1. Marek PISZCZEK, 2. Aleksandra KUCHARCZYK-DRAB

ORCID: 1. 0000-0003-3376-6427; 2. 0000-0002-3821-5918

DOI: 10.15199/48.2025.04.38

Design and installation of a vision measurement system for automating unloading in a warehouse using virtual technologies

Projekt i instalacja systemu pomiaru wizyjnego do automatyzacji rozładunku w magazynie z wykorzystaniem technologii wirtualnych

***Abstract**. Vision measurement systems are increasingly applied in warehouse automation. Measurement-oriented systems have a significant impact on the potential offered by these solutions. This article presents an original approach to the design process and initial implementation of such a vision-based measurement system in a facility dedicated to unloading long-length items using an overhead crane. Special attention is given to the virtual technology tools employed, which proved highly useful in both the design and installation phases.*

***Streszczenie.** Systemy wizyjne coraz częściej znajdują zastosowanie w automatyzacji magazynów. Bardzo istotny wpływ na oferowany potencjał tych rozwiązań mają systemy o charakterze pomiarowym. W artykule zaprezentowano autorskie podejście do procesu projektowania i wstępnego uruchomienia takiego wizyjnego systemu pomiarowego na hali zajmującej się rozładunkiem elementów dłużycowych przy użyciu suwnicy. Szczególną uwagę zwrócono na użyte wirtualne narzędzia, które okazały się bardzo użyteczne zarówno w procesie projektowania jak i instalacji.*

Keywords: vision measurement system, virtual technology tools, Automated Storage and Retrieval Systems.
Słowa kluczowe: wizyjny system pomiarowy, narzędzia technologii wirtualnych, automatyzacja magazynu.

Introduction

Vision measurement systems are increasingly adopted across various sectors of the economy. Machine Vision (MV) solutions, for example, contribute to enhancing quality control effectiveness, particularly within production and logistics industries **[1].** The Automated Imaging Association (AIA) defines machine vision as encompassing both industrial and non-industrial applications in which a combination of hardware and software provides operational guidance to equipment based on image capture and processing. The most common applications of vision systems in warehouses and sorting facilities include package and goods' identification. Equally significant are improvements in warehouse automation, defined as the process of moving inventory to, from, and within such facilities with minimal human intervention **[2].** The adoption of load-handling automation solutions in warehouse facilities is growing in popularity. High demand for these solutions has driven the development of new types of Automated Storage and Retrieval Systems (AS/RS) **[3].** The mentioned solutions also align with Industry 5.0 principles, emphasizing a human-centered approach by supporting human operators in technological processes. This approach contrasts with Industry 4.0, which primarily envisioned replacing human functions with advanced equipment and technical systems **[4,5].** Due to the great interest in such solutions, research is currently being conducted to improve the accuracy of positioning autonomous stations as well as the methods of their calibration **[6][7][8][9].**.The potential implementation of vision-based solutions aligns well with the needs of Warehouse Management Systems (WMS), which offer comprehensive software, hardware, services, and support for managing product movement within warehouses and optimizing space utilization **[10].** A key advantage of using vision systems in automation processes includes cost savings, as well as increased efficiency and safety in the workplace.

The general concept presented in this article pertains to a system for the automated unloading of a trailer filled with long-length elements, utilizing an overhead crane. The implementation efforts are currently underway at Sikla Polska Sp. z o. o. www.sikla.pl. At present, unloading is conducted in a "hybrid" system, meaning that the retrieval and placement of long elements from storage locations are largely automated, whereas trailer loading operations predominantly require the involvement of a skilled operator. The proposed automation concept involves the use of a dedicated vision-based measurement system, capable of achieving the required measurement accuracy (within the trailer's operational area) to the precision of a few centimetres. This system would autonomously determine 3D coordinates for the crane within its own coordinate system, thus enabling the automation of various load-handling tasks without operator intervention. Consequently, the operator's role would be primarily to oversee the correctness of these processes and to manage any "problematic" situations as they arise.

The figures below (Figures 1 and 2) provide a visualization of the main unloading aspects, including: (1) load identification by the vision system, (2) estimation of the load coordinates for retrieval, (3) transmission of data to the crane, (4) load retrieval, and (5) placement in the designated storage location.

