RESEARCH

The evaluation of hand-crafted and learned-based features in Terrestrial Laser Scanning-Structure-from-Motion (TLS-SfM) indoor point cloud registration: the case study of cultural heritage objects and public interiors

Jakub Markiewicz[1], Patryk Kot[2*], Łukasz Markiewicz[3] and Magomed Muradov[2]

Abstract

Modern technologies are commonly used to inventory different architectural or industrial objects (especially cultural heritage objects and sites) to generate architectural documentation or 3D models. The Terrestrial Laser Scanning (TLS) method is one of the standard technologies researchers investigate for accurate data acquisition and processing required for architectural documentation. The processing of TLS data to generate high-resolution architectural documentation is a multi-stage process that begins with point cloud registration. In this step, it is a common practice to identify corresponding points manually, semi-manually or automatically. There are several challenges for the TLS point cloud processing in the data registration process: correct spatial distribution, marking of control points, automation, and robustness analysis. This is particularly important when large, complex heritage sites are investigated, where it is impossible to distribute marked control points. On the other hand, when orientating multi-temporal data, there is also the problem of corresponding reference points. For this reason, it is necessary to use automatic tie-point detection methods. Therefore, this article aims to evaluate the quality and completeness of the TLS registration process using 2D raster data in the form of spherical images and Affine Hand-crafted and Learned-based detectors in the multi-stage TLS point cloud registration as test data; point clouds were used for the historic 17th-century cellars of the Royal Castle in Warsaw without decorative structures, two baroque rooms in the King John III Palace Museum in Wilanów with decorative elements, ornaments and materials on the walls and flat frescoes, and two modern test fields, narrow office, and empty shopping mall. The extended Structure-from-Motion was used to determine the tie points for the complete TLS registration and reliability analysis. The evaluation of detectors demonstrates that for the test sites exhibiting rich textures and numerous ornaments, a combination of AFAST, ASURF, ASIFT, SuperGlue and LoFTR can be effectively employed. For the point cloud registration of less textured buildings, it is advisable to use AFAST/ASIFT. The robust method for point cloud registration exhibits comparable outcomes to the conventional target-based and Iterative Closest Points methods.

Keywords Affine 2D hand-crafted detectors, Cultural heritage, Interiors, Feature based matching, Learned-based detectors, Reliability assessment, TLS registration

---

*Correspondence:
Patryk Kot
p.kot@ljmu.ac.uk
Full list of author information is available at the end of the article

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 2 of 26

Introduction

Modern measurement technologies such as terrestrial laser scanning (TLS) are commonly applied to register, preserve, protect and monitor different engineering objects [1], perform structural health monitoring [2–5], assist with construction management [6], carry out three-dimensional (3D) model reconstruction [7], monitor deformation of structures [8–13] and significantly assist with preservation and safeguarding of cultural heritage objects and sites [14–21], owing to its accurate data acquisitions and processing, which is required to generate the documentation such as 3D models, vector drawings or other architectural documentation [22–26]. The acquisition and processing of point clouds from terrestrial laser scanners is a multi-step process consisting of (1) Survey planning, (2) Field operation, (3) Data preparation, (4) Data registration, (5) Data processing, and (6) Quality control and delivery [27]. Planning of the optimal TLS positions and target locations depends on the surveying area and the design consideration of the project. Based on the adopted data orientation method, these target locations might be natural points that are detected in the point cloud or specific signal points in the form of black and white chessboards, retroreflective points, or spheres with a known radius (Fig. 1). Since the TLS point clouds are collected in the local reference system, it is required to perform the registration step (first step of the TLS point cloud processing methodology), allowing to transform point clouds into the assumed reference system [28].

For large and complex objects and sites, obtaining data from multiple TLS positions and transforming them into the defined reference system is required, as a single position will not provide the significant data needed for an accurate model generation. The transformation into the defined reference system relies on detecting corresponding points, shapes or features in at least two-point clouds, and the exterior orientation parameters are obtained for each scan. These parameters determine the spatial location of the central point of the scanner system in the assumed reference system together with three rotation angles, which are then used to transform the point cloud [29].

In literature, many investigations address the problem of TLS point cloud registration in the context of the effectiveness, efficiency and robustness of this process [30–34] and divide these methods into two main groups depending on the amount of the input data—pairwise or multiview registration [2]. Most of these algorithms are the coarse–fine-strategy [35, 36], which assumes that (1) in the first step—the translation and rotation parameters are approximated [28] and (2) in the final step—fine registration is performed by algorithms such as normal distribution transform (NDT) algorithm and its variants [37–39] or Iterative Closest Points (ICP) algorithm or its variants [38, 40]. A review of the commonly used methods for TLS registration can be found in the article [41].

Several challenges are encountered during data registration in Terrestrial Laser Scanning (TLS) point cloud processing. These challenges pertain to ensuring the accurate spatial distribution of data, addressing control point identification, enhancing automation in the process, and conducting robustness analysis. This becomes especially critical when examining extensive and intricate

(a)   (b)

Fig. 1 **a** The example of the artificial targets, **b** registration between two scanned positions [27]

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 3 of 26

heritage sites where the deployment of marked control points is unfeasible. Furthermore, in the case of multi-temporal data alignment, the issue of establishing correspondences between reference points also arises. Consequently, automatic tie-point detection methods are necessary to mitigate these challenges effectively.

This paper aims to present the possibility of using the TLS-SfM method for the orientation of point clouds from terrestrial laser scanning of the interiors of historic and public buildings. This research compares the utilisation of selected 2D hand-crafted and learned methods for finding tie points. This article presents the effectiveness of different algorithms (AFAST, ASIFT, ASURF, LoFTR, SuperGlue and KeyNet with AffiNet and HardNet) in the point detection step with extended quality and robustness analysis based on the reliability assessment. The interiors of historical 17th-century basements at the Royal Castle in Warsaw without decorative structure (Test Site I and II), the Museum of King Jan III's Palace at Wilanow with decorative elements, ornaments, and materials on walls (Test Site III) and flat frescos (Test Site IV), narrow office (Test Site V) and shopping mall (Test Site VI), were selected for this study. For such objects, the distribution of the signalised points utilised in the data registration process may not be possible owing to the inability to distribute it on historical wall fragments, the deployment of tripods that would have the effect of obscuring the objects under development and the spatial distribution of points (caused by the complex shapes of the objects being developed), which would affect the accuracy of registration and error detection according to robustness theory.

The method for point cloud registration is based on intensity rasters (together with a depth map) and an extended Structure-from-Motion (TLS-SfM) approach. The advantage of the method for point cloud registration over the Target-based method is that more automatically detected tie points are used for orientation with better spatial distribution and robust outliers' detection regarding the reliability theory. The Iterative Closest Points (ICP) method is based on the point-to-point and point-to-plane approaches, which require clouds to be pre-oriented when connecting point clouds to guarantee the final registration's correctness. In the TLS-SfM approach, such a condition is unnecessary since the selection and elimination of tie points are utilised in a two-step manner through descriptor matching and geometrical verification based on the RANSAC algorithm.

This article is divided into five main sections. Sect. "Principle of work" presents the fundamental principles of the hand-crafted and learned feature detectors and descriptors. Sect. "Methodology" contains a description of the test sites and the approach used. Sect. "Results and discussion" presents the results of the detector assessments, and Sect. "Conclusion" concludes the proposed study, highlighting the advantages and limitations of using different affine 2D detectors and future work approaches.

Principle of work

TLS point cloud registration

Several methods of TLS data registration exist, which may be generally divided (followed by Vosselman and Maas [42] proposed definitions) into target-based and feature-based [16, 43–49]. The TLS data registration methods are generally based on the corresponding features between two or more datasets. Still, the main differences might be seen in determining and matching these corresponding points. Despite the existence of two different approaches to the determination of tie points, to define the relationship between the local instrument and the global reference system, Eq. (1) is used:

$$
\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = M_{ij} * \begin{bmatrix} x_{ij} \\ y_{ij} \\ z_{ij} \end{bmatrix} + \begin{bmatrix} X_j^c \\ Y_j^c \\ Z_j^c \end{bmatrix}
$$

$$
M_{ij} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} cos\varphi cos\kappa & -cos\varphi sin\kappa & sin\varphi \\ cos\omega sin\kappa + sin\omega sin\varphi cos\kappa & cos\omega cos\kappa - sin\omega sin\varphi sin\kappa & -sin\omega cos\varphi \\ sin\omega sin\kappa - cos\omega sin\varphi cos\kappa & sin\omega cos\kappa + cos\omega sin\varphi sin\kappa & cos\omega cos\varphi \end{bmatrix}
$$

(1)

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 4 of 26

In this Equation, the coordinates of the object points (reference points) correspond to the vector $\left(X_i\ Y_i\ Z_i \right)^T$, points in the local (scanner) coordinate system and are represented by the vector $\left(x_{ij}\ y_{ij}\ z_{ij} \right)^T$, the scanner position $\left(X_i^c\ Y_i^c\ Z_i^c \right)^T$ scanner rotation M_{ij}(three Euler angles ω, φ, κ that are used to construct the rotation matrix).

The least-square estimation is required to determine the exterior orientation parameters for the oriented point. Teunissen [50] used the well-known Gauss-Markow linear model (a linearised form of the nonlinear input relationships), which is also used in the TLS/photogrammetric bundle adjustment process [51]. To determine the normal equation matrix and vector, the least-square adjustment is used with the following analytic form (Eqs. 2, 3, 4):

The selection and arrangement of tie points in the point cloud orientation process play a crucial role. When considering the possibilities of using tie points in the data orientation process, it is essential to consider their use in accuracy and detecting, locating, and eliminating outliers that may occur during the adjustment process. Reliability theory deals with diagnosing outliers in observations and datasets used in the alignment process [52–57].

In this article, the proposed reliability approach will compensate for the orientation quality based on the local reliability criteria, which enables determining if the pair of tie points is correctly matched. The proposed quality assessment method will focus on the RMSE on control and check points evaluation and consider the points' spatial distribution. Based on the least square method

$$y + e = Ax;\ e \sim (0, C_e)$$

$$A = \begin{bmatrix} x_1\ y_1\ z_1 & 0\ 0\ 0 & 0\ 0\ 0 & 1\ 0\ 0 \\ 0\ 0\ 0 & x_1\ y_1\ z_1 & 0\ 0\ 0 & 0\ 1\ 0 \\ 0\ 0\ 0 & 0\ 0\ 0 & x_1\ y_1\ z_1 & 0\ 0\ 1 \\ x_2\ y_2\ z_2 & 0\ 0\ 0 & 0\ 0\ 0 & 1\ 0\ 0 \\ 0\ 0\ 0 & x_2\ y_2\ z_2 & 0\ 0\ 0 & 0\ 1\ 0 \\ 0\ 0\ 0 & 0\ 0\ 0 & x_2\ y_2\ z_2 & 0\ 0\ 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_m\ y_m\ z_m & 0\ 0\ 0 & 0\ 0\ 0 & 1\ 0\ 0 \\ 0\ 0\ 0 & x_m\ y_m\ z_m & 0\ 0\ 0 & 0\ 1\ 0 \\ 0\ 0\ 0 & 0\ 0\ 0 & x_m\ y_m\ z_m & 0\ 0\ 1 \end{bmatrix} \quad x = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \\ a_{33} \\ X^c \\ Y^c \\ Z^c \end{bmatrix} \quad y = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ X_2 \\ Y_2 \\ Z_2 \\ \vdots \\ X_m \\ Y_m \\ Z_m \end{bmatrix} \quad e = \begin{bmatrix} e_{X_1} \\ e_{Y_1} \\ e_{Z_1} \\ e_{X_2} \\ e_{Y_2} \\ e_{Z_2} \\ \vdots \\ e_{X_m} \\ e_{Y_m} \\ e_{Z_m} \end{bmatrix} \tag{2}$$

$$A^T P A x = A^T P y \tag{3}$$

$$P = C_y^{-1} \tag{4}$$

where: A—coefficient matrix (m×n) (m—number of observational equations,—*m*number of unknowns), $rank(A) = u$ (full rank);—*x*parameter vector (n×1);—*y* observation vector (m×1) (uncorrelated observations); C_e—observation error covariance matrix (m×m) (positively determined) is also the observation result covariance matrix, i.e., $C_e \equiv C_y$.

(Eqs. 2 and 3), the formula for local reliability criteria is determined (Eq. 5), which is called the "disturbance-response" dependency and is one of the basic elements of reliability theory:
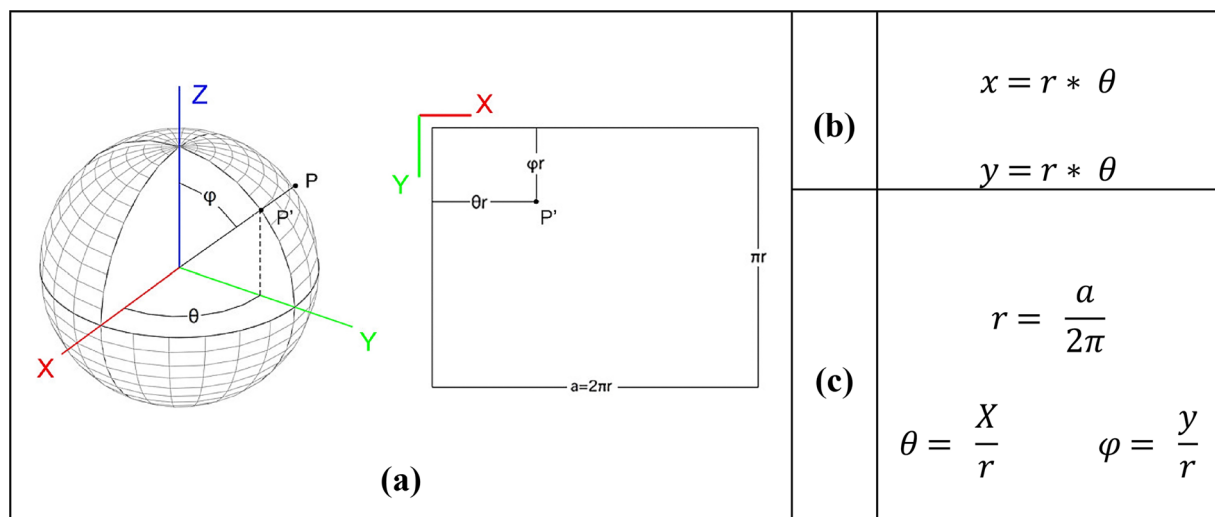
$$v = -Ry$$
$$R = I - A\left(A^T A\right)^{-1} A^T \tag{5}$$

where: R—reliability matrix of the tie points; I—identity matrix, A—coefficient matrix based on the tie points.
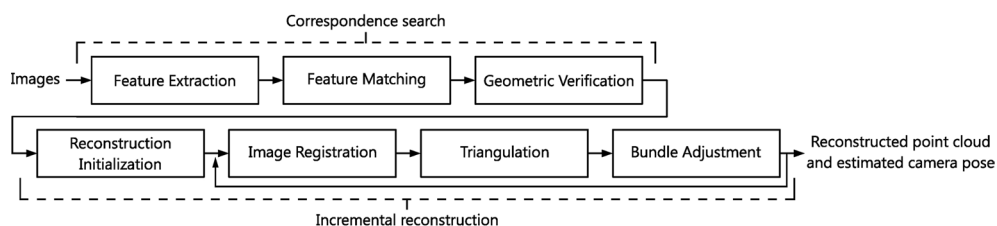
Fig. 2 Incremental SfM methodology [59]

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 5 of 26

Fig. 3 Relation between spherical coordinates and coordinates on spherical photographs **a** Graphical representation of the relation between polar coordinates measured and the raster image in spherical projection [63], **b** formula for recalculation of polar coordinates to spherical projection, and **c** formula for recalculation of x,y spherical projection onto polar coordinates

The analysis of the internal reliability factors, based on the diagonal value of the matrix R (an orthogonal projection operator), that values are between $<0,1>$. It is stated that if: (1) $\{R\}_{ii} = 0$ the tie point is uncontrolled by other points; (2) $\{R\}_{ii} = 1$ the tie point is fully controlled by other points; (3) $\{R\}_{ii} > 0.5$ tie point (in relation to other points) is well distributed regarding reliability theory. This method is very useful for automatically analysing and selecting the TLS point registration [58].

