

Fusing tie points' RGB and thermal information for mapping large areas based on aerial images: A study of fusion performance under different flight configurations and experimental conditions

Yu Hou^{a,*}, Rebekka Volk^b, Meida Chen^{a,c}, Lucio Soibelman^a

^a Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, 3620 S. Vermont Ave., Los Angeles, CA 90089-1453, USA

^b Institute for Industrial Production, Karlsruhe Institute of Technology, Karlsruhe, Baden-Wuerttemberg, Germany

^c Institute for Creative Technologies, University of Southern California, 12015 E. Waterfront Dr., Los Angeles, CA 90094, USA

A B S T R A C T

Keywords:

Thermal mapping

Large scale photogrammetry

Tie points

Data fusion

UAS data collection

Three-dimensional thermal mapping from aerial images can be used in energy audits. Tie points that define the location of object points in a 3D space for reconstructing a 3D model, also include thermal information, which plays an important role in energy audits. However, it is often harder and less accurate to extract common features and determine tie points from low-resolution thermal images. It is more effective and accurate to use high-definition RGB images to determine tie points and fuse the RGB and thermal information. In this study, we investigate how to utilize high-definition RGB images that allow for more accurate tie point detection, how different flight configurations affect tie point data fusion, and how tie point data fusion performance can be improved. We propose a tie points' thermal and RGB data-fusion framework to create district-level thermal mapping to solve such problems. This paper aims to evaluate how different flight configurations affect the results of the proposed data fusion approach. Flight configurations include different camera altitudes (60 m and 35 m), distinct camera angles (45 degrees and 30 degrees), diverse flight path designs (mesh grid and Y path), and various building styles (campus buildings and city buildings).

We find the following results in this paper: (1) higher flight altitude is not suggested for our data fusion approach; (2) a 30-degree thermal camera angle is suggested for roof inspection, while a 45-degree thermal camera angle is suggested for façade inspection when using the tie point data fusion approach; (3) a Y flight path performs better than a mesh grid path; and (4) our tie point data fusion approach performs better in traditional European city buildings than in modern campus buildings. We also demonstrate that pixels in the thermal images' central area can more accurately represent thermal information than pixels around the image edges for tie point data fusion. Additionally, our studies show that images taken at the edges of mapping areas have more errors. Thus, it is crucial to enlarge the survey area to obtain more accurate possible results.

1. Introduction

The reduction of greenhouse gas emissions and energy consumption is currently a high priority for many countries and organizations. The European Union (EU) has set a goal of cutting at least 40% of greenhouse gas emissions (from 1990 levels) and increase energy efficiency by 32.5% more before 2030 [1–3]. Individual countries also have set their own goals. The U.K. passed *The 2008 Climate Change Act* [4], one of the world's first comprehensive frameworks of laws on climate change [5]. Germany also drafted *The Climate Action Law* in 2019 to guarantee that the nation achieves its national and European climate targets [6]. As a

part of the law, their *Climate Action Plan 2050* calls for making the nation's building stock largely climate neutral [6,7]. These actions are a clear sign that measures are being taken to improve energy efficiency.

As buildings comprise up 35% of Germany's total energy consumption [8], auditing energy performance and improving energy efficiency to reduce consumption have become increasingly important. To save energy use in buildings, it is important to understand the building envelope that separates the conditioned and unconditioned building environment [9,10]. In terms of a single building, an efficient envelope can maintain the conditioned environment. For example, an envelope with fewer leaks reduces the heating or cooling load and prevents such

* Corresponding author.

E-mail addresses: yuhou@usc.edu (Y. Hou), rebekka.volk@kit.edu (R. Volk), meidache@usc.edu (M. Chen), soibelman@usc.edu (L. Soibelman).

loss through walls and roofs. It has been demonstrated that increasing the efficiency of building envelopes is a low investment and high-return plan [11].

However, auditing energy performances and improving energy efficiency should not be limited to single buildings. Although it is common to deploy a district heating network for residential and commercial buildings, disadvantages exist due to issues related to the topologies of such networks. Plants generate heat and supply it to subnetworks, while several groups of buildings connected to these sub-networks consume heat. Buildings connected to different sub-networks are not influenced by each other, but buildings connected to the same subnetwork are dependent. Thus, energy performance audits require a global view and must take entire communities into consideration. Many approaches improve the performance of energy audits and save building energy, such as fan pressurization (blower door test), ultrasound, and thermography, [12–15]. However, unfortunately, these approaches have some limitations. Fan pressurization requires people to vacate indoor rooms to conduct experiments, which is impractical for occupied rooms. Ultrasound cannot precisely determine where energy loss comes from. Thermography is a method used to detect surface temperature values of building envelopes to measure energy loss [16–18]. Researchers have also used thermal images to reconstruct 3D thermal mapping models to allow users to comprehensively audit, simulate and make decisions regarding a building’s energy performance. However, it is not possible to conduct energy audits on a group of several individual buildings in an entire district within a short time using existing approaches. Additionally, certain buildings areas are difficult to reach, such as roofs and facades of high-rise buildings. Therefore, to broaden the survey scope and reduce labor costs, this paper introduces a study that proposes to deploy automated fly-past surveys. In the proposed approach, unmanned aircraft systems (UASs), also known as drones, deploy mounted infrared thermal (IRT) cameras to collect data. UAS images are then used to reconstruct 3D models with photogrammetry technology, which allows city and building managers to fully audit an entire district, construct 3D as-is building energy models, and conduct energy simulations and diagnostics [19].

Currently, technologies that utilize optical images (RGB images) acquired from UASs to reconstruct 3D point clouds or mesh models in large areas are available, but the technologies that utilize thermal images for model reconstruction leave much to be desired. Thermal photogrammetry faces several challenges for large areas in particular: (1) a lower resolution of thermal images [20,21], (2) a trade-off between accuracy and efficiency [22], (3) the effect of environmental conditions for data collection [23,24]; and (4) the object-to-camera distance. Many studies have planned to improve the performance of thermal mapping model recreation by avoiding the shortcomings of thermal images [24,25], but these studies have split and separately processed RGB information from thermal information and thus failed to make full use of high-definition RGB images that allow for more accurate tie point detection and high-quality point cloud/mesh model reconstruction during the mapping process. Several benefits of leveraging high-resolution RGB images exist to improve the performance of thermal model reconstruction. First, the same or a smaller number of high-resolution RGB images can boost the performance of thermal photogrammetry compared to the use of only thermal images [24]. A reduction in the number of required images decreases both the data collection and processing time, which can better meet the efficiency requirements of energy audits for an entire district-level investigation. Second, the ambient temperature has less influence on RGB photogrammetry such that the amount of sunlight and shade do not dramatically change. It has been demonstrated that some building components may have wider-ranging colors in RGB images but single thermal values [22,26]. The single thermal value of a large building component (e.g., a flat roof) is a considerable problem for model reconstruction. A monotonous thermal value can cause large undifferentiated areas resulting in reconstructions that lack useful information. A sparse point cloud model cannot provide

enough information for an energy audit. Fortunately, RGB images usually can reconstruct a dense point cloud model. With these facts, we leveraged high-resolution RGB images to improve the performance of thermal model reconstruction for the approach proposed in this paper.

To reconstruct a 3D model, existing photogrammetry algorithms detect all potential feature points in images. A common feature point that can be clearly identified in two or more images can then define the location of a 3D object point called a reference point or a tie point. Tie points are extremely important because they contain essential feature information for reconstructing a 3D mapping model, such as geographic coordinates and color information of objects. Tie points can also include thermal information when reconstructing a thermal mapping model. However, the drawback of thermal images is that low resolution compared to RGB images makes tie points’ thermal information detection more difficult [27]. Therefore, studies on fusing tie points’ thermal and RGB information are vital. Our study uses high-resolution RGB images to create a mapping model and then add thermal information to the tie points in the RGB mapping model. Thus, the tie points are detected by high-resolution RGB images and include thermal information. This tie point data fusion procedure might be impacted by data collection approaches such as flight altitudes, camera angles, and flight patterns, and it produces various results when surveying a range of buildings. The study described in this paper examines these factors and analyzes how various UAS flight configurations influence the performance of tie point data fusion and how data fusion performances differ when surveying distinct building styles. This paper also aims to provide suggestions for data collection and strategies for data processing to improve the tie point data fusion approach.

The rest of this paper is organized as follows. In Section 2, related work is summarized. In Section 3, proposed research methods are presented. The results of four different experiments under varying conditions with test factors are described in Section 4, followed by their evaluation and discussion in Section 5 and the conclusion and future work in Section 6.

2. Related work

2.1. Data collection approaches

In the past few years, aerial surveys have dominated the market in auditing large areas. For example, Iwaszczuk et al. 2016 used oblique airborne thermal infrared images taken from helicopters to extract and register textures onto a premade 3D BIM model [28,29]. Martinez et al. 2006 [30] used autonomous helicopters to audit building energy. However, aerial surveys, including those facilitated by helicopters and airplanes, require a large amount of labor and time. In addition, buildings in investigated areas may be vulnerable under helicopters and airplanes. Today, unmanned aircraft systems (UASs) have become one of the most popular tools for data collection in the field of 3D capturing of buildings and construction site surveying for both indoor and outdoor structures [31–33]. For example, Kim et al. 2019 [31] operated drones over a construction site to efficiently gather data. Hamledari et al. 2017 [33] implemented computer vision algorithms on drone images to automatically detect components of indoor partitions such as studs, insulation, electrical outlets, and the state of drywall work. This information allowed updating of the 4D BIM schedule and progress information.

