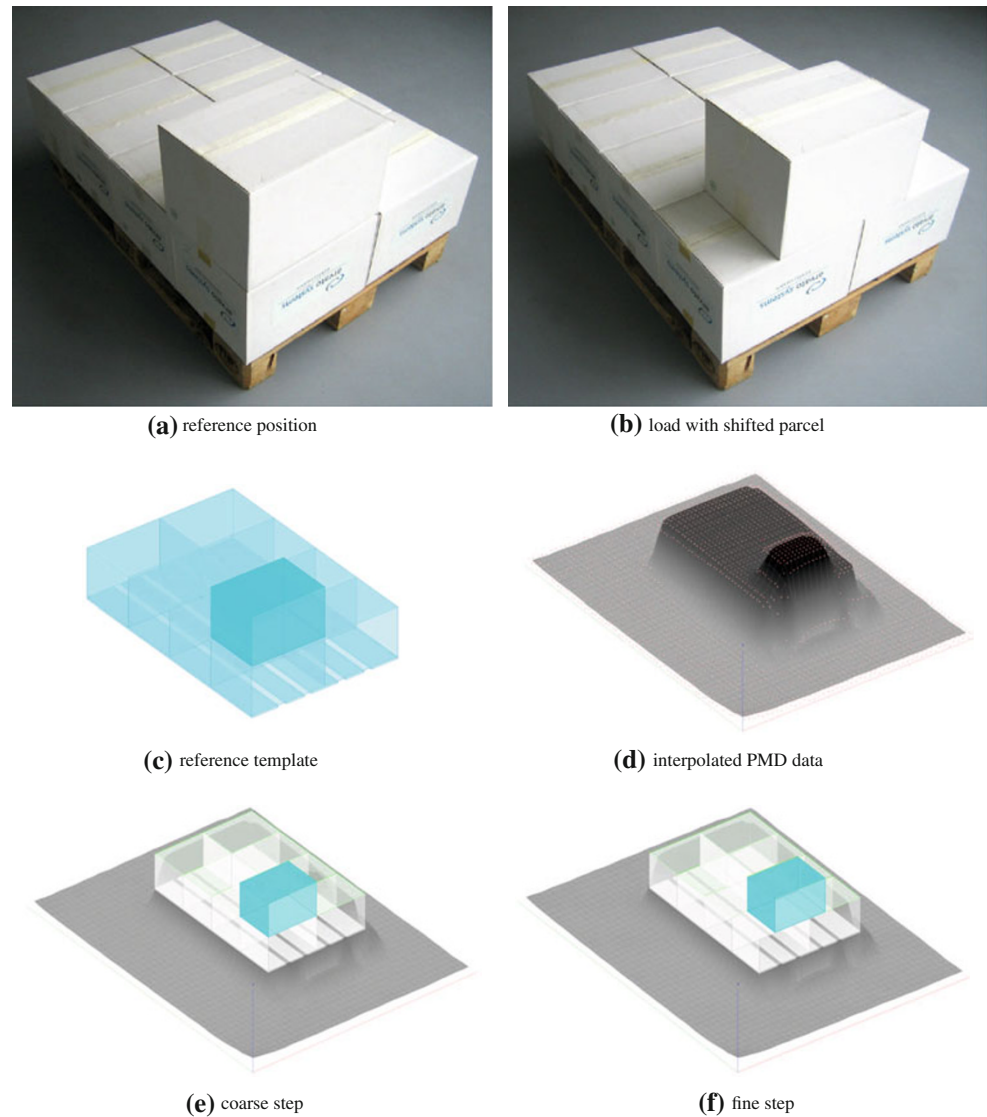
Figure 13 visualizes the detection process in direct relationship to the real scene and the raw data of the PMD camera. The detection algorithm uses the known reference template (cf. Fig. 13c). The raw data (point cloud) of the PMD camera are shown in Fig. 13d as an interpolated surface. The detection process is accomplished in two steps as described in Sect. 3. In the coarse step (cf. Fig. 13e), the reference template as a whole is adapted into the point cloud while the fine step (cf. Fig. 13f) fits each particular element of the template into the local point cloud.

The experimental results of detecting a selected parcel of a specified material type on the first, second and third layer of the euro pallet are visualized in Fig. 14 in terms of box-whisker-plots. Each plot embraces thirty independent calculations. In cases of translations ($x$- and $y$-axis), each plot represents the absolute deviation of the barycenter of the real parcel position to the barycenter of the detected parcel position. In cases of rotations (about $z$-axis), each plot represents the absolute deviation of the rotation about

$z$-axis of the real parcel to the detected parcel. The average deviation of the detected parcel position and the real parcel position is in cases of translation in average between 10 and 15 mm. In consideration to the fact that the image size of the middle pixel of a PMD sensor in a distance of $2-3$ m to the ground plane is approximately $22 \times 22–33 \times 33$ mm the detection result are good for the low resolution ($64 \times 50$ pixels) of the PMD camera. In the case of rotations, the average deviation is in most cases under 4 deg. The small rotation of 5 deg was in most cases not detected which depends on the low resolution of the camera. Larger rotations are detected in all cases. An important fact is, the ignoring small deviations, the detection accuracy does not depend on the material of the test objects—the approach is regardless of the color of the parcel (white or brown) and of variants of additional material (cf. Fig. 12c).