TLS point featured-based cloud registration
The current state-of-the-art approach for TLS data registration is based on two main methods, namely, (1) point-based (provided control points/markers) and (2) feature-based methods [42]. One of the feature-based methods is Structure-from-Motion (SfM), which is carried out in the following steps: (1) feature extraction; (2) feature matching; (3) geometric verification; (4) reconstruction initialisation; (5) image registration; (6) triangulation, and (7) bundle adjustment (Fig. 2). To generalise, the SfM approach might be divided into two main parts: the correspondence search phase (1–3) and iterative reconstruction phase (4–6) [46, 59–62].

The classical SfM uses the group of collected images. Still, in the case of TLS registration, the point cloud should be converted into the spherical raster based on the cartographical Equation (Eqs. 6, 7, 8). Referring to Fig. 3, a TLS with the panoramic architecture acquire the spherical coordinate observation defined as a ρ—the measured distance between the object and scan position, θ—horizontal direction and ϕ—vertical (elevation) angle.

These values might be expressed concerning the Euclidean coordinate system (Eqs. 6, 7, 8):

$$\rho_{ij} = \sqrt{x_{ij}^2 + y_{ij}^2 + z_{ij}^2} \tag{6}$$

$$\theta_{ij} = \arctan\left(\frac{y_{ij}}{x_{ij}}\right) \tag{7}$$

$$\alpha_{ij} = \arctan\left(\frac{z_{ij}}{\sqrt{x_{ij}^2 + y_{ij}^2}}\right) \tag{8}$$

A spherical image (for which the raster grey-level value assumes the laser beam reflectance intensity value) is used, together with the map of depth (i.e., the distance to the analysed object), for TLS data orientation. This point cloud representation is applied and implemented in many commercial software tools [46, 64–67]. The main advantage of that data representation is the possibility of using raw data with the highest resolution and without the interpolation of new values of pixel coordinates. It is also possible to generate an intensity raster of any resolution, and this can be done by converting new pixel values based on the formulas shown in Fig. 3.

To compare the points in different rasters, it is necessary to determine the invariant features. The detection and description of features for each characteristic point are essential for the process of detection of homologous points because the final points' recognition as tie points

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 6 of 26

is carried out by matching their relative descriptors in the process of data orientation. There are two approaches usually applied: (1) the Approximate Nearest Neighbour-Based Point Matching [68] and (2) Brute Force matching [69].

The fundamental principles of the 2D feature

Feature detection (also called extraction) is the first and the most essential step in the SfM methodology that relies on the detectors. The key extraction principle is to recognise each raster data (image from a group of processed images) and a group of characteristic points (also called keypoints) based on the local characteristic of the intensity. For feature extraction, different methods and algorithms can be used, such as point detectors [70], line detectors [71] or blob detectors [72], which affects the robustness of the detected features and efficiency of the matching method.

Those features should have the following properties that allow to determine the characteristics of the detector "[73]: (1) *Repeatability*—the possibility to detect a high percent of the features possible to recognise the scene part visible in both images taken under different viewing conditions; (2) *Distinctiveness/informativeness*—the intensity patterns used for detecting points should show a lot of variations; (3) *Locality*—the neighbourhood used to determine the point should be local in order to reduce the probability of occlusions and invariant of the photometric and geometric deformations; (4) *Quantity*—a number of the detected features that should be sufficiently large and allow to detect features even on small objects (however, number of keypoints depends directly on the application); (5) *Accuracy*—definition of the quality and possibility of feature localisation in regards to the scale-space and photometric and geometrical distortions; (6) *Efficiency*—determination of the required time for feature detection (important in the time-critical applications)".

At present, there are two distinct approaches for detecting keypoints in images. The first approach involves utilising a group of hand-crafted algorithms, such as Scale-Invariant Feature Transform (SIFT) introduced by Lowe [74] and Speeded Up Robust Features (SURF) proposed by Bay and Ess [75]. The second approach, a learned-based feature extraction approach, employs methods such as SuperGlue or LoFTR. Hand-crafted detectors operate by detecting keypoints based on the grayscale gradient values in the local neighbourhood, using either blob detectors like SIFT, SURF, or CenSurE, or corner detectors like FAST introduced by Rosten and Drummond [76] and BRISK proposed by Leutenegger et al. [77] [REF], which compare grayscale

differences with the analysed pixel. Point and blob detectors found wide application in the orientation of point clouds from terrestrial laser scanning [63]. The advantages of using point and blob detectors are (1) the speed of detection and match of tie points—might be extracted very efficiently, (2) the accuracy of localisation and scale-invariant, (3) stability over varying viewpoints and (4) the accuracy of TLS data registration [73, 78]. One of the significant limitations of these detectors is that they were designed to use images projected in the central projection. Such an approach assumes that standard image deformations might be expected. For this reason, using spherical rasters from point cloud conversions can result in significant deformations that contribute to problems concerning explicit identification and matching keypoints [19, 46, 63, 79]. This problem can be solved in two ways: (1) using different mapping representations (i.e., "virtual image", orthoimage or Mercator representation) [60, 63] or (2) adding an affine component to the detectors [80].

In recent years, novel learning-based solutions have been developed to overcome the limitations of hand-crafted methods. These solutions encompass various approaches. The first approach, known as "detect-then-describe," involves using a learned detector and descriptor, which can either be fully learned or combined with hand-crafted and learning-based methods. Notable works in this domain include Barroso-Laguna et al. [81], Verdie et al. [82] for the detector, Ebel et al. [83], and Mishchuk et al. [84] or the descriptor.

The second approach, "end-to-end," aims to jointly optimise the entire pipeline to extract sparse image correspondences. Examples of end-to-end methods include SuperPoint, introduced by DeTone et al. [85]; SuperGlue, proposed by Sarlin et al. [86]; and DISK, presented by Tyszkiewicz et al. [87]. These end-to-end methods have been utilised to enhance both the repeatability and reliability of keypoints, leading to improved success rates in image matching and more accurate pose estimation, as demonstrated by Remondino [88].

More recently, researchers such as Choy et al. [89], Rocco et al. [90], and Li et al. [91] introduced a new approach, "end-to-end detector-free local feature matching methods." These methods eliminate the feature detector phase and directly generate dense descriptors or dense feature matches. Notably, Sun et al. [92] introduced the LoFTR approach, which builds upon the Transformer architecture proposed by Vaswani et al. [93]. In contrast to the sequential process of image feature detection, description, and matching, LoFTR establishes pixel-wise dense matches at a coarse level and subsequently refines these matches at a fine level.

Markiewicz *et al. Heritage Science* (2023) 11:254

Page 7 of 26

The feature description, matching and images registration

To match characteristic points in several photographs, it is necessary to describe their features based on their neighbourhood [72]. This is carried out by descriptors, which enable the determination of the invariant features that form the basis for comparing points in different photographs. The characteristic points' descriptions can be unified using one descriptor for each detector. For that purpose, the operations of the SIFT descriptor were utilised [72]. The operations of the SIFT descriptor consist of two stages: (1) calculation of the gradient (scale) and orientation of each point within the neighbourhood of a key point and (2) determination of a 128-element vector of features (a descriptor). The Gaussian images are used to determine the orientation of keypoints, which corresponds to the scale of a given keypoint. For each image point, the gradient module and orientation are calculated. The keypoints' features are measured in relation to the determined orientation, which results in the description being independent of the rotation. The SIFT algorithm considers the gradient module and orientation within the neighbourhood of 16×16 for a given keypoint. Then, this area is divided into regions of 4×4 size, in which the resultant orientation histograms are re-created. The consequent gradient module for eight orientations is determined within each area based on the particular points of the modules. Thus, the point feature descriptor is a vector of $4 \times 4 \times 8 = 128$ elements. The vector is normalised to reduce the influence of illumination. The next stage of considering points as tie points in image data orientation is their relative matching. In this article, the Approximate Nearest Neighbourhood-Based Point Matching [60] was used. At the end of the final iterative,

the bundle adjustment process relies on the methodology described in subSect. "TLS point cloud registration".

Methodology

Selected test site

The proposed method for automatic Terrestrial Laser Scanning data registration that involves the use TLS-SfM with hand-crafted and learned features to detect non-signalised tie points on point clouds was tested at six different sites, namely historic 17th-century basements at the Royal Castle in Warsaw without decorative structure (Test Site I and II), Museum of King Jan III's Palace at Wilanów with decorative elements, ornaments, and materials on walls (Test Site III) and flat frescos (Test Site IV), narrow office (Test Site V) and shopping mall (Test Site VI).

The Test Sites I and II are constructed of bricks filled with mortar. It has an irregular shape with a ceiling in the form of arches, with a maximum height of approximately 3.2 m and a minimum of about 2.1 m. Due to its historical character and the prevailing humidity conditions, the part of the room has damp walls and fragments of bricks crumble, making it impossible to place the signalled control points on the object. On the other hand, it was impossible to place the points on tripods because of the size and dimensions of the individual rooms. If the target-based methodology is implemented, it will increase the number of required scanner positions, leading to inaccurate point cloud registration.

Both Test Sites were marked with check points (that were not used for orientation parameters determination but were used for the independent quality assessment), which were placed at different heights. All points
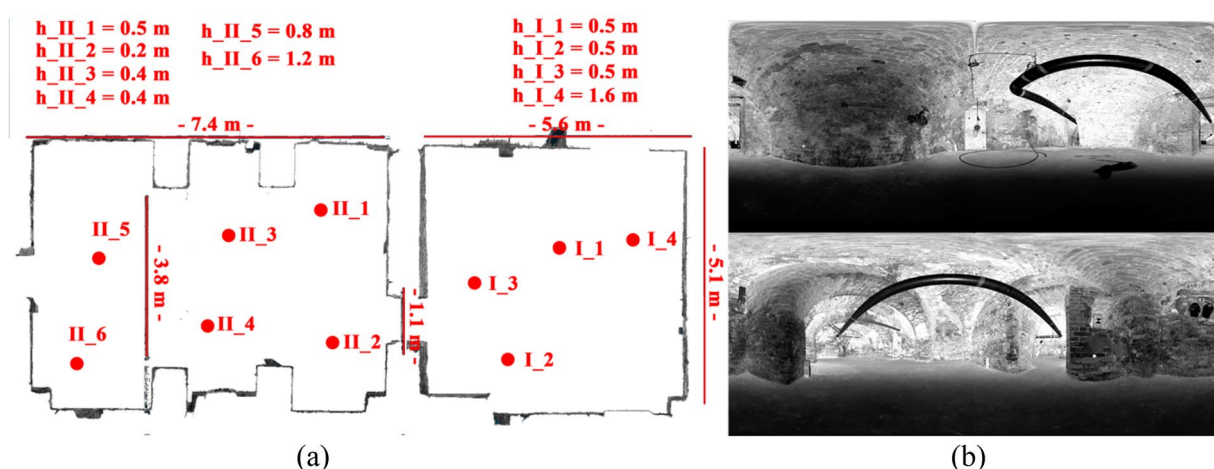

Fig. 4 **a** The floor plan with marked dimensions and Terrestrial Laser Scanning (TLS) positions (red dots). Each name of the laser scanner position contains the name of the selected test site (I and II) and specified id (1, 2, 3, 4, etc.). For each TLS position, the height (h) was also defined as **b** a spherical map of point clouds for each Test Site