Another popular method is the use of mobile mapping systems (MMSs) and scanning from a vehicle, known as an unmanned ground vehicles (UGVs) [19,34,35], and researchers have also combined UGV and UAS techniques [36]. Park et al. 2019 [36] demonstrated that construction site areas were difficult to explore from the ground level. They used a UAS over a construction site to generate a model that then created an optimal path for a mobile robot. Their approach was designed to reduce human interference in comprehensive data collection and to frequently monitor and update to provide the most effective decision-

making. For data collation in which an UGV is not feasible, portable mobile mapping (PMMS), such as GeoSlam Zeb-Revo has proved to be more effective. Dewez et al. 2017 [37] used a Zeb-Revo handheld mobile laser scanner to map underground cavities, and Rodriguez-Martin et al. 2018 [38] used a Zeb-Revo to map 3D models in hospital machine rooms.

2.2. Photogrammetry and thermal mapping modeling

Imagery data collected through UASs or UGVs can be processed by photogrammetry to reconstruct a 3D mapping model that uses a series of paired images for further analysis. Finding correspondence is the first step in photogrammetry, a process that detects common features in distinct but overlapped images. Common features are attributes that describe the same objects in distinct images. There are two widely used algorithms for common feature detection, scale-invariant feature transform (SIFT) [39] and speeded-up robust features (SURF) [40]. The SIFT algorithm recognizes an object in one of the reference images and compares its features to others to locate matching features based on the Euclidean distance of their feature vectors. SURF conducts the same process of locating candidate matching features, but it is based on different calculations. Although SIFT is superior to SURF when images have different scales, SURF has a better result when images are blurry. Additionally, SURF is three times faster than SIFT [41]. After candidate-matching features are found, a random sample consensus (RANSAC) [42] is utilized to remove incorrectly matched outliers. Next, detected and filtered feature points determine corresponding points in the 3D space. Since these 3D points tie the images together, they are defined as tie points. The final step involves structure from motion (SfM) [43–45] in which a 3D spatial point cloud model is reconstructed from 2D reference images. If a mesh model is needed, the triangulation algorithm can convert the given point cloud model into a consistent polygonal (mesh) model [46].

UASs- and UGV-based RGB and thermal photogrammetry have been widely used. Related works are summarized in Table 1 and show two types of studies: mapping from 2D data and from 3D data. Mapping RGB or thermal models from 2D data refers to using a series of paired images RGB or thermal images to reconstruct 3D models by photogrammetry algorithm. Mapping from 3D data refers to using both a LiDAR (3D laser scanning) tool and an RGB or thermal camera. As Laguela Lopez et al. 2014 [47] described, upper façade parts and roofs cannot be investigated using UGVs. Additionally, fewer studies have focused on thermal mapping modeling for energy audits at a large district level using UASs.

Thermal mapping photogrammetry allows researchers to detect thermal leakages [24,25,63,64], and they have used quantitative and qualitative approaches in their studies. Quantitative approaches have included (1) R-value and U-value measurements [65], (2) moisture content identification [66]; and (3) a calculation of the percentage of areas with thermal anomalies. Qualitative approaches have included (1)

Table 1
3D Model reconstruction.

Data sources	Mapping from 2D data	Mapping from 3D data
UASs	RGB model (Clarkson et al. 2017 [48]; Chen et al. 2019 [49]; Chen et al. 2020 [50]; Park et al. 2019 [36]) Thermal model (Iwaszczuk et al. 2016 [34]; Maset et al. 2017 [51]; Bayomi et al. 2019 [52]; Hou et al. 2021 [53])	RGB model (Kim et al. 2013 [53]; Xu et al. 2015 [54]; Campos-Taberner et al. 2016 [55]; Castagno et al. 2018 [56]) Thermal model (Ilehag et al. 2018 [57])
UGVs	RGB model (Yamaguchi et al. 2018 [58]; Park et al. 2019 [36]) Thermal model (Lopez et al. 2014 [47]; Lin et al. 2019 [59])	RGB model (Chen et al. 2018 [60]) Thermal model (Laguela et al. 2011 [61]; Borrmann et al. 2012 [62];

classification of walls, glazing, and windows; (2) thermal bridge identification; (3) air leakage inspection; (4) moisture inspection; and (5) HVAC and electrical systems presentation [24]. However, all of these studies have been limited to single buildings, and they failed to consider tradeoffs between efficiency and accuracy when auditing building energy using UASs for an entire district. To improve the performance and efficiency of thermal 3D mapping model reconstruction, it is necessary to consider the several issues summarized in Table 2. Researchers are aware of these facts as well as the differences between RGB cameras and thermal cameras. As a result, they have tested various factors related to data collection, such as different flight paths, camera angles, and flight altitudes that influence the performance of reconstructing thermal mapping models [22]. However, the traditional thermal mapping approach only uses thermal images without the full benefit of high-resolution RGB images to support modeling. Many researchers are exploring an effective method to fuse both thermal and RGB information.

2.3. Data fusion approaches

Research conducted on the fusion of thermal information and RGB information has produced three primary methods to perform geometric registration between thermal images and 3D models: 2D-2D (image to image) matching, 2D-3D (image to model) matching, and 3D-3D (model to model) matching [59]. Related studies are as follows:

2.3.1. 2D-2D

Ham et al. 2013 [27] tested several traditional algorithms, including SIFT, Affine-SIFT (ASIFT) and SURF to register visible images through thermal images captured from different cameras. Although these algorithms were commonly used for thermal-thermal or visible-visible pairs, their studies showed that the traditional algorithms performed relatively poorly for thermal-visible pairs. To improve performance, Weinmann et al. 2014 proposed using a shape-based matching algorithm for 2D-2D image registration [68]. González-Aguilera et al. 2012 [69] extracted

Table 2

Issues may improve the performance and efficiency of thermal model reconstruction.

Number	Issue	Description
1	Camera resolution	Higher resolution thermal cameras, that perform better than 640×512 , are not broadly available in the retail market. Lower resolution cameras are unable to capture sufficient thermal information and do not enable algorithms that could detect common features from overlapped images to reconstruct high-definition 3D models.
2	Distance and angle	The distance between the object and the cameras [13], as well as the incident angle (the angle between a ray reflected on a surface and the normal line), and image viewing angle should also be considered [67]. This is necessary because when taken from distinct angles, thermal images facing the same target might capture different temperature values and introduce potential errors [22].
3	Experiment time	The time and duration of thermal cameras use could influence results. Many studies have suggested that data collection should be conducted in the early morning and late afternoon, and specifically on cloudy days [23,24].
4	Flight time	Similar to the way in which an RGB model photogrammetry performance is influenced by light and shade, thermal model photogrammetry performance is influenced by objects' temperature values. RGB images representing building's appearance cannot be changed over short time spans, but thermal images representing the surface temperatures of buildings can change in short time spans. Thus, relevant thermal images must be taken as quickly as possible.

common features in the thermal images and the range images (converted from a LiDAR point cloud model) and registered thermal images (2D) with range images (2D).

2.3.2. 2D-3D

Weinmann et al. 2014 conducted a random sample consensus (RANSAC) for registering 2D-3D correspondences [68]. Their research showed relatively good performance indoors. Iwaszczuk et al. 2016 used many airborne thermal images to perform linear feature matching when conducting 2D-3D fusion and considered the uncertainty of 3D model reconstruction. Lin et al. 2019 [59] pointed out the lack of research focusing on data acquisition that uses a handheld thermal camera and focused on the fusion of thermal images with RGB 3D point cloud models for façade energy audits. The authors then proposed an approach to reduce thermal disagreements, caused by misalignments, for the same object points provided by several associated images. Their methods showed good results for both geometric registration accuracy and thermal registration accuracy. Lin et al. 2018 [70] also focused on fusion of selected thermal textures with 3D mesh models. Their research demonstrated that the Gaussian filter performed better than a texture voting strategy. However, the authors' handheld approach was not feasible for mapping a larger area. Lagüela et al. 2013 [71] registered a 3D point cloud model with 2D images based on linear 3D features from 3D model and 2D line segments from images.

2.3.3. 3D-3D

Wiens et al. 2019 [72] summarized two important tasks when processing point cloud data, surface estimation, and point cloud registration. They proposed a new approach, the likelihood method, which was considered competitive with the popular iterative closest point (ICP) method. The authors' method reduced the predictive mean squared error by 18%. Hoegner et al. 2016 and Stilla et al. 2009 [63,73] focused on registering a thermal point cloud onto an RGB point cloud, thus meeting the prerequisite that thermal and RGB point clouds should register at the closest points. However, it was difficult to minimize the distance between two point clouds since the thermal image had lower resolution when compared with the RGB images. Additionally, two reconstructions for thermal point clouds and RGB point clouds were based on two mutually exclusive series of thermal images and RGB images respectively, which could not guarantee the registration accuracy for two mutually independent point clouds. Alternatively, Truong et al. 2017 [74] used a fixed relative pose relation between RGB cameras and thermal cameras as a reference point for the ICP to enhance the registration between the thermal point cloud and the RGB point cloud.

3. Research methods

3.1. Research architecture

This study framework consisted of four steps: (1) data collection, (2) camera calibration and image registration, (3) data fusion process, and (4) evaluation. Fig. 1 illustrates the research method workflow.

Our first step was to investigate the influence of four factors related to distinct flight configurations of data collection and influences of diverse building styles on our data fusion approach. Specifically, the goal was to explore the conditions under which the data fusion approach performs better. These four factors included the three factors related to flight configurations: (1) different camera altitudes, (2) distinct camera angles, and (3) diverse flight paths and (4) a factor related to diverse building styles that refers to various architectural styles of buildings in urban areas (see Fig. 1, Step 1). After running these data collection conditions using the four factors, we compared and analyzed the performances of the data fusion approach using various testing factors.