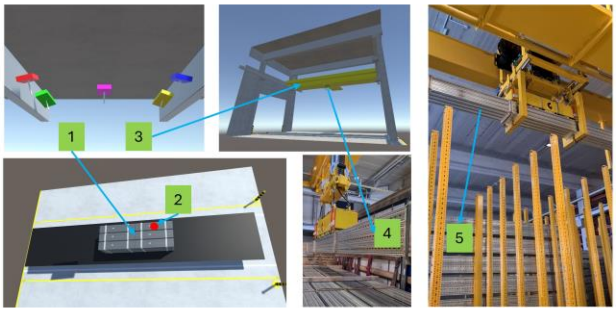

Fig. 1. Key aspects of an automated semi-trailer unloading system

Supplementary information related to unloading, including an example type of long-length element (in the form of a steel profile), methods of combining these elements into larger packages, and visual monitoring of each unloading phase, is presented in Figure 2.

The primary consideration for this defined automation process is the availability of a vision measurement system with appropriate specifications. The design and implementation of such a vision system form the central topic of this article.

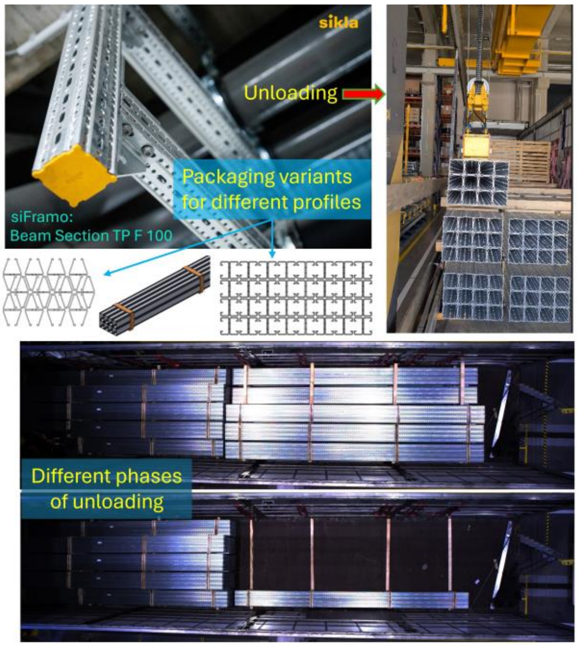

Fig. 2. Visualisation of the key elements involved in unloading a semi-trailer filled with a steel construction component

From an analytical perspective, the task of the vision measurement system is to estimate the spatial coordinates of a target point P_w (the object of interest) through triangulation of homologous 2D points p_i of this object, visible in images from at least two cameras. The goal of triangulation is to determine a solution for point P (i.e., the best estimate of its value P_w) that minimizes reprojection error across all views. The parameters K_i i (R_i, t_i) denote, respectively, the intrinsic matrix and the extrinsic matrix (rotation and translation) of the i-th cameras.

$$P_w = \mathrm{argmin}_P \sum_i |p_i - K_i[R_i|t_i]P|^2 \tag{1}$$

In a multi-camera vision system, the triangulation process is generalized by incorporating multiple views, thereby enhancing triangulation accuracy. When devices/systems with non-overlapping coordinate systems (e.g., the vision system and the overhead crane) collaborate, an additional transformation must be determined. This issue can be addressed using various techniques, including Procrustes statistical analysis, which is widely applied in computer vision and 3D shape analysis. Knowing the coordinates $X_i = \{ P_{w,j}, \ldots, P_{w,N}\}$ in the original system (i.e., the vision system) and all transformation parameters - translation t, rotation R and scaling s (determined, for instance, via singular value decomposition - SVD) - the coordinates Y_i in the second coordinate system (i.e., the crane) can be determined. Here, ϵ_i represents the errors arising from the limited accuracy of the input data upon which the transformation was based.

$$Y_i = sRX_i + t + \epsilon_i \tag{2}$$

Naturally, the general algorithm described above for determining object positions within the crane's coordinate system (based on image data from cameras) requires a series of preliminary analyses. The most critical of these include the calibration of the vision system and the optimization of camera placement within the vision system.