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 8 of 26
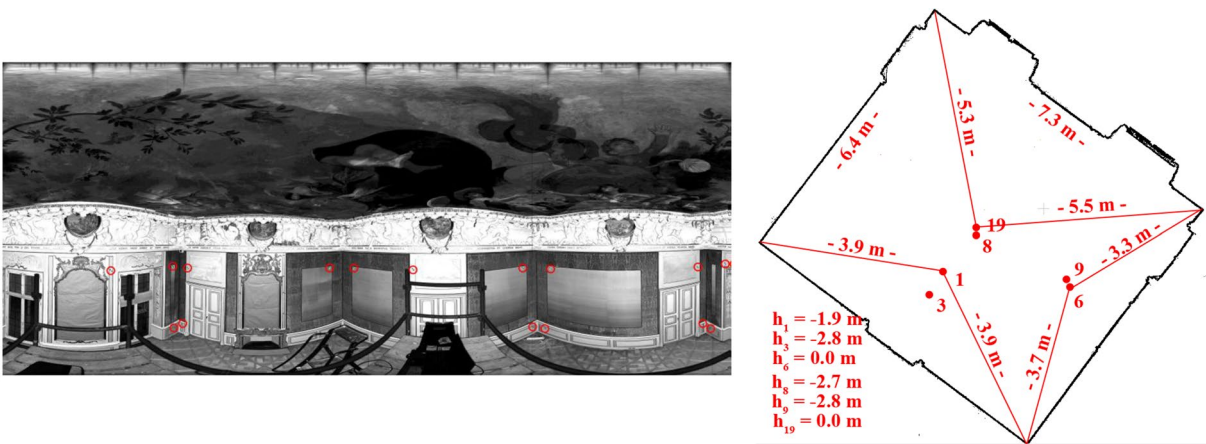
Fig. 5 The point cloud in the spherical projection of Test Site III with marked points (red circles) [63]
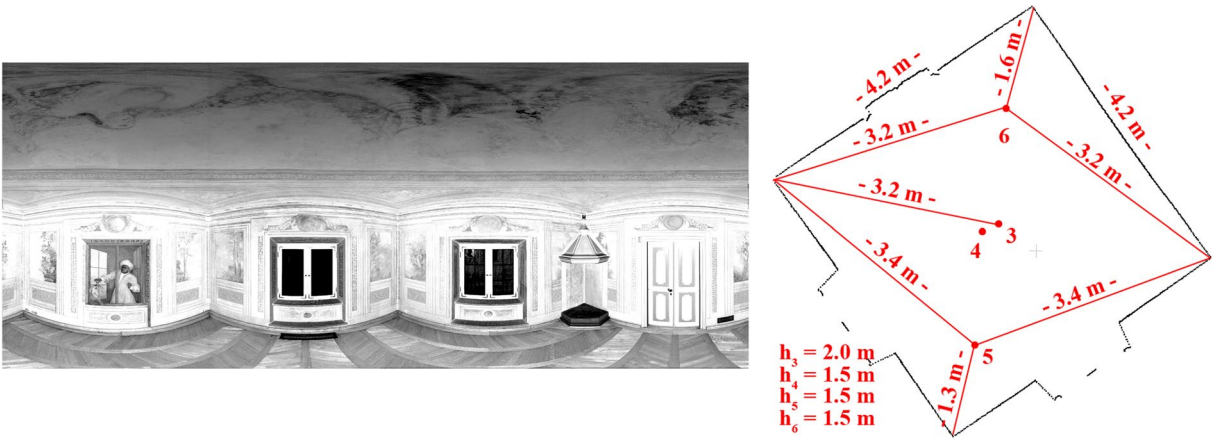
Fig. 6 The point cloud in the spherical projection of Test Site IV without marked points [63]
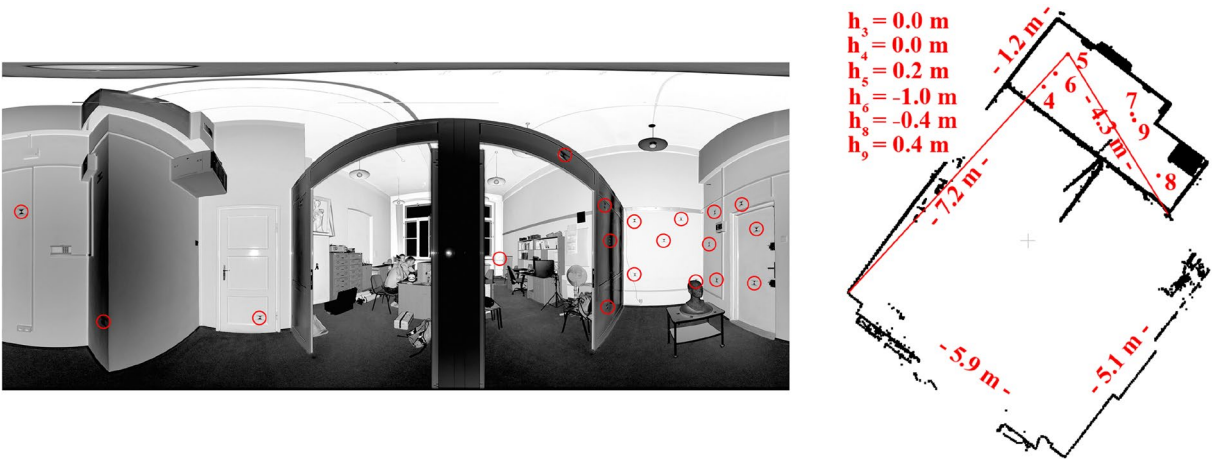
Fig. 7 The point cloud example in the spherical projection of Test Site V with marked check points (red circles) [63]

were measured with Total Station Leica TCRP 1202 with angular accuracy 2 s., linear accuracy 2 mm + 2 ppm. TLS data used in this work was acquired by phase-shift scanners Z + F 5006 h (Test Sites I, II and IV–VI) and Z + F 5003 from different positions and heights with an angular resolution 360°/320° and point resolution 6.3 mm/10 m (Test Site III). Figure 4 presents the floor plan with marked dimensions for Test Sites I and II, including Terrestrial Laser Scanning (TLS) positions and marked reference points.

The Test Site I is a regular-shaped facility with dimensions of approx. 5.6 m × 5.1 m. A ventilation pipe runs through the centre of the room (halfway up the room) and is used to dehumidify the room, which limits the placement of the scanner stations. It was necessary to increase the number of scanner positions used for a full Test Site inventory and the number of marked control points. The Test Site II has dimensions of 7.4 m × 5.1 m and is divided by curves at 1/3 and 2/3 of the distance. In addition, it has recesses and long windowpanes. Therefore, increasing the number of signalised points and scanner positions was necessary, which resulted in some points not being visible on all scans.

Test Sites III and IV are two decorated historical chambers at the Museum of King Jan III's Palace at Wilanów. Test site III: "The Queen's Bedroom" was characterised by geometric complexity in the form of rich ornaments, bas-reliefs, and facets. Moreover, mirrors in golden frames, decorative fireplaces, fabrics, etc., hung on the walls (Fig. 5). Test Site III is dimensions are approximately 6.4 m × 7.3 m × 5.3 m.

Figure 5 presents the distribution of scanner positions and the scanning distances. Five out of six scans were acquired with the selected fragment of a chamber (the incomplete extent). The seventh scan (acquired with the full angular resolution) was applied as the reference scan. Sixteen marked points were distributed over the test site

(considered as check points in further analyses), which were used for TLS data orientation.

Test site IV: "The Chamber with a Parrot" is characterised by the small number of ornaments and the lack of bas-reliefs, facets, or fabrics on the walls. In this Test Site, the walls were painted with patterns, which imitated spatial effects. Figure 6 presents the distribution of scanner positions and scanning distances, where the first scan was considered the reference scan. Due to the restriction on placing marked points on historical surfaces, automatically detected points defined as check points were used for the accuracy analysis. The dimensions of Test Site IV are approximately 4.2 m × 4.2 m × 2.6 m.

The Test Site V is the office room at the main hall of Warsaw University of Technology. The smooth walls characterise the selected Test Site without the texture; lamps and power wires were on the ceiling, and the floor was covered with dark carpet. Figure 7 presents the distribution of scanner stations and scanning distances. The dimensions of the office room are approximately 7.4 m × 5.9 m × 4.5 m.

The Test site VI is the "Empty shopping mall". The walls of the room were smooth, without texture. Lamps, electric wires, and an air-conditioning system were on the ceiling; the floor was concrete. Figure 8 presents the distribution of scanner stations and scanning distances. Scan three was used as the reference scan, and eight marked points were distributed over the test site (considered as check points in further analyses), which were used for TLS data orientation. The dimensions of the Test Site VI are approximately 21.5 m × 7.1 m × 6.3 m.

The TLS-SfM approach

The approach based on a modified SfM algorithm was used to register the TLS-derived point. Figure 9 shows a schematic of the data processing using the TLS-SfM method.
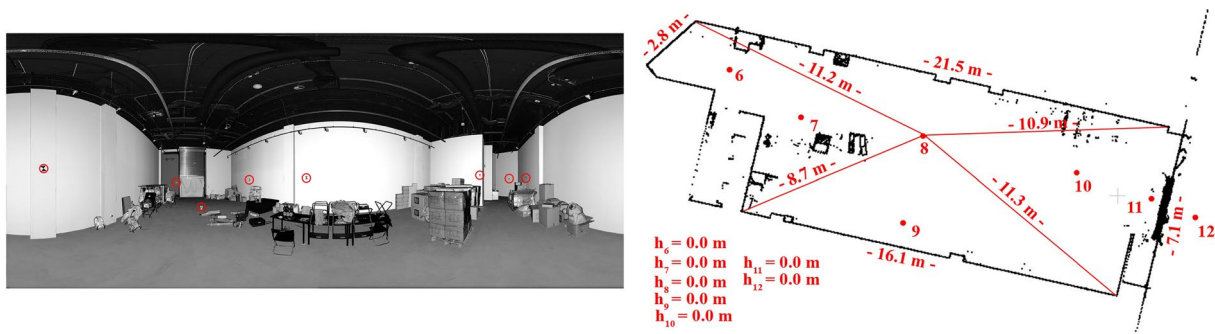
Fig. 8 The point cloud example in the spherical projection of test site VI with marked check points (red circles) [63]
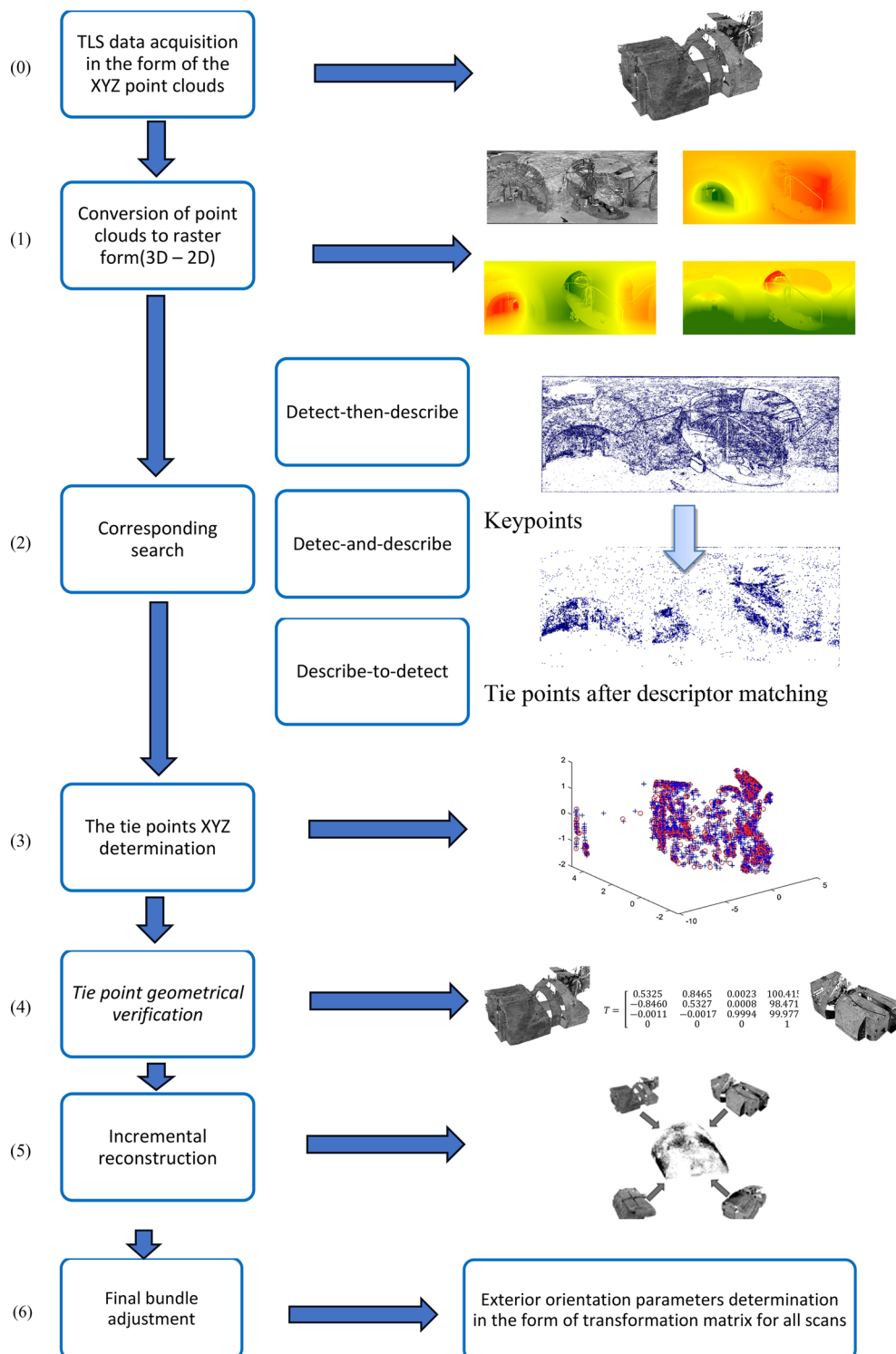
Fig. 9 Workflow of the proposed TLS-SfM point cloud registration approach

The TLS-SfM method is a multi-stage approach that consists of the following steps:

1) Conversion of point clouds to raster form (3D-2D).

   To convert point clouds to raster form, unprocessed raw data was selected to generate rasters with the maximum possible resolution (for each raster) and do not require interpolating the coordinate values for pixels. The mathematical relationship between cartesian and spherical coordinates was described by Fangi [44]. The data conversion from 3 to 2D consisted of converting the coordinates of the points from Cartesian to spherical based on Eqs. 6, 7, 8. The x and y coordinates in the raster area correspond to the values of the vertical and horizontal angles, respectively, and the intensity of the laser beam reflection and the X, Y, and Z coordinates of the points, respectively, are used to assign grey level values of the new raster. As a result of this step, 4 rasters are generated for each point cloud.

2) Corresponding search

   In the proposed TLS-SfM method, the process of finding tie points (feature detection and description) has been implemented using detect-the-describe, detect & describe (end-to-end) and describe-to-detect (end-to-end detector-free local feature matching methods) approaches. The detect-than-describe approach used a two-stage data transformation based on affine-based feature point detection and feature description using a descriptor. For both cases, both hand-crafted and learned-based algorithm approaches were used. A detailed description of the algorithms used is presented in subSect. "Overview of the investigated algorithms and evaluated criteria". This step is performed for all possible pairs of rasters. To determine these pairs, the methods of permutations without repetitions are used:

   $$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{9}$$

   where: k=2 (a pair of scans), n—the number of all scans.
   Descriptor matching (for detect-than-describe and end-to-end methods) is performed using the Approximate Nearest Neighbourhood-Based Point Matching algorithm and L2 distance metrics.

3) Tie points XYZ determination

   The 2D coordinates of the pre-matched tie points detected on the intensity rasters were used to interpolate the coordinates of the XYZ points. The X, Y and Z rasters generated in the first data processing step were used for this purpose, respectively. The bilinear method was used as the interpolation method.

4) Tie point geometrical verification

   The geometrical verification of the detected tie points (based on 3D coordinates, performed in the iterative process (RANSAC method) with the following assumptions—full registration (the accuracy on control and check points do not exceed 5 mm and covariances factors are higher than 0.5), initial registration used for final registration bases on the ICP (threshold 10 mm) and non-registration (values on control and check higher than 10 mm). The output of this data processing step was (1) the set of correct tie points, (2) the linear RMSE value of the scan pair match, (3) the number of tie points and (4) approximate transformation parameters.