Furthermore, accurate $z$-values are essential for the de-palletizing process. Large deviations according the $z$-value may cause a damage of the euro pallet load by the de-palletizing robot. The threshold concerning the $z$-value should lie under 1 cm which satisfies the height of the foam buffer of the robot gripper. In conjunction with the accuracy of the $z$-values, the discussion of the PMD

**Fig. 13** Visualization of the detection process for evaluation. **a** According to the reference position **b** the parcel on 2nd layer is shifted. **c** Based on the reference template and **d** the PMD data, the actual position is detected in two steps: **e** the coarse step and **f** the fine step

(a) reference position

(b) load with shifted parcel

(c) reference template

(d) interpolated PMD data

(e) coarse step

(f) fine step

camera typical artifacts is important. The non-systematic noise can be compensated by a temporal filter which is supported by the PMD camera (O3D201) we used. Many literature sources [20, 30, 57, 35] discuss the systematic artifacts in detail. The manufacturers reacted with a technological review in order to compensate some of the systematic artifacts. The temperature of the used PMD camera (O3D201) stays constant after a operation duration of ten minutes. Hence, the thermal drift does not affect the measurements. Furthermore, the reflectivity-related error seems to be internally compensated the PMD camera we used. We demonstrate this fact in Fig. 15. The four different types of parcels were placed side by side, white parcels with the dimensions $400 \times 300 \times 260$ mm, and brown parcels with the dimensions $380 \times 280 \times 280$ mm in three different variants, (1) without additional materials, (2) with plastic envelope, (3) with plastic envelope and label. The $z$-values of the PMD camera in terms of a point

cloud were recorded. Besides the fact that the parcels do not fit perfectly the dimensions, the $z$-values of the brown parcels diverge in few millimeters only. The $z$-value of the white parcel differs in approximately 5 mm from the real $z$-value and is the impact of another systematic artifact as mentioned in the following. Figure 16 visualizes in terms of four box-whisker-plots the deviation regarding to the $z$-value of the real object's dimension and the raw data of the PMD camera. Each box-whisker-plot represents the measure of a particular $z$-value, 144 mm (euro pallet), 404 mm (1st layer), 664 mm (2nd layer), 924 mm (3rd layer). All four plots have in common that the raw data of the PMD camera underlies a random noise in a range of approximately 5 mm. Interpreting the diagram notices that the deviation depends on the distance of the measured object. It seems to be a systematic relationship between the deviation error and the measured distance called the wiggling error [20, 35] which is caused by the non-ideal
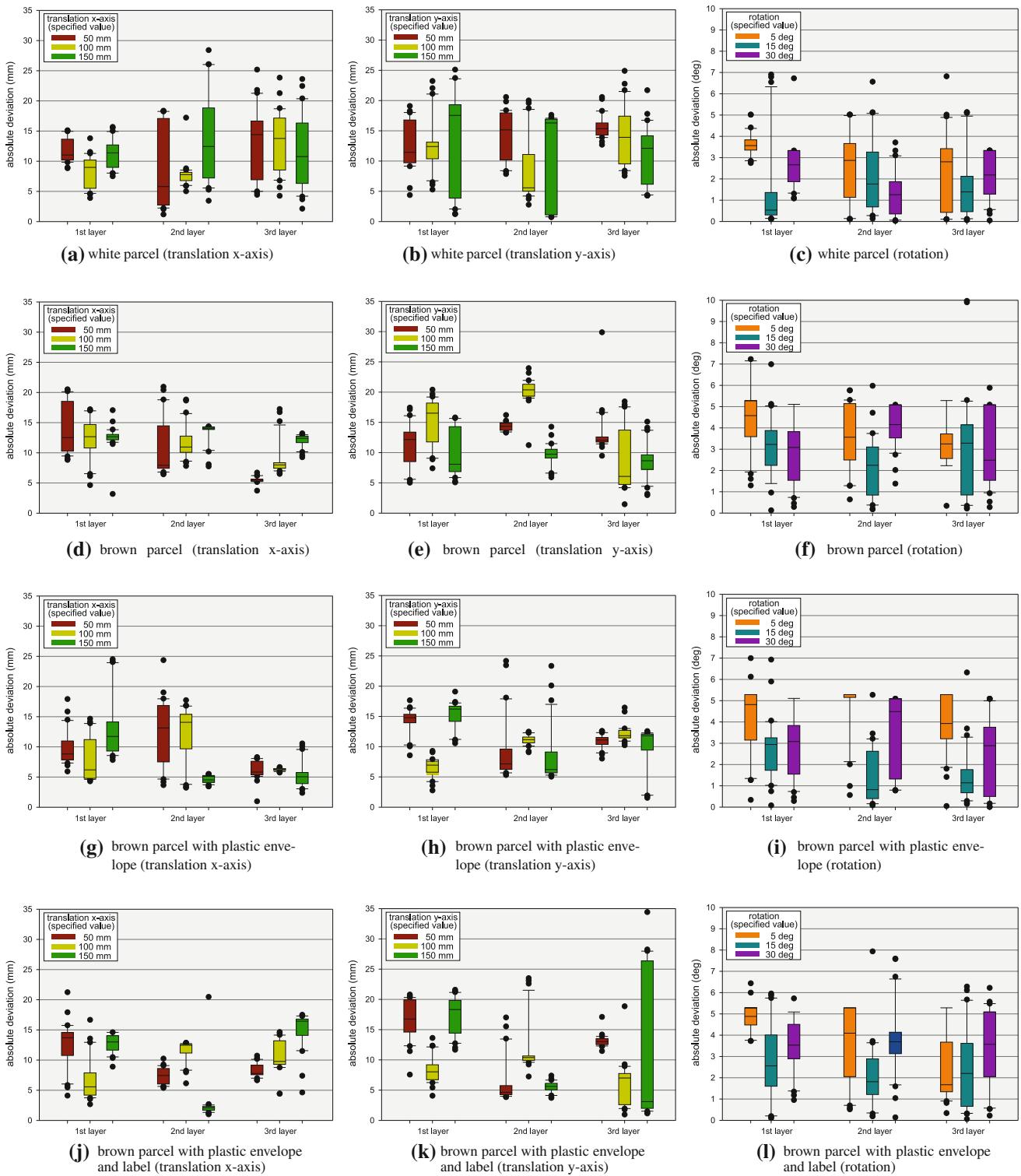
**Fig. 14** Visualization of the accuracy of the detection process when shifting and rotating one specified parcel at each layer. The *box-whisker-plots* represent the deviation between the detected to the real package position

