Pose Determination by Single Image-Assisted Total Station for Timber Construction Elements

Bestimmung der Pose von Holzbauteilen mit Video-Tachymeter

Stanislav Shevchuk, Gabriel Kerekes, Volker Schwieger

In modern construction culture sustainable materials (e.g., timber) and innovative automatized construction technologies are often used. One method is the automatic assembly of timber cassettes by robotic construction cranes, e.g., spider cranes. In order to assist the construction processes, high-precision tracking and pose determination of construction components in near real time are required. The studied approach was based on the application of image-assisted total stations using open-source computer vision algorithms. The use of ArUco fiducial markers was chosen as the simplest and most reliable support for object recognition, tracking and pose determination. This method provides 6DoF object pose determination from a single image. The approach was verified using a Trimble Ri, which uses a coaxial camera and a custom camera-to-telescope conversion model. Experiments were conducted while the timber cassettes were static as well as moving. The achieved pose accuracy finally varies from the range of millimeters to centimeters, depending on the distance between the object and the total station and the angle between the optical axis and the imaged object.

Keywords: Image assisted total station, object recognition, fiducial markers, 6DoF pose determination

Die moderne Baukultur bedient sich häufig nachhaltiger Materialien (z.B. Holz) als auch innovativer automatisierter Bautechnologien. Eine Methode ist die automatische Montage von Holzkassetten mithilfe von Roboterbaukränen, z.B. Spinnenkränen. Um die Bauprozesse zu unterstützen, sind eine hochpräzise Positionsbestimmung, Lagebestimmung und Verfolgung der Bauteile in nahezu Echtzeit erforderlich. Der entwickelte Ansatz basiert auf der Nutzung eines Video-Tachymeters (in der Kombination mit Open-Source-Computer-Vision-Algorithmen). Die Verwendung von ArUco-Markern wurde als einfachstes und zuverlässigstes Hilfsmittel zur Objekterkennung, Verfolgung und Posenbestimmung gewählt. Diese Methode ermöglicht die Bestimmung der 6DoF-Position eines Objekts aus einem einzigen Bild. Der Ansatz wurde mit einem Trimble Ri verifiziert, der eine koaxiale Kamera und ein spezielles Kamera-zu-Teleskop-Konvertierungsmodell verwendet. Die Experimente wurden sowohl mit statischen als auch mit bewegten Holzkassetten durchgeführt. Die erzielte Posengenauigkeit variiert je nach Abstand zwischen dem Objekt und der Totalstation und in Abhängigkeit von dem Winkel zwischen optischer Achse und abgebildetem Objekt. Sie erreicht abschließend Werte von Millimetern bis zu Zentimetern.

Schlüsselwörter: Videotachymeter, Objekterkennung, ArUco-Marker, 6DoF-Posenbestimmung

DOI: 10.14627/avn.2026.2.1

1 INTRODUCTION

As part of the research presented earlier in the articles /Lerke & Schwieger 2021/, /Lauer et al. 2023/, the application of robotic total station networks used for the automated assembly of structures from timber cassettes by robotic cranes has been proposed. Up to now, the total station network was used to determine the pose of a spider crane vacuum gripper, tracking a reflector prism which was mounted on it. The cassettes cannot be modified with prisms for their direct tracking with robotic total stations. They have different shapes, and each time the vacuum gripper picks them up with a slightly different offset and rotation, which makes accurate pose determination impossible with this approach. The task can be accomplished using image-assisted total stations (IATSs) available on the market /Paar et al. 2021/, /Zschiesche 2023/, or custom-made systems combining a standard total station with a digital camera module /Walser 2004/, /Reiterer & Wagner 2012/. Such systems are already being successfully adapted for various engineering applications, e.g., geomonitoring /Schlüter & Zschiesche 2024/.

To implement object positioning with an image-assisted total station (IATS), three main tasks need to be performed:

- Recognition of construction components in the image. This step requires a robust and fast algorithm, capable of providing a near real time solution for recognizing the object.
- Tracking of a selected construction component using the total station's sighting axis. This requires precise conversion of image pixels to horizontal and vertical angles of the IATS.
- Six degrees of freedom (6DoF) pose determination (coordinates and orientation) of a construction component using photogrammetric and computer vision methods. For this purpose, it is necessary to define a precise conversion model from object points in the image (e.g., recognized feature points) to external cartesian coordinate system in which the total station is set. The resulting coordinates of the construction component's center and its rotation matrix must be presented in this coordinate system.

To avoid additional uncertainties from camera-to-telescope coordinates conversion, the linear offsets of the coordinate system origins and angular difference between collimation axis and camera axis must be calibrated using a model of camera-to-telescope coordinate system linkage.

This paper covers the case study using a single coaxial IATS, specifically the Trimble Ri, and shows the current possibilities and limitations of such a system used together with a down-scaled timber cassette model.

2 OBJECT TRACKING AND POSITIONING WITH IATS

2.1 IATS Camera Purposes

Robotic total stations (including IATSs) design typically involves automatic tracking and positioning of a reflector, using a number of methods /Erhart & Lienhart 2017/. There is no option for tracking a custom object (without a prism) included in an IATS which are available on the market nowadays. However, the visual user interface of some total stations allows for camera-assisted aiming, with internal pixel-to-gon conversion algorithms. Examples of such implementations can be found in /Erhart & Lienhart 2017/, /Zhang et al. 2021a/.

In this article, a concept for IATS with coaxial cameras is considered. This camera placement is based on the alignment of its optical axis with the collimation axis of the total station. In this case, the camera sensor is not placed directly on the collimation axis and uses a Porro prism to deflect the rays towards it.

For calibration of these systems, providing camera-to-telescope coordinates conversion, the following values must be taken into account /Zhang et al. 2021a/, /Vogel 2006/:

- Intrinsic parameters of the camera: focal length, position of the principal point, distortion of the lens;
- Camera sensor mounting pose: Porro prism angular displacement, image plane tilting, image plane rotation, offset between camera coordinate system origin and telescope origin.

Inaccuracies in these values lead to additional aiming errors. For these purposes a reliable mathematical model linking the coordinates of the image and the total station must be used.

2.2 Coaxial Camera Image-to-TS Geometric Model

In /Zhang et al. 2021a/ a mathamatical model of coaxial camera and total station measurements conversion is considered for reliable pixel-to-gon conversion. In order to obtain vertical and horizontal angles of point P in the total station coordinate system, several angle transformations must be done. The coordinate systems and values which are necessary to consider within the conversion as well for position determination by image are shown in *Fig. 1*. This figure extends the model given in /Zhang et al. 2021a/ adding linear displacements, but is saving its main idea. The improvements made it possible to use the model not only for pixel-to-gon conversion, but also for 3D coordinates conversion from camera coordinate system to telescope coordinate system, and then to an external (superior) one.

In *Fig. 1* the following coordinate systems and their connection parameters are shown:

- Superior (world) coordinate system O_wXYZ is a coordinate system used for outputting resulting values of coordinates and angles. Axis O_wZ points to zenith, O_wX and O_wY lie in horizontal plane.
- Total station coordinate system $OX_{TS}Y_{TS}Z_{TS}$ with origin O in the center of total station. Axis OZ_{TS} points to zenith, OX_{TS} lies on horizontal plane pointing to the horizonal angle zero direction, OY_{TS} finalizes the right-handed coordinate system. The system connected with superior coordinate system O_wXYZ via stationing parameters such as linear coordinates of total station center X_0, Y_0, Z_0 (corrected on TS height) and its zero direction azimuth H_0 around axis *OZ_TS*. Here, the horizontal plane is assumed to be without a tilt. For accurate application of the model, the compensator

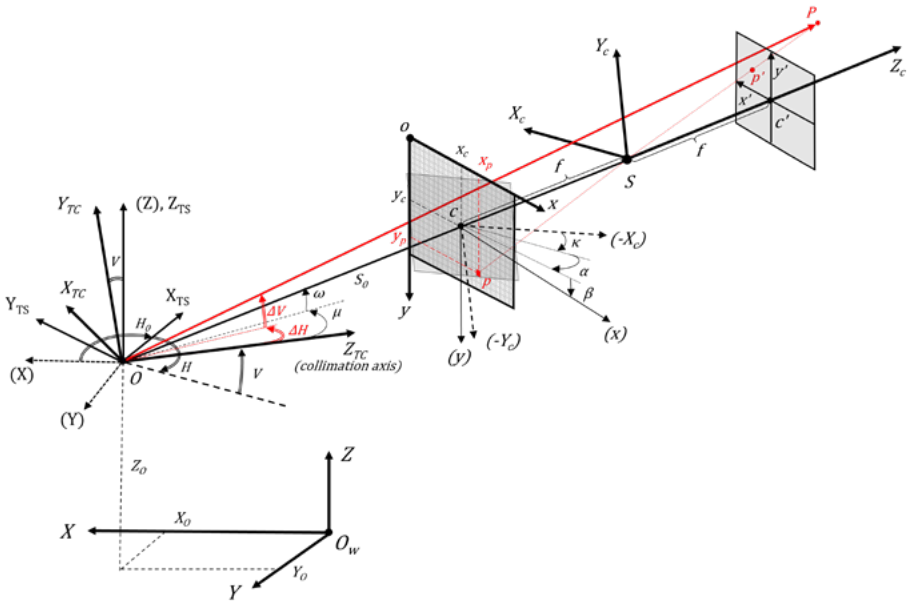

Fig. 1 | Elements of the relationship between the coordinate systems of image and total station

readings must be also taken into account. The influence of compensator readings is considered in /Walser 2004/.

- The telescope coordinate system $OX_{TC}Y_{TC}Z_{TC}$ has the same origin with total station coordinate system. OZ_{TC} is a collimation axis of the instrument which points along the direction of horizontal and vertical angles (H, V) from total station axes in horizontal and vertical planes. This coordinate system is right-handed as well.
- The coaxial camera coordinate system $SX_CY_CZ_C$ is a right-handed coordinate system with origin in central projection origin point (exposure station) S. The orientation of the coordinate system corresponds the telescope coordinate system orientation additionally turned by angles of porro prism deviation from designed position μ and ω in horizontal and vertical planes (collimation and elevation of origin). The distance between origin of central projection S and TS center O equals S_0 along axis SZ_C.
- The negative image plane coordinate system cxy is generated. The coordinate system of the coaxial camera $SX_CY_CZ_C$ must be rotated by π around axis SZ_C and moved to the image principal point c and additionally turned by the camera rotation angle κ around axis SZ_C and then the camera sensor matrix tilting the longitudinal and transverse angles α and β must be taken into account for projection in the image plane. Point P is projecting in the image plane as p. The distance between principal point c and origin of central projection S equals the focal length of the camera f taken in opposite direction along axis SZ_C.
- Finally, the origin of digital image o as top-left corner of it must be applied, so it can be shifted from c knowing camera intrinsic elements x_c, y_c in the image plane converted into pixels.
- Instead of the negative image plane a conceptual positive image plane $c'x'y'$ can be introduced. This plane corresponds to the negative image plane rotated by π around the optical axis SZ_C, with the principal point c' located at a distance equal to the foval length f on the opposite side of the projection center S along the axis SZ_C. Projection of point P on this plane will be at point p'. Within this paper this system is not used, but it is published by, e. g., /Erhart & Lienhart 2017/.