5) Incremental reconstruction

   The *Incremental reconstruction* process starts with selecting the reference scan to which the other point clouds will be registered. To do this, the pair of point clouds for which the highest number of tie points was first detected is selected. From this pair of points, the point cloud with more connections to the other scans is selected. To match the remaining pairs of scans, the process is performed iteratively according to the following steps:

   (a) Localise a new pair of scans to the current pre-registered point clouds,
   (b) Compute the approximate point clouds registration parameters,
   (c) Find correspondence points on multiple point clouds,
   (d) Repeat steps a-c until all pairs of scans have been added.
       The result of this stage is an approximation of the mutual orientation parameters and all possible connections between point clouds.

6) Final bundle adjustment
   A final bundle adjustment is based on early iterative matching of the point clouds to the reference scan. This involves determining the orientation elements of the point clouds with simultaneous filtering of outlier

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 12 of 26

observations based on RMSE error values and reliability coefficients. In addition, based on the measured control points, it is possible to orient the point clouds to the reference coordinate system. As a result of the TLS-SfM process, point cloud orientation elements are obtained in the adopted reference system.

Overview of the investigated algorithms and evaluated criteria

This study investigates the quality improvement and completeness of the TLS registration process using 2D raster data and affine-detectors. To compare and verify the results of the point cloud registration, based on the selected hand-crafted and learned features, the multistage TLS-SfM registration methodology was followed.

(1) Hand-crafted affine detectors, namely, corner detector (AFAST) and blob detectors (ASURF and ASIFT), were tested. The use of affine in feature point detection involves two steps: (a) multiple virtual image generation (which includes the skew, tilt, and rotation) to simulate the influence of the affine and (b) for each virtual image, apply the detector:

- FAST (Features from Accelerated Segment Test) [76] utilises corner keypoints in images to detect by comparing the brightness intensities of pixels in a circular neighbourhood around each pixel of interest. The technique will classify the pixel as a corner depending on the neighbourhood's brightness and number of contiguous pixels and then to the central pixel using a threshold value. The FAST corner detector is based on a decision tree structure that allows for quick evaluation of the pixel intensities, making it suitable for real-time applications.

- SIFT (Scale-Invariant Feature Transform) [74]—the purpose of SIFT is to detect and describe distinctive image keypoints. The advantage of this technique is its invariant nature to the scale changes, rotations, and changes in illumination, which makes it robust to variations in image conditions. The working principle of the SIFT algorithm is identifying stable keypoints using a scale-space representation of the image and applying a Difference of Gaussians (DoG) operator to detect local extrema. These keypoints are then described based on their surrounding gradient orientations, resulting in highly distinctive and invariant feature descriptors.

- SURF (Speeded-Up Robust Features) [75] offers faster computation. It provides robustness against image transformations by utilising integral images to efficiently calculate various image filters, such as the Haar wavelet responses, which capture both local intensity and orientation information. SURF detects keypoints by identifying locations with extreme responses in scale-space and orientation.

(2) Authors implemented the learned-based features:

- SuperGlue [94] for reliable correspondence between keypoints across different images. Unlike traditional hand-crafted methods, SuperGlue predicts the matching likelihood and establishes matches directly from the input data. It consists of two main components: (1) a learned embedding network and (2) a geometric verification module. The embedding network is used to map keypoints from two images into a shared feature space, where their similarity is measured. The geometric verification module uses the learned embeddings to estimate a geometric transformation between the keypoints and refine the matches. SuperGlue can leverage rich contextual information and handle challenging scenarios such as occlusions and viewpoint changes owing to jointly learning feature representation and the matching process.

- LoFTR (Local Feature Transformer) is an end-to-end detector-free local feature-matching method introduced by Sun et al. [92]. LoFTR creates dense pixelwise correspondences between images using a Transformer-based architecture. LoFTR directly predicts dense correspondences without needing a feature detector, unlike traditional approaches that require separate stages for feature detection, description, and matching. It operates in two steps: (1) coarse matching and (2) fine matching. LoFTR employs a selfattention mechanism in the coarse matching stage to allow each pixel to attend to its neighbours and capture their contextual information to create a pixelwise dense matching. The coarse matching stage is used to provide the initial estimation of correspondences. LoFTR uses a hierarchical refinement network to refine the initial matches in the matching stage. This network takes the initial correspondences and iteratively refines them by considering local spatial relationships and context. LoFTR improves the accuracy and reliability of the correspondences by iteratively refining the matches. LoFTR's Transformerbased architecture captures long-range dependencies and global contextual information, enhancing the quality of the dense correspondences. This approach eliminates the need for explicit feature detection and produces dense descriptors directly, leading to improved matching performance.

- KeyNet detector + AffNet + HardNet descriptor (later called KeyNetAffine)—is a combined handcrafted and learned method to detect features. KeyNet is a state-of-the-art keypoint detector [81] that leverages deep learning techniques to detect

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 13 of 26

Table 1 Metrics for evaluating the hand-crafted and learned-based features

Metrics	Description
Number and tie point distribution	The high number of points can lead to detecting the correct points used in the registration step. Additionally, it also affects the possible determination of tie points between multiple scans and the final robustness of the registration
The completeness of data registration	The data registration is understood as the ability to orientate all pair scans to each other with a minimum number of connections and determine the robustness and effectiveness of the TLS-SfM approach
The registration accuracy	The registration accuracy determines the final quality of matching between points clouds on marked check points. The final step of documentation generation results in the final accuracy of 3D models and documentation
The reliability assessment	The reliability assessment makes it possible to assess the correctness of the geometric distribution of the tie points in a fully automatic manner. By meeting the minimum requirements for the values of these coefficients, it is possible to assess whether a network of tie points is robust
The distance between a pair of point clouds	The analysis of the distance between point clouds allows independent control of the accuracy of matching whole and fragmented point clouds. Such a metric enables the accuracy of the fit to be assessed for objects for which it is not possible to distribute marked reference points. This allows the results to be compared independently with stat-of-the-art approaches, namely iterative closest point and target-based

distinctive image keypoints. KeyNet utilises a convolutional neural network (CNN) architecture, which is trained on large-scale datasets with annotated keypoints. KeyNet, to maximise the detection accuracy and robustness, identifies salient and repeatable keypoints, which allows for optimising the network parameters. This detector is highly adaptable to diverse image conditions due to excellent handling of variations in scale, rotation, and illumination, demonstrating outstanding performance in keypoint-based applications, namely, image matching, object recognition, and visual tracking. The HardNet is a feature descriptor used in computer vision applications, particularly for matching and recognition tasks. The HardNet descriptor [84] is designed to capture and encode distinctive information from image patches, making it robust to variations in scale, rotation, and lighting conditions. The descriptor is computed by extracting local patches around keypoints and encoding them into fixed-length feature vectors. HardNet can handle challenging scenarios, such as significant viewpoint changes and occlusions, owing to focusing on the most informative and discriminative patches. HardNet utilises a Siamese neural network architecture that learns to optimise the feature representation for improved matching accuracy. During training, pairs of matching and non-matching patches are used to learn discriminative feature embeddings.

To evaluate the accuracy of TLS point cloud on learned-based methods, it was decided to use those approaches trained on images depicting historical buildings and architectural objects (for LoFTR—MegaDepth, SuperGlue and KeyNetAffine—PhotoTurism, respectively).

The additional retrained learned-based descriptors were chosen due to the desire to test ready-made solutions and compare them with hand-crafted methods. The quality improvement and completeness of the TLS registration process were compared against several metrics presented in Table 1.

Results and discussion
Automatic pairwise point cloud registration- accuracy evaluation

To assess the detector's or affine-detector's applicability in the TLS registration process, the accuracy of the orientation of all possible overlapping pairs of scans from different heights and distances from scanned surfaces was analysed. The results are presented in Table 2 and marked in colour: (1) green—the complete registration with the X, Y and Z with RMSE ≤ 0.005 m and covariance factor > 0.5; (2) orange—preliminary orientation; obtained parameters should be treated as the initial parameters for Iterative closest Point (ICP) registration and (3) red—no registration because the points were not well distributed and/or the RMSE < 0.01 m and/or covariance < 0.5. Additionally, due to the processing of point clouds of wall fragments (rather than the entire room) on Test Site III, it was decided to mark "x" pairs of scans that do not overlap.

The results in Table 2 show that only AFAST (point detector) and ASIFT (blob detector) allow for correct registration of all pairs of scans for all test sites. The remaining algorithms should be analysed individually for each test site. The LoFTR approach obtained the worst results: for Test Site I, only 1 of 6; Test Site II, 0 of 15; Test

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 14 of 26

Table 2 The accuracy of the TLS registration for detectors and a-detectors

Test Site I- Basement at the Royal Castle — detectors: AFAST, ASIFT, ASURF, SuperGlue, LoFTR, KeyNetAffine (columns 2, 3, 4; rows 1–3)

Test Site II- Basement at the Royal Castle — detectors: AFAST, ASIFT, ASURF, SuperGlue, LoFTR, KeyNetAffine (columns 2, 3, 4, 5, 6; rows 1–5)

Test Site III – "The Queen's Bedroom" — detectors: AFAST, ASIFT, ASURF, SuperGlue, LoFTR, KeyNetAffine (columns 2, 3, 4, 5, 6; rows 1–5)

Test Site IV – "The Chamber with a Parrot" — detectors: AFAST, ASIFT, ASURF (columns 2, 3, 4); SuperGlue, LoFTR, KeyNetAffine (columns 4, 5, 6; rows 1–3)

Test Site V - The office room — detectors: AFAST, ASIFT, ASURF, SuperGlue, LoFTR, KeyNetAffine (columns 2, 3, 4, 5, 6, 8, 9; rows 1–8)

Test Site VI – Empty shop (shopping mall) — detectors: AFAST, ASIFT, ASURF, SuperGlue, LoFTR, KeyNetAffine (columns 2, 3, 4, 5, 6, 7; rows 1–6)

Site III, 0 of 9; Test Site IV 6 of 6; Test Site V, 0 of 28 and Test Site VI 0 of 20 pairs of scans were correctly oriented (full orientation). For the other learned-based approaches for point detection, significantly better results were obtained. In the case of the SuperGlue detector for Test Site I, 2 of 6; Test Site II, 11 of 15; Test Site III, 8 of 9; Test Site IV, 6 of 6; Test Site V, 24 of 28 and Test Site VI 6 of 21 pairs of scans were correctly registered. With the KeyNetAffine, it was possible to register all pairs of scans from Test Site IV, 5 of 6 pairs of scans for Test Site I, 12 of 15 for Test Site II, 1 of 9 for Test Site III, 16 out of 28 for Test Site V and 3 of 21 for Test Site VI.

When the multi-position TLS point clouds are registered, not only the percentage of the correctly aligned

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 15 of 26

Table 3 The number of all tie points used in the full bundle adjustment and points for cases for which full bundle adjustment was impossible (marked with a cross)

Detector	Test Site I	Test Site II	Test Site III	Test Site IV	Test Site V	Test Site VI
AFAST	70,977	108,218	5144	146,662	37,215	877
ASIFT	60,308	123,808	3392	106,386	14,597	496
ASURF	27,871	92,659	1184	128.360	13,290	648
SuperGlue	3024	6518	1916	15,796	8268	1264[x]
LoFTR	3295	1359	293[x]	16,907	324[x]	0
KeyNetAffine	2726	7682	1838	6938	7387	176[x]

point cloud is necessary, but also the possibility of a global registration for all possible point clouds. The full registration (based on results of full and preliminary pair of scans orientation) for Test Site I, II, III, IV and V. For Test Site VI, it was impossible to perform the multi-position registration. The incompleteness of a pair of scan registrations for Test Site I and IV might affect the robustness of the global adjustment and approximately equivalence redundancy of the tie point on point clouds.

The hand-crafted detectors are the potential solution to overcome the problems mentioned above. Table 2 shows that the full multi-stage registration was conducted for Test Sites I–V. The worst results were obtained for Test Site VI, for which full registration was only possible with the ASIF and AFAST detectors.

The analyses of the performance of point/blob detectors and a-detectors on test fields characterised both by different textures, structures, numbers, and decorations and by scanner positions to varying distances from walls and heights demonstrated that:

- Using the LoFTR approach, it was not possible to correctly register point clouds obtained by scanner positions, for which corresponding fragments were measured at significantly different angles to the normal vector surface (i.e., acute angles to the normal vector surface) and for significantly different distances from the scanner position. This influenced the occurrence of significant "distortions" in the spherical projection caused by the cartographic conversion of the 3D data from the 2D form.
- Hand-crafted algorithms allow more resistant tie points to be detected, which translates into more correctly oriented scan pairs. The SIFT and SURF algorithms are based on greyscale gradients, making them scale-invariant and more robust. The performance difference is based on using a filter (CenSurE and SiFT—Laplasian centre-surround and Difference of Gaussian algorithms, respectively) and a Hessian (SURF and

Difference of Boxes detector). For this reason, with these detectors, it was possible to detect a higher number of correctly matched keypoints, which affected the higher number of correctly registered pairs of scans.
- Applying affine significantly improved the quality of the TLS point cloud pairwise and multi-stage registration. The use of ASIFT and AFAST allowed the orientation of point cloud pairs, necessary for final multi-position registration, for all Test Sites. This is also noticeable when applied to the KeyNetAffine approach. Compared to other learned-based methods, it was possible to orient more pairs of scans with a wide baseline (Test Site I, II and VI). For the orientation of short baseline pairs of scans characterised by high distortion (Test Site III and V), significantly better results were obtained for the SuperGlue approach.

The number of detected and matched keypoints after the final bundle adjustment

The number of tie points obtained after the full bundle adjustment process was analysed to assess the influence of the hand-crafted and learned features in the TLS registration process and the selection of the appropriate features. Table 3 presents the number of all tie points used in the full bundle adjustment and points for cases for which full bundle adjustment was impossible (marked with a cross).