demodulation of the modulated light. Despite the fact that the deviation error lies under the required threshold of 1 cm, one of the future steps is the detection of the wiggling error function in order to optimize the detection quality. Another significant systematic artifact is the offset error depending on the integration time. We experimentally adjust the integration time in our special scenarios in order to minimize the offset error. The detection of the

offset error in general cases is also one of the our future steps.

The evaluation results confirm the potential of our detection algorithm for automatic unloading systems in praxis. The detection results with the current low-cost camera meet the accuracy requirements from logistical point of view as the critical $z$-value lies under 1 cm. The planar deviation in $x$- and $y$-axis allows more tolerance. The grip point of a particular parcel is defined by the barycenter of its top layer. Since the dimensions of parcels in practice a deviation in $x$- and $y$-axis under 20 mm cause no damage of the load due to unexpected collisions and are acceptable. The usage of higher resolution cameras as the Cam Cube ($200 \times 200$ pixels) will further increase the accuracy. Due to our concept of a low-cost-system, Cam Cube is not integrated in our project yet. But in future, falling prices for higher resolution cameras are expected.
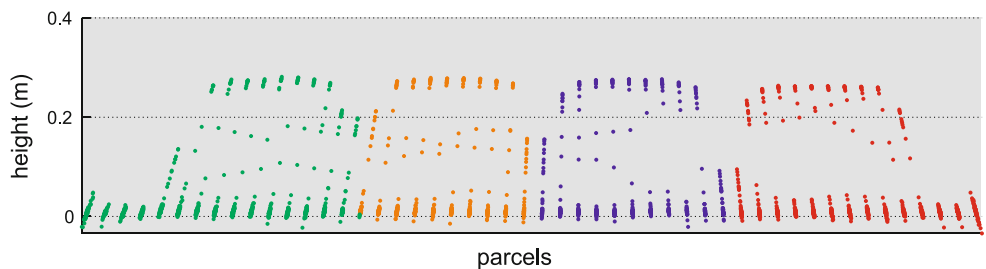
### 4.2.2 Runtime