Linear displacements (coordinates X_0, Y_0, Z_0, focal length f and camera origin displacement from TS center S_0) are not used in this conversion considered in /Zhang et al. 2021a/, but is necessary to use for point positioning position determination by the image.

2.3 Pixels-to-Angles Equations for Non-Prism Object Tracking

In order to convert coordinates of point P projected in the image as $x_p\, y_p$ to horizontal and vertical angles H_P and V_P in the total station coordinate system the following conversions have to be implemented according to model /Zhang et al. 2021a/:

- Image correction for distortion, using for example radial and tangential models /Luhmann et al. 2020/:

$$\begin{aligned} x_p' &= \Delta x_p\left(1+k_1r^2+k_2r^4+\cdots+k_nr^{2n}\right) \\ &\quad +2p_1\Delta x_p\Delta y_p+p_2\left(r^2+2\Delta x_p^2\right), \\ y_p' &= \Delta y_p\left(1+k_1r^2+k_2r^4+\cdots+k_nr^{2n}\right) \\ &\quad +p_1\left(r^2+2\Delta y_p^2\right)+2p_2\Delta x_p\Delta y_p, \\ r^2 &= \Delta x_p^2+\Delta y_p^2. \end{aligned} \tag{1}$$

x_p', y_p' are coordinates of a point P in the image corrected for distortion in pixels, k_1, …, k_n are radial distortion coefficients of order from 1 to n, obtained during calibration; p_1, p_2 are tangential distortion coefficients (if matrix tilt is identified by angles α and β, they can be neglected), obtained during calibration; Δx_p, Δy_p are coordinates of the point in the image taking into account the position of the principal point c_x, c_y:

$$\begin{aligned} \Delta x_p &= x_p - c_x, \\ \Delta y_p &= y_p - c_y. \end{aligned} \tag{2}$$

Image plane rotation and tilting compensation and conversion to camera coordinate system, assuming that the focal length f adequately replaces the principal distance c according to /Luhmann et al. 2020/:

$$\begin{bmatrix} X_p^C \\ Y_p^C \\ Z_p^C \end{bmatrix} = \boldsymbol{R}_{\text{img}\to c} \cdot \begin{bmatrix} x_p' \\ y_p' \\ -f \end{bmatrix} = \boldsymbol{R}_X(\alpha) \cdot \boldsymbol{R}_Y(\beta) \cdot \boldsymbol{R}_Z(\kappa + \pi) \cdot \begin{bmatrix} x_p' \\ y_p' \\ -f \end{bmatrix}. \quad (3)$$

$\boldsymbol{R}_{\text{img}\to c}$ is a compound rotation matrix, which consist of multiplied rotation matrixes $\boldsymbol{R}_X(\alpha)$, $\boldsymbol{R}_Y(\beta)$, $\boldsymbol{R}_Z(\kappa + \pi)$ around axis *X*, *Y* and *Z* formed with rotation angles α and β and on angle $(\kappa + \pi)$ correspondently, π is added because a negative image plane is used; X_p^C, Y_p^C and Z_p^C are coordinates of point *p* in the camera coordinate system. For this conversion the depth coordinate is not required, so the point is taken as projected to the image plane using negative value of the focal length.

- Conversion to telescope coordinate system taking porro prism deviation from its designed position into account:

$$\begin{bmatrix} X_p^{\text{TC}} \\ Y_p^{\text{TC}} \\ Z_p^{\text{TC}} \end{bmatrix} = \boldsymbol{R}_{C\to\text{TC}} \cdot \begin{bmatrix} x_p^C \\ y_p^C \\ z_p^C \end{bmatrix} = \boldsymbol{R}_X(\mu) \cdot \boldsymbol{R}_Y(\omega) \cdot \begin{bmatrix} x_p^C \\ y_p^C \\ z_p^C \end{bmatrix}. \quad (4)$$

$\boldsymbol{R}_{C\to\text{TC}}$ is a compound rotation matrix, which consist of multiplied rotation matrixes $\boldsymbol{R}_X(\mu)$, $\boldsymbol{R}_Y(\omega)$ around axis *X* and *Y* formed with rotation angles μ and ω; X_p^{TC}, Y_p^{TC} and Z_p^{TC} are coordinates of point *p* in telescope coordinate system.

Conversion from telescope coordinate system to total station coordinate system:

- Conversion from telescope coordinate system to total station coordinate system:

$$\begin{bmatrix} X_p^{\text{TS}} \\ Y_p^{\text{TS}} \\ Z_p^{\text{TS}} \end{bmatrix} = \boldsymbol{R}_{\text{TC}\to\text{TS}} \cdot \begin{bmatrix} x_p^{\text{TC}} \\ y_p^{\text{TC}} \\ z_p^{\text{TC}} \end{bmatrix} = \boldsymbol{R}_Z^{\text{T}}(H + \pi/2) \cdot \boldsymbol{R}_Y^{\text{T}}(V + \pi/2) \cdot \begin{bmatrix} x_p^{\text{TC}} \\ y_p^{\text{TC}} \\ z_p^{\text{TC}} \end{bmatrix}. \quad (5)$$

$\boldsymbol{R}_{\text{TC}\to\text{TS}}$ is a compound rotation matrix, which consist of multiplied transposed rotation matrixes $\boldsymbol{R}_Z^{\text{T}}(H + \pi/2)$, $\boldsymbol{R}_Y^{\text{T}}(V + \pi/2)$ around axis *Z* and *Y* formed with measured vertical and horizontal angles *H* and *V* of collimation axis; X_p^{TS}, Y_p^{TS} and Z_p^{TS} are coordinates of point *p* in total station coordinate system.

- Finally, the vertical and horizontal angles H_P and V_P on point *P* can be calculated as:

$$H_P = H + \Delta H = \arctan\left(-\frac{Y_p^{\text{TS}}}{X_p^{\text{TS}}}\right);$$
$$V_P = V + \Delta V = \arcsin\left(\frac{Z_P^{\text{TS}}}{\sqrt{\left(x_p'\right)^2 + \left(y_p'\right)^2 + f^2}}\right). \quad (6)$$

ΔH, ΔV are differences between direction to point *P* and current orientation of the collimation axis.

- Additionally, to convert H_P to the superior coordinate system H_0 has to be added.

Using Eqs. (*1*) to (*6*) the total station can automatically track the object using the coordinates of its center in the image. Coordinates in the images in the Eqs. (*1*) to (*3*) must be converted from pixels to meters, knowing the size of the camera sensor matrix element.

This conversion provides an opportunity to calculate directions to any point in the image. If the point relates to a recognized object, it can be tracked with the IATS using this approach.

2.4 Object Pose Determination Using a Single IATS Image

The model was extended for object position determination using IATS images via close-range photogrammetry approach.

For a point *p* in the image (e.g., center of the object recognized using computer vision techniques), the central projection formulas can be used for its position conversion in the superior coordinate system by a single image using the relationship of the point in the image and in space /Luhmann et al. 2020/:

$$\boldsymbol{X}_P = \boldsymbol{X}_S + m \cdot \boldsymbol{R} \cdot \boldsymbol{x}_p' \quad (7)$$

$\boldsymbol{X}_P$ is a vector which contains coordinates of point *P* in the superior coordinate system; $\boldsymbol{X}_S$ is a vector which contains coordinates of point origin of camera central projection *S* in the superior coordinate system; *m* is a scaling coefficient, which depends on distance to object and focal length; $\boldsymbol{R}$ is an image-to-world rotation matrix (camera extrinsic rotation matrix); $\boldsymbol{x}_p'$ is a vector of coordinates of the point in the image corrected for distortion and principal point position (see Eq. (*1*)) with $-f$ as *z* value.

For known distance *D* from central projection origin to the point (e.g., using known object or marker size) the equation will be:

$$\boldsymbol{X}_P = \boldsymbol{X}_S + D \cdot \frac{\boldsymbol{R} \cdot \boldsymbol{x}_p'}{\left\| \boldsymbol{R} \cdot \boldsymbol{x}_p' \right\|}. \quad (8)$$

Or with the vectors expanded:

$$\begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix} = \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} + D \cdot \frac{\boldsymbol{R} \cdot \begin{bmatrix} x_p' \\ y_p' \\ -f \end{bmatrix}}{\left\| \boldsymbol{R} \cdot \begin{bmatrix} x_p' \\ y_p' \\ -f \end{bmatrix} \right\|}. \quad (9)$$

Applying the model of coaxial camera to TS relationship, rotation matrix $\boldsymbol{R}$ can be written as:

$$\boldsymbol{R} = \boldsymbol{R}_{\text{TS}\to\text{World}} \cdot \boldsymbol{R}_{\text{TC}\to\text{TS}} \cdot \boldsymbol{R}_{\text{C}\to\text{TC}} \cdot \boldsymbol{R}_{\text{img}\to c} \quad (10)$$

$\boldsymbol{R}_{\text{TC}\to\text{TS}}$, $\boldsymbol{R}_{\text{C}\to\text{TC}}$, $\boldsymbol{R}_{\text{img}\to c}$ are rotation matrixes from Eqs. (*3*), (*4*), and (*5*) and $\boldsymbol{R}_{\text{TS}\to\text{World}}$ is a transposed rotation matrix of total station coordinate system around axis *Z* with angle of it stationing H_0 for conversion of its orientation to the superior (world) coordinate system (*Fig. 1*):

$$\boldsymbol{R}_{\text{TS}\to\text{World}} = \boldsymbol{R}_Z^{\text{T}}(H_0). \quad (11)$$

Meanwhile point *S* coordinates can be taken as total station coordinates in superior coordinate system corrected with the distance to the camera center within the telescope along the collimation axis:

$$\begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \boldsymbol{R} \cdot \begin{bmatrix} 0 \\ 0 \\ S_0 \end{bmatrix}. \quad (12)$$

According to the model following values must be available for the calculations:

- Camera intrinsic parameters f, x_c, y_c and distortion coefficients k_1, …, k_n, p_1, p_2 in this case are provided by the total station manufacturer,
- for pixel-to-meter conversion the size of sensor and its resolution must be known (from total station camera specifications),
- image tilting and rotation α, β, κ and porro prism deviation elements angles μ and ω and distance from total station center to camera projection origin S_0 (from manufacturer only),
- angles H and V of total station collimation axis orientation (total station actual measurements),

total station stationing values (coordinates of station X_0, Y_0, Z_0 and initial azimuth H_0),

- Coordinates of the point p in the image, and distance D to the corresponding real point P from the camera projection origin S.

The distance to the object D can be determined by the ratio of the object size in the image and in space. In addition, to determine the orientation of an object in three-dimensional space from a single image, a homography based on analysing the perspective of the object projection can be used to provide 6DoF pose determination by a single image /Hartley & Zisserman 2003/, /Wubben et al. 2019/, /Manawadu & Park 2024/, /Xu et al. 2009/.

Additionally, it is important to mention, that the quality of the image should allow to determine the measured points unambiguously: no blurring, sufficient sharpness and contrast should be ensured. For these purposes, the construction site must be illuminated and the lens systems of the total stations must be correctly adjusted according to the distance to the object (or corrected during tracking).