The number of used tie points presented in Table 3 indicated that hand-crafted detectors recorded the highest number of keypoints for all Test Sites apart from Test Site VI, for which the SuperGlue approach detected the most points. When considering the ratio of the number of points detected by the hand-crafted versus learned-based approach, it can be concluded that 26 times more were detected for Test Site I (AFAST—KeyNetAffine), 91 for Test Site II (ASIFT—LoFTR), 2.8 for Test Site III (AFAST—KeyNetAffine), 21 for Test Site IV (AFAST—KeyNetAffine), and 5 for Test Site V (AFAST—KeyNetAffine). Due to the lack of full bundle adjustment of all

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 16 of 26
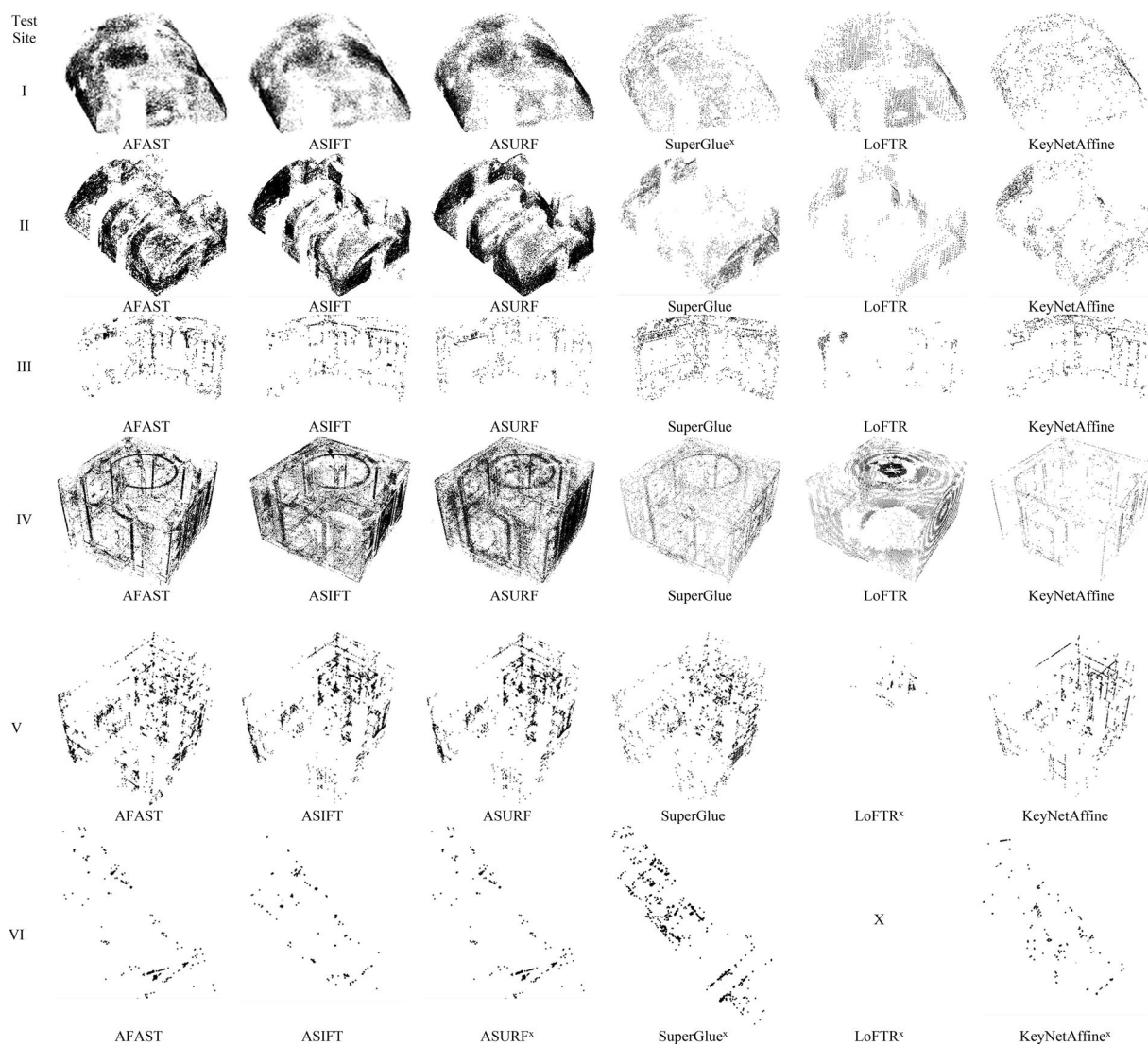
Fig. 10 The tie points distribution used for TLS point cloud registration for each method

scans using Learned features, it was impossible to calculate the points ratio for the two approaches.

The analyses presented in Table 3 also show that, on average, the most tie scores were detected for AFAST and the least for KeyNetAffine. The significant difference in the number of points detected for the two approaches for Test Site I, II and IV is due to the characteristics of the sites. Test Site I and II is a historic brick cellar with an arched ceiling, and Test Site IV is a room with paintings imitating the spatial effect. For this reason, hand-crafted detectors, notably the AFAST detector (due to its mode of operation), detect significantly more points than other Test Sites characterised by less such unambiguous detail.

Their spatial distribution should also be considered when assessing the quality of the tie points used in the

bundle adjustment process. This is crucial, as it impacts the quality of registration and the accuracy of the entire process. Figure 10 shows the distribution of points used in full bundle adjustment and points for cases for which full bundle adjustment was impossible (marked with a cross).

The analysis shows that despite the lower number of tie points detected by Learned-based methods compared to Hand-crafted detectors, their placement guarantees a correct point cloud registration. As with the number of points analysed, the distribution of points should be assessed independently for each Test Site:

- Test Site I—The points detected by the hand-crafted detectors for all detectors have a similar spatial dis-

tribution. Noticeably, the issues are clustered in the lower part of the room and the middle of the ceiling. An uneven distribution characterises points detected using the LoFTR algorithm, and an increased density of points on wall sections is noticeable. For KeyNetAffine, the points are evenly distributed, and unlike for LoFTR, there are no areas with a significantly higher point density. When analysing the results for SuperGlue, there is a significant density of points in one part of the basement due to the inability to detect tie points on the minimum number of pairs of scans mortising full bundle adjustment.

- Test Site II—The distribution of scores for all methods is similar for Test Site I. For the hand-crafted algorithms, the most points (highest density) were detected and used on the two walls visible on all scans. Significantly fewer points are on the ceiling, and the highest density was obtained in the central part of the basement. The best results were obtained for the ASURF, AFAST and ASIFT algorithms. For the learned-based algorithms, the best distribution of points (both points were on the ceiling and the walls) while maintaining a similar density for the entire basement was obtained for KeyNetAffine and the worst for LoFTR, for which points were mainly distributed on the walls in groups of different thicknesses. For the SuperGlue method, most points were distributed on the walls mapped on all scans and a

small number on the ceiling. However, it should be emphasised that the number and distribution of points detected by the learned-based methods allowed the correct registration of all point clouds.

- Test Site III—For Test Site III, which contains rich ornaments, bas-reliefs, and facets, the distribution of tie points was similar for all hand-crafted and learned-based methods except for the LoFTR algorithm. In summary, it can be concluded that the best distribution was obtained for points detected using the SuperGlue approach.
- Test Site IV—As for Test Sites I and II, in this case, a higher point density for points detected by hand-crafted methods. For this type of algorithm, it is noticeable that there is a higher point density for areas where there is a more significant change in grey degree gradients. For this reason, these points are not evenly distributed throughout the study area. For learn-based methods (SuperGlue and KeyNetAffine), the distribution of points is more even than for hand-crafted methods. As for the previous Test Sites of the learned-based algorithm group, the most points were detected using the SuperGlue approach, the least using LoFTR.
- Test Site V—In the case of an office room test field characterised by a lack of diverse texture and equipped with furniture and office equipment, the number, density, and distribution of tie points were

Table 4 Comparison of results of TLS joint/full registration method for all scans and the target-based registration method with reliability assessment for all Test Sites

RMSE on marked check points [mm]						
Detector	Test Site I	Test Site II	Test Site III	Test Site IV	Test Site V	Test Site VI
AFAST	3.4	3.9	2.3	2.5	2.0	9.0
ASIFT	1.8	2.3	1.9	2.8	1.8	4.9
ASURF	3.7	5.0	2.5	2.6	2.1	x
SuperGlue	x	4.1	4.7	2.6	1.9	x
LoFTR	4.2	10.5	x	2.7	x	x
KeyNetAffine	3.6	2.3	5.5	2.5	2.4	x
Target-based method	3.5	4.4	5.7	2.8	1.3	3.8
The reliability indices—minimum values						
Detector	Test Site I	Test Site II	Test Site III	Test Site IV	Test Site V	Test Site VI
AFAST	0.94	0.98	0.66	0.99	0.94	0.10
ASIFT	0.98	0.78	0.58	0.99	0.65	0.60
ASURF	0.97	0.99	0.94	0.99	0.93	x
SuperGlue	x	0.28	0.51	0.98	0.87	x
LoFTR	0.76	0.59	x	0.94	x	x
KeyNetAffine	0.51	0.89	0.86	0.98	0.86	x
Target-based method	0.35	0.20	0.22	0.23	0.29	0.28

x too low number of tie points

similar for the AFAST, ASIFT, ASURF, SuperGlue and KeyNetAffine algorithms. As for the previous Test Sites, the worst results were obtained for the LoFTR-based approach, for which all point clouds could not be registered.

- Test Site VI—An analysis of the distribution of tie points detected on the empty shopping mall scans shows that only hand-crafted ASIFT and AFAST detectors could orient all point clouds. This was due to the conversion of the 3D data to 2D and the influence of the presence of significant distortion in the image. Considering that points were searched on wide-based point clouds, applying the abovementioned methods allowed the detection of an adequate number of points evenly distributed over the entire study area. Comparing the results for points detected on rasters generated from pairs of scans with smaller baseline between point clouds and less distortion, the use of learned-based methods allowed the detection of a more significant number of correctly detected tie points. For this reason, when planning a survey of this type of object, it is crucial to decide whether to make fewer point clouds and use affine-detector-based hand-crated methods or to add several scanner stations to reduce the baseline between point clouds and use learned-based algorithms.

The comparison with the current state-of-the-art methods
To assess the accuracy and correctness of the presented approach for point cloud orientation based on affine-detectors and point clouds converted to raster form, it was decided to compare point clouds with the commonly used approach based on signalised control points (target-based registration) implemented in Z+F LaserControl software [47] and the Iterative Closest Points (ICP) method implemented in the open-source CloudCompare [48].

The target-based
The target-based method relies on the marked points and is commonly applied for TLS point cloud registration. These points should be evenly distributed across the investigated object. To compare results from the feature-based registration method with "normal" and affine detectors, the obtained results were compared with the TLS target-based registration from Z+F LaserControl software. To automatically analyse the influence of the geometrical point distribution with reliability assessment, the values of the covariance factors were compared. Results are shown in Table 4.

Results presented in Table 4 show that the differences between the RMSE values on marked check points (obtained from multi-position TLS registration) depend on Test Sites.

- For Test Site I, significantly higher accuracy of full-bundle adjustment can be observed on points detected with the ASIFT detector compared to the commonly used Target-based approach. The linear RMSE value was 2 times lower (1.8 mm). For the other algorithms, the linear RMSE values were simi-
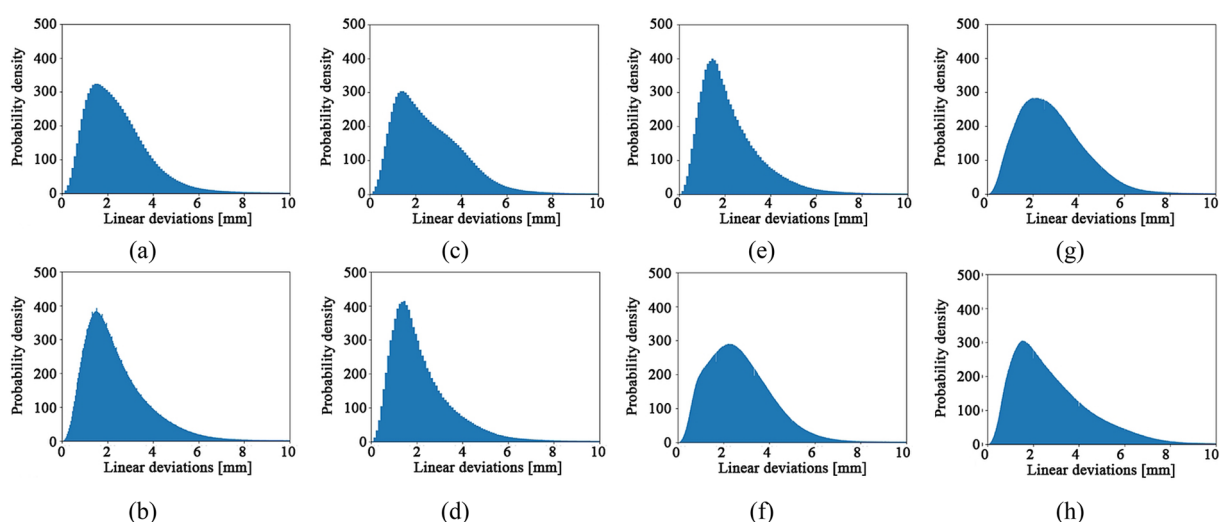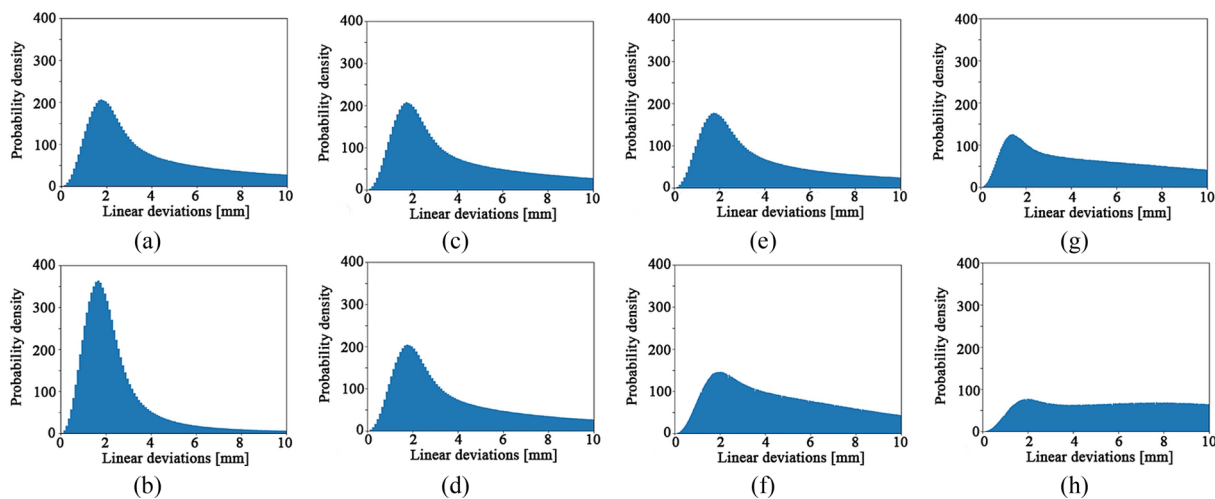