The second step was designed to pre-process the images taken before they were used for reconstructing models. In this step, both thermal images and RGB images needed to be undistorted to reduce the errors

during the 3D model reconstruction. After removing the distortion, we registered and overlapped the thermal images onto the RGB image so that each pixel in the thermal images could find its related pixel in the other RGB images for photogrammetry and further the data fusion process.

In the third step, we introduced the proposed data fusion approach procedure. The RGB images were utilized to reconstruct high-performance 3D models using photogrammetry algorithms. At the same time, the tie points were recorded while the models were reconstructed. Since the second step included the process using algorithms to overlap thermal images onto RGB images, the tie point RGB information was replaced by thermal information in corresponding overlapped thermal images. As shown in Fig. 1, the table in Step 3 is an example. The table headings record each tie point's space coordinates, X, Y, and Z; color information, R, G, and B; and the thermal information, which is represented in grayscale. This procedure allowed for the construction of models with data fusion of tie points' thermal and RGB information. Next, the procedure was repeated for all different conditions using the four test factors.

In the fourth step, we introduced methods to evaluate the data fusion approach under varying conditions with various testing factors. The evaluation methods included (1) simple average value, (2) earth mover's distance, and (3) a mean square errors (MSE) evaluation for images. These evaluations were based on examining errors that were potentially introduced by simultaneously replacing RGB pixels with thermal pixels. The first method calculated the mean value of the potential errors, the second method analyzed the frequency distribution of the errors, and the third method evaluated the errors from the thermal image perspective. All three evaluations were used to examine the experiments conducted under varying conditions to determine how well the data fusion approach performed in each of the four factors. We introduced each step in the following paragraphs.

3.2. Data collection

The UASs used in this study consisted of three parts: (1) the main body, (2) the data collection system, and (3) the controller. First, the main body consisted of an aircraft and a gimbal. Second, a FLIR DUO Pro R camera was used in the data collection system. This camera has both a thermal lens and an optical RGB lens in a single integrated package. The RGB image has a 4000×3000 resolution with a field of view (FOV) of $56^\circ \times 45^\circ$ (horizontal x vertical), and the thermal image has a 640×512 resolution with a FOV of $45^\circ \times 37^\circ$. One of the benefits of using the FLIR DUO Pro R is that both thermal and RGB images can be obtained with one flight task, and the gap between thermal lens and RGB lens is shorter than if using two separate cameras, which can reduce the uncertainties for image registration. Algorithms have been developed and used to overlap thermal images onto the RGB images for image registration. The third part was the controller that remotely met all experiment requirements for the data collection, as it had been used to operate GPS systems in both aircraft and cameras. The GPS system was not only used for automatic flight, but also to record georeferencing image data. Such data allowed the 3D reconstruction software to precisely re-calculate image positions and reduce errors [75,76].

The data collection step utilized UAS hardware to explore the influence of various factors related to distinct flight configurations of data collection and diverse building styles on the data fusion approach. There were four factors of data collection: (1) different camera altitudes, (2) distinct camera angles, (3) diverse flight paths, and (4) various building styles in urban areas. First, the camera altitude test included 35 m and 60 m. Although the altitude should be higher than the highest buildings in the tested area. It should not be too high, as the results could be inaccurate. Therefore, the altitudes were tested at 35 m and 60 m, which were 1.5 times and 3 times higher than the highest building over the tested area. Second, the camera angles determined how much information could be detected from the buildings, especially the façade

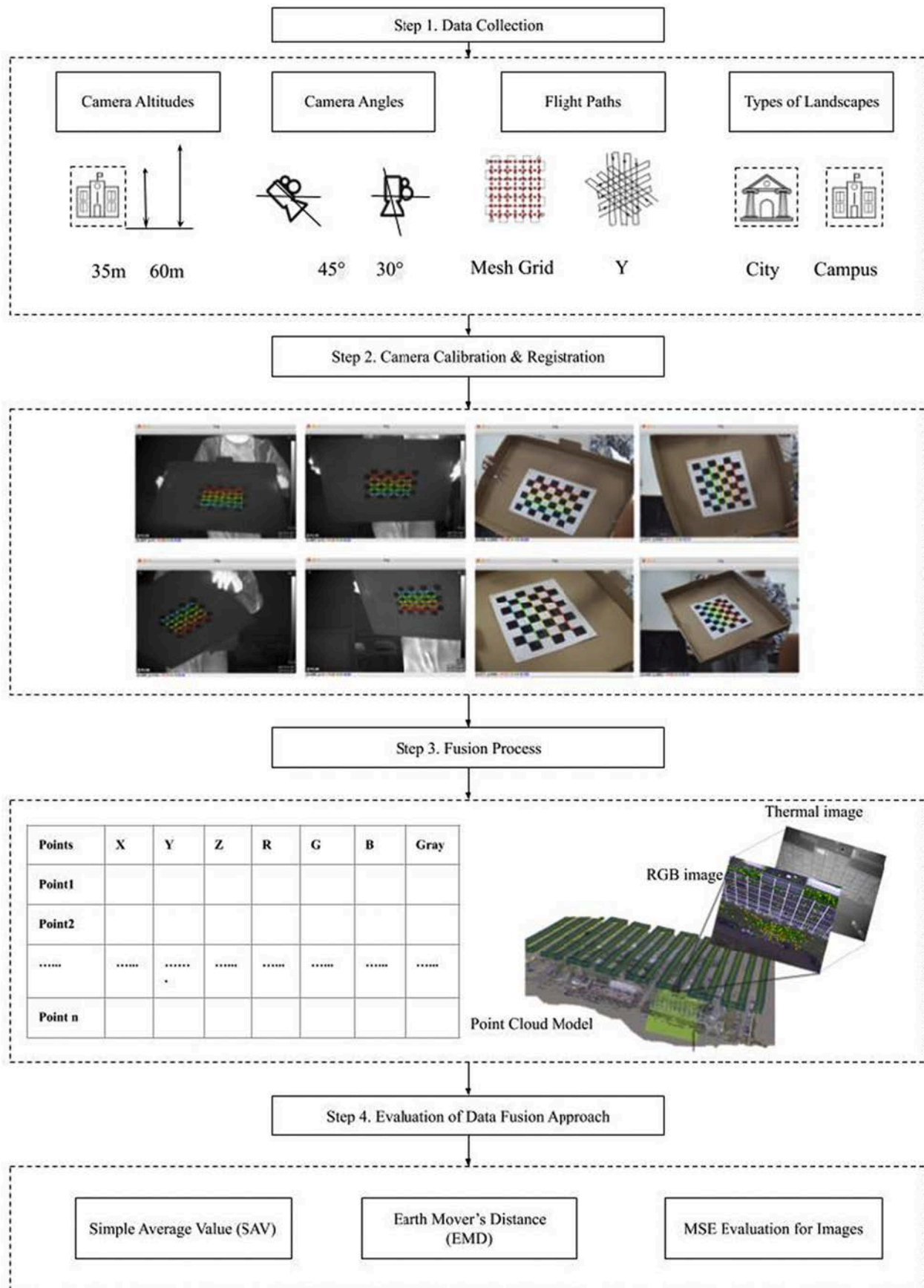


Fig. 1. Research method workflow.

information. In this study, two distinct camera angles of 45 and 30 degrees were tested. The angles were between the camera view line and the vertical axis (see Fig. 1, Step 1). Third, the diverse flight path included the mesh grid and Y path. The commonly used flight path design was a mesh grid in which the drones flew facing four directions, north, south, east, and west. A new contrasting flight path, the Y path, was also used in which the drone flew in only three directions for example, north, southeast, and southwest with the same front and side overlap parameters. It was confirmed that the Y path could save flight time when compared with the mesh grid while producing the same data quality [22]. Finally, the fourth comparison was the data collection conducted on various building styles. We collected data on a campus in Germany that included offices, laboratories and workshop buildings separated by lawns, roads, parking lots, and wide sidewalks that were relatively far from each other. In addition, other data collection was conducted within a dense urban city district in Germany where the buildings were four-to-six stories high and built close together.

Images taken from the varying conditions with the four various factors were used to conduct the data fusion approach, and the performances of the data fusion approach were compared and analyzed. The four experiments presented in Section 4 were tested under varying

conditions defined by the abovementioned four factors.

3.3. Camera calibration and image registration

The images taken during the data collection were pre-processed in this step, including (1) camera calibration, and (2) image registration, before they were used for reconstructing models. Camera calibration was necessary because of distortions in the thermal and RGB images that affected the 3D model reconstruction quality. In addition, image registration was important because thermal images were necessary to accurately replace RGB images in the proposed data fusion approach.