Efficiency is a key criteria in the field of automatic de-palletizing according to time and consequently cost economy. The diagram shown in Fig. 17 illustrates the measures of runtime that depend on the one hand on the number of layers and on the other hand on the number of parcels on each visible layer. While the runtime of the coarse step mostly is independent of the number of parcels, the runtime of the whole calculation step including the fine step shows a linear behavior according to the number of parcels. Overall, in most cases, the detection of a euro
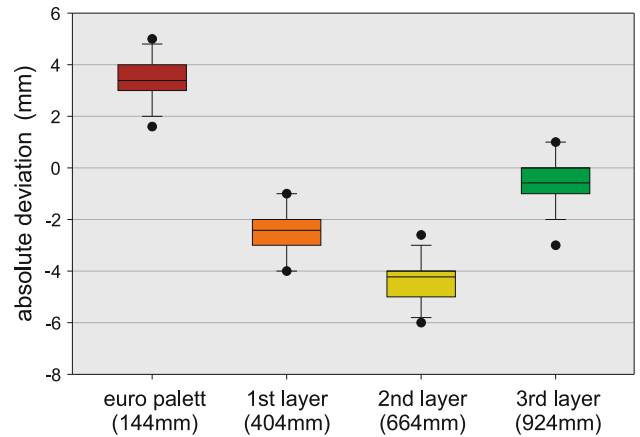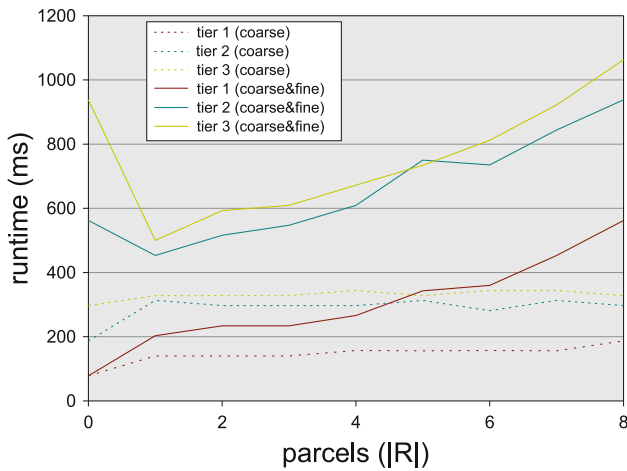


**Fig. 16** Visualization of the deviation regarding the $z$-value of the real object's dimension and the raw data of the PMD camera (*middle line* represents the average deviation)

pallet load requires in average 400 ms and in worst case below 1,100 ms. Hence, the presented approach reaches real-time requirements.

## 5 System comparison—economical competitiveness

Finally, the new approach has to be compared with existing solutions in terms of cost, flexibility, accuracy, robustness and performance to prove the economical potential. For this reason, a utility-cost-analysis has been performed [58, 48]. This type of multi criteria analysis works with quantified parameter values for different technical alternatives. In

**Fig. 15** The test objects with different material properties **a** are visualized as point cloud **b** in orthogonal view projection ($x$-$z$ plane)



**(a)** Test objects with different material properties and the dimensions $380 \times 280 \times 280$ mm (brown parcels) and $400 \times 300 \times 260$ mm (white parcel).



**(b)** Orthogonal view projection ($x$-$z$ plane) of the point cloud generated with the PMD camera (O3D201).

**Fig. 17** Visualization of the runtime regarding the number of layers and the number of parcels on each visible layer

Table 3, several solutions are compared with the presented approach. The columns represent the set $\mathcal{A} = \{$"A1", "A2", "PMD Sensor", "B2+B3", "B1+B2", "B2+B3"$\}$ of the different solutions, an example for every solution is given in the references. The number of elements in the set $\mathcal{A}$ is defined as $|\mathcal{A}|$. On the left side of Table 3, a set of five criteria $\mathcal{C} = \{$"Cost", "Flexibility", "Accuracy", "Robustness", "Performance"$\}$ with a certain quantifier $q_c \in \mathbb{R}$ and a different number of sub criteria $i = 1, \ldots, N$ are listed. The degree of fulfillment $d_{a,c} \in \mathbb{R}$ with indices $a = 1, \ldots, |\mathcal{A}|$ and $c = 1, \ldots, |\mathcal{C}|$ for every sub criteria

$p_{a,c}^N$ on a scale from 0 (insufficient) to 4 (ideal) is defined in the solution columns. Finally, an overall utility value $uv_a \in \mathbb{R}$ can be calculated for every solution. The alternative with the highest value represents the best approach. The following equation shows the calculation of the utility value $uv_a$:

$$uv_a = \sum_{c=1}^{|\mathcal{C}|} q_c \cdot d_{a,c}. \tag{23}$$

Further the degree of fulfillment is given by

$$d_{a,c} = \frac{p_{a,c}^1 + p_{a,c}^2 + \cdots + p_{a,c}^N}{N \cdot p_{\max}} \quad \text{for } c = 2, \ldots, |\mathcal{C}| \tag{24}$$

for every criteria except the cost effectiveness. In case of cost $d_{a,c}$ with $c = 1$ is calculated as follows

$$d_{a,1} = \frac{p_{a,1} - \min(p_{a,1})}{\max(p_{a,1}) - \min(p_{a,1})}. \tag{25}$$