Fig. 11 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site I: **a** target-based method, **b** ICP point-to-point, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine
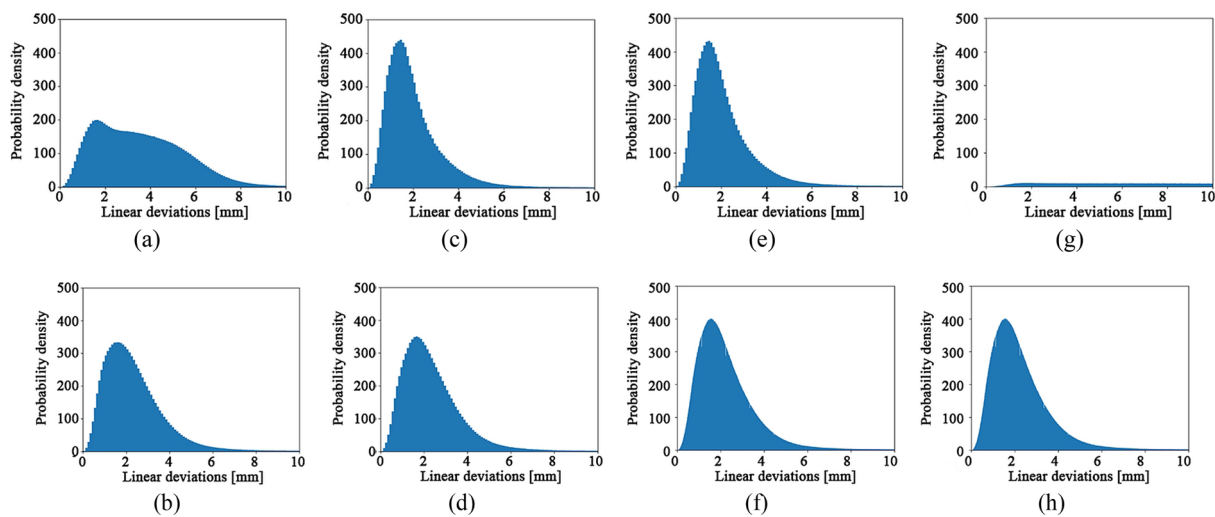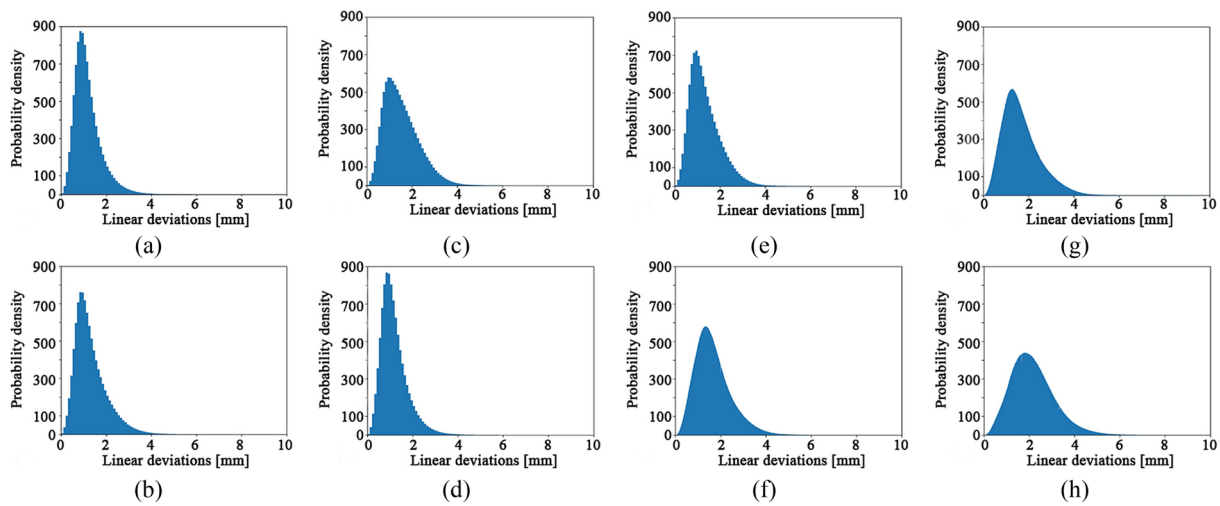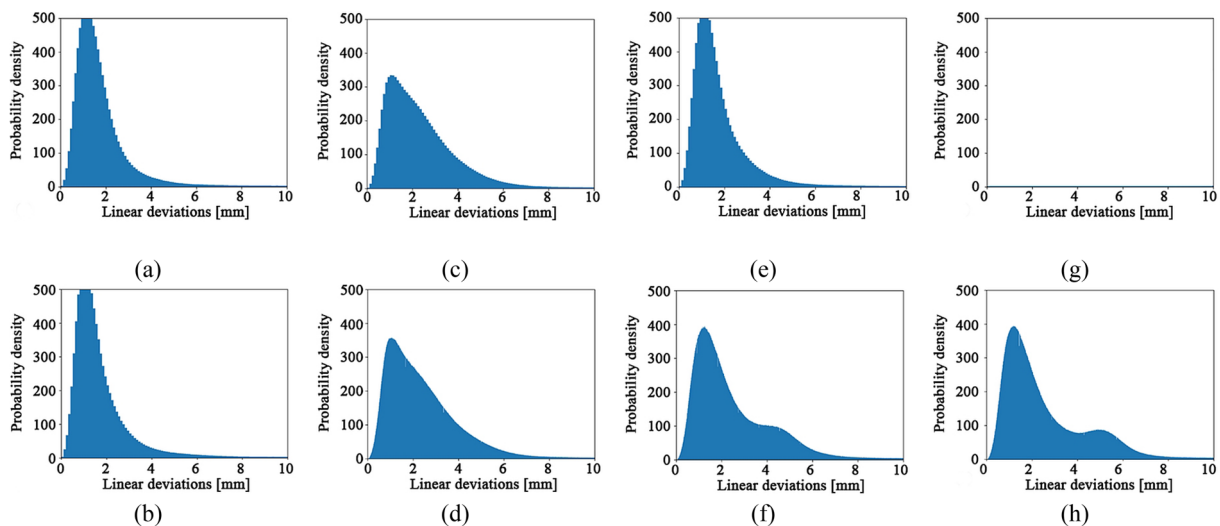
Fig. 12 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site II: **a** target-based method, **b** ICP point-to-point, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine

Fig. 13 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site III: **a** target-based method, **b** ICP point-to-point, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine

lar to those of the Target-based approach and were for AFAST—3.4 mm, ASURF—3.7 mm, KeyNetAffine—3.6 mm and Target-based 3.5 mm, respectively. For the LoFTR-value method, the RMSE was 4.2 mm (0.7 mm higher) than the Target-based approach. The significant impact of using a Hand-crafted detector can be seen by analysing the minimum covariance factors. This contributed to fulfilling the network's controllability condition and improving the geometric distribution of tie points for the minimum values (above 0.5, which is the threshold value). There is a noticeable increase in values from 0.35 for Target-

based to 0.94 for AFAST, 0.98 for ASIFT, 0.97 for ASURF, 0.76 for LoFTR and 0.51 for KeyNetAffine.

- For Test Site II, varying linear RMSE values are evident. The best results were obtained on points detected with ASIFT and KeyNetAffine—linear RMSE values of 2.3 mm—2 times lower than for Target-based. For AFAST and SuperGlue, the linear RMSE values are lower than for Target-based. Only for ASURF, which is 0.6 mm higher than Target-based and LoFTR—6.1 mm. Analysing the values of the minimum reliability indices, as for Test Site I, a significant increase in their values (which translates into a better geometric distribution and resistance

Markiewicz *et al. Heritage Science*   (2023) 11:254
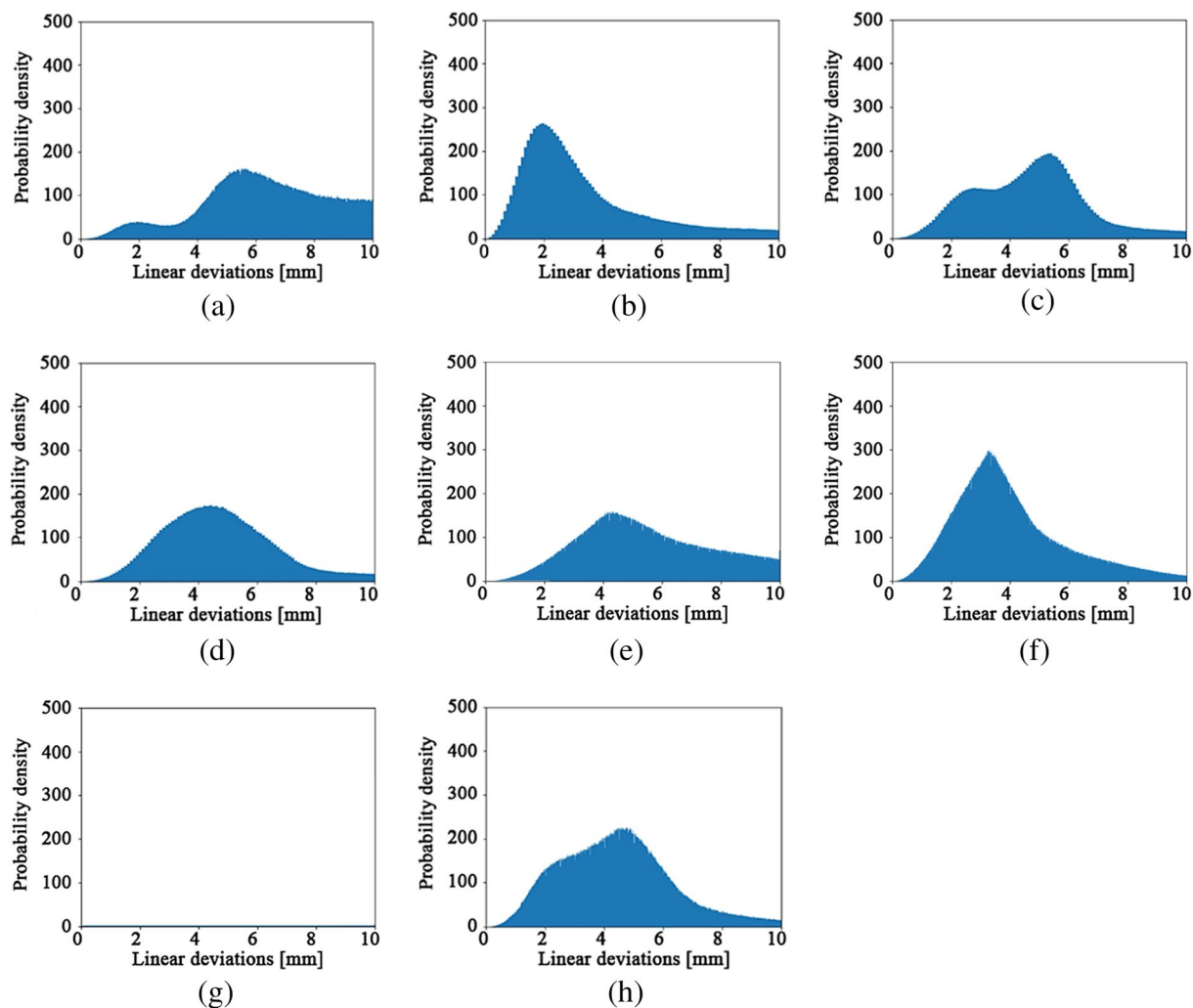
Page 20 of 26

Fig. 14 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site IV: **a** target-based method, **b** ICP point-to-point, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine

Fig. 15 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site V: **a** target-based method, **b** ICP point-to-point, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine

to the influence of outliers) for all methods except SuperGlue.

- In the case of Test Site III, the RMSE's deviation on detectors is approximately 2 times lower than Target-based (5.7 mm) for Hand-crafted detectors and similar to Target-based (but still lower) for Learned-based approaches. The covariance factor for the Hand-crafted method is in the range of 0.58–0.98, for Learned-based methods in the range of 0.51–0.86 and for target-based is 0.22. As mentioned, full registration for all scans with the LoFTR algorithm was impossible.

- For Test Site IV, both Hand-crated and Learned features provided comparable results; therefore, it is difficult to judge if it is necessary to use the Learned-based method, as the obtained mean RMSE values for detectors and target-based method are similar. The minimum covariance factors values (about 0.98) are about 4.5 times better than the target-based method (0.23).

- For Test Site V, similar results for Hand-crafted (2.5 mm–2.8 mm) and Learned-based methods (1.9 mm–2.4 mm) but slightly worse than Target-based (1.3 mm). The minimum covariance factor for both methods is in the range of 0.65–0.94, and for

Fig. 16 The probability density histogram of linear deviations between the worst oriented pair of scans for Test Site VI: **a** target-based method, **b** ICP point-to-point—CloudCompare, **c** AFAST, **d** ASIFT, **e** ASURF, **f** SuperGlue, **g** LoFTR, **h** KeyNetAffine

target-based is 0.29. In this case, orienting the point clouds using LoFTR-detected points was also impossible.

- Completing the multi-station registration scans for Test Site VI was impossible due to the challenge of finding the corresponding points for Hand-crafted and Learned methods, except ASIFT and AFAST. Comparing values of RMSE, similar values can be seen for ASIFT and target-based methods. However, the AFAST detector demonstrated approximately 2–2.5 times worse performance. The min covariance factors for the AFAST, ASIFT and target-based methods were 0.1, 0.60 and 0.28, respectively.

Iterative closest points (ICP)

To assess the accuracy of TLS data registration using affine-detectors, the results were compared with the point-to-point ICP method using open-source CloudCompare software, commonly used in point cloud registration. The quality of point cloud matching was assessed by analysing the linear distance between pairs of point clouds. Point cloud resampling was performed with a fixed distance (1 mm) between points. Figures 11, 12, 13, 14, 15, 16 show the example of the worst scenario for all Test Sites. Each figure contains 8 histograms showing the probability density function of linear deviations between point clouds using the target-based method, the ICP point-to-point, Hand-crafted detectors (AFAST, ASIFT, ASURF) and Learned-based features (SuperGlue, LoFTR and KeyNetAffine).

Based on the analysis of the results for Test Site I (Fig. 11), it can be seen that results obtained from ASIFT,

ASURF, SuperGlue, LoFTR, Target-based and ICP methods are similar to a chi-square distribution. Still, better results are obtained from the detector-based approach.

For Test Site II (Fig. 12), all histogram shapes except the Target-based and LoFTR methods are similar to a chi-square distribution. The distance for 95% of the points for the Target-based method algorithm does not exceed 6 mm. The histogram peak of probability density histogram of linear deviations between the worst oriented pair of scans by LoFTR shows that deviations are higher than 10 mm and registration was performed incorrectly.

Test Site III's best point cloud matching results were obtained for the ICP-based approach (Fig. 13b). The results obtained from Hand-crafted detectors (Fig. 13c–e) are similar to those obtained from target-based registration (Fig. 13a). The peaks of histograms are approximately 2 mm. The shapes of the linear deviations histograms for Learned-based approaches (Fig. 13f–h) are "flat", indicating more significant errors in deviations between point clouds than for Hand-crafted methods.