3 USING FIDUCIAL MARKERS FOR OBJECT RECOGNITION, POSITIONING AND TRACKING

For positioning and tracking of an object without using reflectors it is necessary to recognize its specific points in the image. If the recognized points' coordinates are known in the coordinate system of the object, it is possible to reconstruct its 6DoF pose using these coordinates.

Recognizing objects in an image is one of the central tasks of computer vision, so there is a large number of algorithms which solve this problem. Such algorithms can be roughly be divided into classical algorithms such as feature-matching discriminators, fiducial markers etc., and algorithms based on machine learning/deep learning /Diwakar & Raj 2022/, /Trigka & Dritsas 2025/, /Padilla et al. 2021/.

Nowadays, there has been serious progress in the improvement and application of deep-learning algorithms: they can be effectively applied in near real time and estimate the position and orientation of objects /Padilla et al. 2021/, /Song et al. 2017/, /Kodnongbua et al. 2023/, /Mao et al. 2025/. However, these algorithms are complex and not transparent as classical ones /Zhang et al. 2021b/. In addition, the similarity of timber construction components to each other, the necessity to create a large training data set and the requirements for computing power make it difficult to implement such algorithms in the robotized assembly technology.

At the same time, given the simple shape of building elements and the need for their unambiguous identification and pose determination in near real time, it was decided to use the simplest object recognition algorithms based on the application of fiducial markers.

There are several kinds of fiducial markers /Beligiray et al. 2019/, /Yu et al. 2021/, /DeGol et al. 2017/, /Ulrich et al. 2023/, including those available in open source libraries. Among them, the ArUco markers /Zakiev et al. 2019/, /Poroykov et al. 2020/ were chosen. This algorithm has the following advantages:

- unambiguous identification of a marker by its encoded number;
- high precise rate (close to 100 %) of marker recognition /Kalaitzakis et al. 2021/, /Zakiev et al. 2019/, /Wieczorek et al. 2020/;
- high performance allows the algorithm to be used in near real time;
- low sensitivity to marker tilt and complete indifference to its rotation /Zakiev et al. 2019/, /Poroykov et al. 2020/, /Wieczorek et al. 2020/;
- simplicity of temporary fixation of the marker on a timber construction component;
- the algorithm provides the coordinates of the corners for each recognized marker in the image with subpixel accuracy;
- if some markers are placed on common surface their poses can be calculated together using additional geometric constraints, if their relevant positions are well-known;
- implementation of the ArUco in the OpenCV library free of charge, the community keeps improving the algorithm /Sardar 2025/, /Oščádal et al. 2020/, /Berral-Soler et al. 2024/.

However, in order to use this algorithm for pose determination of a construction component within the developed application, the following requirements must be met:

- each marker must be placed on a flat surface;
- for transition from coordinates of corner points of markers to the position of the object, their object coordinates should be known in the object coordinate system as well;
- brightness, contrast and sharpness of the image must be sufficient for stable marker recognition;
- the distance and the physical size of markers together must provide their visible size in the image, sufficient for recognition (e. g., for marker with code 7 × 7 sections it's size on an image must be greater than 7 × 7 pixels at least 2–3 times);
- for continuous tracking of objects, markers should be attached to the edges of the object, ensuring visibility of at least one marker from most of possible directions.

After each marker i a recognition, getting its corners coordinates in the image (x_i, y_i), the coordinates X_i, Y_i, Z_i and rotation matrix $\boldsymbol{R}_i$ for each marker in the camera coordinate system must be determined analytically using Eq. (*9*). The equations require the distances to each marker, which can be roughly calculated knowing the size of the marker in the image and in real life.

For this purposes an algorithm based on perspective-and-point (PnP) method can be used. The chosen implementation in OpenCV is considered in /Hartley & Zisserman 2003/, /Lepetit et al. 2008/, /Gao et al. 2003/. The resulting position and rotation accuracy depends on the size of the pixels and contain recognition ambiguities, which makes single-image approach less accurate than stereopair processing. Moreover, within the same approach the angular orientation for each marker can be evaluated from corner position analysis, knowing the form of the marker and choosing the most relevant orientation for their current projection in the image. In case of subpixel accuracy of point recognition and proper object size in the image, the accuracy better than 1 gon can be expected /Luhmann 2009/, /Dammert et al. 2025/.

However, OpenCV evaluates the pose and orientation in the camera coordinate system only. As a result $(X_i, Y_i, Z_i)_C$ and rotation matrix $(\boldsymbol{R}_i)_C$ must be converted to the superior coordinate system. For this the Eqs. (*4*), (*5*) and (*9*) can be summarized as:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} + \boldsymbol{R}_{\mathrm{TS}\to\mathrm{World}} \cdot \boldsymbol{R}_{\mathrm{TC}\to\mathrm{TS}} \cdot \boldsymbol{R}_{C\to\mathrm{TC}} \cdot \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}_C . \tag{13}$$

$$\boldsymbol{R}_i = \boldsymbol{R}_{\mathrm{TS}\to\mathrm{World}} \cdot \boldsymbol{R}_{\mathrm{TC}\to\mathrm{TS}} \cdot \boldsymbol{R}_{C\to\mathrm{TC}} \cdot (\boldsymbol{R}_i)_C .$$

(X_i, Y_i, Z_i) are marker *i* center coordinates in the superior coordinate system; $\boldsymbol{R}_i$ is the rotation matrix of matrix *i* in the superior coordinate system.

Using the coordinates and orientation matrixes for each marker *i* the coordinates of corner $j \in [1 \cdots 4]$ for each marker $(X_{i,j}, Y_{i,j}, Z_{i,j})$ can be calculated:

$$\begin{bmatrix} X_{i,j} \\ Y_{i,j} \\ Z_{i,j} \end{bmatrix} = \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} + \boldsymbol{R}_i \cdot \begin{bmatrix} x_j \\ y_j \\ 0 \end{bmatrix}_m . \tag{14}$$

$(x_j, y_j)_m$ are coordinates of each marker corner in the marker coordinate system: $[-l/2, l/2]$, $[l/2, l/2]$, $[l/2, -l/2]$, $[-l/2, -l/2]$ where *l* is marker side length in real world.

This approach (using isolated pose analysis for each marker) gives lower accuracy than an alternative one, which was used in current study. Taking into account the flat shape of timber construction components, these additional geometrical constraints were used: markers were grouped by flat surfaces, using the coplanarity for iterative PnP approach, implemented in OpenCV /Lepetit et al. 2008/, /Gao et al. 2003/.

Knowing the corners *j* of each marker *i* position, reduction from marker position to the object center $(X_{obj}, Y_{obj}, Z_{obj})$ and its orientation $\boldsymbol{R}_{obj}$ can be done solving the equations:

$$\begin{bmatrix} X_{i,j} \\ Y_{i,j} \\ Z_{i,j} \end{bmatrix} = \boldsymbol{R}_{obj} \begin{bmatrix} X_{i,j} \\ Y_{i,j} \\ Z_{i,j} \end{bmatrix}_{CAD} + \begin{bmatrix} x_{obj} \\ y_{obj} \\ z_{obj} \end{bmatrix} . \tag{15}$$

$(X_i, Y_i, Z_i)_{CAD}$ are coordinates of the ArUco marker (or each corner of it) in object coordinate system. The equations can be solved using singular value decomposition (SVD) /Umeyama 1991/ or OpenCV PnP algorithms /Lepetit et al. 2008/. As a result $(X_{obj}, Y_{obj}, Z_{obj})$ and its orientation $\boldsymbol{R}_{obj}$ can be calculated.

4 EXPERIMENTS

4.1 Parameters of Used IATS

Within the research this model was applied to the Trimble Ri robotic total station. The instrument specifications related to its coaxial camera are summarized in *Tab. 1* and the scheme of camera installation in the head of the total station is shown in *Fig. 2* /Bellmann 2023/.

All the optical sensors, including electronic distance meter (EDM) are functioning using a common compound lens system. The camera uses the same lens system as a telescope, which reduces the field of view in case of changing lens position (e.g., while distance measurement, etc.). Resolution 1920 × 1080 pixels and high frame rate (15 frames per second) make the camera usable for near real time object position determination and tracking.

The total station messages contain the necessary information, including camera calibration data and camera setup parameters, given by the manufacturer. The parameters are actual for basic lens position only (focused on infinite, no zoom) which is about 8.32 mm.

Due to instable lens position, camera sensor matrix was generated by the manufacturer values. Additionally TS camera distortion was defined with the Agisoft Metashape calibration tool or OpenCV algorithm /Abeho et al. 2024/.

Parameter	Value
Lens system	Continuous focus
Crosshair type	Digital, superimposed
Field of view	2 gon to 12 gon
Resolution of video stream	960 × 540 or 1920 × 1080 pixel

Tab. 1 | Trimble Ri Camera specifications /Trimble 2022/

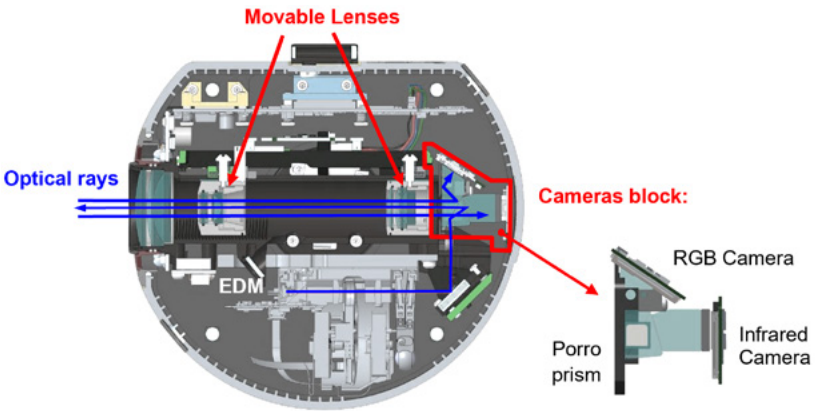

Fig. 2 | Trimble Ri camera setup (adapted after /Bellmann 2023/)

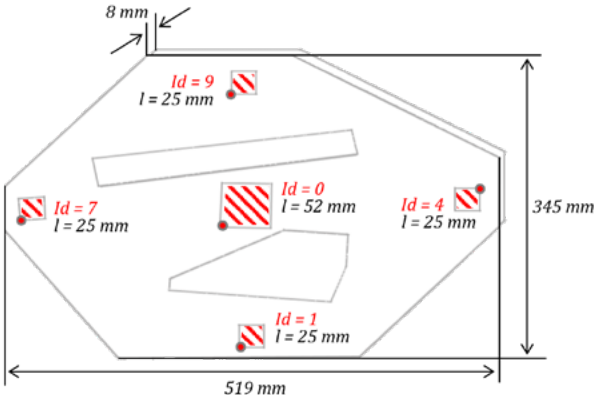

Fig. 3 | Test-object for the experiments with shown ArUco markers used for pose determination

4.2 Software and Test-object

The method considered in the article was implemented as a custom Python application, utilizing OpenCV libraries for markers recognition and image processing. The application contains an interface module for connection with Trimble Ri IATS.