Hence, the cost of the single solutions gets a value for $d_{a,c}$ between 0 and 1. The used criteria of the set $\mathcal{C}$ can be defined as:

- *Cost* including sensor and hardware cost without handling device and software license fee (most prices are estimations)
- *Flexibility* including range of detectable object (size, shape, color, material, etc.) and range of object setup (pile height, sensor distance, number of sensors),
- *Accuracy* including xy-axis, z-axis and rotation,

**Table 3** Utility-cost-analysis

| Criteria | Quantifier | Alternatives | | | | | |
|---|---|---|---|---|---|---|---|
| | | A1: centering frame [31] | A2: proximity sensors in gripper [42] | PMD sensor | B2+B3: 3d sensor, force sensor [25] | B1+B2: 2 camera pairs and structured light [56] | B2+B3: 2 webcams and 2 kinect sensors [38] |
| Cost | 0.4 | 0.0 | 1.00 | 0.94 | 0.71 | 0.64 | 0.93 |
| Cost in Euro | | 15,000 | 1,000 | 1,800 | 5,000 | 6,000 | 2,000 |
| Flexibility | 0.1 | 0.25 | 0.25 | 0.63 | 0.88 | 0.75 | 0.63 |
| Objects | | 1 | 1 | 2 | 3 | 3 | 3 |
| Pallet | | 1 | 1 | 3 | 4 | 3 | 2 |
| Accuracy | 0.2 | 0.83 | 0.67 | 0.58 | 0.58 | 0.75 | 0.75 |
| xy-axis | | 4 | 3 | 2 | 2 | 3 | 3 |
| z-axis | | 2 | 3 | 3 | 3 | 3 | 3 |
| Rotation | | 4 | 2 | 2 | 2 | 3 | 3 |
| Robustness | 0.2 | 1 | 0.75 | 0.75 | 0.75 | 0.625 | 0.375 |
| Light conditions | | 4 | 3 | 3 | 3 | 2 | 2 |
| Industrial environment | | 4 | 3 | 3 | 3 | 3 | 1 |
| Performance | 0,1 | 0,5 | 0,25 | 0,75 | 0,5 | 0,75 | 0,75 |
| Cycle time | | 2 | 1 | 3 | 2 | 3 | 3 |
| Utility value | | 0.38 | 0.70 | 0.75 | 0.66 | 0.63 | 0.68 |

**Fig. 18** Sensitivity analysis for several focal points. All scenarios excluding application have a quantifier of 0.6 for the focus and 0.1 for the other criteria. Within the parameter from 0 to 1, the certain numbers represent the utility value of the different solutions. A fulfillment of 1 represents the optimal solution
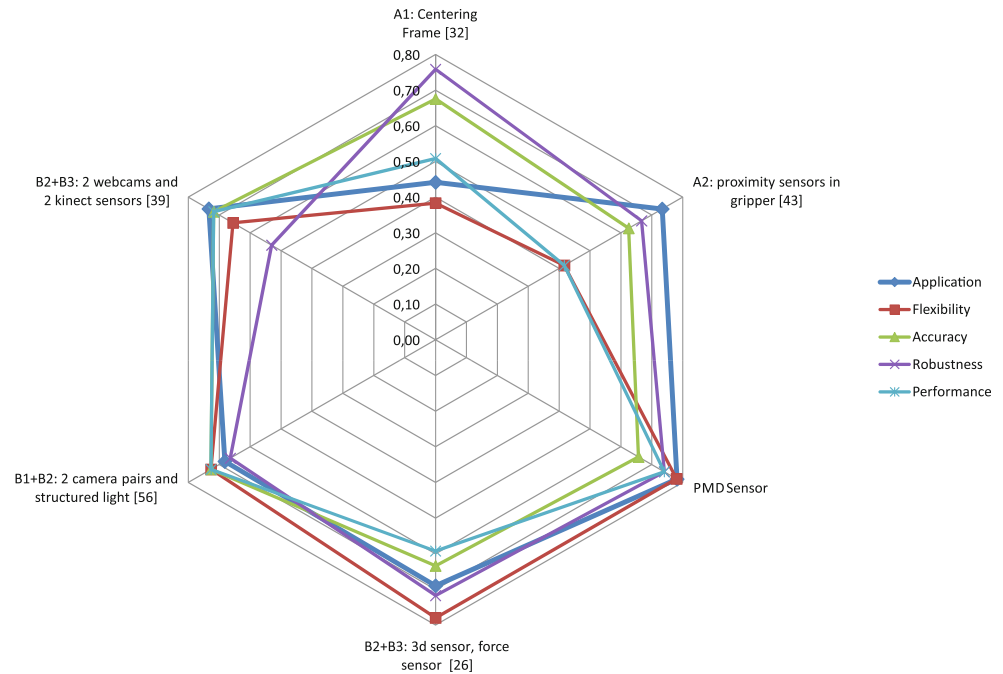


– Robustness including sensitivity against light conditions and suitability for daily use in an industrial environment, and

– *Performance* with the cycle time for detection.