First, we used camera calibration to correct distorted images and improve the performance of the photogrammetry. Due to tangential and radial distortion, images could display objects incorrectly. This problem could be solved through camera calibration that estimates intrinsic and extrinsic parameters. Assuming that P was a point in the 3D space and p was a point in an image, formula (1) explained the projection from 3D point P to 2D point p . In Eq. (1), K was called the camera intrinsic matrix, which denoted transformation from the 3D camera coordinates system to the 2D image coordinates system. It was written as Eq. (2), whereby α , k_f , β , l_f , k and l dealt with pixel unit transformation, and θ was the

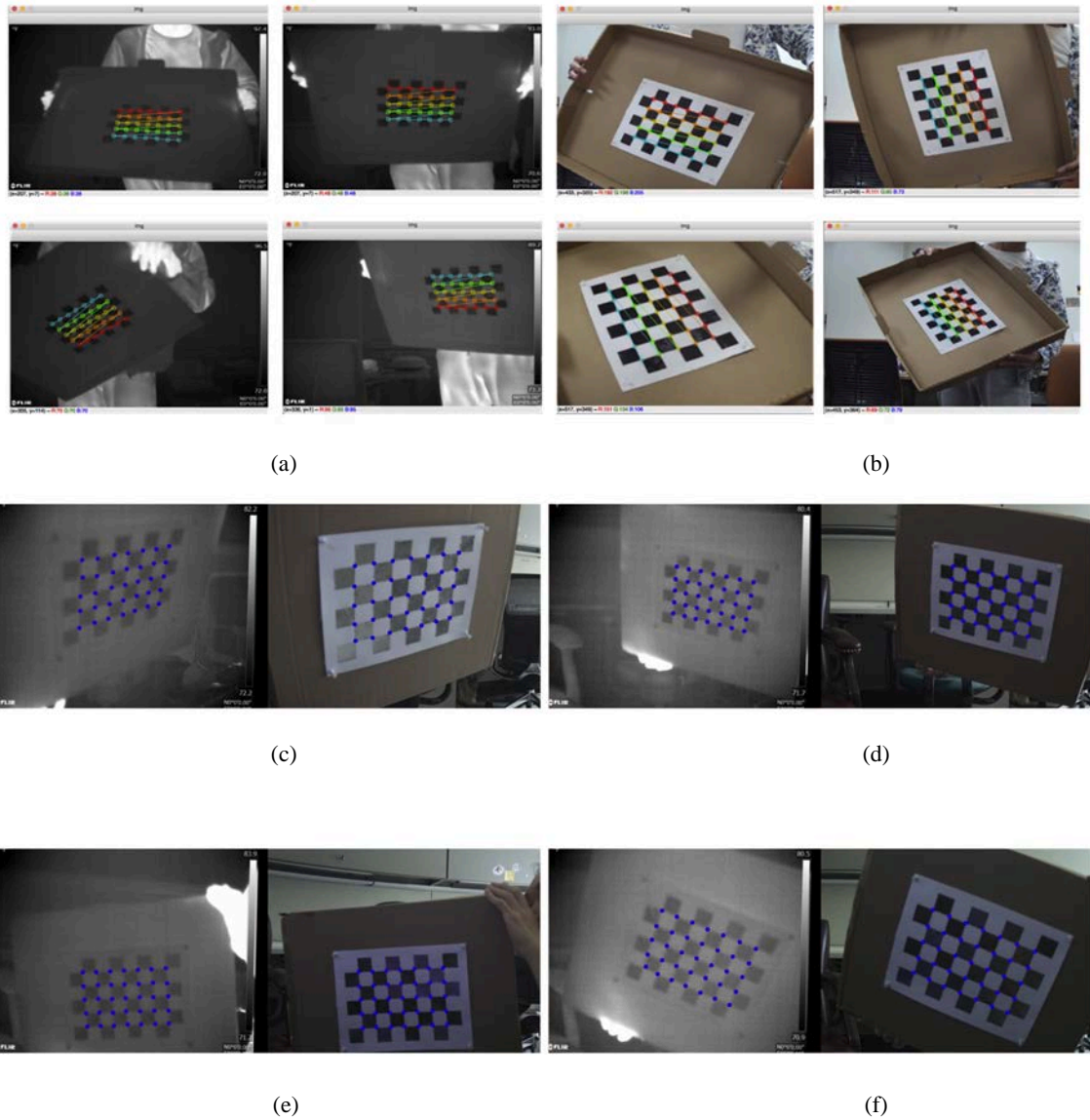


Fig. 2. Camera calibration and image registration.

angle between the x-axis and y-axis in the image coordinate system. The center of an image did not need to be at (0, 0); it could have been at position (x_0, y_0) . In formula (1), R and t were the extrinsic parameters, which denoted the rotation and transformation from the 3D world coordinates system to a 3D camera coordinates system [77,78].

$$p = K [R | t] P \quad (1)$$

$$K = \begin{bmatrix} \alpha & \alpha \cot \theta & x_0 \\ 0 & \frac{\beta}{\sin \theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Distortions of RGB and thermal images obtained from the cameras were automatically corrected by calculated intrinsic and extrinsic parameters. Fig. 2 (a) and (b) show the process of calibrating the camera. The thermal camera captures different levels of infrared light. Materials with different emissivity emit different levels of infrared light for detection by thermal cameras, therefore the chessboard made of a hollow tin foil sheet in cardboard was used for camera calibration, resulting in the solid areas being darker than the hollow areas in images. Fig. 2 (a) shows the camera calibration on the thermal images, and Fig. 2 (b) shows the camera calibration on the RGB images.

Second, image registration allowed thermal images to be aligned onto RGB images at the pixel level, which accurately replaced corresponding RGB pixels with thermal pixels in the data fusion process. Although the thermal and RGB images were taken from the same angles, altitude, and at the same time, the thermal and optical lens packed in the FLIR Duo Pro R did not perfectly overlap, and thus the image registration between the two associated thermal and RGB images was still necessary. As shown in Fig. 2 (c)–(f), common feature points (highlighted using

blue dots) on the chessboard detected from thermal and RGB images could be used for image registration. Feature points on the thermal images (images on the left of Fig. 2 (c)–(f)) were in one matrix space and corresponding feature points on the corresponding RGB images (images on the right of Fig. 2 (c)–(f)) were in another matrix space. The transformation between the two matrix spaces could be calculated and then used to transform all thermal image matrix spaces to RGB image matrix spaces for registration.

3.4. Data fusion process

We took images under varying conditions in the first step and pre-processed images in the second step to reconstruct models and detect tie points. To reconstruct a 3D model, the algorithm detected all potential feature points in all images as presented in Section 2. Tie points that were clearly identified in two or more images defined the location of an object point in a 3D space. There are several well-established software programs to process photogrammetry including Pix4D, Agisoft, and DroneDeploy. However, none of these programs provided an application programming interface (API) that supported a user's extended development. Fortunately, ContextCapture from Bentley [79] recorded all generated intermediate files for extended programming and thus enabled the algorithm to record all the tie points' locations in the 3D space and the source images where they were detected. This function was leveraged to fuse thermal information with RGB information and to evaluate the tie point data fusion performance.

First, we assumed there were (m) visible RGB images $\{V^{(1)}, V^{(2)}, \dots, V^{(m)}\}$ and correspondingly m thermal images $\{T^{(1)}, T^{(2)}, \dots, T^{(m)}\}$ taken with a drone. Visible images were then used to reconstruct a 3D point cloud model similar to the blue building icon shown in Fig. 3 by

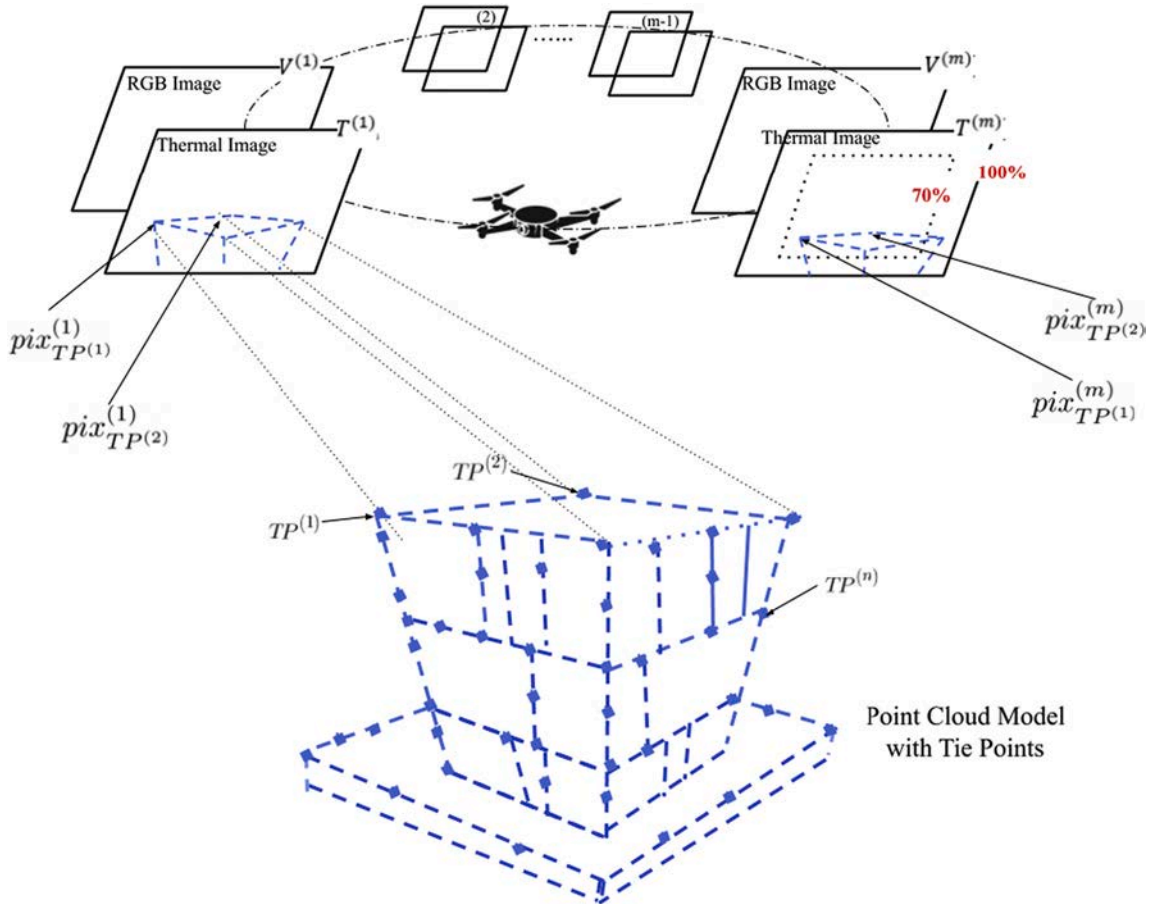


Fig. 3. Data fusion procedure.