As an object for static and kinematic tests instead of the real timber cassettes, a down-scaled (approx. 5 times smaller as real timber cassette) timber test-object with ArUco markers was used, shown on *Fig. 3*. A corresponding CAD model was prepared as a digital model of the object.

In the figure for each marker *id* is an ArUco identification number; *l* is side length; red circles show first corner of the marker ($j = 1$). Coordinates of ArUco marker corners were measured manually with ruler, their positions can be assumed as measured with a few (1 – 2) millimeter accuracy.

4.3 Static experiment

4.3.1 Concept of the Experiment

Evaluation 6DoF pose of an object with the ArUco-based algorithm using Eqs. (*1*) to (*15*) has been tested for static object positions as well as for near real time kinematic tracking mode.

Static experiment included measurements for different distances and different angular orientation of the object. For each pose the

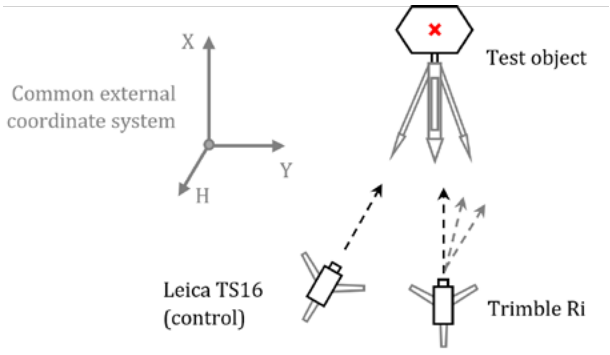

Fig. 4 | Experimental setup for static aiming and pose determination by ArUco markers evaluation

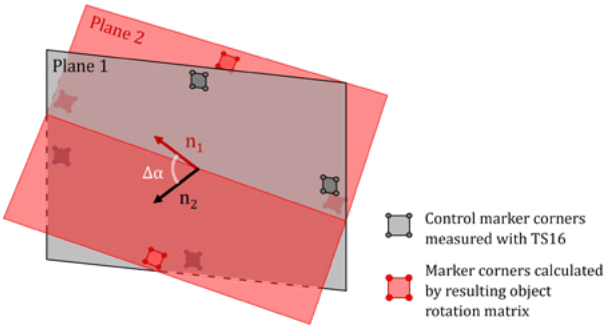

Fig. 5 | Orientation accuracy evaluation as an angle $\Delta\alpha$ between two planes

ground truth control measurements were performed with a Leica TS16 total station which was stationed in the same coordinate system, *Fig. 4*.

In order to evaluate orientation as well each pose was also analyzed using the coordinates of marker corner points measured with Leica TS16. Due to multiple representations of orientation by the matrix as well as Euler angles, the accuracy evaluation cannot be done with the direct comparison of their values. In order to do it, for each measurement two planes were calculated: one (reference) plane taken from TS16 measurements of all marker corners (which used as a ground truth) and the second calculated from all visible markers' corners positions calculated with a resulting rotation matrix for each frame. The angle between the normal vectors of the two planes were taken as the resulting orientation difference, as shown in *Fig. 5*.

4.3.2 Distance Influence Evaluation

The first part of the experiment included 5 poses, with a distance of approximately 2.9 m to 8.3 m. The main goal was to evaluate how critical the influence of the distance is both on object recognition and the accuracy of its obtained 6DoF pose. The object orientation for this part of the experiment was close to the normal of the line of sight of the total station.

For all the images the total station has been used in a common lens position with a focal length 8.3 mm. This lens position provided the maximal field of view. The examples of processed images are shown on *Fig. 6*. For each pose the first image was made with the object placed directly in the center of the image, for the others the images are closer to image corners. The results of the analysis are given in *Tab. 2*.

For the distances longer than 5.7 meters the recognition rate (when at least one ArUco marker was recognized) was less than 50 % (2 frames of 6 for pose 1 and 3 frames of 7 for pose 2). All the other poses from 3 and further had at least one recognized marker for each frame. This is caused by the small size of the markers: e. g., for the distances about 8 meters the biggest marker size was smaller than 22 × 22 pixels, when the others were smaller than 11 × 11 pixels each which made them unrecognizable.

The aiming accuracy (ΔH and ΔV values) by flat image was about 0.1 gon which satisfied the aiming requirements (keeping the object in the field of view). The coordinates of the object center varied from pose to pose with a high dependence on the distance. The main

Pose	Frame	Approximate distance	Recognized markers	Corner points	Aiming		Pose determination			
					ΔH	ΔV	ΔX	ΔY	ΔZ	$\Delta \kappa$
1	1	8.3 m	1 of 5	4	–0.01 gon	0.07 gon	191.8 mm	–114.7 mm	–20 mm	21.3 gon
	2		1 of 5	4	0.01 gon	0.12 gon	197.7 mm	–119.5 mm	–14.1 mm	19.1 gon
			RMS		**0.01 gon**	**0.1 gon**	**194.7 mm**	**117.1 mm**	**17.3 mm**	**20.2 gon**
2	1	7.1 m	1 of 5	4	–0.04 gon	–0.11 gon	–110.7 mm	113.9 mm	2.6 mm	14.8 gon
	2		1 of 5	4	–0.06 gon	–0.10 gon	–185.1 mm	159.8 mm	–2.2 mm	17.2 gon
	3		1 of 5	4	–0.10 gon	–0.14 gon	–210.7 mm	169.6 mm	11.4 mm	17.9 gon
			RMS		**0.07 gon**	**0.12 gon**	**174.1 mm**	**149.7 mm**	**6.9 mm**	**16.7 gon**
3	1	5.7 m	3 of 5	12	0.06 gon	0.10 gon	–27.3 mm	71.1 mm	15.3 mm	5.6 gon
	2		3 of 5	12	0.07 gon	0.11 gon	–46.5 mm	86.2 mm	16.9 mm	7.8 gon
	3		5 of 5	20	0.06 gon	0.12 gon	8.0 mm	43.7 mm	12.8 mm	2.2 gon
	4		1 of 5	4	0.16 gon	0.21 gon	–178.8 mm	167.4 mm	24.0 mm	57.8 gon
	5		2 of 5	8	0.12 gon	0.11 gon	53.5 mm	3.4 mm	9.1 mm	36.8 gon
	6		3 of 5	12	0.07 gon	0.09 gon	31.7 mm	24.3 mm	8.8 mm	3.7 gon
	7		3 of 5	12	0.04 gon	0.10 gon	–50.0 mm	91.5 mm	16.9 mm	1.7 gon
	8		4 of 5	16	0.07 gon	0.16 gon	16.8 mm	34.6 mm	14.9 mm	8.6 gon
			RMS		**0.09 gon**	**0.13 gon**	**72.1 mm**	**81.1 mm**	**15.5 mm**	**24.7 gon**
4	1	4.1 m	5 of 5	20	0.06 gon	0.10 gon	39.2 mm	2.2 mm	4.3 mm	1.3 gon
	2		5 of 5	20	0.05 gon	0.09 gon	39.4 mm	1.6 mm	4.2 mm	1.2 gon
	3		5 of 5	20	0.14 gon	0.10 gon	32.3 mm	1.2 mm	4.6 mm	1.1 gon
	4		1 of 5	4	0.03 gon	0.02 gon	19.5 mm	22.1 mm	1.2 mm	3.4 gon
	5		3 of 5	12	0 gon	0.10 gon	45.9 mm	0.9 mm	4.5 mm	8.0 gon
	6		4 of 5	16	0.03 gon	0.15 gon	41.3 mm	2.1 mm	6.8 mm	1.8 gon
	7		4 of 5	16	0.08 gon	0.06 gon	32.8 mm	5.0 mm	1.4 mm	1.3 gon
	8		3 of 5	12	0.03 gon	0.06 gon	54.8 mm	–9.9 mm	0.7 mm	3.4 gon
	9		5 of 5	20	0.08 gon	0.15 gon	43.2 mm	–4.3 mm	6.7 mm	1.6 gon
			RMS		**0.07 gon**	**0.10 gon**	**39.8 mm**	**8.4 mm**	**4.4 mm**	**3.3 gon**
5	1	2.9 m	4 of 5	16	0.07 gon	0.08 gon	8.8 mm	–0.4 mm	–2.4 mm	3.2 gon
	2		5 of 5	20	0.07 gon	0.09 gon	8.3 mm	0.0 mm	–2.6 mm	1.6 gon
	3		5 of 5	20	0.06 gon	0.07 gon	7.8 mm	0.7 mm	–2.5 mm	2.2 gon
	4		5 of 5	20	0.02 gon	0.16 gon	7.6 mm	4.4 mm	0.4 mm	1.6 gon
	5		5 of 5	20	0.12 gon	0.15 gon	5.5 mm	–0.1 mm	0.5 mm	2.1 gon
	6		4 of 5	16	0.06 gon	0.01 gon	4.9 mm	4.1 mm	5.2 mm	13.3 gon
	7		5 of 5	20	0.03 gon	0.13 gon	7.3 mm	4.5 mm	0.2 mm	2.3 gon
	8		5 of 5	20	0.11 gon	0.08 gon	6.2 mm	–0.6 mm	–3.5 mm	3.0 gon
	9		5 of 5	20	0.06 gon	0.06 gon	8.6 mm	0.2 mm	–3.8 mm	3.4 gon
	10		5 of 5	20	0.09 gon	0.11 gon	8.0 mm	–1.0 mm	–2.0 mm	1.4 gon
			RMS		**0.08 gon**	**0.10 gon**	**7.4 mm**	**2.4 mm**	**2.8 mm**	**4.8 gon**

Tab. 2 | Aiming and object 6DoF pose determination accuracy for distances

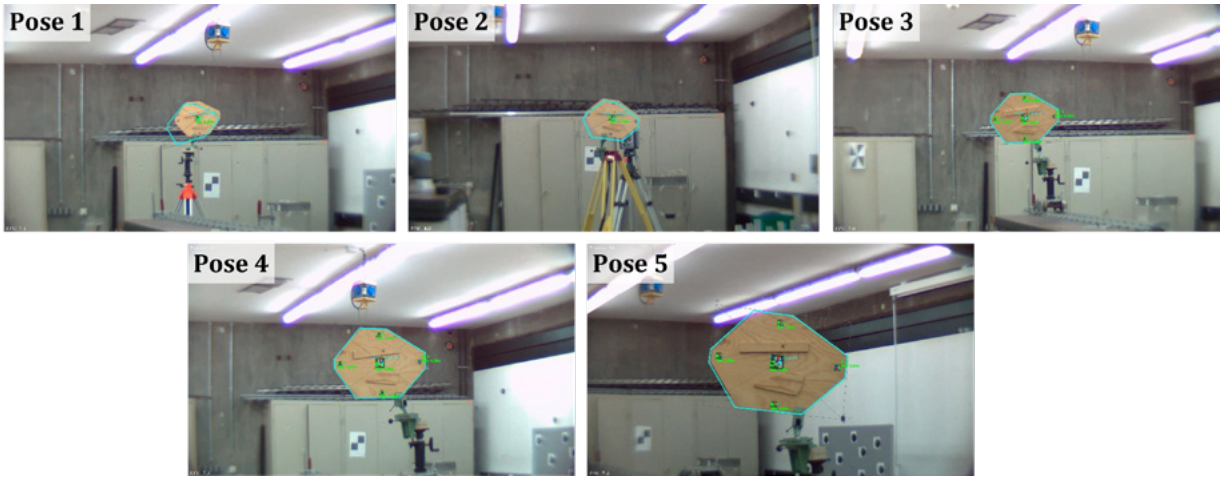

Fig. 6 | The examples of the processed images taken from different distances

influence is caused by the depth (coordinate *Z* of the image coordinate system) uncertainty. Since the coordinate system of the image was converted to the superior coordinate system this uncertainty distributed to all the axes. The vertical angles for each frame were close to horizon, which made the influence of image depth distributed mostly between *X* and *Y* coordinates of the superior coordinate system. The larger distance gave the lower accuracy (up to decimeters) and the smaller ones gave several milimeters as it was expected.