As a result for a specific scenario of depalettizing [shown in Table 3 with a priori knowledge with focal points on cost ($q_1 = 0.4$), accuracy ($q_3 = 0.2$) and robustness ($q_4 = 0.2$)], the presented solution has the highest utility value $uv_3 = 0.75$. Compared with other solutions costs are limited to one PMD sensor and an evaluation unit (about 1,800€), flexibility could be ranked as middle with a wide range of detectable materials and surfaces but limited by package size because of the low resolution of the sensor, accuracy is rather low but adequate for the application, robustness is quite high because of a single industrial sensor, and performance is high with cycle time under 1 second (with potential to reduce it significantly).

In order not to limit the perspective to a certain application, different focuses have been investigated in a sensitivity analysis (cf. Fig. 18). For this reason, all scenarios, excluding application, have a quantifier of 0.6 for the focus and 0.1 for the other criteria. Finally, the presented solution is quite stable to changing focuses. Therefore, the presented approach has one of the highest utility values in most cases.

The main drawback of the used methodology is the partial not quantifiable definition of the degree of fulfillment. Hence to assure a serious comparison, it is necessary to define a specific test scenario under uniform evaluation conditions. Depending on the market situation, volume discount and especially the specific application of the user,

cost, flexibility, accuracy, robustness and performance are not comparable from the data sheet. For this reason, a benchmark test of existing market solutions is planned in the near future at the Fraunhofer-Institute facilities.

## 6 Conclusions

In this study, a novel approach for the detection of parcel loading positions on a pallet was presented. Based on the predetermined model of the loading situation and the raw data of the PMD camera, the innovative registration algorithm produces accurate results regarding the position detection. The detection phase of the algorithm consists of the two detection steps, the coarse step and the fine step while a final verification step serves the correctness of detection results. Overall, this new approach constitutes a competitive alternative in comparison with existing complex and costly 3D-vision systems and reduces from an economic point of view the costs of contour checking in the automated de-palettizing process by concurrently moderate runtime (less than 1 s on the given test system). The presented stable test set-up in combination with a robust PMD camera suits for a static installation in an industrial environment, for example, with belt conveyors.

Future work will focus, in particular, on the integration of the algorithm into a real automated robot application for de-palettizing in order to improve the practical suitability. Quantifiable values of the new approach and existing solutions will be compared within a benchmark test under uniform conditions. Further capabilities lie on the one hand

in the calibration of the PMD camera and the identification of the systematic error functions and on the other hand in increasing the accuracy, for example, by a fusion of the PMD depth data with the gray scale image. Another challenge is the detection and estimation of flying pixels which distort the detection. Using a high-resolution sensor is also possible which will provide more accurate results but may lead to an increase of runtime which can be compensated by mapping the calculation on the today highly parallel graphics processor units (GPUs).

After all the presented approach is not limited to the presented use case of de-palletizing, other possible application areas are: incoming goods inspection, automated guided vehicles (AGV) and adjustment of speed of, for example, a rack feeder.

