

Article

Chances and Challenges: Transformation from a Laser-Based to a Camera-Based Container Crane Automation System

Johannes Benkert *, Robert Maack  and Tobias Meisen * 

Chair for Technologies and Management of Digital Transformation (TMDT), University of Wuppertal, Rainer-Grüenter-Str. 2, 42119 Wuppertal, Germany; robert.maack@uni-wuppertal.de

* Correspondence: johannes.benkert@uni-wuppertal.de (J.B.); meisen@uni-wuppertal.de (T.M.)

Abstract: In recent years, a steady increase in maritime business and annual container throughput has been recorded. To meet this growing demand, terminal operators worldwide are turning to automated container handling. For the automated operation of a crane, a reliable capture of the environment is required. In current state-of-the-art applications this is mostly achieved with light detection and ranging (LiDAR) sensors. These sensors enable precise three-dimensional sampling of the surroundings, even at great distances. However, the use of LiDAR sensors has a number of disadvantages, such as high acquisition costs and limited mounting positions. This raises the question of whether the LiDAR systems of automated container terminals (ACT) can be replaced with cameras. However, this transformation is not easy to accomplish and is explored in more depth in this paper. The field of camera-based container automation presented in this publication is largely unexplored. To the best of our knowledge, there is currently no automated container terminal in real-world operation that exclusively uses cameras. This publication aims to create a basis for further scientific research towards the goal of a fully camera-based container automation. Therefore, the authors present a narrative review providing a broad overview of the mentioned transformation, identifying research gaps, and suggesting areas for future research. In order to achieve this, this publication examines the fundamentals of an automated container terminal, the existing automation solutions and sensor technologies, as well as the opportunities and challenges of a transformation from LiDAR to camera.



Citation: Benkert, J.; Maack, R.; Meisen, T. Chances and Challenges: Transformation from a Laser-Based to a Camera-Based Container Crane Automation System. *J. Mar. Sci. Eng.* **2023**, *11*, 1718. <https://doi.org/10.3390/jmse11091718>

Academic Editors: Nam Kyu Park and Mihalis Goliass

Received: 27 July 2023

Revised: 21 August 2023

Accepted: 28 August 2023

Published: 31 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: container handling; harbor; machine learning; deep learning; camera

1. Introduction

In recent years, the annual container throughput has steadily increased. According to the United Nations' annual review of maritime transport, 80% of the volume of global trade worldwide is shipped by sea [1]. Therefore, a central element of global container transport is the handling of containers at the port. First, the containers are unloaded from the ship using a ship-to-shore (STS) crane. Subsequently, with the help of a yard crane, they are temporarily stored in the container stacks so that they can finally be transported further by ship or overland [2]. In order to manage the above-mentioned processes, terminal operators all over the world rely on automated container handling [3].

In order to improve safety and increase the handling speed, automation is the key approach that terminal operators are focusing on, with the goal of full automation. In this context, it is of central importance that the environment of the crane can be almost completely detected and observed with sensors. The sensor signals are used to validate process steps, derive control signals, and improve safety. Nowadays, the necessary capturing of the environment is carried out partly with cameras and mostly with LiDAR sensors, because this sensor type enables a precise three-dimensional measurement even at long distances and without additional illumination.

However, laser systems are much more expensive than cameras and their use is therefore only worthwhile at large terminals with a high container throughput. In addition, the sensitive laser systems can only be mounted at mechanically less-stressed mounting positions, which leads to process-related disadvantages, like strong occlusion and large distances to the relevant objects.

The continuous development of deep learning methods for processing image information has led to the fact that many automation tasks such as people detection can be controlled with cameras [4–6]. This paper, therefore, discusses the question of what opportunities and risks arise from the transformation from a laser- to a camera-based container-handling process. Hereby, special attention is paid to the necessary process steps for locating objects. This publication is structured as follows: First, the typical structure of a container terminal is outlined, followed by a presentation of the typical container-handling process. Then, the currently used sensors are presented and finally, the mentioned chances and challenges are discussed. A central advantage in this regard is the reduced hardware costs (see Section 6.1.1), which could enable cost-effective localisation of containers and make further expansion stages possible—for example, the energy-optimised travel path of the crane. On the other hand, a central disadvantage is that camera systems do not provide intrinsic depth information.

2. Application Field: Automated Container Terminal

In this section, we describe the typical structure of a modern container terminal and the individual container-handling steps, in order to understand the overall container flow and the current degree of automation. This analysis begins at the moment when the container ship has entered the port and is securely anchored at the quay wall. The focus of this work is on the container flow, which is realised with the help of container cranes. In addition to container handling, LiDAR systems are also successfully used in the port environment for other tasks such as ship berthing [7,8] or on autonomous vehicles [9] in the terminal. However, these LiDAR systems are not directly related to the container flow and are therefore not the subject of this work.

2.1. General Port Structure

The layout of the terminal is based on the four major process steps that are required for container handling: the loading from the ship to the landside, the internal transport of containers, the intermediate storage of containers in the stack, and further transport on the landside [2]. The first area is called the sea side and is typically covered by so-called ship-to-shore cranes. The intermediate storage of containers takes place in the yard area, where stacking cranes are used for container handling. The area of internal horizontal transport is located in between. Straddle carriers, trucks, or automated guided vehicles (AGVs) transport the containers between the individual cranes. If the container is further transported by land, the freight is loaded onto trucks or train wagons in the last area (see Figure 1).

2.2. General Container Movement

The process of general container movement starts when the ship docks at the quay wall and is secured for the further loading and unloading process. Subsequently, STS cranes are placed over the corresponding container bays and the crane boom is lowered. The exact position of the individual cranes is determined with the help of the terminal operating system (TOS). This is a central logistics system that knows the current and future position of each individual container. With the help of this system, it is determined when and where each container is loaded. This information is used to generate jobs, which the crane driver or the automation system receives. A job describes which container has to be grabbed next, where it is currently located, and where it is ultimately placed [10]. The combination of all jobs describes the entire loading and unloading process starting on the ship, through horizontal transport, up to storage in the stack (yard area).

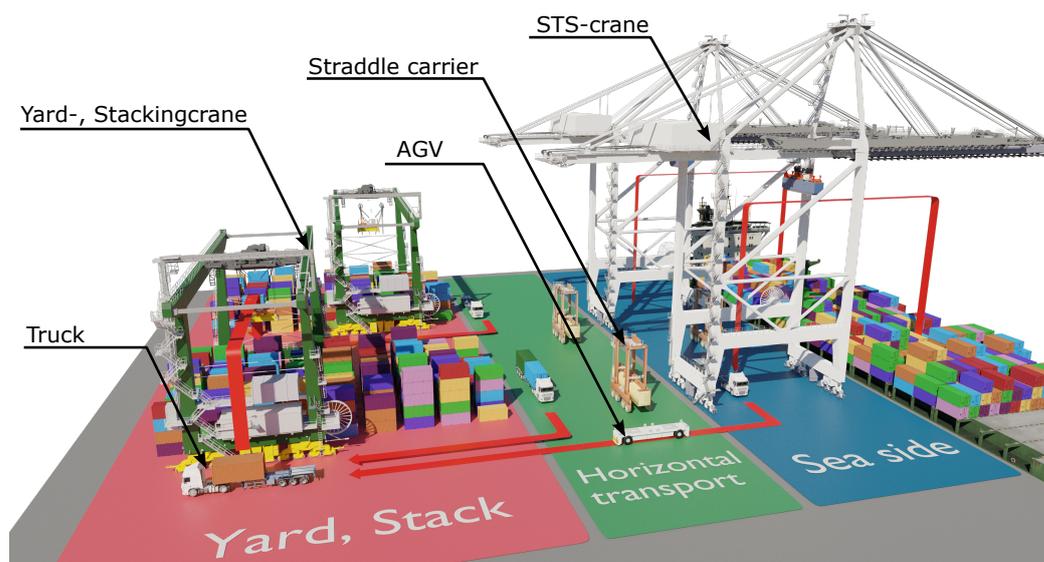


Figure 1. Typical terminal structure: sea side, horizontal transport, and container stacks with yard cranes. The container is picked up on the ship using an STS crane and horizontal transport vehicles like trucks, straddle carriers, or AGVs bring them to the stack area.

In the first process step, the containers are transported from the ship to the landside with the help of STS cranes. This step is performed manually by crane operators, who pick the containers in the vessel and bring them to the horizontal transport vehicles. These vehicles are trucks, straddle carriers, or AGVs, which transport the containers to the yard area.