Orientation determination accuracy was highly dependent on distance. In case of multiple markers on the same surface were available the accuracy (without taking outliers into account) was better as 3 gons. For longer distances and less markers available the accuracy reduces significantly.

4.3.3 Orientation Influence

The second part of the experiment included five poses placed on a distance about 2.8 meters from the total station with different object angular orientations, as shown on *Fig. 7*. The results are demonstrated in *Tab. 3*.

Different orientations influenced the pose determination differently. For example, the worse results were obtained for pose 9 (both for orientation and position), which had an extreme orientation. For all the poses the level of 3–4 cm of positioning and 0.7–2.5 gon of orientation errors were reached.

For the coordinates which had insignificant depth influence (axis *Z*) the accuracy was much higher, about a few millimeters.

The object orientation did not have visible influence on aiming accuracy, keeping it about 0.1 gon (excepting the outliers).

In *Tab. 3* it is also visible that the best solutions were obtained predominantly for the first frames of each pose. The measurements were always made for objects fixed close to the center of the image. The images for this lens position (which provided the maximal field of view) had different quality closer to edges and in the center, which led to lower recognition probability for the markers which were situated in the blurred zones with erroneous point-to-plane solution, giving outlying pose determination errors, as is demonstrated in *Fig. 8*.

One of the possible opportunities to reduce these effects and increase positioning accuracy was using the different lens positions (optical zoom), which significantly decreases the field of view, making this approach more difficult to use for aiming.

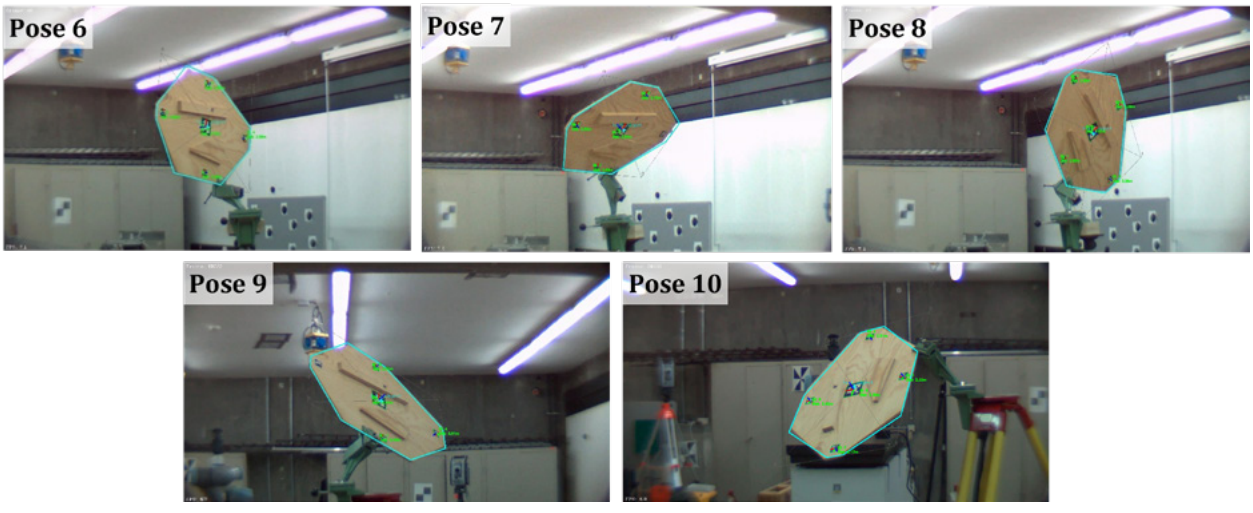

Fig. 7 | The examples of the processed images with different object orientation

Pose	Frame	Recognized markers	Corner points	Aiming		Pose determination			
				ΔH	ΔV	ΔX	ΔY	ΔZ	$\Delta\kappa$
6	1	5 of 5	20	0.07 gon	0.06 gon	17.7 mm	–10.0 mm	–4.8 mm	0.6 gon
	2	5 of 5	20	0.07 gon	0.06 gon	17.4 mm	–9.6 mm	–4.7 mm	0.7 gon
	3	5 of 5	20	0.07 gon	0.06 gon	17.9 mm	–10.2 mm	–4.8 mm	0.7 gon
	4	3 of 5	12	0.07 gon	0.01 gon	19.2 mm	–11.5 mm	–7.7 mm	0.6 gon
	5	4 of 5	16	0.08 gon	0.17 gon	16.8 mm	–9.1 mm	–0.5 mm	0.7 gon
	6	4 of 5	16	0.15 gon	0.08 gon	13.6 mm	–9.7 mm	–4.4 mm	0.6 gon
	7	4 of 5	16	0.03 gon	0.13 gon	17.0 mm	–6.3 mm	–1.9 mm	0.5 gon
	8	3 of 5	12	0.96 gon	0.03 gon	–10.1 mm	–37.8 mm	–7.0 mm	0.9 gon
	9	4 of 5	16	0.07 gon	0.07 gon	17.5 mm	–9.8 mm	–4.5 mm	0.7 gon
	10	3 of 5	12	0.08 gon	0.16 gon	17.3 mm	–9.5 mm	–0.7 mm	0.7 gon
		RMS		**0.31 gon**	**0.09 gon**	**16.6 mm**	**15 mm**	**4.7 mm**	**0.6 gon**
7	1	5 of 5	20	0.09 gon	0.07 gon	9.7 mm	–4.1 mm	–2.8 mm	0.9 gon
	2	4 of 5	16	0.08 gon	0.07 gon	10.7 mm	–5.1 mm	–3.0 mm	0.9 gon
	3	4 of 5	16	0.09 gon	0.06 gon	11.3 mm	–5.7 mm	–2.8 mm	1.0 gon
	4	1 of 5	4	–0.33 gon	–0.47 gon	–55.9 mm	94.6 mm	11.7 mm	1.7 gon
	5	1 of 5	4	0.15 gon	0.13 gon	0.2 mm	4.8 mm	0.1 mm	1.0 gon
	6	4 of 5	16	0.07 gon	0.14 gon	13.3 mm	–7.8 mm	0.0 mm	0.9 gon
	7	2 of 5	8	0.08 gon	0.03 gon	5.9 mm	2.1 mm	–4.4 mm	1.1 gon
	8	2 of 5	8	0.08 gon	0.08 gon	7.3 mm	2.0 mm	–5.4 mm	1.1 gon
	9	2 of 5	8	0.12 gon	0.18 gon	15.5 mm	–13.0 mm	0.4 mm	1.0 gon
		RMS		**0.14 gon**	**0.18 gon**	**21.0 mm**	**32.1 mm**	**4.8 mm**	**1.1 gon**
8	1	4 of 5	16	0.08 gon	0.08 gon	19.5 mm	–14.6 mm	–0.5 mm	0.6 gon
	2	4 of 5	16	0.08 gon	0.08 gon	20.5 mm	–15.9 mm	–0.8 mm	0.7 gon
	3	4 of 5	16	0.08 gon	0.08 gon	20.7 mm	–16.4 mm	–0.8 mm	0.6 gon
	4	2 of 5	8	0.07 gon	0.16 gon	19.1 mm	–13.6 mm	2.2 mm	0.8 gon
	5	2 of 5	8	0.15 gon	0.05 gon	19.2 mm	–18.3 mm	–2.5 mm	0.8 gon
	6	4 of 5	16	0.01 gon	0.04 gon	17.2 mm	–5.9 mm	–1.3 mm	0.9 gon
	7	3 of 5	12	0.01 gon	0.10 gon	14.8 mm	–1.5 mm	2.2 mm	0.6 gon
	8	2 of 5	8	0.15 gon	0.06 gon	14.2 mm	–11.8 mm	–1.5 mm	0.5 gon
	9	2 of 5	8	0.02 gon	0.06 gon	6.7 mm	9.0 mm	1.8 mm	0.8 gon
	10	4 of 5	16	0.12 gon	0.04 gon	21.7 mm	–19.5 mm	–3.0 mm	0.7 gon
		RMS		**0.09 gon**	**0.08 gon**	**17.9 mm**	**13.8 mm**	**1.8 mm**	**0.7 gon**
9	1	4 of 5	16	–0.01 gon	–0.05 gon	0.8 mm	17.1 mm	1.3 mm	2.5 gon
	2	3 of 5	12	–0.05 gon	–0.16 gon	–4.2 mm	21.4 mm	–2.1 mm	2.2 gon
	3	3 of 5	12	–0.12 gon	–0.29 gon	1.2 mm	19.0 mm	–5.5 mm	2.3 gon
	4	3 of 5	12	0.01 gon	0.07 gon	94.0 mm	–26.7 mm	–4.8 mm	2.7 gon
	5	4 of 5	16	–0.02 gon	–0.18 gon	–2.0 mm	18.8 mm	–3.3 mm	2.2 gon
		RMS		**0.06 gon**	**0.18 gon**	**42.1 mm**	**20.9 mm**	**3.7 mm**	**2.4 gon**
10	1	5 of 5	20	0.07 gon	0.09 gon	7.0 mm	0.2 mm	–2.0 mm	0.9 gon
	2	5 of 5	20	0.07 gon	0.09 gon	19.0 mm	–11.7 mm	–3.2 mm	0.9 gon
	3	5 of 5	20	0.07 gon	0.09 gon	19.9 mm	–12.7 mm	–3.3 mm	0.9 gon
	4	4 of 5	16	0.01 gon	0.12 gon	19.5 mm	–12.2 mm	–3.2 mm	0.8 gon
	5	4 of 5	16	0.11 gon	0.06 gon	19.1 mm	–7.9 mm	–1.1 mm	0.6 gon
	6	4 of 5	16	0.07 gon	0.04 gon	20.8 mm	–16.8 mm	–5.0 mm	0.9 gon
	7	3 of 5	12	0.09 gon	0.16 gon	21.9 mm	–15.0 mm	–5.5 mm	0.8 gon
	8	3 of 5	12	0.14 gon	0.09 gon	18.4 mm	–11.9 mm	–0.6 mm	0.8 gon
	9	4 of 5	16	0.04 gon	0.15 gon	15.2 mm	–11.2 mm	–3.8 mm	0.6 gon
		RMS		**0.08 gon**	**0.10 gon**	**18.4 mm**	**12.0 mm**	**3.4 mm**	**0.8 gon**