ContextCapture. All visible images were used to reconstruct (n) tie points in a 3D point cloud file, noted $\{TP^{(1)}, TP^{(2)}, \dots, TP^{(n)}\}$, blue solid dots shown in Fig. 3, and every tie point in a 3D point cloud contained some pixel information from various images, noted point-image pairs $\{TP^{(1)} : (V_1, T_1), \dots, (V_{k^{(1)}}), T_{k^{(1)}})\}, \{TP^{(2)} : (V_1, T_1), \dots, (V_{k^{(2)}}), T_{k^{(2)}})\}, \dots, \{TP^{(n)} : (V_{any\ number}, T_{any\ number}), \dots, (V_{k^{(n)}}), T_{k^{(n)}})\}$. Since every tie point had a different total number of associated images, we used $k^{(1)}, k^{(2)}, \dots, k^{(n)}$ to distinguish this number. Also, different tie points might be associated with the same or various images, so we assumed that associated images for $TP^{(n)}$ started with any number.

ContextCapture processed and provided several intermediate files that consisted of (1) tie points' positions, (2) visible source images' relevant position for the tie points, and (3) pixel positions and color information for their corresponding tie points that were stored in an XML file. Later, these files facilitated the replacement of all RGB images used for reconstructions with registered thermal images taken from the same position with the same angle. Next, visible color information for each tie point was fused with corresponding registered thermal information from source images. Finally, each tie point received different thermal values from various relevant thermal images. These relevant thermal images' corresponding RGB images facilitated the reconstruction of tie points. The irrelevant RGB images that did not contribute to reconstructing tie points, as well as the irrelevant thermal images, were automatically eliminated by ContextCaptures.

An example of the tie points is represented as $\{TP^{(1)} : (V_1, T_1), \dots, (V_{k^{(1)}}), T_{k^{(1)}})\}$. $TP^{(1)}$ was reconstructed by $k^{(1)}$ number of visible images $\{V_1, \dots, V_{k^{(1)}}\}$. If we replaced all visible images with thermal images $\{T_1, \dots, T_{k^{(1)}}\}$, the points would receive different thermal values from the relevant thermal images. Thermal values were pixels in thermal images, noted as $\{pix_{TP^{(1)}}^{(1)}, \dots, pix_{TP^{(1)}}^{(k^{(1)})}\}$. Therefore, the mean value $\bar{x}_{TP^{(1)}}$ and the standard deviation (SD) value, $SD_{TP^{(1)}}$ of received thermal values for $TP^{(1)}$ were calculated, as elaborated in Eqs. (3) and (4). The mean value $\bar{x}_{TP^{(1)}}$ was defined as the final thermal value (FTV) of the tie point. Each tie point received various thermal values from various relevant thermal images. However, it was important to point out that the various relevant thermal values for one tie point are supposed to be identical, because the tie point represents the same object. In other words, a tie point's final thermal value (FTV) should have been the same as the thermal values it received from thermal images, noted as $\bar{x}_{TP^{(1)}}$.

$pix_{TP^{(1)}}^{(1)}, \dots, pix_{TP^{(1)}}^{(k^{(1)})}$. Additionally, a tie point's $SD_{TP^{(1)}}$ was supposed to be 0.

$$\bar{x}_{TP^{(1)}} = \frac{\sum_{i=1}^{k^{(1)}} pix_{TP^{(1)}}^{(i)}}{k^{(1)}} \quad (3)$$

$$SD_{TP^{(1)}} = \sqrt{\frac{\sum_{i=1}^{k^{(1)}} \left| pix_{TP^{(1)}}^{(i)} - \bar{x}_{TP^{(1)}} \right|^2}{k^{(1)}}} \quad (4)$$

Finally, the mean values $\left\{ \bar{x}_{TP^{(1)}}, \bar{x}_{TP^{(2)}}, \dots, \bar{x}_{TP^{(n)}} \right\}$ and SD values $\{SD_{TP^{(1)}}, SD_{TP^{(2)}}, \dots, SD_{TP^{(n)}}\}$ for all n tie points were calculated, and a graph of the frequency distribution of the SD values was generated. This tie point data fusion approach was repeated, and multiple SD distributions under varying conditions with various factors were obtained.

3.5. Evaluation

We selected three widely used evaluation methods: (1) the simple average value (SAV), (2) the earth mover's distance (EMD), and (3) the mean square errors (MSE) evaluation for images. SAV is an essential statistical method, EMD is commonly used to measure distance between two distributions, and the MSE is generally used to assess the average of squares of errors. There are other methods measuring the differences between two distributions, such as the Kullback-Leibler divergence

(KLD) and the Jensen–Shannon divergence (JSD). However, KLD was not symmetric and could not be strictly considered as a metric [80,81], while JSD was not sensitive to small changes in datasets [80]. Therefore, EMD was used in this study.

Since multiple SD distributions were obtained under varying conditions, these selected evaluation methods were all used to explore different performances of the tie point data fusion approach under varying conditions.

3.5.1. Simple average value (SAV)

Performance of the data fusion approach was first evaluated through a frequency distribution analysis of the SD values by calculating the simple average value (SAV) of all SD values. As shown in Eq. (5), we had (n) tie points and corresponding (n) SD values noted as $\{SD_{TP^{(1)}}, SD_{TP^{(2)}}, \dots, SD_{TP^{(n)}}\}$. $\bar{\sigma}$ represents the average of all SD values.

$$\bar{\sigma} = \frac{\sum_{i=1}^n SD_{TP^{(i)}}}{n} \quad (5)$$

The calculated value was defined by Eq. (5) as a test value, $\bar{\sigma}_{test}$. As shown in Fig. 3, each tie point's SD value was calculated based on pixels from various images. For example, $SD_{TP^{(1)}}$ for $TP^{(1)}$ was calculated based on the corresponding $k^{(1)}$ number of pixels in $k^{(1)}$ number of various images (one pixel per image). Each tie point's corresponding pixel in various images referred to the same object. Therefore, the $k^{(1)}$ number of pixels had the same thermal value, to be precise, the ideal value for $SD_{TP^{(1)}}$ was supposed to be 0. Similarly, other ($n-1$) SD values $\{SD_{TP^{(2)}}, \dots, SD_{TP^{(n)}}\}$ were also 0. Furthermore, the ideal value $\bar{\sigma}_{ideal}$ for the SD values' distribution should have also been 0. Unfortunately, due to errors caused by the distance between cameras and objects and by acquisition angles, the thermal images presented different values for the same object. Thus, $\bar{\sigma}_{test}$ was not 0. Comparing $\bar{\sigma}_{test}$ to $\bar{\sigma}_{ideal}$, smaller $\bar{\sigma}_{test}$ demonstrated that the method by which we replaced RGB information with thermal information in a model performed better under corresponding conditions.

3.5.2. Earth Mover's distance (EMD)

In this evaluation method, the frequency (probability) SD values' distribution obtained from Section 3.4 is defined as a test distribution, and the distribution consisting of ideal SD values is the ideal distribution. Ideal distribution is described as all tie points individually receiving equal temperature values from thermal images, meaning that all points respectively had an SD value of 0. As shown in Fig. 4, the distribution in (a) is one possible test distribution while (b) is the ideal distribution. The horizontal axis represents all observed possible SD values in the distribution, and the vertical axis represents the number of tie points with the given SD values.

This evaluation measuring the discrepancy (similarity) between the two distributions is different than the first evaluation presented in Section 3.5.1. It requires that the two distributions have the same frequency (probability) space, which means that as long as observed possible SD values in the test distribution have a non-zero value of number of tie points having the given SD values, the corresponding SD values in the ideal distribution should also have a non-zero value of number of tie points. However, the ideal distribution just had the frequency (probability) of a 0 SD value, and the number of tie points with other observed possible SD values was null. As shown in Fig. 4 (b), the red arrow indicates a bar showing the number of tie points with an SD value of "0." The ideal distribution needed a slight adjustment to fill the vacancy of the observed possible SD values in which the number of tie points was null. Thus, in order to use the evaluation methods introduced in this section, we had to maintain the same frequency (probability) space between the test distribution and the ideal distribution. The adjustment allowed a frequency of every ideal distribution's possible SD value, except the SD value of "0," to add one. As shown in the detailed image in the rectangle in Fig. 4 (b), the number of tie points that had

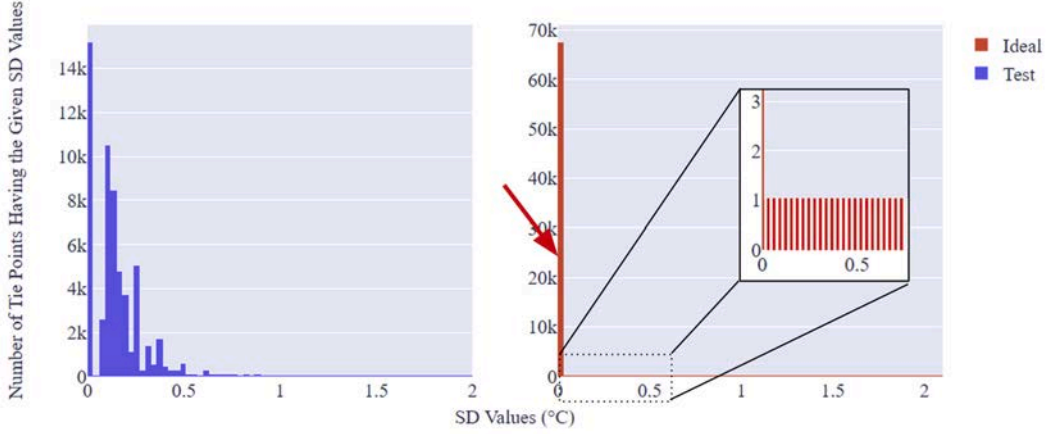


Fig. 4. Test distribution and ideal distribution.

observed possible SD values (except the SD value of 0) were all 1 s (non-negative as expected) instead of 0 s. This idea was to steal the frequency mass, which is similar to the add-one smoothing statistical method (also called Laplace smoothing) [82,83]. The test SD values' distributions conducted under varying conditions were compared with the ideal SD values' distribution to determine the performances.