Based on the analysis of the results for Test Site IV (Fig. 14), the results obtained from all methods (except KeyNetAffine) are similar to a chi-square, which were obtained by the Target-base and ICP point-to-point approaches. Despite not obtaining a chi-square distribution for the KeyNetAffine methods, it should be considered that the scans were oriented correctly as, for 95% of the points, the distance does not exceed 4 mm, which does not exceed a scanning point resolution of 6 mm/10 m.

Results obtained for Test Site V (Fig. 15) show that point clouds were oriented correctly using algorithms based on Hand-crafted detectors. In contrast, for the ASIFT detector, the distribution of values takes the shape of a chi-square distribution and coincides with histograms obtained for the target-based and ICP methods. Similar to the results obtained for Test Site IV (not chi-square distribution of other detectors) for the Learned-based approach, the deviations of 95% of the points do not exceed 6 mm, which does not exceed a scanning point resolution of 6 mm/10 m.

The worst results for comparing point cloud distances were obtained from empty shop using the Target-based method (Fig. 16a). This was due to the 12 mm/10 m scanning resolution, which translated into point cloud density and the ability to identify signalised points. For this reason, it is recommended to use the ICP method, which allows for the correct orientation of the data. Despite this, the probability density histogram of linear deviations between the worst oriented pair of scans shows that the distances between clouds do not exceed the accepted scanning resolution of 12 mm/10 m, which can be considered an acceptable registration result.

In summary, the data orientation results presented using an affine-detector allow robust registration, and choosing the ASIFT detector allows for complete data registration.

Conclusion

This article evaluated the quality improvement and completeness of the TLS registration process using 2D raster data from spherical images and Hand-crafted and Learning features in the multi-stage TLS point cloud registration. For this study, to compare and verify the detectors and A-detectors, the Royal Castle in Warsaw without decorative structure (Test Site I and II), Museum of King Jan III's Palace at Wilanow with decorative elements, ornaments, and materials on walls (Test Site III) and flat frescos (Test Site IV), narrow office (Test Site V) and shopping mall (Test Site VI) were used. The performed experiments demonstrated that:

- The proposed TLS point cloud registration approach is a fully automatic solution independent of the object's interior type.
- The selection of a suitable detector should depend on the test site being measured. In the case of cultural heritage interiors (characterised by a good texture and number of ornaments), it is possible to use both Hand-crafted detectors AFAST, ASURF, ASIFT and Learned-based SuperGlue and LoFTR. For the point cloud registration of public buildings, it is recommended to use detectors such as AFAST or ASIFT. On the other hand, using the ASIFT detector allowed for point cloud registration regardless of the geometry dependencies between individual scans and the test field being developed.
- It is recommended to use the ASIFT or AFAST detector for TLS point cloud registration because these detectors could perform the multi-station registration at all Test Sites. Another solution might be to consider increasing the number of posts to minimise significant deviations on spherical images and use Learned methods, namely SuperGlue and KeyNetAffine.
- The use of the affine hand-crafted detectors allows for detecting the high number of tie points, improving the accuracy and completeness of the TLS registration process compared to the learning-based approach. The number of ties detected increased for cultural heritage sites by 21–91 times and for public objects by about 2.8–5 times.
- In analysing the accuracy of point cloud orientation on signalised check points, two cases should be considered separately, i.e., decorated rooms and public facilities. For decorative sites, the smaller values can be observed for linear RMSE errors

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 23 of 26

for hand-crafted features (values approximately 2 times smaller) than those obtained by the Target-based approach and similar to Target-based values for the Learned-based approach. When comparing the results obtained for public interiors, it can be observed that similar accuracies to the target-based method were obtained for hand-crafted features and learned-based (where it was possible to register all scans). That proves that using a-detectors for point cloud orientation is correct and reasonable.

- For low internal reliability indices, we have relatively low controllability of observations and thus low detection of outliers at the reference points. An important consideration is the number of points and their geometric distribution. In the target-based method, it is challenging to distribute many points and sometimes even impossible, while in the feature-based approach, a large number of points are automatically detected. A large number of points distributed over the entire surveyed object allows for relative control of points and the correct removal of outliers.

- By analysing the internal reliability indices, using a-detectors allows for increased controllability of points and the detection of outliers in the dataset. This fulfilled the network's controllability condition, with 0.5 being the acceptable threshold value. Comparing results obtained from Hand-crafted and Learned features with values obtained for the points detected with the Target-based method, it can be observed that for Test Site I, the minimum value is 0.51–0.97, while for the target-based method, the minimum is 0.35. For Test Site II, the minimum is between 0.59 and 0.98 (only for SuperGlue is 0.28), while for the target-based method, the average is 0.20. For Test Site III, the average minimum covariance factors values (0.71) are about 3.2 times better compared to the target-based method (0.22); for Test Site IV, the minimum covariance factors for the targets-based method is 0.23 and about 4 times worse than the detector-based method. In the case of Test Site V, the minimum covariance factor for the detector-based method is in the range of 0.65–0.94, while for the target-based method, it is 0.29 and for the Test Site VI, the minimum covariance factors are 0.10, 0.60 and 0.28 for AFAST, ASIFT and target-based, respectively.

- The proposed robust method for point cloud registration based on intensity rasters (together with a depth map) and affine-detectors allows us to obtain similar results as commonly used target-based and Iterative Closest Points methods. The advantage of the proposed approach for point cloud orientation over the Target-based method is that more automatically detected tie points are used for orientation, with better spatial distribution and robust outliers detection regarding the reliability theory. When registering point clouds using the ICP method, the clouds must be pre-oriented, as this guarantees the correctness of the final registration. In the affine-detectors approach, such a condition is not required since the selection and elimination of tie points are utilised in a two-step manner through descriptor matching and geometrical verification based on the RANSAC algorithm.

- The obtained TLS registration results based on learned-based methods (on data trained on the images by the authors of the solutions) attest to high performance and use in data orientation. To further improve the accuracy and completeness of the data orientation on objects with poorer texture and less ornamentation (Test Sites V and VI), the authors plan to prepare a test dataset based on intensity rasters based on TLS point clouds.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Faculty of Geodesy and Cartography, Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland. [2]Built Environment and Sustainable Technologies (BEST) Research Institute, Liverpool John Moores University, Liverpool L3 3AF, UK. [3]Faculty of Mechatronics, Warsaw University of Technology, św. Andrzeja Boboli 8, 02-525 Warsaw, Poland.

References
1. Mukupa W, Roberts GW, Hancock CM, Al-Manasir K. A review of the use of terrestrial laser scanning application for change detection and deformation monitoring of structures. Surv Rev. 2016. https://doi.org/10.1080/00396265.2015.1133039.

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 24 of 26

2. Dong Z, Yang B, Liang F, Huang R, Scherer S. Hierarchical registration of unordered TLS point clouds based on binary shape context descriptor. ISPRS J Photogramm Remote Sens. 2018;144:61–79. https://doi.org/10.1016/j.isprsjprs.2018.06.018.

3. Vacca G, Mistretta F, Stochino F, Dessi A. Terrestrial laser scanner for monitoring the deformations and the damages of buildings. Int Arch Photogramm Remote Sens Spat Inf Sci. 2016;XLI-B5:453–60. https://doi.org/10.5194/isprs-archives-XLI-B5-453-2016.

4. Rashidi M, Mohammadi M, Sadeghlou Kivi S, Abdolvand MM, Truong-Hong L, Samali B. A decade of modern bridge monitoring using terrestrial laser scanning: review and future directions. Remote Sens. 2020;12:3796. https://doi.org/10.3390/rs12223796.

5. Wang W, Zhao W, Huang L, Vimarlund V, Wang Z. Applications of terrestrial laser scanning for tunnels: a review. J Traffic Transp Eng. 2014;1(5):325–37. https://doi.org/10.1016/S2095-7564(15)30279-8.

6. Bosché F. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. Adv Eng Informatics. 2010;24(1):107–18. https://doi.org/10.1016/j.aei.2009.08.006.

7. Lu-Xingchang, Liu-Xianlin. Reconstruction of 3D model based on laser scanning. In: Zlatanova, S, Coors, V editors. Innovations in 3D Geo information systems. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg; 2006. p. 317–32. https://doi.org/10.1007/978-3-540-36998-1_25.

8. Truong-Hong L, Lindenbergh R. Measuring deformation of bridge structures using laser scanning data. In: 4th Jt Int Symp Deform Monit. Athens, Greece; 2019. https://pure.tudelft.nl/ws/portalfiles/portal/86837891/2019_JISDM_LTH_RL_Final.pdf. Accessed 18 Nov 2023.

9. Truong-Hong L, Lindenbergh R. Inspecting structural components of a construction project using laser scanning. In: Ungureanu L-C, Hartmann T, editors. EG-ICE 2020 Work Intell Comput Eng Proc. Universitatsverlag der TU Berlin; 2020. p. 352–62. http://resolver.tudelft.nl/uuid:2fd1d6b6-81a6-4ab1-81a8-2c2fb85708a9. Accessed 18 Nov 2023.

10. Truong-Hong L, Laefer DF, Hinks T, Carr H. Combining an angle criterion with voxelization and the flying voxel method in reconstructing building models from LiDAR data. Comput Civ Infrastruct Eng. 2013;28:112–29. https://doi.org/10.1111/j.1467-8667.2012.00761.x.

11. Siwiec J, Lenda G. Integration of terrestrial laser scanning and structure from motion for the assessment of industrial chimney geometry. Measurement. 2022;199:111404. https://doi.org/10.1016/j.measurement.2022.111404.

12. Matwij W, Gruszczyński W, Puniach E, Ćwiąkała P. Determination of underground mining-induced displacement field using multi-temporal TLS point cloud registration. Measurement. 2021;180:109482. https://doi.org/10.1016/j.measurement.2021.109482.

13. Chen X, Ban Y, Hua X, Lu T, Tao W, An Q. A method for the calculation of detectable landslide using terrestrial laser scanning data. Measurement. 2020;160:107852. https://doi.org/10.1016/j.measurement.2020.107852.

14. Abbate E, Sammartano G, Spanò A. Prospective upon multi-source urban scale data for 3D documentation and monitoring of urban legacies. ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci. 2019;XLII-2/W11:11–9. https://doi.org/10.5194/isprs-archives-XLII-2-W11-11-2019.

15. Arif R, Essa K. Evolving Techniques of Documentation of a World Heritage Site in Lahore. ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci. 2017;XLII-2/W5:33–40. https://doi.org/10.5194/isprs-archives-XLII-2-W5-33-2017.

16. Cipriani L, Bertacchi S, Bertacchi G. An optimised workflow for the interactive experience with Cultural Heritage through reality-based 3D models: cases study in archaeological and urban complexes. ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci. 2019;XLII-2/W11:427–34. https://doi.org/10.5194/isprs-archives-XLII-2-W11-427-2019.

17. Heras V, Sinchi E, Briones J, Lupercio L. Urban heritage monitoring, using image processing techniques and data collection with terrestrial laser scanner (TLS), case study Cuenca-Ecuador. Int Arch Photogramm Remote Sens Spat Inf Sci. 2019;XLII-2/W11:609–13. https://doi.org/10.5194/isprs-archives-XLII-2-W11-609-2019.

18. Kot P, Markiewicz J, Muradov M, Lapinski S, Shaw A, Zawieska D, et al. Combination of the photogrammetric and microwave remote sensing for cultural heritage documentation and preservation—preliminary results. Int Arch Photogramm Remote Sens Spat Inf Sci. 2020;XLIII-B2-2:1409–13. https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1409-2020.

19. Markiewicz J, Łapiński S, Kot P, Tobiasz A, Muradov M, Nikel J, et al. The quality assessment of different geolocalisation methods for a sensor system to monitor structural health of monumental objects. Sensors. 2020;20(10):2915. https://doi.org/10.3390/s20102915.

20. Wojtkowska M, Kedzierski M, Delis P. Validation of terrestrial laser scanning and artificial intelligence for measuring deformations of cultural heritage structures. Measurement. 2021;167:108291. https://doi.org/10.1016/j.measurement.2020.108291.

21. Giżyńska J, Komorowska E, Kowalczyk M. The comparison of photogrammetric and terrestrial laser scanning methods in the documentation of small cultural heritage object—case study. J Mod Technol Cult Herit Preserv. 2022. https://doi.org/10.33687/jmtchp.001.01.0013.

22. Tobiasz A, Markiewicz J, Lapinski S, Nikel J, Kot P, Muradov M. Review of methods for documentation, management, and sustainability of cultural heritage: case study: museum of King Jan III's Palace at Wilanów. Sustainability. 2019;11(24):7046. https://doi.org/10.3390/su11247046.

23. Gonizzi Barsanti S, Remondino F, Visintini D. 3D surveying and modeling of archaeological sites-some critical issues. ISPRS Ann Photogramm Remote Sens Spat Inf Sci. 2013;II-5/W1:145–50. https://doi.org/10.5194/isprsannals-II-5-W1-145-2013.

24. Markiewicz JS, Podlasiak P, Zawieska D. A new approach to the generation of orthoimages of cultural heritage objects-integrating TLS and image data. Remote Sens. 2015;7(12):16963–85. https://doi.org/10.3390/rs71215869.

25. Lewińska P, Róg M, Żądło A, Szombara S. To save from oblivion: comparative analysis of remote sensing means of documenting forgotten architectural treasures—Zagórz Monastery complex, Poland. Measurement. 2022;189:110447. https://doi.org/10.1016/j.measurement.2021.110447.

26. Kuzyk Z. The use of modern measurement methods in the inventory of endangered cultural heritage objects in Lviv. J Modern Technol Cult Herit. 2023. https://doi.org/10.33687/jmtchp.002.02.0019.

27. Van Genchten B. Theory and practice on terrestrial laser scanning. Learn tools Adv three-dimensional Surv risk Aware Proj. 2008. pp. 1–241. https://lirias.kuleuven.be/retrieve/122640. Accessed 18 Nov 2023.

28. Xu Y, Boerner R, Yao W, Hoegner L, Stilla U. Pairwise coarse registration of point clouds in urban scenes using voxel-based 4-planes congruent sets. ISPRS J Photogramm Remote Sens. 2019;151:106–23. https://doi.org/10.1016/j.isprsjprs.2019.02.015.

29. Habib A, Detchev I, Bang K. A comparative analysis of two approaches for multiple-surface registration of irregular point clouds. Int Arch Photogramm Remote Sens Spat Inf Sci ISPRS Arch. 2010;38:61–6.