## References

1. ABB (2010) ABB TrueView—Vision Systeme (Applikationsausrüstung). http://www.abb.de. Last visited 2012-01-15
2. Almaaitah A, Ali K, Hassanein HS, Ibnkahla M (2010) 3d passive tag localization schemes for indoor rfid applications. In: IEEE international conference on communications (ICC), pp 1–5
3. Athalye A, Savic V, Bolic M, Djuric P (2011) A radio frequency identification system for accurate indoor localization. In: IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 1777–1780
4. Bechteler T, Yenigün H (2003) 2-D localization and identification based on SAW ID-tags at 2.5 GHz. IEEE Trans Microwave Theory Techniques LNCS 51(5):1584–1590
5. Besl P, McKay H (1992) A method for registration of 3-D shapes. In: IEEE transactions on pattern analysis and machine intelligence, pp 239–256
6. Binar Elektronik AB (2010) SensActive—Bin picking. http://www.sensactive.com/index-de.htm. Last visited 2012-07-06
7. BMBF (2012) 02PB3101, Verbundprojekt Generisches Entladesystem (GENESYS). Tech. rep
8. Boehnke K (2007) Object localization in range data for robotic bin picking. In: Automation science and engineering, 2007. CASE 2007. IEEE international conference on, pp 572–577
9. Boughorbel F, Zhang Y, Kang S, Chidanbaran U, Abidi B, Koschan A, Abidi M (2003) Laser ranging and video imaging for bin picking. Assembl Autom 23:53–59
10. Bouzakis A, Weigelt S, Overmeyer L (2010) RFID-assistierte Stückguterfassung und -handling anhand eines kognitiven Robotersystems. Logist J Proc 6(1):1–11
11. Buchholz D, Winkelbach S, Wahl FM (2010) RANSAM for industrial bin-picking. In: ISR/ROBOTIK. VDE Verlag, pp 1–6
12. Choi J, Lee H, Lee SC, Lee DH (2011) Reducing localization ambiguity of immobile passive UHF RFID tagged physical objects. In: Ubiquitous robots and ambient intelligence (URAI), 2011 8th international conference on, pp 35–40
13. Choi JS, Lee H, Engels D, Elmasri R (2012) Passive UHF RFID-based localization using detection of tag interference on smart shelf. IEEE Trans Syst Man Cybern Part C Appl Rev 42(2):268–275
14. Corban M (2006) "Sehende" Roboter setzen sich durch. FM—Das Logistik Magazin 22–25
15. David HA, Hartley HO, Pearson ES (1954) The distribution of the ratio, in a single normal sample, of range to standard deviation. In: Biometrika 41:482–493. Oxford University Press, Oxford
16. De Meyer A (1998) Manufacturing operations in Europe: where do we go next? Eur Manag J 16(3):262–271
17. Echelmeyer W, Kirchheim A, Wellbrock E (2008) Robotics-logistics: challenges for automation of logistic processes. Autom Logist 1–3:2099–2103
18. FANUC (2012) iRVISION. http://www.fanucrobotics.de. Last visited 2012-07-06
19. ischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24(6):381–395
20. Foix S, Alenya G, Torras C (2011) Lock-in time-of-flight (ToF) cameras: a survey. Sens J IEEE 11(9):1917–1926
21. Fuchs S, Haddadin S, Keller M, Parusel S, Kolb A, Suppa M (2010) Cooperative bin-picking with time-of-flight camera and impedance controlled DLR lightweight robot III. In: IROS, pp 4862–4867
22. Ghita O, Whelan P (2003) A bin picking system based on depth from defocus. Mach Vis Appl 13:234–244
23. Ghobadi S, Hartmann K, Weihs W, Netramai C, Loffeld O, Roth H (2006) Detection and classification of moving objects-stereo or time-of-flight images. In: Computational intelligence and security. IEEE, Center for Sensor Systems, University of Siegen, pp 11–16
24. Ghobadi S.e.a. (2006) Classification of 3D moving objects using support vector machines and 3D-time of flight camera. In: 3rd international conference on artificial intelligence in engineering and technology, pp 224–229
25. Intermodalics (2012) http://www.intermodalics.eu/vision-robotics. Last visited 2012-07-04
26. Kaiser B (2009) Bildverarbeitung für ein generisches Entladesystem. Ph.D. thesis, Karlsruher Institut fr Technologie
27. Kaiser R, Baier G (2009) Neue Lösungen zum Einsatz von Robotik bei der Handhabung konplexer Güter. In: Roboter in der Intralogistik, pp 91–105. Fraunhofer IPA Technologieforum F 191
28. Katsoulas D, Kosmopoulos D (2001) An efficient depalletizing system based on 2D range imagery. Robot Autom 1:305–312
29. Keysystech (2012) Robotvision. http://www.keysystech.de/de/html/robotvision.html. Last visited 2012-01-15
30. Kolb A, Barth E, Koch R, Larsen R (2009) Time-of-flight sensors on computer graphics. In: Proceedings of Eurographics (State-of-the-Art Report)
31. Kuka-Robotics (2012) Depalletizing and palletizing of banana cartons. http://www.kuka-robotics.com/germany/de/solutions/solutions_search/L_R162_Depalletizing_and_palletizing_of_banana_cartons.htm. Last visited 2012-07-04
32. Kuka-Robotics (2012) Palettieren von Molkereiprodukten. http://www.kukarobotics.com. Last visited 2012-01-15
33. Kuka-Robotics (2010) KUKA.PalletTech. http://www.kuka-robotics.com. Last visited 2012-01-15
34. Linde H (2006) On aspects of indoor localisation. Ph.D. thesis, Technical University Dortmund
35. Lindner M, Kolb A (2007) Calibration of the intensity-related distance error of the PMD TOF-Camera. In: SPIE: intelligent robots and computer vision XXV, vol 6764, pp 6764–35
36. Mamaev I, Safronov K, Tchouchenkov I, Wörn H (2008) Real-time framework for bin-picking problem using advantages of PMD-technology. In: Proceedings of the 10th international workshop on computer science and information technologies (CSIT'2008)