Vision system design - tools and their functionalities

Designing multi-camera vision measurement systems can be accomplished using various types of software tools, ranging from commercial software dedicated to computer vision (e.g., Halcon) to scientific environments (e.g., Matlab) with specialized toolboxes (Image Processing and Computer Vision Toolbox) and extending to open-source libraries (e.g., OpenCV) and general-purpose real-time simulators (e.g., Unity3D). In their work, the authors utilized the OpenCV library and Matlab environment for image processing and analysis. Unity, in turn, was primarily employed to support the vision system design process as a simulation platform capable of synthesizing high-resolution, realistic images. Besides VR solutions, AR aspects were also incorporated, particularly to support the installation process of the physical vision system. In the developed simulation-measurement setup for designing and deploying vision systems (Figure 3), the Unity3D platform interacts with the Python environment, with communication facilitated through sockets.

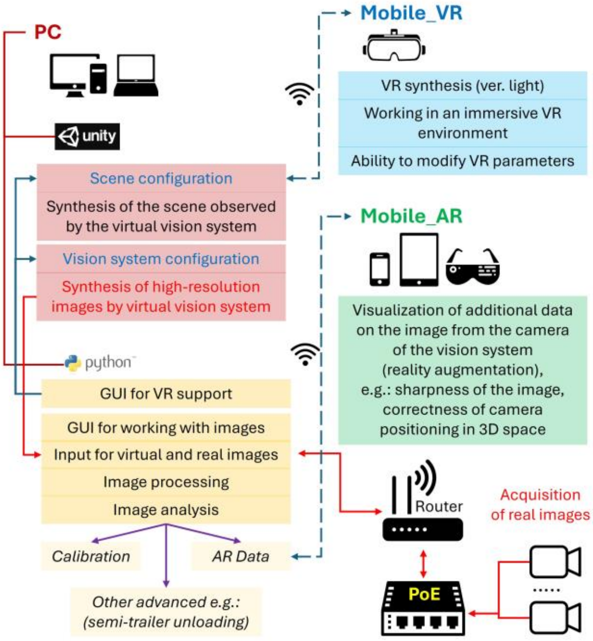

Fig. 3. Simulation and measurement stand (functionalities and components)

The primary operating mode with the simulator is executed on a desktop PC, where both the VR application (synthesizing the virtual scene and images from virtual cameras) and a complementary Python application (serving as a graphical interface and module for image processing, metadata synthesis, and analysis) are installed. In an extended variant, the virtual environment can be launched on mobile VR goggles, allowing immersion into the synthesized world of the virtual warehouse. This setting enables activities such as virtual calibration of the vision system, viewed from a first-person perspective, with interaction possible with select elements of the environment. Working within a virtual object allows not only acceleration of numerous design tasks but also facilitates verification of essential actions required for the installation and calibration of the physical vision system.

From the perspective of describing the physical vision system, two main components are critical: the observed scene and the image acquisition setup with appropriate parameters. To construct the virtual measurement scene depicted in Figure 4, data regarding the warehouse geometry (where unloading occurs) and key mechanical elements of its equipment (in this case, the overhead crane) were utilized. Based on this information, a three-dimensional model of the facility was developed. Additionally, a trailer and various load configurations were modelled in the virtual environment.

The second component developed was the virtual vision system (Figure 4). The proposed solution offers considerable flexibility, including the number of cameras supported, options for camera positioning on the modelled facility, and modification of each camera's parameters, such as sensor resolution, physical dimensions, and the focal length of the optics used.