After the vehicles have arrived at the corresponding stack, the yard crane automatically localises the load on the vehicle, grabs the container, and places it in the stack. The container remains there until it is transported further either by another ship or by land. To enable this onward transport, the container is either brought back to the STS crane or taken to the landside, where it is further transported by train or road.

2.3. Mechanical Components

In the previous section, we introduced the general loading and unloading process on a terminal. As described, the container is lifted and set down by a crane at several points. In order to accomplish this, a number of mechanical components are required on the crane and container. Thereby, the outlined mechanical properties represent a distinctive feature of automated cranes compared to other robots. In addition, the mechanical structure of the crane imposes requirements on the measurement precision that must be achieved through sensor systems.

Figure 2 shows the most important mechanical components of a container crane. The first one is the gantry drive. This drive allows the crane to move parallel to the container ship or stack [11]. Due to the high mass that must be moved in the gantry direction, this is the slowest available travel direction. A movement perpendicular to the gantry movement is made possible by the trolley drive. The trolley is a carriage mounted on rollers, which is used together with the gantry drive for the exact positioning of the gripping tool (the spreader) above the desired container. The hoist is used to lift or lower the spreader. In order to lift a container, a mechanical connection between the container and spreader is required. Therefore, the spreader as well as trucks and other mounting elements on the ship are equipped with so-called twistlocks [12]. A twistlock is a rotating cone that engages with the corresponding counterparts on the container (corner casts). By turning the cone, a firm mechanical connection is established. The twistlocks and corner casts have a standardised size so that the containers can be loaded using uniform equipment worldwide. Figure 2 also visualises the mechanical components and the dimensions of the slotted hole of a corner cast, which are specified in ISO-1161 [13]. These dimensions determine the required

positioning accuracies of an automated container crane. Only when the crane is positioned with an accuracy of roughly ± 2 cm can the twistlocks slide into the corresponding slots of the corner casts and thus, the crane is able to pick up the container.

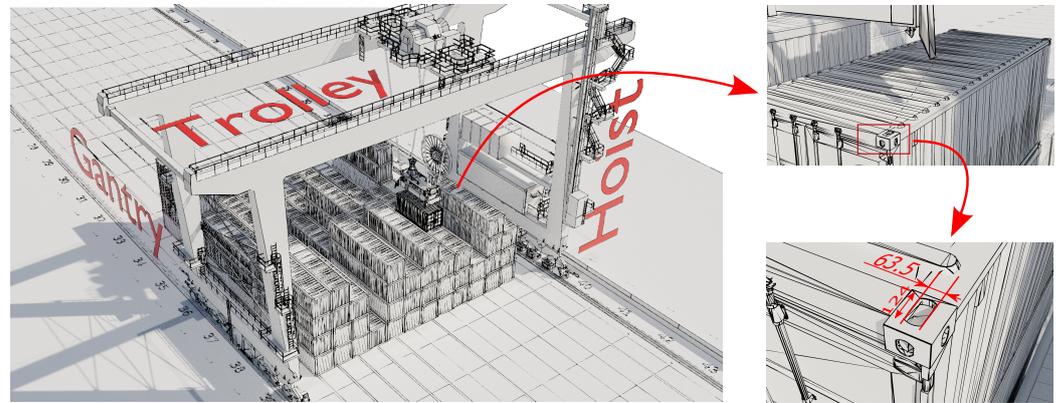


Figure 2. Important mechanical components on a container crane. Gantry, trolley, and hoist drives are used to position the spreader. On the containers, corner casts and twistlocks are needed to build up a mechanical connection to the spreader. The corner cast's slotted hole has a size of 124 mm \times 63.5 mm, which is defined in ISO-1161 [13].

2.4. Supporting Automation Systems

In the complete loading chain, there are various points where the crane operator is either supported by an automation system or where the loading is even carried out automatically. In order to better understand which areas are already being automated today and what potential this holds for the future, the existing automation solutions are briefly presented below and summarised in Table 1.

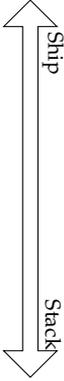
Nowadays, the loading process at the ship is mostly performed manually, since on the one hand the required measuring accuracies cannot be maintained due to the long distances between spreader and trolley, and on the other hand because there are many people in the vicinity of the container loading and thus, a high potential for injury exists [14]. Nevertheless, automation solutions support the crane operator in controlling the crane. Active sway control prevents the load from swaying when the crane is moving. In order to handle external sway disturbances—such as wind loads—the sway controller uses an observer. This is mostly a camera that detects the sway, so that the sway controller can counteract the imposed sway accordingly.

In order to enable a fast transition from the crane to other transport vehicles, these vehicles are pre-positioned so that the loading area is precisely under the crane's set-down position and the vehicle does not need to be maneuvered during the loading process. For this purpose, a laser system measures the position of the vehicle and transmits positioning instructions either to a driver or to the control system of the crane. Such systems are available on the ship-to-shore (STS) crane as well as on the stacking cranes.

Due to the smaller crane sizes in the yard area and the associated reduced measuring distances, the degree of automation is much higher for stacking cranes. After the automatic placement of the vehicles, the container is grabbed to be carried to the correct position in the stack. The set-down position is measured with the help of a laser system and the container is set down exactly at the desired position. In this way, accidental overturning of container stacks can be prevented [15].

To ensure that no unwanted collisions occur during travel and when setting down the load, the travel path of the crane is continuously monitored by the previously mentioned laser system. If a collision is imminent, the speed is first reduced and the crane is stopped completely if necessary. The main automation tasks and their corresponding challenges are summarised in Table 1.

Table 1. Main automation tasks and the corresponding challenges in an automatic container terminal.

Crane Vehicle	Task	Current Technology	Main Challenges	
	STS	Pick container on ship	Manually realised	<ul style="list-style-type: none"> • Long distances • Mech. deformation • Target (container on ship) is moving
	Horizontal transport	Align vehicle under crane	LiDAR-based	<ul style="list-style-type: none"> • Different vehicle types
	Yard crane	Collision prevention while moving container	LiDAR-based	<ul style="list-style-type: none"> • Low latency • Exact 3D position required
	Yard crane	Automatic stacking	LiDAR-based	<ul style="list-style-type: none"> • Occlusion • Limited field of view
All	Remote Operation	Camera-based	<ul style="list-style-type: none"> • Low latency • Remote operator requires a good overview 	

Throughout the logistics chain, there are other automation solutions that monitor and control the process. These include, for example, the validation of the container number, monitoring for damaged containers, and the detection of dangerous goods symbols [16]. As we focus on localising objects (vehicles, trucks) using cameras, these systems are important, but nevertheless will be ignored in this publication, as long as they are not influencing the positioning process.

3. Required Capabilities for Automated Container Handling

In the previous chapter, we generally introduced the crane environment and outlined that the loading of containers is supported by automation systems in many places and in specific situations occurs completely automatically. To achieve this, the system must evaluate the current loading operation in various ways. In this section, we will present the three main categories into which individual tasks can be assigned. This categorisation illustrates the overarching task types that a future camera system has to master.

3.1. Scene Classification

In various process steps, the situation below the sensor must be assessed. For example, it must be checked whether the right container is grabbed by the spreader, if there are still twistlocks attached (see Figure 3a), and whether a vehicle is ready to load. In these situations, the environment is recorded and the situation is assigned to one of several predetermined categories.

3.2. Static Object Detection

Another central task in crane automation is to locate objects. Those objects can be different containers, vehicles, or other obstacles in the area. If neither the objects nor the sensor is moving, this is static object detection. A typical example of this is to locate containers in the stack (see Figure 3b). The objects must first be recorded and then their exact position must be localised in the sensor coordinate system. The recorded position is subsequently transformed into a crane coordinate system so that the crane automation software can work with the transmitted position values. To be able to use the determined position for the automation chain, the position is defined by translation and rotation. In total, the position of the object (container, vehicle, . . .) has six degrees of freedom.