Tab. 3 | Aiming and 6DoF pose determination accuracy for different orientation angles

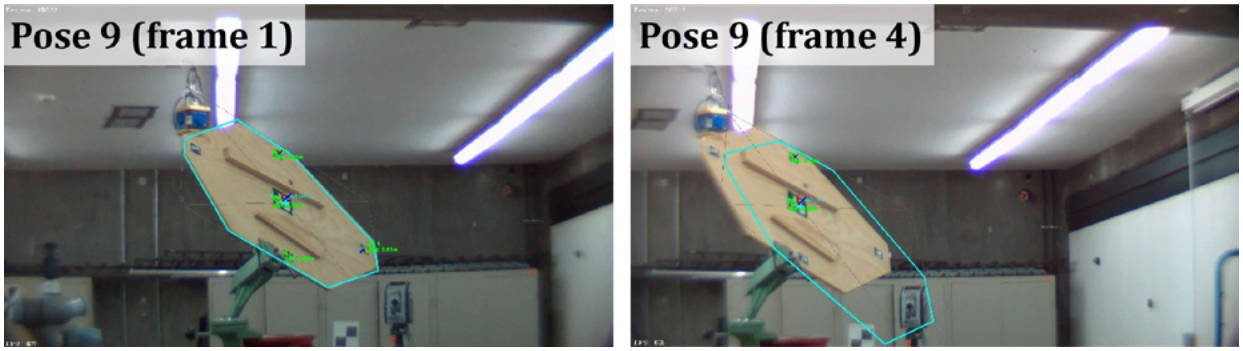

Fig. 8 | Object recognition issues close to the edges of the image with initial lens position

4.3.4 Lens Position Changing (Optical Zoom) Influence

The tests were repeated for the same poses and distances with the total station automatic optical zoom (which changed the lens position in the telescope). Trimble Ri automatically calculates camera intrinsic elements and value of S_0, giving them in calibration messages, which makes possible to use the images for any possible lens position.

In *Fig. 9* the frames for some poses with optical zoom are demonstrated.

The results of accuracy evaluation for the zoomed frames are shown in *Tab. 4*.

All the zoomed images had better image quality without blurred parts. The size of markers provided better recognition and geometry for pose analysis. Recognition rate as well as the resulting accuracy significantly increased, reducing the pose determination error to few centimeters level. Coordinate Z for the most of images showed accuracy under 2 mm. Orientation accuracy increased significantly as well, standing on level of 1 gon for the most cases, except for pose 5.

Obviously, aiming accuracy (ΔH and ΔV values) was also higher, providing level of 0.02 to 0.03 gon for vertical and horizontal angles correspondingly. Meanwhile for tracking in the zoomed mode the object must be moved slower in order to keep it stable in the field of view.

Fig. 9 | Examples of processed TS images taken with automatic optical zoom

Pose	Focal length	Distance	Recognized markers	Corner points	Aiming		Pose determination			
					ΔH	ΔV	ΔX	ΔY	ΔZ	$\Delta\kappa$
1	38.1 mm	8.3 m	5 of 5	20	–0.01 gon	0.01 gon	52.8 mm	–31.5 mm	–6.8 mm	0.7 gon
2	26.1 mm	7.1 m	5 of 5	20	0.00 gon	0.04 gon	25.9 mm	–4.0 mm	2.0 mm	1.2 gon
3	26.1 mm	5.7 m	5 of 5	20	–0.01 gon	0.04 gon	26.5 mm	–4.5 mm	1.9 mm	1.2 gon
4	22.9 mm	4.1 m	5 of 5	20	0.00 gon	0.03 gon	23.7 mm	–8.5 mm	0.1 mm	0.7 gon
5	22.9 mm	2.8 m	3 of 5	12	0.03 gon	0.02 gon	0.3 mm	8.9 mm	1.7 mm	2.6 gon
6	22.9 mm	2.7 m	5 of 5	20	0.02 gon	0.02 gon	9.0 mm	–5.0 mm	0.0 mm	1.0 gon
7	22.9 mm	2.9 m	3 of 5	12	0.02 gon	0.02 gon	4.6 mm	3.9 mm	3.7 mm	0.3 gon
8	22.9 mm	2.9 m	4 of 5	16	0.01 gon	0.03 gon	21.6 mm	–16.8 mm	–1.9 mm	1.2 gon
10	22.9 mm	2.8 m	3 of 5	12	0.01 gon	0.02 gon	19.3 mm	–14.0 mm	–0.6 mm	1.0 gon
			RMS		**0.02 gon**	**0.03 gon**	**25.1 mm**	**13.7 mm**	**2.9 mm**	**1.3 gon**

Tab. 4 | Aiming and 6DoF position determination accuracy for different lens positions

4.4 Kinematic Experiment

Kinematic positioning evaluation tests were focused on the stability of the pose determination from a single image during the construction process.

The experiment was made from two points of view, approximately parallel and perpendicular to line of track. In order to at least partly isolate the depth coordinate influence for the analysis the coordinate system was changed between start and end points and the instrument was set quasi-perpendicular and quasi-parallel to the track, *Fig. 10*.

The setup 1 quasi-perpendicular to the track provided better control of height and depth coordinates. The object was moving approximately perpendicular to the visual axis of the total station, which made axis *Y* lying approximately along the depth direction (side track). The distance from TS to the test object was approximately 2.5 meters. The setup 2 provided better evaluation of transverse coordinate and height stability. The local axis *Y* and the line of sight of the total station were approximately collinear to the depth axis, as well as the direction of object movements. Distance varied between 2.3 (end point) and 4.1 meters (start point). Track lengths for both setups were approximately 1.8 meters. *Fig. 11* and *Fig. 12* show plots of the coordinates and orientation analysis for both parts of the experiment.

In order to choose the best object tracking strategies the experiment was made with different velocities, including static position which was included into the first test. Stop-and-go approach was also tested for test 1 first forward track. *Tab. 5* contains the values

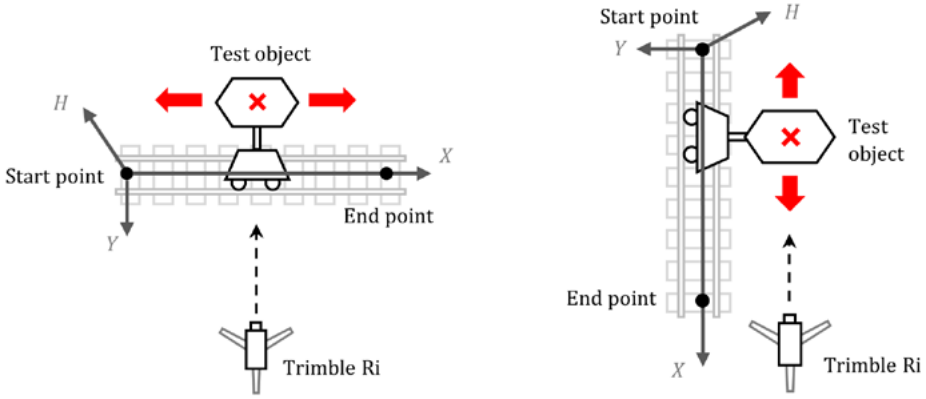

Fig. 10 | Kinematic experiments: setup 1 (left) and setup 2 (right)

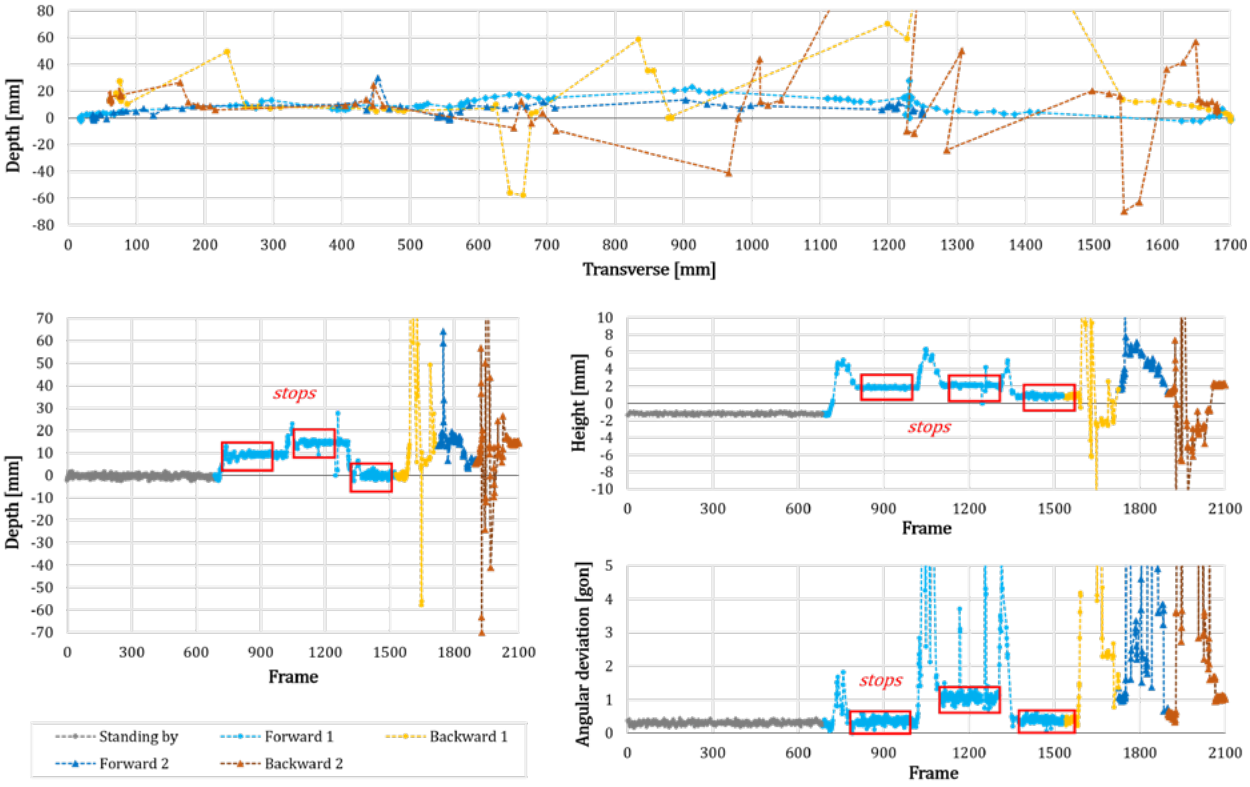

Fig. 11 | Side trajectory (setup 1) pose deviations from straight line and mean orientation