The distribution comparison metric used in this study was EMD [84]. Intuitively, two distributions can be perceived as a collection of blue bins for test distribution, and a collection of red bins for ideal distribution (see Fig. 4). Bins or class intervals were equal parts into which a fixed total range of a data set is divided. If the test distribution contained a mass of earth fully spread in these blue bins and the other as a collection of empty red bins in which each bin's capacity was the height of the bin, the EMD then measured the least amount of work needed to fill the red bins with earth from blue bins. This explains why we adjusted the ideal distribution's space, since the plan was to have the same number of blue and red bins.

Assume that the test and ideal SD values' distributions were A and B. A had m bins noted as $A = \{(a_1, w_{a1}), (a_2, w_{a2}), \dots, (a_m, w_{am})\}$ in which a_i was the blue bin and w_{ai} represented the weight of earth in the blue bins; whereas B had n red bins with $B = \{(b_1, w_{b1}), (b_2, w_{b2}), \dots, (b_n, w_{bn})\}$ in which b_i was the red bin and w_{bi} represented the capacity of red bins. Imagine moving a mass of earth from one of the A's bins into one or some of B's bins, and let $D = [d_{i,j}]$ as the distance between bins a_i and b_j . Now, the one-time earth transportation was defined as $F = [f_{i,j}]$ in which $f_{i,j}$ was the weights of earth moved between a_i and b_j . The cost of this transportation was the product of weight $f_{i,j}$ and distance $d_{i,j}$. The goal was to minimize the overall cost of moving all earth from blue bins to red bins, as shown in Eq. (6).

$$\min \sum_{i=1}^m \sum_{j=1}^n f_{i,j} d_{i,j} \quad (6)$$

However, the Eq. (6) was not unlimited. It was subject to the following constraints, from Eq. (7) to Eq. (10). First, transportation could only flow from distribution A to B and not vice versa, as shown in Eq. (7).

$$f_{ij} \geq 0 \quad 1 \leq i \leq m, 1 \leq j \leq n \quad (7)$$

The amount of earth moved from the A bin to the B bin could not exceed the capacity of the corresponding blue bins, as shown in Eq. (8).

$$\sum_{j=1}^n f_{ij} \leq w_{ai}, \quad 1 \leq i \leq m \quad (8)$$

The amount of earth moved from the A bin to the B bin could not exceed the capacity of the corresponding red bins, as shown in Eq. (9).

$$\sum_{i=1}^m f_{ij} \leq w_{bj}, \quad 1 \leq j \leq n \quad (9)$$

The total amount of earth moved from all A bins to all B bins could not exceed the minimum value between the total bin capacity of A and B, as shown in Eq. (10). This amount was the total flow.

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} \leq \min \left\{ \sum_{i=1}^m w_{ai}, \sum_{j=1}^n w_{bj} \right\} \quad (10)$$

Once the abovementioned linear optimization problem was solved, the optimal flow $F = [f_{i,j}]$ was found, and the EMD was defined as normalizing the work by the total flow, as shown in Eq. (11).

$$EMD(A, B) = \frac{\sum_{i=1}^m \sum_{j=1}^n f_{i,j} d_{i,j}}{\sum_{i=1}^m \sum_{j=1}^n f_{i,j}} \quad (11)$$

A higher EMD value score represents a larger difference between the test and the ideal distribution. Therefore, EMD could be used to compare different test distributions under varying conditions with the ideal distribution in order to determine which test distribution is more similar to the ideal distribution, thus indicating under which test condition the tie points fusion method performs better.

3.5.3. MSE evaluation for images

In contrast to the first two evaluation methods that focused on evaluating tie points, the third evaluation method focused on evaluating thermal images. Thermal images were compared with their relevant tie points in the 3D model reconstructed by RGB images. In the data fusion approach (Section 3.4), each tie point received various thermal values from the k relevant thermal images among all m thermal images for example, $\{TP^{(1)} : (V_1, T_1), \dots, (V_k, T_k)\}$ for the first tie point of the n tie points received thermal values from $k^{(1)}$ relevant thermal images. Thus, the mean value was defined as the final thermal value (FTV) for example, $\bar{x}_{TP^{(1)}}$ for the first tie point calculated in Eq. (12). Each tie point in tie point data fusion had its final thermal value (FTV).

$$\bar{x}_{TP^{(1)}} = \frac{\sum_{i=1}^{k^{(1)}} \text{pix}_{TP^{(1)}}^{(i)}}{k^{(1)}} \quad (12)$$

The tie point received several thermal values from the pixels in several images, and in reversing that process, several such pixels in one thermal image can find several related tie points in a 3D model. This is a many-to-many relationship. The differences between the thermal values of pixels in the thermal images and their relevant tie points' final

thermal value (FTV) are due to the data fusion method in varying conditions. Such differences can be calculated in Eq. (13), which is called the mean square error (MSE). If the FTV is a measurement for a tie point, the MSE is a measurement for an image. A detailed explanation of the MSE for one thermal image is as follows. As previously mentioned, the tie point data fusion approach was evaluated in different experiments under varying conditions. Assume each experiment had m thermal images and thus m MSE values. Assume for every MSE value of every thermal image, it had l relevant tie points. For example, $MSE^{(1)}$ had $l^{(1)}$ relevant tie points that corresponded to $l^{(1)}$ pixels in this image; pix was then used to denote the value of a pixel in a thermal image. As shown in Fig. 3, for example, $pix_{TP^{(1)}}^{(1)}$ represented a pixel in the first thermal image with a relevant tie point $TP^{(1)}$. Similarly, $pix_{TP^{(2)}}^{(1)}$ represented another pixel in the first thermal image with a relevant tie point $TP^{(2)}$. Finally, pixels, $\{pix_{TP^{(1)}}^{(1)}, pix_{TP^{(2)}}^{(1)}, \dots, pix_{TP^{(l^{(1)})}}^{(1)}\}$, in the first thermal image can be used to calculate $MSE^{(1)}$, as defined in Eq. (13).

$$MSE^{(1)} = \frac{1}{l^{(1)}} \sum_{i=1}^{l^{(1)}} \left[pix_{TP^{(i)}}^{(1)} - \bar{x}_{TP^{(i)}} \right]^2 \quad (13)$$

There were m MSE values for each experiment. The average of the total MSE values for one experiment, \overline{MSE} , maximum and minimum value among all m MSE values MSE_{max} and MSE_{min} could be calculated by Eqs. (14), (15), and (16). If MSE is a measurement for one image, then MSE_{max} , MSE_{min} , and \overline{MSE} are measurements for one experiment.

$$\overline{MSE} = \frac{1}{m} \sum_{i=1}^m MSE^{(i)} \quad (14)$$

$$MSE_{max} = \max\{MSE^{(1)}, MSE^{(2)}, \dots, MSE^{(m)}\} \quad (15)$$

$$MSE_{min} = \min\{MSE^{(1)}, MSE^{(2)}, \dots, MSE^{(m)}\} \quad (16)$$

In summary, the datasets were run with various tested factors discussed in Step one through the approach explained in Section 3.4. The various tested factors included: (1) altitude comparison (60 m and 35 m), (2) camera angle comparison (45 degrees and 30 degrees), (3) flight path design comparison (mesh grid and Y path), and (4) comparison of various building styles in urban areas (a campus with sparsely spaced buildings and a dense city area). The three evaluation methods included: (1) the simple average value, (2) the earth mover's distance, and (3) an MSE evaluation for images. These methods were used to measure and compare the performance of the proposed tie point data fusion approach under the four different data acquisition testing factors previously described.

4. Experiments

IRT for energy audits usually has a requirement for temperature differences, which is at least 10 °C (18 °F) between indoor and outdoor [13]. In order to meet this requirement, we conducted our research on university campuses and in a city center in the winter in Karlsruhe, Germany. Houses are well insulated in Germany. The average recommended temperature there was 17 °C (63 °F) for indoor living, and the outdoor temperature was 5 °C (23 °F) at 7 AM when we conducted our experiments. This temperature difference met the requirement as previously described. All experiments we conducted in this study are explained in the following sections. Each of the four different experiments corresponds with one of the four test data collection factors as stated in the previous paragraph. After conducting each experiment, the performances were evaluated.

4.1. Experiment 1: altitude comparison

We tested two different altitudes, 35 m and 60 m, in experiment 1. As described in the research method section, we calculated the first two

evaluation methods, the SAV and EMD values, for these two flight altitudes and the results are summarized in Table 3. In Table 3, the condition of each experiment with various test factors is abbreviated in the header. For example, "Campus_45°_Mesh_35m" represents the corresponding experiment that was conducted on a campus with a camera angle of 45 degrees, a flight altitude of 35 m, and a mesh grid flight path. Therefore, "Campus_45°_Mesh_35m" (Column 1) and "Campus_45°_Mesh_60m" (Column 2) were two experiments conducted and abbreviated as "35 m" and "60 m" in experiment 1. The rest of the columns are for additional experiments described in Sections 4.2, 4.3, and 4.4.

The first and second columns show that using an altitude of 60 m to collect data for tie point data fusion can introduce errors because of the higher SAV and EMD values. A possible explanation is that the higher a drone flies, the further a thermal sensor is from the objects, and thus, accurately detecting thermal values becomes more difficult.