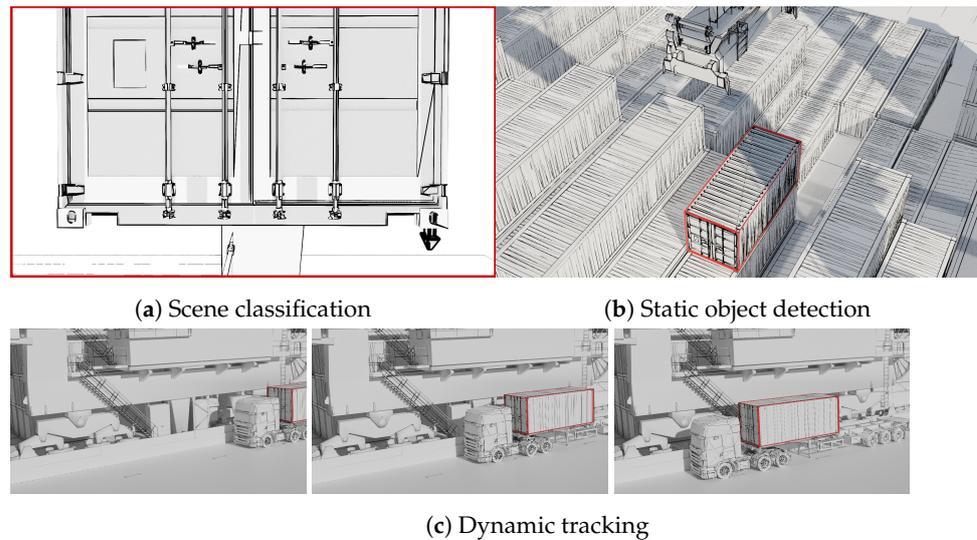


Figure 3. There are various problem classes for automating a crane: (a) Classification—a label is assigned to the entire detected area. (b) Static object detection—an object is localised in a static environment. (c) Dynamic object detection—a moving object is detected and tracked in several successive captures.

3.3. Dynamic Object Detection (Tracking)

Following on from the static object detection, dynamic object detection follows. The main task of a container terminal is to exchange containers. Accordingly, there are not only static, but also dynamic objects that need to be recorded. The dynamic results either from an external movement of the objects or an internal movement of the sensor if it is mounted on a movable component of the crane. A dynamic capture of the position is required, for example, when the vehicles are automatically positioned under the crane, as shown in Figure 3c.

Another example of dynamic object tracking is collision protection. Here, the crane's movement is monitored and the distance to close objects is continuously tracked. In this example, however, the objects themselves are static and the sensor moves.

4. Sensor Usage during Container Handling

In the previous sections, the automation tasks were first presented in general and then divided into the three main categories (classification, static object recognition, and dynamic object recognition). To deal with all these different tasks, it is necessary to capture the environment. This is carried out using LiDAR and camera systems. In Section 2.4, the existing automation systems were presented in general. In this section, the usage of camera and LiDAR sensors will be presented in detail. Based on this, we subsequently define the potential field of application for camera-based solutions. Additionally, we present which process steps are currently already being solved with cameras. This opens up the potential to use existing hardware for multiple tasks.

4.1. LiDAR Usage

4.1.1. Pick/Place Container on Vehicle

In order to be able to transport the containers between the cranes, they are loaded onto horizontal transport vehicles (especially trucks). The loading point between the crane and the truck takes place at a defined point. For this reason, the measuring system is permanently installed near this loading point. This mounting position is selected so that the laser system is as close as possible to the loading point. As a consequence of the reduced distance between the sensor and the measured object, a dense point cloud is captured, which helps precisely localise the loading area on the vehicle [17–19].

4.1.2. Pick/Place Container in Stack

Most automation nowadays takes place in the area of the yard crane. Containers are automatically stored in the stack and remain there until they are transported further. For this purpose, the setting-down position on the ground or on another container roof must be precisely measured. Similar steps are required to grab a container in the stack. Therefore, the target container must be precisely measured. After gathering the exact 3D pose, the crane position is slightly adjusted so that the spreader is exactly above the target container. Finally, the spreader is lowered and the container is precisely picked or placed at the desired position [20–22]. The distances between the trolley (sensor mounting position) and container are much smaller on yard cranes than on ship-to-shore cranes. Therefore, the sensor setup is less affected by mechanical deformation that arises from lifting heavy loads (see Section 6.1.5).

4.1.3. Crane Movement in Stack or Over Ship

A three-dimensional survey of the environment is not only required for grabbing and setting down containers. Even while the crane is moving, the direction of travel is continuously monitored for possible obstacles, so that the crane can be stopped in time to prevent an impending collision. This surveillance is mainly carried out using a 2D-LiDAR sensor facing in the traveling direction [23].

4.2. Camera Usage

Not only lasers, but also cameras, as already mentioned, are occasionally used in automation solutions. This type of sensor is mostly implemented where there is interaction with humans, the intended problem can be solved in a two-dimensional image, or color information has an important meaning. Cameras are inexpensive and have a compact design and no moving components inside. Therefore, they can withstand high mechanical loads. For this reason, they are used in various places in the crane environment.

4.2.1. Remote Operation

Container automation continues to advance. Nevertheless, disruptions occur again and again in the process flow, causing the automation to stop. In addition, there are laws in certain regions of the world that prohibit the automation of certain processes. This is often the case when human life would be in danger [24]. To deal with these difficulties, automated cranes are equipped with cameras that allow them to be controlled remotely by a human [25]. The camera positions are selected in such a way that the operator has the best possible view of the situation and the further path of movement of the container at hand. This is achieved by mounting the cameras on the spreader and thus, close to the respective loading situation. Second, in addition to static cameras, pan-tilt-zoom (PTZ) cameras are used whose field of view can be adjusted depending on the situation.

4.2.2. Container Number and Damage Check

Loading and unloading operations are optimised via the terminal operating system (TOS) so that unnecessary movements of the crane and other vehicles are reduced. To validate the actual database state of the TOS, cameras are used during the loading process to check whether the numbers printed on the containers and transport vehicles correspond to the numbers stored in the system. In addition, permanently installed cameras are used to check whether any damage to the containers is present [26,27].

4.2.3. Measuring Spreader Position

The load lifted by the crane and the ropes between the spreader and trolley form a mechanical pendulum that is oscillated by the movement of the crane. In addition, external factors such as wind can exert a force on the pendulum. In order to be able to reduce and compensate for the unwanted deflection of the load, the exact position of the spreader

must be known. With the help of a camera, the current deflection and skew of the load is calculated and, if necessary, compensated for [28,29].

4.3. Sensor Functionality

The previous chapter has shown that a wide range of automation solutions exist in the port environment. These solutions are mainly based on LiDAR and camera sensors. In order to better compare the individual sensor types and their advantages and disadvantages, we will subsequently introduce their characteristics and functioning.

4.3.1. 3D-LiDAR Technology

In recent years, LiDAR technology has been continuously developed and improved and today forms an important basis for the field of autonomous driving in order to record the environment in three dimensions [30]. As described in Section 4.1, 3D-LiDAR sensors are also used in the crane environment.

The basic measuring principle of a LiDAR sensor is based on the emission and reception of a light pulse. The acquired object distance is then computed with the help of the known speed of light. In order to be able to measure an entire plane with this method, the measuring beam is continuously deflected via a mirror or prism (see Figure 4). By simultaneously measuring the distance and the angular increment, the positions can be specified in a polar coordinate system (2D-LiDAR). For many container-automation scenarios it is not sufficient to just sample the world in a two-dimensional plane. Especially for complex tasks such as the localisation of containers in 3D space, a 3D sampling of the environment is required and therefore, 3D-LiDAR units are used. In the harbour, there are mainly two types of 3D-LiDAR units: mechanical and solid 3D-LiDAR systems.

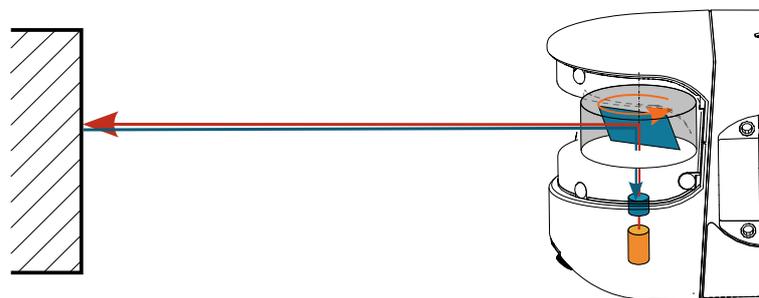


Figure 4. Working principle of a 2D-LiDAR sensor.

Mechanical 3D-LiDAR

A mechanical 3D-LiDAR sensor further extends the idea of a 2D-LiDAR sensor by rotating the whole 2D-LiDAR along a second axis that is orthogonal to the rotation axis of the mirror. As a consequence, two angles are measured continuously and are then combined with the captured distance (see Figure 5). These polar coordinates (φ, θ, d) can be transformed to a precise position in Cartesian coordinates (x, y, z) [31]. A major advantage of this system is the variable point density depending on the application. The slower the 2D-LiDAR sensor is swiveled, the denser the resulting point cloud in the swivel direction. In addition, only individual segments can be measured precisely. However, this measuring principle also has disadvantages. Due to the large size, mounting is not possible everywhere. In addition, the high mass of the 2D-LiDAR sensor cannot be accelerated as much as desired without negatively affecting the measurement quality due to overshoots. Thus, only slow swiveling is possible, which leads to a low sampling rate.