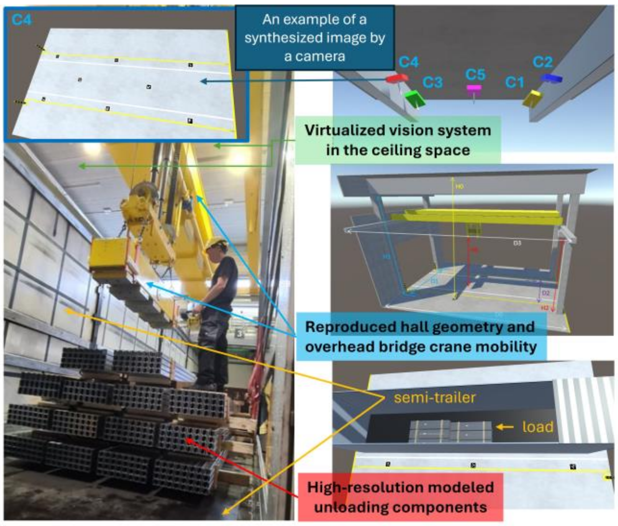

Fig. 4. Virtual representation of measurement space, key objects for unloading and vision system with synthesized image

By modeling 2D graphical markers (ArUco) and a calibration board (chessboard pattern), the system enables interaction with virtual objects commonly used alongside physical vision systems, as shown in Figure 5.

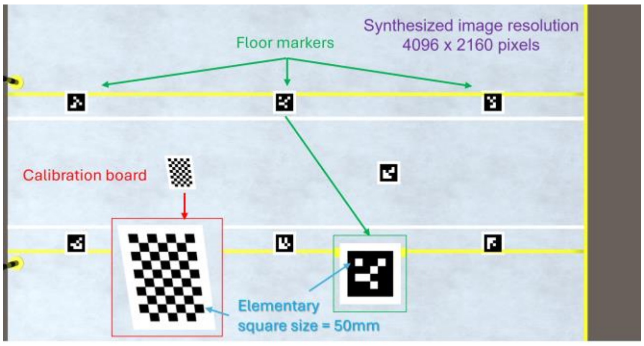

Fig. 5. Basic measurement elements implemented in VR dedicated to vision system design - C5 camera view

AR and VR technologies are often used in projects as supporting tools, which is systematically presented in **[11].** The Virtual Tools we used, developed for workstations (both VR/AR, mobile and stationary), proved to be extremely useful at various stages of system design and installation. A summary of available solutions is presented in Figure 6.

After testing different spatial camera configurations, selecting optimal parameters, and performing calibrations, the final spatial configuration of the vision system includes five cameras arranged in a square formation (6m per side) above the trailer, with cameras positioned at each corner and the center. The selected image acquisition components are the Phoenix 8.9MP camera with an IMX267 sensor and a compatible RT-V0828-MPY2 lens with an 8mm focal length.

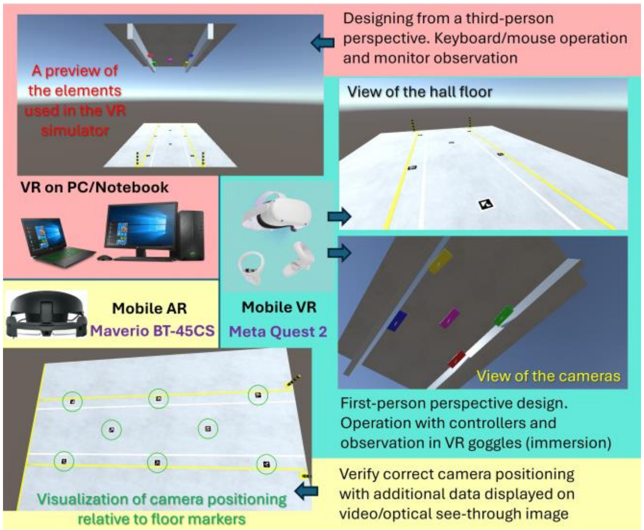

Fig. 6. Examples of the use of virtual tools

Commissioning of the vision measurement system

With the complete hardware specification and spatial configuration of the vision system components (cameras, floor markers, and calibration board) established, it was possible to carry out the installation, calibration, and validation of the proposed solution.

During the installation process:

- optimal depth of field was set for each camera,
- floor markers were securely fixed to the ground,
- cameras were physically mounted and positioned based on VR simulator data, with additional guidance from AR visualizations (Figure 8).

Key elements of the installation process are depicted in Figures 7 and 8..