Fig. 5 shows the SD values' distributions for these experiments, and (a) shows the SD values' distribution for a flight altitude of 35 m. The x-axis in Fig. 5 (a) represents all possible SD values, and the y-axis represented the number of tie points with the corresponding given SD values. For example, 10,494 tie points, the greatest number of tie points, had SD values ranging from 0.3–0.4 in Fig. 5 (a). Fig. 5 (b) shows the same SD distribution as Fig. 5 (a) and is visualized with various colors. As described in the research method, each tie point had a different number of k of related thermal images, for example $TP^{(1)}$ had $k^{(1)}$ number of related thermal images. The color coding changed with k . If number k ranged from 0 to 5, it was blue. If k ranged from 5 to 10, it was red, and if more than 10, it was green. Similarly, Fig. 5 (c) and (d) show the two SD distributions for a flight altitude of 60 m. Comparing Fig. 5 (a) and (c), we observe that the highest numbers of tie points were 10,494 and 3544 for the 35 m experiment and for the 60 m experiment, which respectively fell between 0.3 and 0.4 of possible SD values for the 35 m experiments, and between 0.7 and 0.8 for the 60 m experiment. A distribution of the 35 m visually moved closer to the vertical axis than a distribution of the 60 m. This also demonstrated that the 35 m experiment performed better. The rest of the figures are for additional experiments described in Sections 4.2, 4.3, and 4.4.

4.2. Experiment 2: camera angle comparison

In experiment 2, we examined two distinct camera angles 30 degrees and 45 degrees while all other flight configurations remained the same. The SD distribution of these experiments can be seen in Fig. 5, and the numerical data are summarized in Table 3. In this experiment, the two criteria, SAV and EMD values, for the two different camera angles were not consistent as shown in Table 3. For example, the 30-degree flight configuration (Column 3), when compared to the 45-degree flight configuration (Column 2) had higher EMD values, but lower SAV values. More evaluation criteria were then needed to allow the comparison of the performances for the 30-degree and the 45-degree experiments.

In Fig. 5, (e) and (f) the two distributions for the 30-degree camera angle are shown. Similarly, Fig. 5 (c) and (d) show the two distributions for the 45-degree camera angle. We found that the number of tie points having more than 5 related thermal images ($k \geq 5$) was greater in the 45-degree experiment than in the 30-degree experiment. One possible explanation is that deploying the camera at 45 degrees allowed the detection of more facade details where more feature points can be recognized. The research targets in this experiment were modern buildings with fewer decorations and details on their flat roofs. Since the 3D model and tie points were reconstructed and detected by RGB images, features on facades were more heterogeneous and easier for algorithms to detect when compared to the less-diverse features on flat roofs.

Table 3

Statistics of evaluation criteria of different experiments conducted with varying test factors.

Abbreviation of each experiment	Campus_45°_Mesh_35m	Campus_45°_Mesh_60m	Campus_30°_Mesh_60m	Campus_30°_Y_60m	City_45°_Mesh_60m
Abbr. in Experiment 1	35 m	60 m	–	–	–
Abbr. in Experiment 2	–	45°	30°	–	–
Abbr. in Experiment 3	–	–	Mesh Grid Path	Y Path	–
Abbr. in Experiment 4	–	Campus	–	–	City
SAV	1.36849	1.74714	1.26505	1.15278	1.18603
EMD Value	0.56933	0.8608	0.93317	0.70301	0.58232

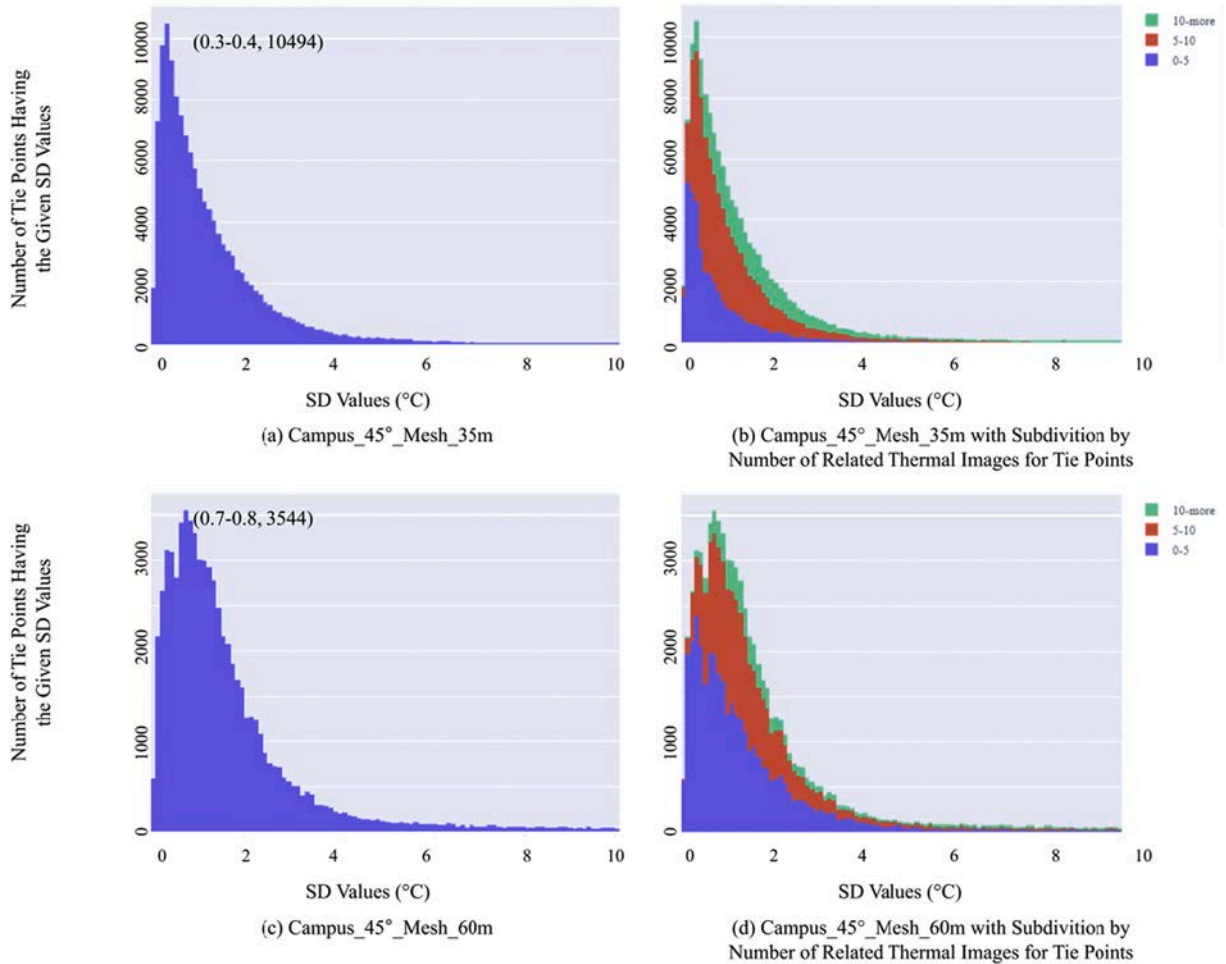


Fig. 5. The distributions of SD values for the different experiments.

4.3. Experiment 3: flight path comparison

In experiment 3, we tested two different flight paths: a mesh grid and a Y path. The evaluation numerical data for these experiments are summarized in Table 3, and the tie points SD distributions are found in Fig. 5. In Table 3, a mesh grid flight path (Column 3) had higher SAV and EMD values. In Fig. 5, (e) and (f) are the two distributions for the mesh grid flight path. Similarly, (g) and (h) are the two distributions for the Y path. Compared with Fig. 5 (e) and (g), we can see highest numbers of tie points are 9253 and 4477 for the mesh grid experiment and the Y path experiment, which respectively fell between 0.3 and 0.4 of possible SD values for the former experiment and between 0.2 and 0.3 for the latter experiment. The distribution of the Y path visually moved closer to the vertical axis than the distribution of the mesh grid. This observation demonstrates that the Y flight path is closer to the ideal distribution.

4.4. Experiment 4: comparison of various building styles in urban areas

In experiment 4, we tested two distinct urban areas with two building types. Our experiments were conducted on a campus area where the modern buildings were constructed relatively far away from each other in a spacious area (shown in Fig. 6 (a)), and in a city area with traditional and closely built European architectures in a dense urban area (shown in Fig. 6 (d)). The magnified images in Fig. 6 (b) and (e) showed tie points with the RGB information reconstructed and detected by the RGB images for the campus experiment and the city experiment, respectively. The magnified images in Fig. 6 (c) and (f) showed tie points fused with thermal information from thermal images for the campus experiment and the city experiment, respectively. The dark purple color represents a lower thermal value, and the lighter yellow color represents a higher thermal value in Fig. 6 (c) and (f).