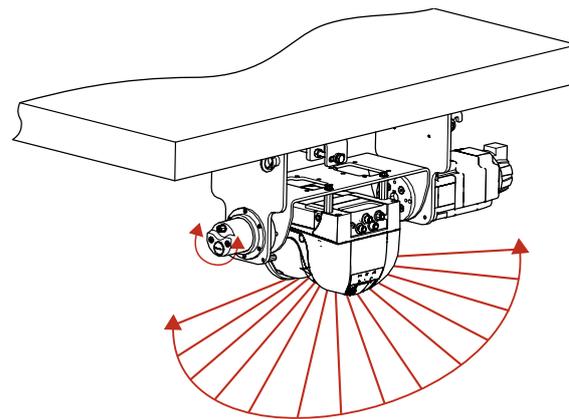


Figure 5. 3D-LiDAR sensor contains a 2D-LiDAR unit which is swiveled along an axis.

Solid LiDAR

To overcome the problems of mechanical sensors, so-called solid LiDAR has been developed. In recent years, the development processes of LiDAR units has progressed and so the dimensions of the individual measurement units can be reduced intensively. As a consequence, it is nowadays possible to place multiple emitter–receiver units closely packed on one chip. The individual units are arranged such that they span a fan of measurements [31]. The most modern sensors of this type have a field of view (FOV) of approx. 40° . The measurement fan consists of up to 128 individual rays. Due to the small design, this measuring fan can be accelerated, so that the fan can be rotated up to 100 times per second [32]. Here, however, the point density is static and dependent on the constant rotation speed of the unit. In addition, the number of fans is still very limited today. This results in a high horizontal, but low vertical resolution. The units are significantly smaller in design. However, the rotating components are sensitive to shocks and vibrations, which means that not every mounting position on the crane is suitable.

4.3.2. Camera Technologies

A successful industrial camera application is dependent on certain conditions. The relevant area must be visible within the camera's field of view. Furthermore, the camera must produce a sharp and adequately bright image so that all the relevant details are discernible. These properties depend on the so-called camera parameters, which are presented below.

In container terminals, the lighting conditions change throughout the day (day to night) and through changing weather conditions. Furthermore, there are also short-term changes in exposure due to clouds and other weather effects. The camera sensitivity refers to the ability of a camera to capture images in low-light conditions. The dynamic range refers to the range of light levels that a camera can capture in a single image, from the darkest shadows to the brightest highlights. A camera with a high dynamic range can capture more detail in both the brightest and darkest areas of an image. This aspect is important for industrial crane applications because there are dark, shadowed areas between tall container stacks. However, there are also metal surfaces that reflect strongly and are almost mirror-like.

As presented in Section 3, a possible application field is the detection of moving objects. Therefore, the frame rate is a crucial factor in selecting an appropriate camera. The frame rate defines how many images are captured per second. An appropriate frame rate depends on the speed of the moving objects or the camera.

From a practical application perspective, the size and robustness of the camera are also important factors. The camera size should be appropriate for the specific application and installation requirements, and it should be able to withstand shocks and vibrations.

In addition to the parameters of a camera mentioned so far, the field of view (FOV) plays a particularly important role for a successful industrial use. The FOV describes the

area that the camera captures. The captured area depends on the extrinsic and intrinsic parameters of the camera.

Extrinsic Camera Parameters

Extrinsic camera parameters describe the location and orientation of the camera in 3D space, with respect to a global coordinate system. For a concrete application in the port environment, this means where the camera is mounted on the crane. The shorter the distance between the camera and the object to be detected, the more detailed the object can be captured. The greater the distance chosen, the more surrounding objects are visible in each frame, providing a better overview of the overall situation. Therefore, a careful balance must be struck between the two options and a suitable compromise must be found. The chosen camera's extrinsic properties are especially important for vision-based measurement tasks, because this relation is an essential basis for mapping image coordinates to their corresponding 3D world coordinates [33].

Intrinsic Camera Parameters

Intrinsic camera parameters refer to the internal characteristics of the camera that affect how the camera captures and projects images. One of these internal parameters is the focal length. The focal length describes the distance between the camera lens and the image sensor. A smaller focal length results in a wider field of view, which means that the camera can capture a larger area but with less detail. On the other hand, a larger focal length results in a narrower field of view, which means that the camera can capture a smaller area with more detail. The level of detail in the image depends not only on the focal length but also on the size and resolution of the camera chip. The sensor size describes the physical size of the chip, which digitizes the projected image. Depending on the resolution (number of pixels in the x and y directions), the pixel size and thus the aspect ratio of the pixels are determined. There are several types of chips used in cameras, including CCD (charge-coupled device) and CMOS (complementary metal-oxide-semiconductor) sensors. Both types of sensors convert light into electronic signals, but they differ in their internal mechanisms and performance characteristics [34].

In order to obtain a sharp and well-illuminated image, additional lenses are required to create the scene projection. However, adding lenses to the camera system causes distortion of the image and blurred edges. These phenomena can be described by lens distortion coefficients.

Fixed Mounted and PTZ Cameras

In crane automation, two different types of cameras are utilised: fixed mounted and PTZ cameras. For automation tasks where the evaluation area is fixed and does not change over time, fixed-mount cameras are used. In these applications, the intrinsic and extrinsic camera parameters do not change. Therefore, the mounting position and the field of view are fixed. This has the great advantage that the system can be carefully measured (calibrated) once at the beginning, and from then on, this static relationship between the 3D world and the sensor projection remains stable and can be used for converting image points to world coordinates.

However, in the area of stacking cranes, the working area cannot always be clearly defined beforehand. There may be high stacks of containers as well as low ones.

In both situations, the container must be clearly visible in the image so that, for example, a remote operator can land the spreader precisely on the container. In such situations, cameras with a variable field of view—such as PTZ cameras—are used. These cameras have an optical zoom and are able to dynamically adjust the current FOV (intrinsic camera parameters). In addition, these cameras can be rotated about two axes, thus repositioning the optical axis (extrinsic camera parameters). This camera type can also cover highly dynamic situations with a variable working area. However, since both internal and external camera parameters change during operation, it is not sufficient to measure

the relationship between the 3D world and the camera image once. Rather, the projection matrix must be continuously adjusted during operation.

4.3.3. Sampling Density

In the preceding sections, the two sensor concepts and their properties were introduced in general. In some process steps, automatic container handling requires that even fine structures (such as twistlocks or corner casts; see Section 2.3) be localised. For this to be possible, these structures must also be resolved by the sensor used. This property is characterised by the so-called sampling density. In order to obtain a better understanding about this parameter, we will compare the sampling density of both sensor technologies based on a real example from the stack area (see Section 2.1). For this purpose, some assumptions are made, which will be presented below and visualised in Figure 6.

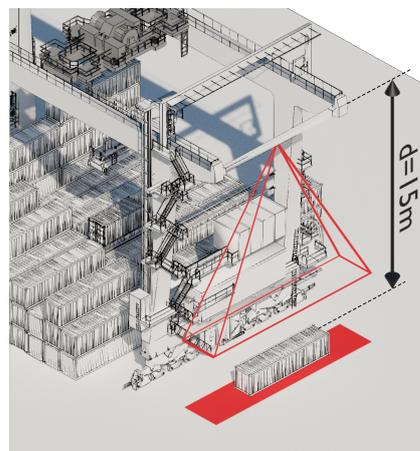


Figure 6. Example: Localising a container below a yard crane.

- A straddle carrier placed a 40 ft container below a stacking crane.
- Its rough position is known, but should be precisely measured by an automation system.
- The area to be considered is thus selected to be larger in each direction by half a container width or length.
- The sensor is mounted at the crane girder, roughly 15 m above the container.
- The sensor must evaluate a field of view of approx. $60^\circ \times 20^\circ$.
- As the sensor and container do not move in this example, it is a static object detection (see Section 3).
- To avoid slowing down the automation process, the evaluation needs to be completed within one second.