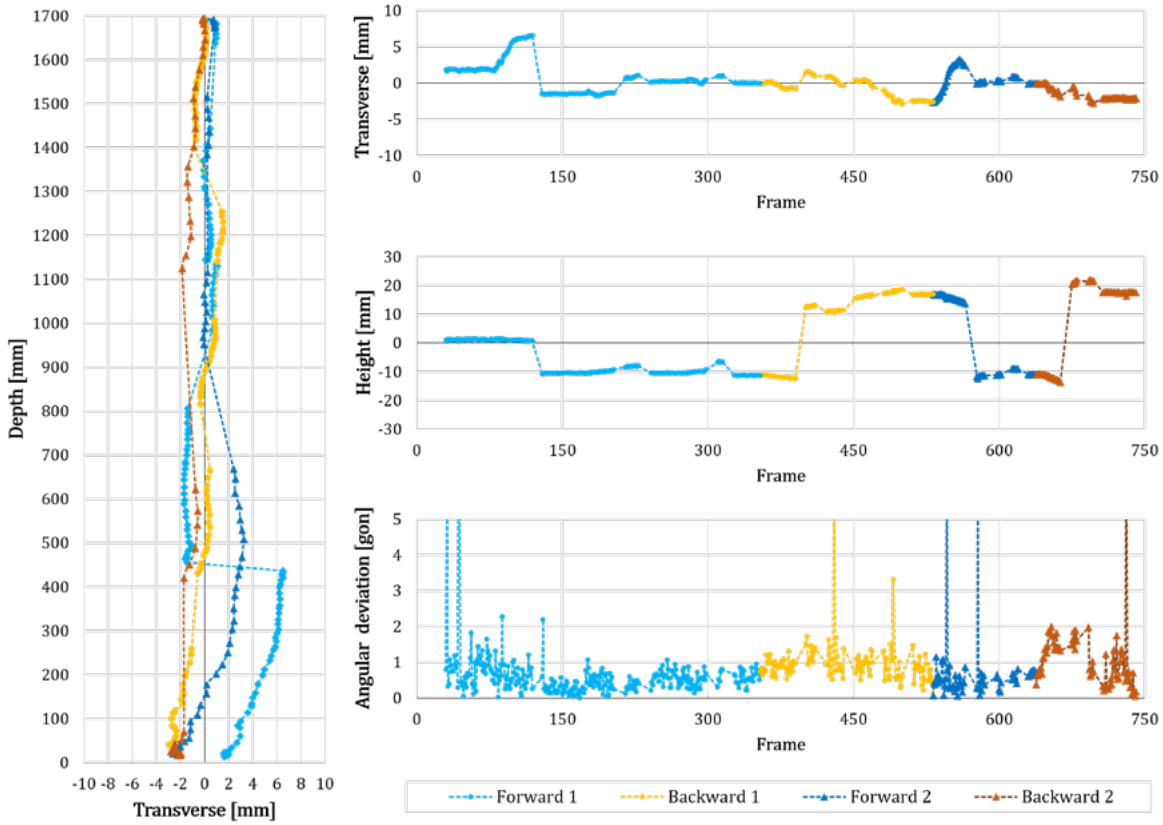

Fig. 12 | Along trajectory (setup 2) pose deviations from straight line and mean orientation

(RMS) of deviations from the straight line and the mean orientation for both tests, summarizing the plots (*Figs. 12*, *13*).

The best accuracy was obtained for static (standing by) trajectory: position deviations of few millimeters and orientation accuracy of 0.3 gon.

During the assembly process there is no opportunity to keep the object standing for the long time, therefore a stop-and-go strategy (with stops each 0.5 – 0.7 meter) was also tested (test 1, forward 1 track). This strategy led to several mm position deviations and 2 gon angular deviation level.

For the kinematic without stops with various velocities the accuracy of the coordinates decreased to few centimeters level and several gons in orientation making this strategy worse for application. Higher velocities led to outliers (up to 1 – 2 decimeters) mostly for the depth coordinate of the camera.

The influence of depth to the coordinates is visible on *Fig. 13,* where the same track is shown from different points of view, demonstrating the depth variations influence, including outlier points.

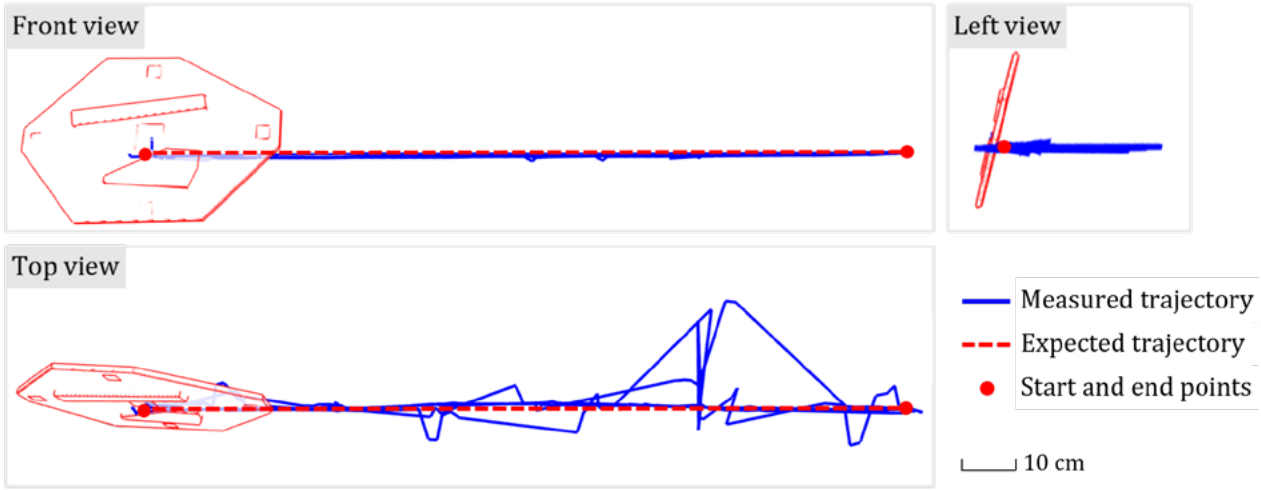

Fig. 13 | Example of trajectory visualization from different points of view (for setup 1)

Trajectory	Mean distance	Mean velocity	Standard Deviation (from straight line/mean pose)			
			Transverse	Depth	Height	Angles
Standing by	2.5 m	0 m/s	0.1 mm	0.7 mm	1.2 mm	0.3 gon
Test 1: Trajectory is perpendicular to the depth axis						
Forward 1	2.5 m	0.12 m/s (between stops)	-	8.6 mm	2.0 mm	2.0 gon
Backward 1		0.12 m/s	-	23.7 mm	5.2 mm	6.4 gon
Forward 2		0.18 m/s	-	16.7 mm	4.7 mm	8.2 gon
Backward 2		0.18 m/s	-	28.8 mm	4.8 mm	9.4 gon
Test 2: Trajectory goes along the depth axis						
Forward 1	2.3 m to 4.1 m	0.06	2.3 mm	–	8.4 mm	1.0 gon
Backward 1		0.12	1.4 mm	–	14.5 mm	1.2 gon
Forward 2		0.2	1.6 mm	–	13.5 mm	1.3 gon
Backward 2		0.2	1.7 mm	–	16.5 mm	1.4 gon

Tab. 5 | Kinematic experiments results

5 CONCLUSION

The implementation of the proposed method for timber construction components 6DoF pose determination using image assisted total stations within the task of the robotic assembly /Lauer et al. 2023/ was developed and tested using a single-image approach.

Within the solution the compound camera-to-TS model (*Fig. 1*) was extended and used together with ArUco fiducial markers for object (timber construction component) recognition. Based on these ideas, a custom-made software powered by OpenCV algorithms was developed in order to evaluate the accuracy of the method.

For the tests the Trimble Ri total station /Trimble 2022/ was used together with a downscaled test-object, which was generated to imitate a real timber cassette (*Fig. 4*). All the necessary values for the camera-to-TS conversion were available from total station messages and/or custom calibration.

Depending on the distances and the number of recognized markers the accuracy level of a few centimeters was reached. The best solutions (under 2 cm, up to several millimeters) were obtained using automatic zoom. The largest influence was caused of uncertainty of depth coordinate definition from a flat image. Orientation accuracy, depending on distance, was reached at the level of few gons. The same accuracy level was obtained for different object orientations.

Kinematical tests showed the same accuracy level for object velocities of about 1 m/s. The best quality can be obtained by stopping the object movement and waiting several seconds for higher solution quantity.

The further development of the timber element tracking with IATS network includes single-image approach as a less precise option (in comparison with stereo approach with multiple images), which can be used for rough position and orientation evaluation while the stereo processing is not available. This approach can be also helpful for sharing object positions within the total station network to prevent object tracking in case of losing the line of sight. The integration of the approach is shown in *Fig. 14*.

The precise positioning may be reached using stereo image processing, which is a different topic which is out of the scope for this

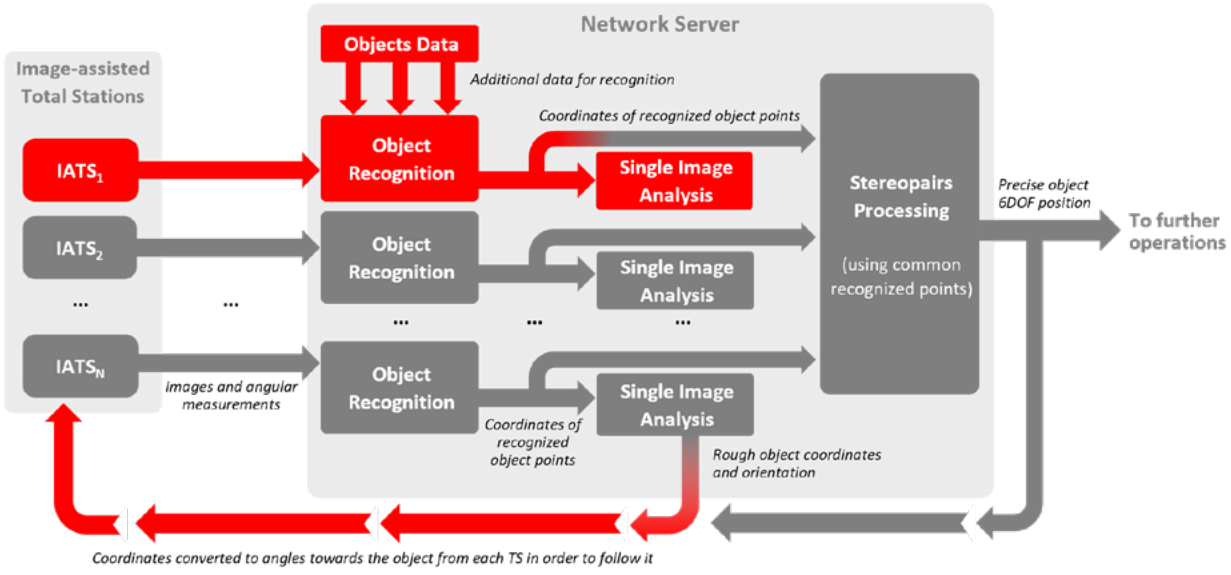

Fig. 14 | Concept of IATS network for tracking and pose determination of construction components with integrated single-image approach

paper and will be considered in further publications. The technique is still in development and results can be improved with further algorithm modifications.

ACKNOWLEDGEMENT

The research published in this article is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2120/1 – 390831618. The authors cordially thank the DFG.

REFERENCES

Abeho, D.; Shoko, M.; Odera, P. (2024): Effects of camera calibration on the accuracy of Unmanned Aerial Vehicle sensor products. In: International Journal of Engineering and Geosciences 9(2024)3, 314–323.