Fig. 7. Commissioning of the vision system on the target measurement stand – *installation*

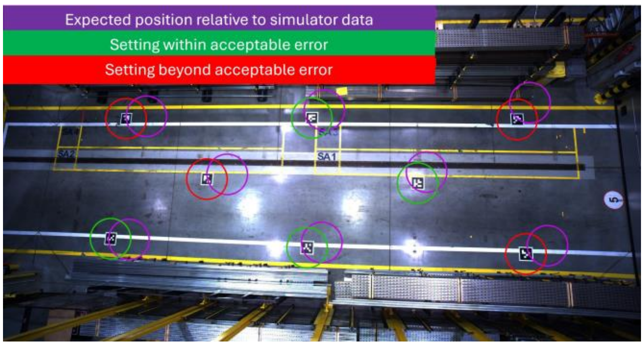

Fig. 8. The process of positioning a camera on a mounting rack with AR support – C1 camera view

Following installation, image data were collected for calibration, and essential data were obtained for coordinate transformation from the vision system's coordinate system to that of the crane. These aspects of the process are illustrated in Figure 9.

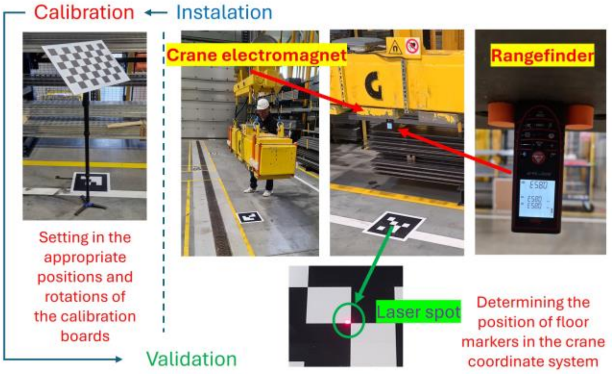

Fig. 9. Commissioning of the vision system on the target measurement stand – *calibration*

For image acquisition during calibration, the calibration board was placed at floor marker locations (with 10 variations incorporating changes in height and board rotation). Reference data for floor marker positions were sourced from the crane controller and a rangefinder attached to the crane's electromagnet (Figure 9).

To validate the installed system, in addition to using 2D floor markers, 3D markers (colored spheres) were introduced to verify the accuracy of 3D position estimation beyond the floor plane (with measurements taken at varying heights). Components used for validating the vision measurement system, along with their example spatial locations and configurations, are shown in Figure 10.

Fig. 10. Commissioning of the vision system on the target measurement stand - *validation*

Virtual and real vision system - comparison of results

A particularly time-intensive aspect of both the simulation work and the setup of the real stand was calibration. However, due to the adopted procedure, it was relatively straightforward to compare the resulting outcomes. Notably, the calibration quality was demonstrated by comparable and consistently low mean reprojection errors of 0.07 and 0.09 pixels (for simulation and real-world measurements, respectively), indicating good calibration accuracy (Figures 11 and 12).

Subsequent figures also illustrate additional validation results, including estimation errors for floor marker positions relative to the reference plane (Figure 13), accuracy of distance measurements using 3D markers (Figure 14), and accuracy of 3D marker position estimation within the crane's coordinate system (Figure 15). These results collectively indicate the robust metrological parameters of the vision system, meeting specified requirements.

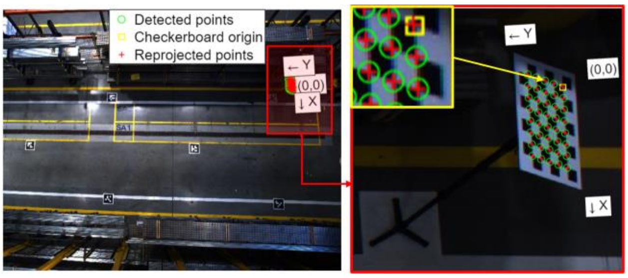

Fig. 11. Visualization of physical camera calibration - detected feature points on the board and reprojected positions of these points after calibration