In order to safely lift a container using a crane, a mechanical connection is made using corner casts and twistlocks (see Section 2.3). The exact position of these corner casts—which have a size of roughly $10\text{ cm} \times 10\text{ cm}$ —is therefore particularly important and will be examined more closely in this example. For this comparison, exemplary representatives of each sensor type, mechanical LiDAR, solid LiDAR, and camera, were chosen.

In this evaluation, it must be taken into account that the information content differs between the LiDAR system and the camera. LiDAR systems provide a position value in space for each measurement, whereas cameras can provide colour information. Thus, this comparison is only valid for use cases in which the intended problem can be solved completely in a two-dimensional projection (without spatial information) or where multiple camera views can be combined to obtain spatial information. Table 2 shows the vertical and horizontal resolution and the respective frequency of the individual sensor. For the mechanical LiDAR, the assumption was made that the swivelling speed was chosen accordingly so that the required area could be scanned in one second.

Table 2. Evaluation of different sensor types to localise a corner cast.

	Model	h_{res}	v_{res}	Frequency	Measurements Per dm^2 ($d = 15\text{ m}$)
Solid LiDAR	Velodyne Puck Hi-Res	0.1°	1.33°	20 Hz	1.09
	Ouster OS1	0.17°	0.351°	20 Hz	2.44
mech. LiDAR	Swivel-Unit with Sick LMS5xx Heavy Duty	0.1667°	0.2°	1 Hz	4.38
	Swivel-Unit with Sick LRS4000	0.04°	0.8°	1 Hz	4.56
Camera	AXIS P3925	0.029°	0.03°	30 Hz	180.16

Looking at the table, one can clearly see that the camera has the highest density of measuring points in this example. In the relevant area of the corner casts, approx. 180 measuring points (pixels) are to be expected. On the other hand, when considering the solid LiDAR technology, it is noticeable that only one or two measuring points can be found in the respective field. To determine the exact position of the container, other features—such as entire areas or edges—must, therefore, be taken into account.

The mechanical LiDAR systems provide a similar density of measuring points. However, these systems have the advantage that the number of measuring points correlates linearly with the swiveling speed. If the measurement time were increased fivefold, the number of measurement points would also increase fivefold. However, the measuring point density would not be evenly distributed, but would only increase in the direction of swivel. To obtain a comparable measurement point density to that of the camera, the measurement duration would have to be increased to more than 30 s, which would be unacceptably slow for any automation process.

For this comparison, a task from the field of static object detection has been chosen. However, as presented in Section 3, there are also situations in which the object and/or camera are in motion. In these cases, dynamic object detection is required, which places even higher demands on the sampling speed, which can only be fulfilled by solid LiDAR and camera systems. In comparison to mechanical LiDAR systems, these sensors can scan the desired object 30 times (camera) or 20 times (solid LiDAR) within the specified one second. This makes it possible to average out measurement errors or to capture dynamics.

5. Related Work

As presented in the previous chapters, container automation nowadays relies on laser-based measurement systems in many areas. In the port environment, the research area of camera-based container automation is relatively new and has been studied only selectively. To best of our knowledge, there is currently no existing literature that examines the transformation from laser-based to camera-based container automation on this scale. Individual publications focus on early camera-based applications, but often only a single area of the container flow is considered. A comprehensive view of the entire process is still missing. On the other hand, cameras have been successfully used for measuring objects in other fields of application for many years. Shirmohammadi et al. generally considered the trend of vision-based measurement (VBM) [35]. They first show the general processing chain: visual sensor, preprocessing, image analysis, measurand identification, measurement, and result. In addition to the process steps considered by the authors, this chain should be supplemented in the crane environment by additional plausibility checks. These additional checks are necessary to increase the safety of the terminal employees. Furthermore, the authors considered what uncertainties can arise from the use of cameras for evaluating real-world scenarios. The following points are the biggest causes of measurement inaccuracies:

poor lighting, changing camera angles, incorrect calibration/gauging, and different camera equipment [35]. The points mentioned are also relevant for the use of cameras in the port environment.

Mi et al. picked up the general considerations regarding VBM and looked at the current development trend towards VBM in ACT. The authors presented camera solutions that are already in use nowadays. In addition, future application fields in the terminal environment were considered, for example, container surface damage recognition, the task of truck positioning, and truck-lifting prevention [36].

Ref. [37] took a closer look at one of the aforementioned use cases: the positioning of trucks under yard cranes (ARMGs). For this purpose, two cameras—which are mounted vertically above the truck lane—are used to determine the exact position of the loaded container on a truck. To localise the exact position, they proposed a multi-stage approach. First, the rough position of the corner casts is determined using an adapted single-shot detector. Then, the detections are preprocessed with classical computer vision methods to compensate for illumination invariances and other disturbances. In the final step, the best rectangular fit around the corner cast is computed [37]. The position and size of the corner hole is then used to compute the offset distance and deflection angle between the container and spreader landing position. Ref. [37] evaluated their approach in a set of experiments and compared the unsupported SSD results against the complete multi-stage approach. The positioning error of the modified SSD detection is about 8.52 px in the gantry and 4.44 px in the trolley direction. Projecting the positions detected in the image back into the real world results in a measurement error of 48.2 mm in the gantry and 25.1 mm in the trolley direction. If these measurement errors are compared with the mechanical properties of a container presented in Section 2.3, the accuracy would not be sufficient to automatically pick up a container. Therefore, the additional processing steps are required that reduce the error to 19.6 mm in the gantry and 14.3 mm in the trolley direction. Consequently, it is possible that the spreader automatically slips into the corner casts and the container can automatically be grabbed.

Based on their previous findings, Zhang et al. focused in their following publication on the realtime three-dimensional attitude positioning when loading containers onto and off of trucks [37]. When loading containers automatically, it is essential to know the exact three-dimensional position of the container in order to avoid dangerous situations, such as the truck lifting up or the container overturning. They used the same basic structure as in the previously presented paper. First, the rough position is determined with the help of a deep learning method and it is then specified using classical computer vision methods. However, the application presented in this paper differs in the mounting position of the cameras and in terms of the required inference speed. In order to achieve a high inference speed, the backbone architecture of the model was adapted. Instead of a VGG16, a ResNet 18 backbone was chosen, which contains significantly fewer parameters [38]. The determined predictions are fed into a detection-based tracking network afterwards, which combines the detection results over time. Finally, the individual results are further improved using classic computer vision methods for detecting the slotted hole in the corner cast. The final investigations show that the spatial position in the gantry and hoist directions can be determined much more precisely than the trolley position. This is due to the fact that the offset of the container in the trolley direction is only determined by the size of the corner cast hole. All in all, the presented approach can determine the three-dimensional position in approx. 80 ms, which enables the desired safety monitoring.

ACTs are a dangerous working environment, as heavy loads are moved. Therefore, it is important to know the exact position of the load. The previously discussed work [37] achieves this by localising the corner casts. Ref. [39] deals with the same problem but by localising the spreader and not the picked-up load. In comparison to the approach of [37], the position of the spreader can be determined even without a picked-up load. In the method presented, the image of a spreader is compared with a three-dimensional wireframe rendering of a 3D triangle mesh model of the spreader. The initial estimated pose

is then continuously shifted so that the visible lines of the wireframe match the straight line segments detected in the original spreader image [39]. With the presented method, the spreader position can be determined with a maximum error of 2.5 m and an average error of 0.5 m. These values are too inaccurate for the exact control of an automation system, but they are sufficient for avoiding rough container collisions and protecting dock workers.

6. Evaluation of Using Cameras Instead of LiDAR Sensors for Automated Container Handling

In Section 2.4, the typical container flow in a terminal was presented. LiDAR sensors are used at different process steps. In addition, cameras for remote control but also for other automation tasks are available (see Section 4), which cover a similar field of view. Application papers in similar and other domains have shown that specific problem types can be solved using cameras and deep learning techniques. In the following, the chances and challenges of replacing LiDAR-driven automation with cameras in the field of ACTs will be discussed. The idea of using cameras instead of expensive LiDAR and radar sensors to save costs is not new. This idea has been driven from the field of autonomous driving in recent years. However, at the time of writing, there has still been no breakthrough in this field. However, it must be considered that the crane environment differs significantly from the application of autonomous driving. The movement possibilities of a crane are limited, the objects to be detected are known in advance, and the environment can be illuminated and adapted if necessary. In relation to the specific application field of container automation, the chances as well as the challenges of this transformation are discussed in the following sections. Finally, the mentioned aspects are evaluated.