37. Mojix (2009) Mojix EPC compliant real-time location system. http://www.mojix.com/products/documents/RTLS_brochure.pdf. Last visited 2012-01-15

38. Motoman (2012) http://www.motoman.com/products/vision/. Last visited 2012-07-04

39. Ni L, Liu Y, Lau Y, Patil A (2003) LANDMARC: indoor location sensing using active RFID. In: Proceedings of the first IEEE international conference on pervasive computing and communications, pp 407–415

40. Nilsson P (2010) 3d-vision guided robotics for material handling. In: 41st international symposium on and 6th German conference on robotics (ROBOTIK), pp 1–7

41. Oh JK, Baek K, Kim D, Lee S (2009) Development of structured light based bin picking system using primitive models. In: IEEE international symposium on assembly and manufacturing (ISAM), pp 46–52

42. Packaging P (2012) RDC-MX. http://www.pearsonpkg.com/products/showProduct/RDC-MX. Last visited 2012-07-04

43. Parlak S, Marsic I (2011) Non-intrusive localization of passive rfid tagged objects in an indoor workplace. In: RFID-technologies and applications (RFID-TA), 2011 IEEE international conference on, pp 181–187

44. Pires JN, Paulo S (2003) High efficient robotic de-palletizing system for the non-flat ceramic industry. Robot Autom 3:3529–3534

45. Pochyly A, Kubela T, Kozak M, Cihak P (2010) Robotic vision for bin-picking applications of various objects. In: Robotics (ISR), 2010 41st international symposium on and 2010 6th German conference on robotics (ROBOTIK), pp 1–5

46. Prasse C, Skibinski S, Weichert F, Stenzel J, Müller H, ten Hompel M (2011) Concept of automated load detection for de-palletizing using depth images and RFID data. In: IEEE international conference on control systems, computing and engineering (ICCSCE 2011)

47. Ringbeck T, Hagebeuker B (2006) Mehrdimensionale Objekterfassung mittels PMD-Sensorik. Sensor Magazin 2:40–42

48. Rinza P, Schmitz H (1977) Nutzwert-Kosten-Analys. VDI-Verlag, Düsseldorf

49. Rusinkiewicz S, Levoy M (2001) Efficient variants of the ICP algorithm. In: Third international conference on 3D digital imaging and modeling (3DIM)

50. Ruske KD, Kauschke P, Reuter J, von der Gracht H, Gnatzy T, Darkow IL (2010) Transportation & logistics 2030—volume 3: emerging markets new hubs, new spokes, new industry leaders? Tech. rep., SMI (Supply Chain Management Institute) and Price Waterhouse Cooper

51. Safronov K, Tchouchenkov I, Wörn H (2008) Hierarchical iterative object recognition method for a PMD-sensor supplied bin-picking system. In: 8th Asia-Pacific conference on control and measurement (APCCM 2008), Harbin, China, pp 80–84

52. Sample A, Macomber C, Jiang LT, Smith J (2012) Optical localization of passive UHF RFID tags with integrated LEDs. In: IEEE international conference on RFID, pp 116–123

53. Scholz-Reiter B, Thamer H, Uriarte C (2011) An approach for 3D object recognition of universal goods. Int J Comput 5(2):218–225

54. Schuh G, Wemhöner N, Friedrich C (2006) Scenario-based life-cycle analysis of manufacturing systems. CIRP J Manuf Syst 35(2):129–135

55. Sinnerbrink R (2008) Effizienzsteigerung in der Logistik—Neue 3D-Kamera optimiert Erkennungssicherheit und Geschwindigkeit. Inspect 4:43–46

56. Universal Robotics (2012) Random-box-mover. http://www.universalrobotics.com/random-box-mover. Last visited 2012-07-03

57. Uriarte C, Scholz-Reiter B, Ramanandan S, Kraus D (2011) Modeling distance nonlinearity in tof cameras and correction based on integration time offsets. In: CIARP, pp 214–222

58. VDI (1998) VDI 2225 Blatt 3; design engineering methodics—engineering design at optimum cost—valuation of costs. Verein Deutscher Ingenieure

59. VDMA (2012) VDMA Forum Intralogistik erwartet für 2011 Umsatzplus von neun Prozent. http://www.vdma.org. Last visited 2012-01-15

60. VMT-Vision (2012) Depalletizing and palletizing with VMT D/P. http://vmt-vision-technology.com/575/Griff-in-die-Kiste.html. Last visited 2012-07-06

61. Zhang B, Skaar S (2009) Robotic de-palletizing using uncalibrated vision and 3D laser-assisted image analysis. Intell Robots Syst 10–15:3820–3825