Bellmann, R. (2023): Trimble Ri White Paper & Benchmarking Report: Geo-Messdiskurs. Trimble Jena GmbH. https://dvw.de/api/assets/downloads/thuringen/gmd2023_03_bellmann_multisensorsystem_trimble_ri.pdf (11.09.2025).

Benligiray, B.; Topal, C.; Akinlar, C. (2019): STag: A stable fiducial marker system. In: Image and Vision Computing 8(2019), 158–169.

Berral-Soler, R.; Muñoz-Salinas, R.; Medina-Carnicer, R.; Marín-Jiménez, M. J. (2024): DeepArUco++: Improved detection of square fiducial markers in challenging lighting conditions. In: Image and Vision Computing 152(2024), 105313.

Dammert, L.; Thalmann, T.; Monetti, D.; Neuner, H.; Mandlburger, G. (2025): Numerical simulation and sensitivity analysis of UAS trajectory determination with robotic total stations. In: Survey Review (2025), 1–17.

DeGol, J.; Bretl, T.; Hoiem, D. (2017): ChromaTag: A Colored Marker and Fast Detection Algorithm. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, 1481–1490.

Diwakar, D. R. (2022): Recent Object Detection Techniques: A Survey. In: International Journal of Image, Graphics and Signal Processing(IJIGSP) 14(2022)2, 47–60.

Ehrhart, M.; Lienhart, W. (2017): Object tracking with robotic total stations: Current technologies and improvements based on image data. In: Journal of Applied Geodesy 11(2017)3, 131–142.

Gao, X.; Hou, X.; Tang, J.; Cheng, H. (2003): Complete solution classification for the perspective-three-point problem. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 25(2003)8, 930–943.

Hartley, R.; Zisserman, A. (2003): Multiple View Geometry in Computer Vision. 2nd Ed. Cambridge University Press.

Kalaitzakis, M.; Cain, B.; Carroll, S.; Ambrosi, A.; Whitehead, C.; Vitzilaios, N. (2021): Fiducial Markers for Pose Estimation. In: Journal of Intelligent & Robotic Systems 101(2021)4.

Kodnongbua, M.; Jones, B.; Bin Safeer Ahmad, M.; Kim, V.; Schulz, A. (2023): ReparamCAD: Zero-shot CAD Re-Parameterization for Interactive Manipulation. In: SIGGRAPH Asia 2023 Conference Papers (SA '23). Association for Computing Machinery, New York, NY, Article 69, 1–12.

Lauer, A. P. R.; Benner, E.; Stark, T.; Klassen, S.; Abolhasani, S.; Schroth, L.; Gienger, A.; Wagner, H. J.; Schwieger, V.; Menges, A.; Sawodny, O. (2023): Automated on-site assembly of timber buildings on the example of a biomimetic shell. In: Automation in Construction 156(2023), 105–118.

Wichmann

Technikwissen punktgenau:

Vermessungskunde für den Planungs-, Bau- und Umweltbereich

- Buch ist für Studierende und Anwender außerhalb der Geodäsie eine verständliche Einführung sowie praxisnaher Leitfaden
- Vermittelt sowohl Grundlagen als auch aktuelle Entwicklungen und Trends
- Anschauliche Darstellung mit vielen Abbildungen und zahlreichen Beispielen
- Ermöglicht so eine eigenständige Umsetzung des Stoffs im Sinne eines echten Lehrbuchs

5., neu bearb. und erw. Auflage 2024
388 Seiten
36,– € (Buch/E-Book)
50,40 € (Kombi)

Preisänderungen und Irrtümer vorbehalten. Sowohl das E-Book als auch das Kombiangebot (Buch + E-Book) sind ausschließlich auf **www.vde-verlag.de** erhältlich.

Bestellen Sie jetzt: (030) 34 80 01-222 oder www.vde-verlag.de/buecher/537746

Werb.-Nr. 2602107

Lepetit, V.; Moreno-Noguer, F.; Fua, P. (2008): EPnP: An Accurate O(n) Solution to the PnP Problem. In: International Journal of Computer Vision 81(2008)2, 155–166.

Lerke, O.; Schwieger, V. (2021): Analysis of a kinematic real-time robotic total station network for robot control. In: Journal of Applied Geodesy 15(2021), Article 3.

Luhmann, T. (2009): Precision potential of photogrammetric 6DOF pose estimation with a single camera. In: ISPRS Journal of Photogrammetry and Remote Sensing 64(2009)3, 275–284.

Luhmann, T.; Robson, S.; Kyle, S.; Boehm, J. (2020): Close-Range Photogrammetry and 3D Imaging. De Gruyter, Berlin/Boston.

Manawadu, M.; Park, S.-Y. (2024): 6DoF Object Pose and Focal Length Estimation from Single RGB Images in Uncontrolled Environments. In: Sensors 24(2024)17, 5474.

Mao, Y.; Jing, J.; Mikolajczyk, K. (2025): OpenDlign: Open-World Point Cloud Understanding with Depth-Aligned Images. In: Proceedings of the 38th International Conference on Neural Information Processing Systems (NIPS '24) 37(2025). Curran Associates Inc., Red Hook, NY, Article 3208, 101144–101167.

Oščádal, P.; Heczko, D.; Vysocký, A.; Mlotek, J.; Novák, P.; Virgala, I.; Sukop, M.; Bobovský, Z. (2020): Improved Pose Estimation of Aruco Tags Using a Novel 3D Placement Strategy. In: Sensors 20(2020)17, 4825.

Paar, R.; Roič, M.; Marendič, A.; Miletič, S. (2021): Technological Development and Application of Photo and Video Theodolites. In: Applied Science 2021(11), 3893.

Padilla, R.; Passos, W. L.; Dias, T. L. B.; Netto, S. L.; da Silva, E. A. B. (2021): A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. In: Electronics (2021)10, 279.

Poroykov, A.; Kalugin, P.; Shitov, S.; Lapitskaya, I. (2020): Modeling ArUco Markers Images for Accuracy Analysis of Their 3D Pose Estimation. In: Proceedings of the 30th International Conference on Computer Graphics and Machine Vision (GraphiCon 2020), Part 2, Short 14-1. https://doi.org/10.51130/graphicon-2020-2-4-14.

Reiterer, A.; Wagner, A. (2012): System Considerations of an Image Assisted Total Station – Evaluation and Assessment. In: allgemeine vermessungs-nachrichten (avn) 119(2012)3, 83–94.

Sardar, A. S. (2025): Real-Time Tracking and Distance Measurement of OpenCV ArUco Marker Using Webcam. In: International Journal of Latest Technology in Engineering Management & Applied Science 14(2025)1, 313–321.

Schlüter, M.; Zschiesche, K. (2024): Bewegungsmuster von Brückenbauwerken bei laufendem Verkehr messen und auswerten. In: allgemeine vermessungs-nachrichten (avn) 131(2024)1, 9–16.

Song, K.; Wu, C.; Jiang, S. (2017): CAD-based Pose Estimation Design for Random Bin Picking using a RGB-D Camera. In: Journal of Intelligent & Robotic Systems 87(2017), 455–470.

Trigka, M.; Dritsas, E. (2025): A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection. In: Sensors 25(2025)1, 214.

BuildingPoint (2026): Trimble Ri Datasheet (EN). https://buildingpoint.com.au/wp-content/uploads/2024/06/trimble-ri-datasheet.pdf (07. 04. 2026)

Ulrich, J.; Blaha, J.; Alsayed, A.; Rouček, T.; Arvin, F.; Krajník, T. (2023): Real Time Fiducial Marker Localisation System with Full 6 DOF Pose Estimation. In: ACM SIGAPP Applied Computing Review 23(2023)1, 20–35.

Umeyama, S. (1991): Least-squares estimation of transformation parameters between two point patterns. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 13(1991)4, 376–380.

Vogel, M. (2006): Vom Pixel zur Richtung: die räumlichen Beziehungen zwischen Abbildungsstrahlen und Tachymeter-Richtungen. PhD Thesis, Technical University of Darmstadt.

Walser, B. (2004): Development and Calibration of an Image Assisted Total Station. PhD thesis, ETH Zurich.

Wieczorek, B.; Warguła, Ł.; Kukla, M.; Kubacki, A.; Górecki, J. (2020): The effects of ArUco marker velocity and size on motion capture detection and accuracy in the context of human body kinematics analysis. In: Technical Transactions (2020), e2020036.

Wubben, J.; Fabra, F.; Calafate, C. T.; Krzeszowski, T.; Marquez-Barja, J. M.; Cano, J.-C.; Manzoni, P. (2019): Accurate Landing of Unmanned Aerial Vehicles Using Ground Pattern Recognition. In: Electronics 8(2019)12, 1532.

Xu, C.; Kuipers, B.; Murarka, A (2009): 3D pose estimation for planes. In: 12th IEEE International Conference on Computer Vision Workshops, ICCV Workshops 2009, Kyoto, Japan, September 27–October 4, 2009. IEEE Computer Society, Washington, D. C., 673–680.

Yu, G.; Hu, Y.; Dai, J. (2021): TopoTag: A Robust and Scalable Topological Fiducial Marker System. In: IEEE Transactions on Visualization and Computer Graphics 27(9), 3769–3780.

Zakiev, A.; Shabalina, K.; Magid, E. (2019): Pilot Virtual Experiments on ArUco and AprilTag Systems Comparison for Fiducial Marker Rotation Resistance. In: Proceedings of International Conference on Artificial Life and Robotics 24(2019), 132–135.

Zhang, L.; Li, C.; Zheng, Y. (2021a): Image-Assisted Total Station Camera Mounting Error Correction Model and Analysis. In: Journal of Surveying Engineering 147(2021)3.

Zhang, Y.; Tino, P.; Leonardis, A.; Tang, K. (2021b): A Survey on Neural Network Interpretability. In: IEEE Transactions on Emerging Topics in Computational Intelligence 5(2021)5, 726–742.

Zschiesche, K. (2023): Methods and algorithms of an imaging total station. Doctoral thesis, University of Freiburg.

AUTHORS

Dr.-Ing Stanislav Shevchuk

UNIVERSITY STUTTGART
INSTITUTE OF ENGINEERING GEODESY (IIGS)

Geschwister-Scholl-Str. 24D | 70174 Stuttgart | Germany
stanislav.shevchuk@iigs.uni-stuttgart.de

Dr.-Ing Gabriel Kerekes

UNIVERSITY STUTTGART
INSTITUTE OF ENGINEERING GEODESY (IIGS)

Geschwister-Scholl-Str. 24D | 70174 Stuttgart | Germany
gabriel.kerekes@iigs.uni-stuttgart.de | ORCID: 0000-0002-9037-0499

Prof. Dr.-Ing Volker Schwieger

UNIVERSITY STUTTGART
INSTITUTE OF ENGINEERING GEODESY (IIGS)

Geschwister-Scholl-Str. 24D | 70174 Stuttgart | Germany
volker.schwieger@iigs.uni-stuttgart.de | ORCID: 0000-0001-9055-9809

Manuskript eingereicht: 16. 12. 2025 | Im Peer-Review-Verfahren begutachtet