6.1. Chances

6.1.1. Cost Reduction

Due to the far distance between the sensor and container as well as the challenging demand regarding measurement precision, the 3D-LiDAR sensors used in the harbour environment are quite expensive. At the time of writing, typical outdoor 3D-LiDAR sensors cost more than USD10,000 [32]. Meanwhile, an outdoor PTZ camera costs several hundred dollars. LiDAR sensors are becoming smaller in size, enabling them to be utilised in new areas such as smartphones. As the market demand increases, the cost of LiDAR sensors will eventually decrease. However, an industrial camera, with a cost of several hundred euros, is still considerably less expensive than a LiDAR sensor. If one also takes into account that some cameras are already needed for remote control and could be reused for other applications, there is a clear cost advantage, even if several of these cameras would have to be mounted.

Due to the reduced hardware costs, automated cranes would become affordable for smaller terminals that currently cannot afford expensive LiDAR sensors. Therefore, these terminals could increase their throughput and safety as well.

In addition, there has been a clear trend towards energy reduction in container terminals in recent years [40]. If the exact spatial locations of the containers can be detected by inexpensive sensors, the travel path of the crane can be optimised, resulting in energy savings.

6.1.2. Mounting Positions

The LiDAR systems used during stacking operations are currently attached to the bottom side of the trolley. Therefore, the system is protected from environmental influences such as heavy rain and is mechanically less stressed. However, this mounting position quickly causes occlusion effects, especially if a container is grabbed by the spreader (see Figure 7). The problem gets worse when neighboring stacks are high and cause additional occlusion effects. As a consequence, the hardware setup is only able to capture a portion of the relevant area. When the load is lowered, it is not possible to further measure the target position. Therefore, the container is placed blindly on the destination for the last few

meters. If mechanical wear and tear leads to an uneven movement of ropes and tires, the container is not placed precisely on the target position.

The problem could be reduced if a greater distance were maintained between the individual container rows. However, this would lead to a lower stack density and thus, lower utilisation of the terminal. Due to their compact design and high mechanical load capacity, cameras can be mounted directly on the spreader. Due to the reduced distance to the target object, occlusion does not occur at all or only occurs much later during the lowering process of the container. The stack density also remains unchanged. Today, spreader cameras are already being used successfully for remote control [41].

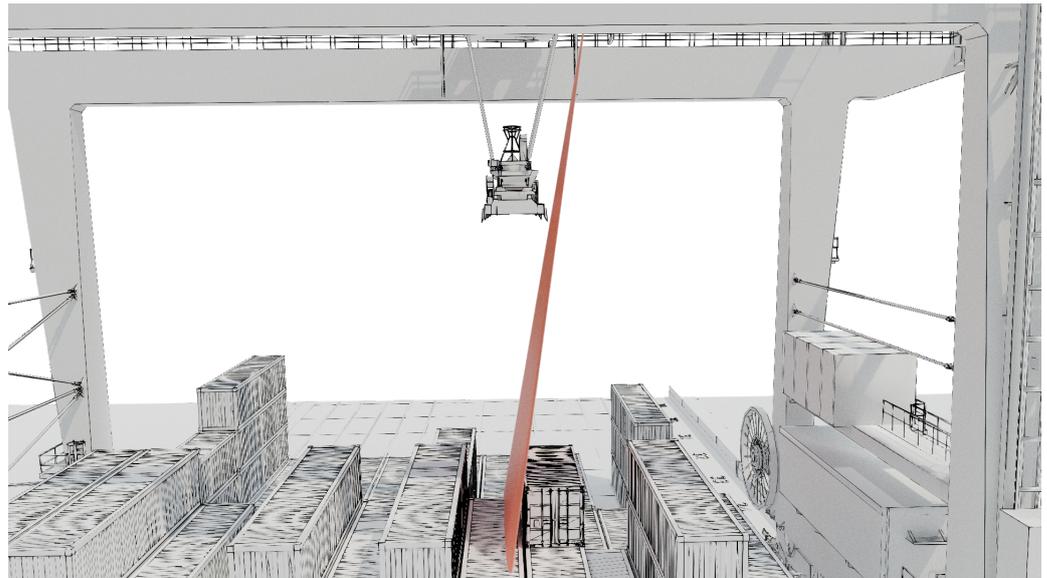


Figure 7. The typical LiDAR mounting position protects the sensor from strong shocks and weather effects. However, occlusion may occur due to the relative position of the objects.

6.1.3. High Sampling Rate and Sampling Density

Over the past decades, the pixel density of camera chips has been steadily increased, resulting in a higher number of pixels. Modern industry cameras offer a 4K resolution (3840×2160 pixels) at a frame rate of 30 FPS. Thus, potentially about 250 million measurement points per second are available for further evaluation [42]. In comparison, a modern solid 360° LiDAR sensor achieves up to 5 million measurement points per second, but distributed over a significantly larger field of view.

One must clearly differentiate between the information content of each sensor type. With LiDAR systems, the spatial position in the x , y , and z directions can be determined for each individual measurement point. In addition, information is obtained about the remission of the object that was hit by the individual measuring beam. In comparison, the pixels of a camera sensor provide different information. Only reduced spatial information can be determined with a single pixel. However, color information about the environment is obtained. In addition, the spatial density is significantly higher. There are fewer unscanned areas and detailed contextual relationships can be more easily determined because of the high scanning density.

6.1.4. Multifunctional Use

In order to be able to intervene in the event of errors, each automatic stacking crane can be controlled remotely. To achieve this, each crane is equipped with additional cameras for remote control. During automatic operation, these cameras are not used and can therefore enable automation solutions. The current FOV is chosen in such a way that a remote operator can easily capture the current situation and therefore safely operate the crane. As a

consequence, the camera FOV often fits quite well for automation tasks as well. In addition, some cameras can be adjusted in their FOV. Therefore, one sensor fulfills multiple functions.

6.1.5. Less Affected by Mechanical Deformations

Heavy weights are loaded with the help of a crane. This means that the crane boom and the whole crane structure can bend under the mechanical stress. This bending is intentional and taken into account in the crane design. However, this leads to a general problem of attaching optical sensors to the crane structure. The field of view slightly changes depending on the crane deformation (extrinsic sensor parameter). If there are large distances between the sensor and acquired object, even small deviations lead to large measuring errors.

Due to the compact and robust design of a camera, these sensors can be placed closer to the desired object and are therefore less affected. Furthermore, there is the possibility to attach cameras to the spreader and directly measure the desired relative offset between the spreader and the target object. Crane structure deformations are compensated for by performing relative measurements.

6.1.6. Complexity during Commissioning and Operation

Our eyes, like cameras, transform our environment into a two-dimensional representation. Therefore, we are accustomed to this representation of our surroundings and capable of assessing its content intuitively. This ability makes it easier to commission camera systems compared to LiDAR systems. The precise adjustment and registration of the sensor system is additionally facilitated by the high pixel density (see Section 6.1.3 for comparison). Due to their low cost, cameras are often already used for other tasks in the port environment. The port operators responsible for commissioning and maintaining the sensors thus often have experience with the sensor technology used, even if there has been no container automation before. The low storage requirements of images compared to LiDAR point clouds also facilitate daily work. It is possible to store images on a large scale and use them to test and further develop the system.

6.2. Challenges

Besides the mentioned chances, there are also some disadvantages regarding camera-based container automation.

6.2.1. Missing Depth Information

The currently used laser systems send a light pulse and measure the time until the reflected light again reaches the sensor. By knowing the time difference, one can compute the distance between the laser and the object. This active sensor technology enables very precise distance measurements even over long distances. Modern LiDAR sensors have a range of up to 200 m with an accuracy of up to 1 cm. Another advantage of this sensor technology is the high sampling frequency. Modern sensors can rotate their reflective unit up to 100 times per second and, therefore, can scan the environment in a two-dimensional scan plane multiple times per second. This is particularly advantageous for collision monitoring.

In a camera system, the three-dimensional world is projected onto a two-dimensional chip (see Section 4.3.2). Therefore, the depth information is lost. For some automation tasks, this only plays a subordinate role, for example, in the case of two-dimensional positioning. In other cases, the third dimension is essential, for example, to prevent collisions.

Depending on the selected application, this problem can be overcome by reconstructing the missing depth information using two cameras or known object sizes as shown in [43]. However, such solutions result in more complex systems and do not provide accurate depth information in all circumstances.

6.2.2. Bad Visibility

LiDAR sensors measure distances by actively emitting and receiving a light impulse. This measurement concept works even in absolute darkness and in poor visibility. However, passive cameras do not actively emit light but just capture the reflected light of the surroundings. In order to use passive cameras for container automation it must, therefore, be ensured that the environment is sufficiently illuminated.