30. Salvi J, Matabosch C, Fofi D, Forest J. A review of recent range image registration methods with accuracy evaluation. Image Vis Comput. 2007;25:578–96. https://doi.org/10.1016/j.imavis.2006.05.012.

31. Tam GKL, Cheng ZQ, Lai YK, Langbein FC, Liu Y, Marshall D, et al. Registration of 3D point clouds and meshes: a survey from rigid to Nonrigid. IEEE Trans Vis Comput Graph. 2013;19:1199–217. https://doi.org/10.1109/TVCG.2012.310.

32. Pomerleau F, Colas F, Siegwart R. A review of point cloud registration algorithms for mobile robotics. Found Trends Robot. 2015;4:1–104. https://doi.org/10.1561/2300000035.

33. Weinmann M. Reconstruction and analysis of 3D scenes. In: Irregularly distributed 3d points to object classes. Springer International Publishing; 2016.

34. Cheng L, Chen S, Liu X, Xu H, Wu Y, Li M, et al. Registration of laser scanning point clouds: a review. Sensors. 2018;18(5):1641. https://doi.org/10.3390/s18051641.

35. Guo Y, Sohel F, Bennamoun M, Lu M, Wan J. Rotational projection statistics for 3D local surface description and object recognition. Int J Comput Vis. 2013;105:63–86. https://doi.org/10.1007/s11263-013-0627-y.

36. Pavlov AL, Ovchinnikov G V., Derbyshev DY, Tsetserukou D, Oseledets I V. AA-ICP: Iterative closest point with anderson acceleration. In: 2018 IEEE Int Conf Robot Autom. 2018. p. 1–6. https://doi.org/10.48550/arXiv.1709.05479.

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 25 of 26

37. Biber P, Straßer W. The normal distributions transform: a new approach to laser scan matching. In: Proc 2003 IEEE/RSJ Int Conf Intell Robot Syst (IROS 2003). 2003. p. 2743–8. https://doi.org/10.1109/IROS.2003.1249285.

38. Das A, Waslander SL. Scan registration with multi-scale k-means normal distributions transform. In: 2012 IEEE/RSJ Int Conf Intell Robot Syst. IEEE; 2012. p. 2705–10. https://doi.org/10.1109/IROS.2012.6386185

39. Takeuchi E, Tsubouchi T. A 3-D scan matching using improved 3-D normal distributions transform for mobile robotic mapping. In: 2006 IEEE/RSJ Int Conf Intell Robot Syst. IEEE; 2006. p. 3068–73. https://doi.org/10.1109/IROS.2006.282246.

40. Tazir ML, Gokhool T, Checchin P, Malaterre L, Tazir ML, Gokhool T, et al. Cluster ICP: Towards Sparse To Dense Registration. In: 15th Int Conf Intell Auton Syst. Baden-Baden, Germany: Springer; 2018. p. 730–47. https://hal.science/hal-01863896/file/LamineT_IAS18_Cluster_ICP_Towards_Sparse_to_Dense_Registration.pdf. Accessed 18 Nov 2023.

41. Dong Z, Liang F, Yang B, Xu Y, Zang Y, Li J, et al. Registration of large-scale terrestrial laser scanner point clouds: a review and benchmark. ISPRS J Photogramm Remote Sens. 2020;163:327–42. https://doi.org/10.1016/j.isprsjprs.2020.03.013.

42. Vosselman G, Maas H-G. Airborne and terrestrial laser scanning. Boca Raton: CRC Press; 2010.

43. Boehler W, Marbs A. Investigating laser scanner accuracy. Int Arch Photogramm Remote Sens Spat Inf Sci. 2003;34:696–701.

44. Lichti D, Stewart M, Tsakiri M, Snow AJ. Benchmark tests on a three-dimensional laser scanning system. Geomatics Res Australas. 2000;72:1–24.

45. Lichti DD, Gordon SJ, Stewart MP, Franke J, Tsakiri M. Comparison of digital photogrammetry and laser scanning. Int Soc Photogramm Remote Sens. 2002; XXXIV part 5. pp. 39–44.

46. Markiewicz JS. The use of computer vision algorithms for automatic orientation of terrestrial laser scanning data. Int Arch Photogramm Remote Sens Spat Inf Sci ISPRS Arch. 2016;XLI-B3:315–32. https://doi.org/10.5194/isprs-archives-XLI-B3-315-2016.

47. Z+F LaserControl, LaserScanning Software. 2023. https://www.zofre.de/laserscanner/laserscanning-software/z-f-lasercontrolr. Accessed 18 Nov 2023.

48. Besl P, McKay N. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 1992;14:239–56. https://doi.org/10.1109/34.121791.

49. Luhmann T, Robson S, Kyle S, Boehm J. Close-range photogrammetry and 3D imaging. Photogramm Eng Remote Sens. 2015. https://doi.org/10.1016/S0099-1112(15)30069-0.

50. Teunissen PJG. Adjustment theory. Delft: Delft Academic Press/VSSD; 2003.

51. Börlin N, Murtiyoso A, Grussenmeyer P, Menna F, Nocerino E. Modular bundle adjustment for photogrammetric computations. ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci. 2018;XLII–2:133–40. https://doi.org/10.5194/isprs-archives-XLII-2-133-2018.

52. Rofatto VF, Matsuoka MT, Klein I, Veronez MR, Bonimani ML, Lehmann R. A half-century of Baarda's concept of reliability: a review, new perspectives, and applications. Surv Rev. 2020;52:261–77. https://doi.org/10.1080/00396265.2018.1548118.

53. Nowak E, Odziemczyk W. Adjustment of observation accuracy harmonisation parameters in optimising the network's reliability. Rep Geod Geoinf. 2018;105:53–9. https://doi.org/10.2478/rgg-2018-0021.

54. Hekimoglu S, Demirel H, Aydin C. Reliability of the conventional deformation analysis methods for vertical networks. FIG XXII Int Congr. Washington; 2002. p. 1–13.

55. Berber M, Dare P, Vaníček P. Robustness analysis of two-dimensional networks. J Surv Eng. 2006;132:168–75. https://doi.org/10.1061/(ASCE)0733-9453(2006)132:4(168).

56. Lichti DD, Pexman K, Tredoux W. New method for first-order network design applied to TLS self-calibration networks. ISPRS J Photogramm Remote Sens. 2021;177:306–18. https://doi.org/10.1016/j.isprsjprs.2021.05.014.

57. Baarda W. A testing procedure for use in geodetic network. Delft: Publications on Geodesy, New Series, Netherlands Geodetic Commission; 1968

58. Markiewicz J, Łapiński S, Bocheńska A, Kot P. The reliability assessment of the TLS registration methods—the case study of the Royal Castle in Warsaw. Int Arch Photogramm Remote Sens Spat Inf Sci. 2021;XLIII-B2-2:855–61. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-855-2021.

59. Bianco S, Ciocca G, Marelli D. Evaluating the performance of structure from motion pipelines. J Imaging. 2018;4(8):98. https://doi.org/10.3390/jimaging4080098.

60. Moussa W. Integration of digital photogrammetry and terrestrial laser scanning for cultural heritage data recording. Univ. Stuttgart. University of Stuttgart, Germany; 2014. https://www.ifp.uni-stuttgart.de/dokumente/Dissertationen/c-725_w.moussa.pdf. Accessed 18 Nov 2023.

61. Urban S, Weinmann M. Finding a good feature detector-descriptor combination for the 2D keypoint-based registration of TIs point clouds. ISPRS Ann Photogramm Remote Sens Spat Inf Sci. 2015;II-3/W5:121–8. https://doi.org/10.5194/isprsannals-II-3-W5-121-2015.

62. Karwel AK, Markiewicz J. The methodology of the archival aerial image orientation based on the SfM method. Sens Mach Learn Appl. 2022. https://doi.org/10.55627/smla.001.02.0015.

63. Markiewicz J, Zawieska D. The influence of the cartographic transformation of TLS data on the quality of the automatic registration. Appl Sci. 2019;9(3):509. https://doi.org/10.3390/app9030509.

64. Wang Z, Claus B. Point based registration of terrestrial laser data using intensity and geometry features. Int Arch Photogramm Remote Sens Spat Inf Sci. 2008;XXXVII-B5:583–9.

65. Barnea S, Filin S. Extraction of objects from terrestrial laser scans by integrating geometry image and intensity data with demonstration on trees. Remote Sens. 2012;4(1):88–110. https://doi.org/10.3390/rs4010088.

66. Markiewicz JS, Kajdewicz I, Zawieska D. The analysis of selected orientation methods of architectural objects' scans. In: Remondino F, Shortis MR, editors. Proc. SPIE 9528, Videometrics, range imaging, and applications XIII, 952805. 2015. https://doi.org/10.1117/12.2184959.

67. Markiewicz JS. The example of using intensity orthoimages in TLS data registration—a case study. Int Arch Photogramm Remote Sens Spat Inf Sci. 2017;XLII-2/W3:467–74. https://doi.org/10.5194/isprs-archives-XLII-2-W3-467-2017.

68. Tran TTH, Marchand E. Real-time keypoints matching: application to visual servoing. In: Proc IEEE Int Conf Robot Autom. 2007. pp. 3787–92. https://inria.hal.science/inria-00350601/PDF/2007_icra_tran.pdf. Accessed 18 Nov 2023.

69. Jakubovic A, Image VJ, Matching F, Matchers O-F. Int Symp ELMAR. IEEE. 2018;2018:83–6. https://doi.org/10.23919/ELMAR.2018.8534641.

70. Harris C, Stephens M. A combined corner and edge detector. Procedings Alvey Vis Conf. 1988;1988:23.1-23.6. https://doi.org/10.5244/C.2.23.

71. Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell. 1986;PAMI-8:679–98. https://doi.org/10.1109/TPAMI.1986.4767851.

72. Lowe DG. Object recognition from local scale-invariant features. Proc Seventh IEEE Int Conf Comput Vis. 1999;2:1150–7. https://doi.org/10.1109/ICCV.1999.790410.

73. Tuytelaars T, Mikolajczyk K. Local invariant feature detectors: a survey. Found Trends® Comput Graph Vis. 2007;3(3):177–280. https://doi.org/10.1561/0600000017.

74. Brown M, Lowe DG. Invariant features from interest point groups. Br Mach Vis Conf. 2002. pp. 656–65.

75. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis. 2004;60:91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94.

76. Rosten E, Drummond T. Machine learning for high speed corner detection. In: Comput Vis -ECCV 2006. 2006;1:430–43. https://doi.org/10.1007/11744023_34.

77. Leutenegger S, Chli M, Siegwart RY. BRISK: binary robust invariant scalable keypoints. In: Proc IEEE Int Conf Comput Vis. 2011. pp. 2548–55. https://doi.org/10.1109/ICCV.2011.6126542.

78. Weinmann M. Visual features—from early concepts to modern computer vision. Berlin: Springer; 2013. https://doi.org/10.1007/978-1-4471-5520-1_1.

79. Markiewicz J, Zawieska D. Analysis of the selection impact of 2D detectors on the accuracy of image-based TLS data registration of objects of cultural heritage and interiors of public utilities. Sensors. 2020;20:3277. https://doi.org/10.3390/s20113277.

80. Yu G, Morel J-M. ASIFT: an algorithm for fully affine invariant comparison. Image Process Line. 2011;1:11–38. https://doi.org/10.5201/ipol.2011.my-asift.

81. Barroso-Laguna A, Riba E, Ponsa D, Mikolajczyk K. Key.Net: keypoint detection by hand-crafted and learned CNN filters. 2019. https://arxiv.org/pdf/1904.00889.pdf. Accessed 18 Nov 2023.

Markiewicz *et al. Heritage Science*   (2023) 11:254

Page 26 of 26

82. Verdie Y, Kwang Moo Yi, Fua P, Lepetit V. TILDE: a temporally invariant learned detector. In: 2015 IEEE Conf Comput Vis Pattern Recognit. IEEE; 2015. pp. 5279–88; https://doi.org/10.1109/CVPR.2015.7299165.
83. Ebel P, Mishchuk A, Yi KM, Fua P, Trulls E. Beyond cartesian representations for local descriptors. 2019. https://arxiv.org/pdf/1908.05547.pdf. Accessed 18 Nov 2023.
84. Mishchuk A, Mishkin D, Radenovic F, Matas J. Working hard to know your neighbor's margins: local descriptor learning loss. 2017. https://arxiv.org/pdf/1705.10872.pdf. Accessed 18 Nov 2023.
85. Detone D, Malisiewicz T, Rabinovich A. SuperPoint: self-supervised interest point detection and description. IEEE Comput Soc Conf Comput Vis Pattern Recognit Work. 2018. pp. 337–49. https://arxiv.org/pdf/1712.07629.pdf. Accessed 18 Nov 2023.
86. Sarlin PE, Detone D, Malisiewicz T, Rabinovich A. SuperGlue: learning feature matching with graph neural networks. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2020. pp. 4937–46. https://arxiv.org/pdf/1911.11763.pdf. Accessed 18 Nov 2023.
87. Tyszkiewicz MJ, Fua P, Trulls E. DISK: Learning local features with policy gradient. Adv Neural Inf Process Syst. 2020. pp. 1–15. https://arxiv.org/pdf/2006.13566.pdf. Accessed 18 Nov 2023.
88. Remondino F, Menna F, Morelli L. Evaluating hand-crafted and learning-based features for photogrammetric applications. Int Arch Photogramm Remote Sens Spat Inf Sci. 2021;XLIII-B2-2:549–56. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-549-2021.
89. Choy CB, Gwak JY, Savarese S, Chandraker M. Universal correspondence network. Adv Neural Inf Process Syst. 2016. pp. 2414–22. https://arxiv.org/pdf/1606.03558.pdf. Accessed 18 Nov 2023.
90. Rocco I, Cimpoi M, Arandjelović R, Torii A, Pajdla T, Sivic J. Neighbourhood consensus networks. Adv Neural Inf Process Syst. 2018. pp. 1651–62. https://arxiv.org/pdf/1810.10510.pdf. Accessed 18 Nov 2023.
91. Li X, Han K, Li S, Prisacariu V. Dual-resolution correspondence networks. Adv Neural Inf Process Syst. 2020. pp. 1–20. https://arxiv.org/pdf/2006.08844.pdf. Accessed 18 Nov 2023.
92. Sun J, Shen Z, Wang Y, Bao H, Zhou X. LoFTR: detector-free local feature matching with transformers. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2021;4:8918–27.
93. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Adv Neural Inf Process Syst. 2017. pp. 5999–6009. https://arxiv.org/pdf/1706.03762.pdf. Accessed 18 Nov 2023.
94. Sarlin P-E, DeTone D, Malisiewicz T, Rabinovich A. SuperGlue: learning feature matching with graph neural networks. 2019. http://arxiv.org/abs/1911.11763. Accessed 18 Nov 2023.

Publisher's Note