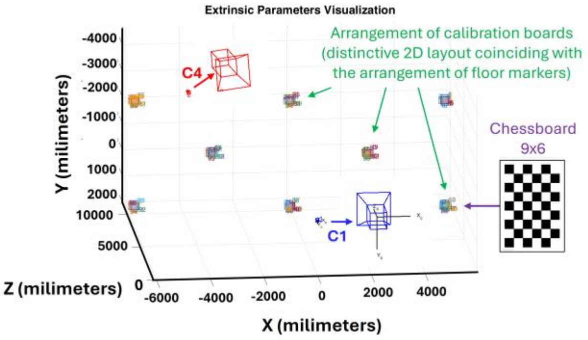

Fig. 12. Stereo cameras calibration results (C1-C4) - 3D visualization

Summary

The topics presented in this article, which combine warehouse resource management, vision-based measurement systems, and virtual technology tools, propose an innovative approach to creating smarter, more efficient, and precise solutions for automation and digital transformation. The use of vision systems (including those with measurement functionality) is a key strategy for achieving these objectives. However, the design, installation, and deployment of advanced (specialized) vision systems is complex, requiring knowledge, experience, and often substantial time investment. This complexity highlights the need for tools that allow designers and installers to streamline the process from the initial need identified by an organization seeking automation to the deployment of a dedicated solution. Examples of such tools based on virtual technologies are presented in this article.

These tools include the ability to simulate vision system operations in a virtualized environment, perform selected design tasks and tests within an immersive VR setting, and use AR equipment and synthesized content in augmented reality mode during installation. Using the proposed tools, the authors designed a measurement vision system and conducted various tests, ultimately enabling its physical implementation (installation and activation) in a long-length item unloading facility. Using VR and AR tools reduced vision system deployment from 5–12 days to 3 by streamlining installation, calibration, and testing. The reason for this is that these technologies minimise the need for physical modifications and iterative adjustments. A comparison of validation results between the virtual and physical vision systems shows significant alignment, demonstrating not only the accuracy of the proposed tools but also their utility, as they enabled preemptive identification of numerous issues that could have been time-consuming and costly to address in real-world conditions. Currently, efforts are underway to integrate the aforementioned vision measurement system with the crane used in the unloading process. Upon completion, the focus will shift to designing vision-based unloading operations, with the authors planning to enhance the current set of tools supporting the design and deployment of vision systems through VR and AR solutions.

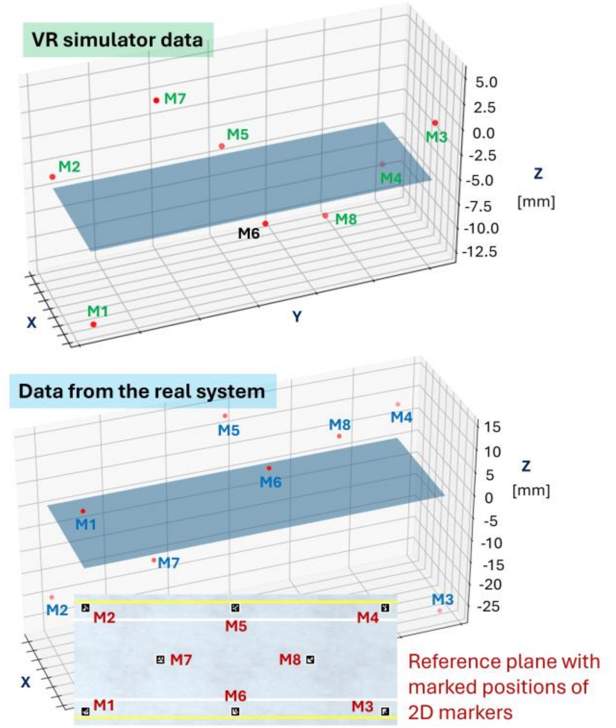

Fig. 13. Estimation of the position of floor markers relative to a reference plane for real and simulator data

***Authors**: dr hab. inż. Marek Piszczek, profesor WAT, Wojskowa Akademia Techniczna, Instytut Optoelektroniki, ul. gen. S. Kaliskiego 2, 00-908 Warszawa, E-mail: rmarek.piszczek@wat.edu.pl; mgr inż. Aleksandra Kucharczyk - Drab, Wojskowa Akademia Techniczna, Instytut Optoelektroniki, ul. gen. S. Kaliskiego 2, 00-908 Warszawa, E-mail: aleksandra.kucharczyk@wat.edu.pl.*