6.2.3. Fixed Field of View and Point Density

As explained in the section on mechanical 3D-LiDAR, a mechanical 3D-LiDAR sensor is a standard 2D-LiDAR unit mounted on a swivel unit. By simultaneously scanning and swiveling, the laser can capture a dense 3D point cloud. The density is increased by reducing the swivel speed. This setup is really beneficial if a varying point density is desired. However, this setup leads to a slow acquisition frequency. The scan procedure for one complete 180° of the swivel unit lasts multiple seconds.

In comparison, a rigid-mounted camera has a predefined FOV with a fixed horizontal and vertical resolution. If a varying FOV in terms of direction and size is needed, the use of PTZ cameras could be considered. Regions of interest can be precisely captured using an optical zoom. However, these cameras are more expensive than rigidly mounted cameras.

6.2.4. Available Datasets

As outlined in the previous sections, there are first approaches striving for a VBM for ACTs. These methods are often trained in a supervised manner and therefore, labeled datasets are required. For other domains such as people or car detection and localisation there are public datasets that can be used for training and benchmarks. Furthermore, these available datasets can be used to pretrain new model architectures.

However, for the application field of container automation there are nearly no publicly available datasets. In order to be able to use deep learning procedures, large data records must be collected and prepared. Data recording efforts can be reduced by generating synthetic datasets, e.g., by utilising simulations or generative methods.

6.3. General Evaluation and Looking Ahead

In the previous sections, the port environment and typical areas of application were presented in detail. Subsequently, the different sensors were presented, and the opportunities and challenges that arise from their use were considered. The key aspects of this comparison are summarised in Table 3.

To elaborate the core findings from the aforementioned, we evaluate these results and take a visionary look into the future. Looking at the left column of Table 3, the clear advantages of using cameras become immediately apparent. Cameras are cost-effective, capable of capturing the environment at a high resolution and frame rate, and robust against shocks. These advantages open up prospective applications and automation fields, such as the complete automation of the water side, which has so far only been partially automated. With reduced costs, it would then also be possible for smaller terminals to automate their processes, achieve a higher throughput, and increase safety in the terminals. In addition, this system can help enable more sustainable loading of containers. Through cost-effective sensors, the spatial position of the containers under the crane can be determined. The determined container heights can then be used to minimise the crane's movement and save energy.

However, the idea of replacing expensive LiDAR sensors with inexpensive cameras is not a new idea in the field of container automation. Rather, this idea originated years ago in the field of autonomous driving and has been extensively researched since then. However, the breakthrough has not yet come, and car manufacturers worldwide continue to rely on a hybrid sensor mix to implement driver-assistance systems in their vehicles. Does this mean that a transformation from LiDAR- to camera-based container automation is hopeless? Looking at the right column of Table 3, we see that the same problems arise in the crane environment as in autonomous driving. The biggest challenge arises from

the lack of depth information and poor visibility in the dark. However, this is precisely where the domain of autonomous driving differs from the port environment. Container cranes are much larger than typical vehicles on the road, making it possible to mount a large number of cameras with varying fields of view. Thus, automated container handling can be performed based on many different perspectives. An autonomous vehicle, on the other hand, can only rely on a cockpit view within its domain. In addition, the variability of objects to be detected is strongly limited and standardised by norms compared to road traffic. Additional floodlights enable bright illumination of the current situation at night.

Table 3. Summary of opportunities and challenges of the transformation from laser-based to camera-based container automation.

Chances	Challenges (Ways Forward)
<p>Cost reduction: Camera sensors are less expensive than LiDAR systems. When existing cameras are used, a special cost advantage is created.</p> <p>Mounting positions: Cameras are small and robust against shocks. Therefore, this sensor can also be mounted in places that are subjected to higher mechanical stresses.</p> <p>Higher sampling rate and density: Cameras provide a higher sampling rate (measurements per second) and a higher sampling density.</p>	<p>Missing depth information: Standard industry cameras do not provide depth information. Therefore, the object position can only be located in a projection and is not directly known in world coordinates.</p> <div style="border: 1px solid black; padding: 2px;"> <p>Ways forward: By combining multiple cameras or using context information and prior knowledge (known object sizes), the missing depth can be compensated for.</p> </div> <p>Bad visibility: In comparison to LiDAR sensors, cameras do not actively emit light for their measurements. For a reliable operation, it is therefore necessary that the environment be adequately illuminated.</p> <div style="border: 1px solid black; padding: 2px;"> <p>Ways forward: Attaching additional lights to the crane.</p> </div> <p>Fixed FOV: The working range of a container crane is very large, and therefore, the potential field of view of a sensor must also be very large. However, rigidly attached cameras have only a fixed, limited FOV.</p> <div style="border: 1px solid black; padding: 2px;"> <p>Ways forward: Attaching multiple cameras to a crane or using cameras with a varying FOV.</p> </div>

Comparing the two application domains, we find that they differ significantly in their complexity. The problems that arise in the intended transformation are valid for both domains, but the container-automation domain facilitates clear solutions to these problems. At the moment, these are just potential strategies that have been selectively evaluated. Although the breakthrough in camera-only autonomous driving is still pending, we are convinced that some of the crane-automation issues can be fully addressed with cameras. Taking into account that significantly more research is being conducted in the field of camera-based deep learning applications compared to LiDAR technology, it can be assumed that further advancements in image processing can be expected within the coming years. Whether all the requirements of a fully autonomous terminal can be realised with these advancements, however, remains to be shown by further scientific investigations.

7. Conclusions

Nowadays, terminal operators worldwide aim for automated container terminals. For an automated crane operation, the three-dimensional position of containers, vehicles, and set-down positions must be determined. This is mainly conducted by utilising LiDAR sensors. This publication discussed the potential to replace the existing laser-based solutions with cameras. The main benefit of this transformation comes from the reduction in hardware costs. Thus, automation solutions would also become viable for smaller terminals, which could also increase their handling rate and safety. Further advantages result from the additional mounting options. On the one hand, the occlusion of the measuring area can

be reduced. In addition, further tasks can be automated that previously had to be carried out manually.

However, the use of cameras also poses some challenges like the fixed field of view. Some first ideas exist for how to cope with these challenges, but further research is required to evaluate these approaches. The most important issue is the lack of depth information. This deficit can be overcome as described by [37] with the utilisation of context information and prior knowledge, for example, by using known object sizes for depth estimates. In addition, triangulation with several cameras is conceivable. However, whether the accuracy is sufficient for the expected automation solutions in every application needs to be investigated.

In summary, this evaluation shows that the transformation from laser- to camera-based crane automation offers important opportunities, but the identified challenges and their possible solutions must be further evaluated.

Author Contributions: Writing—review & editing, J.B., R.M. and T.M. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge support from the Open Access Publication Fund of the University of Wuppertal.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. *Navigating Stormy Waters*; United Nations: Geneva, Switzerland, 2022.
2. Luo, J.; Wu, Y. Scheduling of Container-Handling Equipment during the Loading Process at an Automated Container Terminal. *Comput. Ind. Eng.* **2020**, *149*, 106848. [[CrossRef](#)]
3. Martín-Soberón, A.M.; Monfort, A.; Sapiña, R.; Monterde, N.; Calduch, D. Automation in Port Container Terminals. *Procedia Soc. Behav. Sci.* **2014**, *160*, 195–204. [[CrossRef](#)]
4. Nogueira, V.; Oliveira, H.; Augusto Silva, J.; Vieira, T.; Oliveira, K. RetailNet: A Deep Learning Approach for People Counting and Hot Spots Detection in Retail Stores. In Proceedings of the 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, 28–30 October 2019; pp. 155–162. [[CrossRef](#)]
5. Fujiyoshi, H.; Hirakawa, T.; Yamashita, T. Deep Learning-Based Image Recognition for Autonomous Driving. *IATSS Res.* **2019**, *43*, 244–252. [[CrossRef](#)]
6. Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep Learning—Method Overview and Review of Use for Fruit Detection and Yield Estimation. *Comput. Electron. Agric.* **2019**, *162*, 219–234. [[CrossRef](#)]
7. Chen, C.; Li, Y. Ship Berthing Information Extraction System Using Three-Dimensional Light Detection and Ranging Data. *J. Mar. Sci. Eng.* **2021**, *9*, 747. [[CrossRef](#)]
8. Mentjes, J.; Wiards, H.; Feuerstack, S. Berthing Assistant System Using Reference Points. *J. Mar. Sci. Eng.* **2022**, *10*, 385. [[CrossRef](#)]
9. Vaquero, V.; Repiso, E.; Sanfeliu, A. Robust and Real-Time Detection and Tracking of Moving Objects with Minimum 2D LiDAR Information to Advance Autonomous Cargo Handling in Ports. *Sensors* **2018**, *19*, 107. [[CrossRef](#)]
10. Kim, K.H.; Lee, H. Container Terminal Operation: Current Trends and Future Challenges. In *Handbook of Ocean Container Transport Logistics*; Lee, C.Y., Meng, Q., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 220, pp. 43–73. [[CrossRef](#)]
11. Arena, A.; Casalotti, A.; Lacarbonara, W.; Cartmell, M. Dynamics of Container Cranes: Three-Dimensional Modeling, Full-Scale Experiments, and Identification. *Int. J. Mech. Sci.* **2015**, *93*, 8–21. [[CrossRef](#)]
12. Zhang, P.; Xie, C.; Fei, H. Twist Lock Unlocking Process Research and Unlocking Fixture Design in Container Terminals. In Proceedings of the 2015 4th International Conference on Computer, Mechatronics, Control and Electronic Engineering, Guangzhou, China, 28–29 September 2015. [[CrossRef](#)]
13. *ISO 1161*; Series 1 Freight Containers—Corner and Intermediate Fittings—Specifications Standard. International Organization for Standardization: Geneva, Switzerland, 2020.
14. Vrakas, G.; Chan, C.; Thai, V.V. The Effects of Evolving Port Technology and Process Optimisation on Operational Performance: The Case Study of an Australian Container Terminal Operator. *Asian J. Shipp. Logist.* **2021**, *37*, 281–290. [[CrossRef](#)]
15. Gustafsson, T.; Heidenback, C. Automatic Control of Unmanned Cranes at the Pasir Panjang Terminal. In Proceedings of the International Conference on Control Applications, Glasgow, UK, 18–20 September 2002; Volume 1, pp. 180–185. [[CrossRef](#)]
16. Tang, C.; Chen, P.; Li, Y. Automatic Damage-Detecting System for Port Container Gate Based on AI. In Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition, Xiamen, China, 30 October 2020–1 November 2020; pp. 146–151. [[CrossRef](#)]

17. LASE—Industrielle Lasertechnik GmbH. LaseTPS—Truck Positioning Crane. 2023. Available online: https://lase-solutions.com/wp-content/uploads/2022/03/LaseTPS_-_Truck_Positioning_System.pdf (accessed on 16 January 2023).
18. Siemens, A.G. SIMOCRANE Truck Positioning System (TPS)—Highly Precise Laser Measurement System for Accurate Truck Positioning. 2023. Available online: <https://assets.new.siemens.com/siemens/assets/api/uuid:6286d728-87e9-4c13-9ada-7ae32e282e07/vrtl-b10009-00-7600-144dpi-simocrane-truck-positioning-system.pdf> (accessed on 16 January 2023).
19. QLaserOn. Laser Guided Container Loading and Unloading System. 2023. Available online: https://www.qagetech.com/PDF/QLaserOn_R1_Int.pdf (accessed on 20 January 2023).
20. Siemens, A.G. SIMOCRANE Final Landing System (FLS). 2023. Available online: <https://assets.new.siemens.com/siemens/assets/api/uuid:31c62b34-5ec2-436a-8b51-dad9fa14ab0d/vrtl-b10019-00-7600-144.pdf> (accessed on 16 January 2023).
21. LASE—Industrielle Lasertechnik GmbH. LaseAYC—Automatic Yard Crane. 2023. Available online: https://lase-solutions.com/wp-content/uploads/2022/01/ds_LaseAYC_-_Automatic_Yard_Crane_web.pdf (accessed on 22 January 2023).
22. Blaiklock, P. AUTOMATED STACKING CRANES. 2017. Available online: https://wpassets.porttechnology.org/wp-content/uploads/2019/05/25183601/052-053_3.pdf (accessed on 5 March 2023).
23. PEMA—Port Equipment Manufacturers Association. *Information PAPER—Collision Prevention at Ports & Terminals*; PEMA—Port Equipment Manufacturers Association: Brussels, Belgium, 2023.
24. Ilkova, V.; Ilka, A. Legal Aspects of Autonomous Vehicles—An Overview. In Proceedings of the 2017 21st International Conference on Process Control (PC), Strbske Pleso, Slovakia, 6–9 June 2017; pp. 428–433. [CrossRef]
25. Gattuso, D.; Pellicanò, D.S. Perspectives for Ports Development, Based on Automated Container Handling Technologies. *Transp. Res. Procedia* **2023**, *69*, 360–367. [CrossRef]
26. Hütten, N.; Meyes, R.; Meisen, T. Vision Transformer in Industrial Visual Inspection. *Appl. Sci.* **2022**, *12*, 11981. [CrossRef]
27. Liu, Y.; Li, T.; Jiang, L.; Liang, X. Container-Code Recognition System Based on Computer Vision and Deep Neural Networks. In Proceedings of the Advances in Materials, Machinery, Electronics II: Proceedings of the 2nd International Conference on Advances in Materials, Machinery, Electronics (AMME 2018), Xi’an, China, 20–21 January 2018; p. 040118. [CrossRef]
28. Kawai, H.; Choi, Y.; Kim, Y.B.; Kubota, Y. Position Measurement of Container Crane Spreader Using an Image Sensor System for Anti-Sway Controllers. In Proceedings of the 2008 International Conference on Control, Automation and Systems, Seoul, Republic of Korea, 14–17 October 2008; pp. 683–686. [CrossRef]
29. Kawai, H.; Kim, Y.B.; Choi, Y. Measurement of a Container Crane Spreader Under Bad Weather Conditions by Image Restoration. *IEEE Trans. Instrum. Meas.* **2012**, *61*, 35–42. [CrossRef]
30. Li, Y.; Ibanez-Guzman, J. Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems. *IEEE Signal Process. Mag.* **2020**, *37*, 50–61. [CrossRef]
31. Li, N.; Ho, C.P.; Xue, J.; Lim, L.W.; Chen, G.; Fu, Y.H.; Lee, L.Y.T. A Progress Review on Solid-State LiDAR and Nanophotonics-Based LiDAR Sensors. *Laser Photonics Rev.* **2022**, *16*, 2100511. [CrossRef]
32. Pacala, A. Introducing the OS1-128 Lidar Sensor. 2019. Available online: <https://ouster.com/blog/introducing-the-os-1-128-lidar-sensor/> (accessed on 22 April 2023).
33. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004. [CrossRef]
34. Ohta, J. *Smart CMOS Image Sensors and Applications*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2020.
35. Shirmohammadi, S.; Ferrero, A. Camera as the Instrument: The Rising Trend of Vision Based Measurement. *IEEE Instrum. Meas. Mag.* **2014**, *17*, 41–47. [CrossRef]
36. Mi, C.; Huang, Y.; Fu, C.; Zhang, Z.; Postolache, O. Vision-Based Measurement: Actualities and Developing Trends in Automated Container Terminals. *IEEE Instrum. Meas. Mag.* **2021**, *24*, 65–76. [CrossRef]
37. Zhang, Y.; Huang, Y.; Zhang, Z.; Postolache, O.; Mi, C. A Vision-Based Container Position Measuring System for ARMG. *Meas. Control* **2022**, *56*, 596–605. [CrossRef]
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
39. Lourakis, M.; Pateraki, M. Markerless Visual Tracking of a Container Crane Spreader. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 11–17 October 2021; pp. 2579–2586. [CrossRef]
40. Iris, Ç.; Lam, J.S.L. A Review of Energy Efficiency in Ports: Operational Strategies, Technologies and Energy Management Systems. *Renew. Sustain. Energy Rev.* **2019**, *112*, 170–182. [CrossRef]
41. Motec GmbH. Motec Camera Monitor Systems for Container Handling. 2023. Available online: <https://www.katze.cl/public/pdf/Motec-Brochure-Container-Handling-Port-Logistics.pdf> (accessed on 20 May 2023).
42. Axis Communications. AXIS Q1798-LE Network Camera. 2023. Available online: <https://www.axis.com/dam/public/82/dd/17/datasheet-axis-q1798-le-network-camera-de-DE-384319.pdf> (accessed on 20 May 2023).
43. Mi, C.; Huang, S.; Zhang, Y.; Zhang, Z.; Postolache, O. Design and Implementation of 3-D Measurement Method for Container Handling Target. *J. Mar. Sci. Eng.* **2022**, *10*, 1961. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.