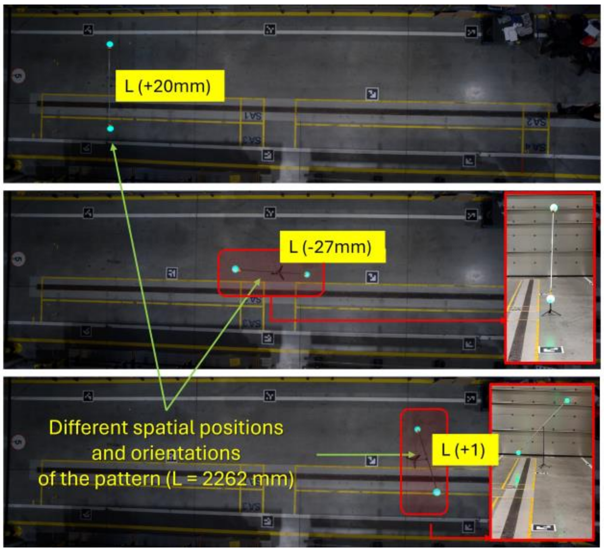

Fig. 14. Test results in the coordinate system of the vision system - verification of length constancy (based on a double set of 3D markers with fixed base) - distance estimation error (Δx) mm.

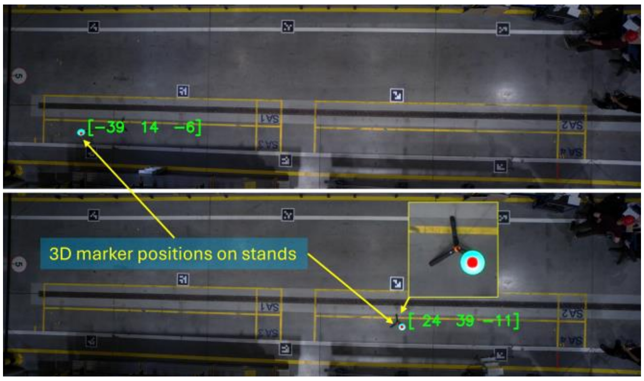

Fig. 15. Test results in the coordinate system of the overhead bridge crane (based on a single 3D marker) - position estimation error [Δx, Δy, Δz] mm

REFERENCES

[1] Reznik N., Innovative Technologies in the Logistics System, *Intellectualization of logistics and Supply Chain Management.,* vol.26, (2024), 64-74
[2] Thubis A.A., Rohman J., Intelligent Warehouse in Industry 4.0—Systematic Literature Review, *Sensors*, 23(8), (2023), 4105
[3] Majeed, H., Mohamad K., Rashid, A., The Automatic Storage and Retrieval System: An Overview, *International Journal of Computer Applications* (2019). Vol. 177. 36-43
[4] Xu X., Lu Y., Vogel-Heuser B., Wang L., *Journal of Manufacturing Systems* (2021) Vol. 61, 530-535
[5] Golovianko M., Terziyan V., Branytskyi V., Malyk D.,Industry 4.0 vs. Industry 5.0: Co-existence, Transition, or a Hybrid, *Procedia Computer Science* (2023), vol. 217, 102-113
6] Zhang Z., A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11(2000), 1330-1334
[7] Sun J., Wang P., Qin Z., Qiao H., Overview of camera calibration for computer vision, *Proceeding of the 11th World Congress on Intelligent Control and Automation*., (2014), 86-92
[8] Rameau F., Park J., Bailo O., Kweon I.S., MC-Calib: A generic and robust calibration toolbox for multi-camera systems, *Computer Vision and Image Understanding*, vol 217 (2022), 103353
[9] Cuevas-Velasquez H., et al, Global Point Cloud Registration Network for Large Transformations, ArXiv, 2403.18040 (2024)
[10] Khan M.G., Huda N.U., Zaman K.K.U., Smart Warehouse Management System: Architecture, Real-Time Implementation and Prototype Design, *Machines*, 10(2) (2022), 150
[11] Liao T., She J., How Does Virtual Reality (VR) Facilitate Design? A Review of VR Usage in Early-Stage Engineering Design, *Proceedings of the Design Society*. 3 (2023), 2115-2124.