The experiment statistics are summarized in Table 3 and their SD

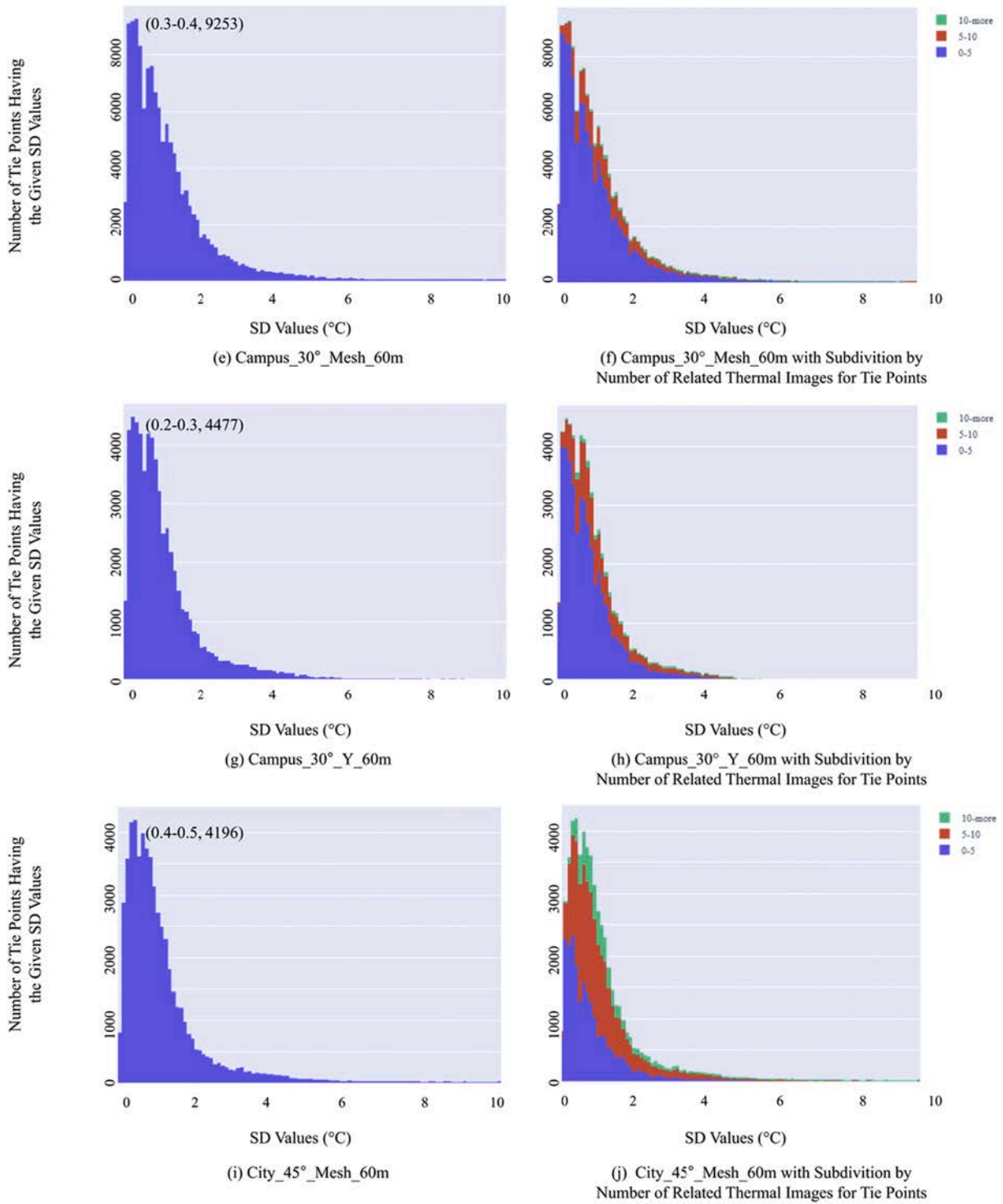


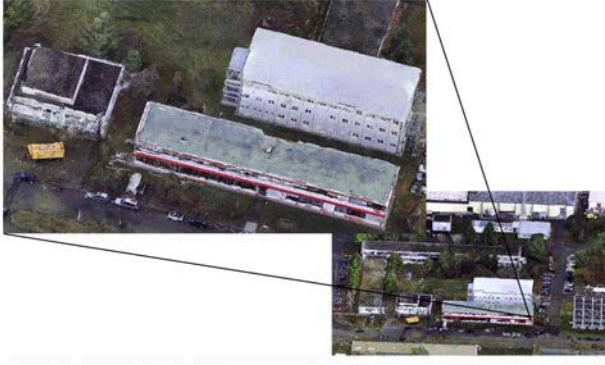
Fig. 5. (continued).

values' distributions are found in Fig. 5. In Table 3, the comparison between the campus area (Column 2) and the city area (Column 5) show that data fusion errors were generally fewer when conducting research in the city area, because of the smaller SAV and EMD values. European traditional buildings in cities are more complex in terms of architectural details than modern campus buildings with features that the algorithms more easily detected. Thus, fewer errors were introduced in the city experiment. In Fig. 5, (c) and (d) show distributions for the campus area. Similarly, Fig. 5 (i) and (j) show distributions for the city area. Distribution for the city experiment in Fig. 5 (i) was closer to the ideal

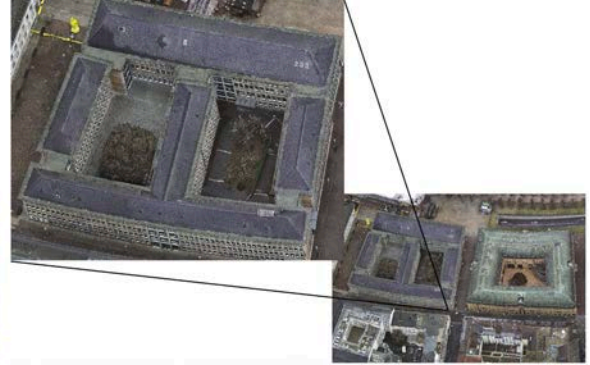
distribution than that of the campus experiment as demonstrated by Fig. 5 (c).

4.5. MSE values

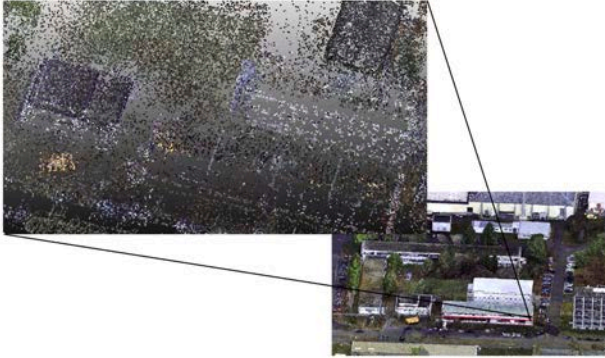
As described in Section 3, each experiment was conducted with m pairs of thermal and RGB images. After we fused the tie points' RGB information with the thermal information, we calculated all m MSE values for the m thermal images, based on the fused tie points. The MSE values of each experiment are summarized in Table 4. According to



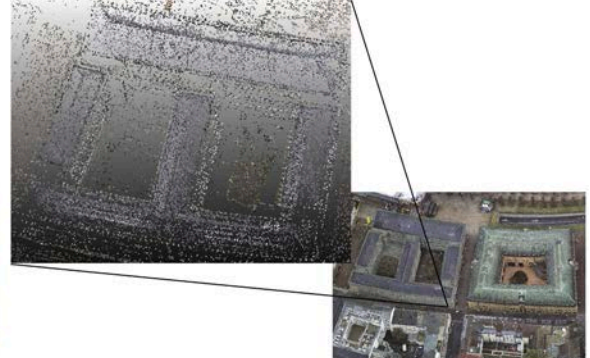
(a) Campus Area: Completed RGB Point Cloud Model



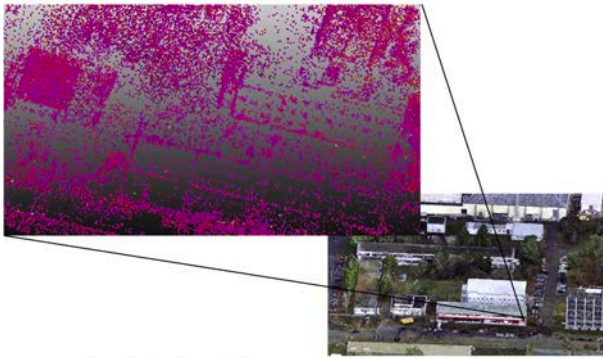
(d) City Area: Completed RGB Point Cloud Model



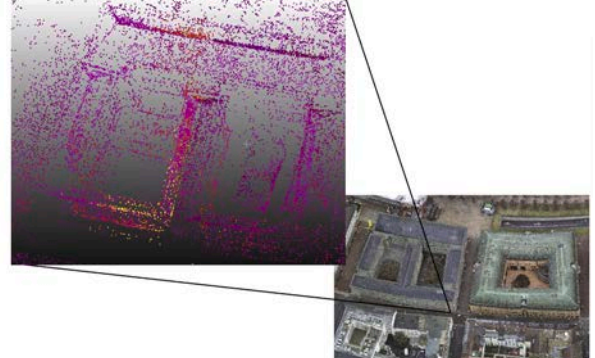
(b) Campus Area: RGB Tie Points Model



(e) City Area: RGB Tie Points Model



(c) Campus Area: Thermal Tie Points Model



(f) City Area: Thermal Tie Points Model

Fig. 6. Various mapping areas.

Table 4

MES evaluation statistics of different experiments conducted with varying test factors.

Abbreviation of each experiment	Campus_45°_Mesh_35m	Campus_45°_Mesh_60m	Campus_30°_Mesh_60m	Campus_30°_Y_60m	City_45°_Mesh_60m
MSE_{max}	221.31286	214.53633	60.29066	632.37457	188.73010
MSE_{min}	0.00754	0.01038	0.02433	0.00779	0.00000
\overline{MSE}	7.73743	13.19546	3.73976	6.32774	6.69587

Table 4, “Campus_45°_Mesh_60m” had the highest \overline{MSE} value and Campus_30°_Mesh_60m had the lowest \overline{MSE} value.

During the collection of the thermal and RGB images pairs, we recorded their GPS information. Therefore, we were able to draw a heatmap that graphically represents the m MSE values. In the heatmaps shown in Fig. 7, the coordinates are the thermal image GPS coordinates

(the gray dots) and the values are the image MSE values. Color coding represents different values. The brighter color, hot yellow spots represent higher MSE values, while the transparent blue color represents lower MSE values. Examples can be seen in Fig. 7. Fig. 7 (a), (b), (c), (d) and (e) are the five MSE heatmaps for the experiments “Campus_45°_Mesh_35m,” “Campus_45°_Mesh_60m,”

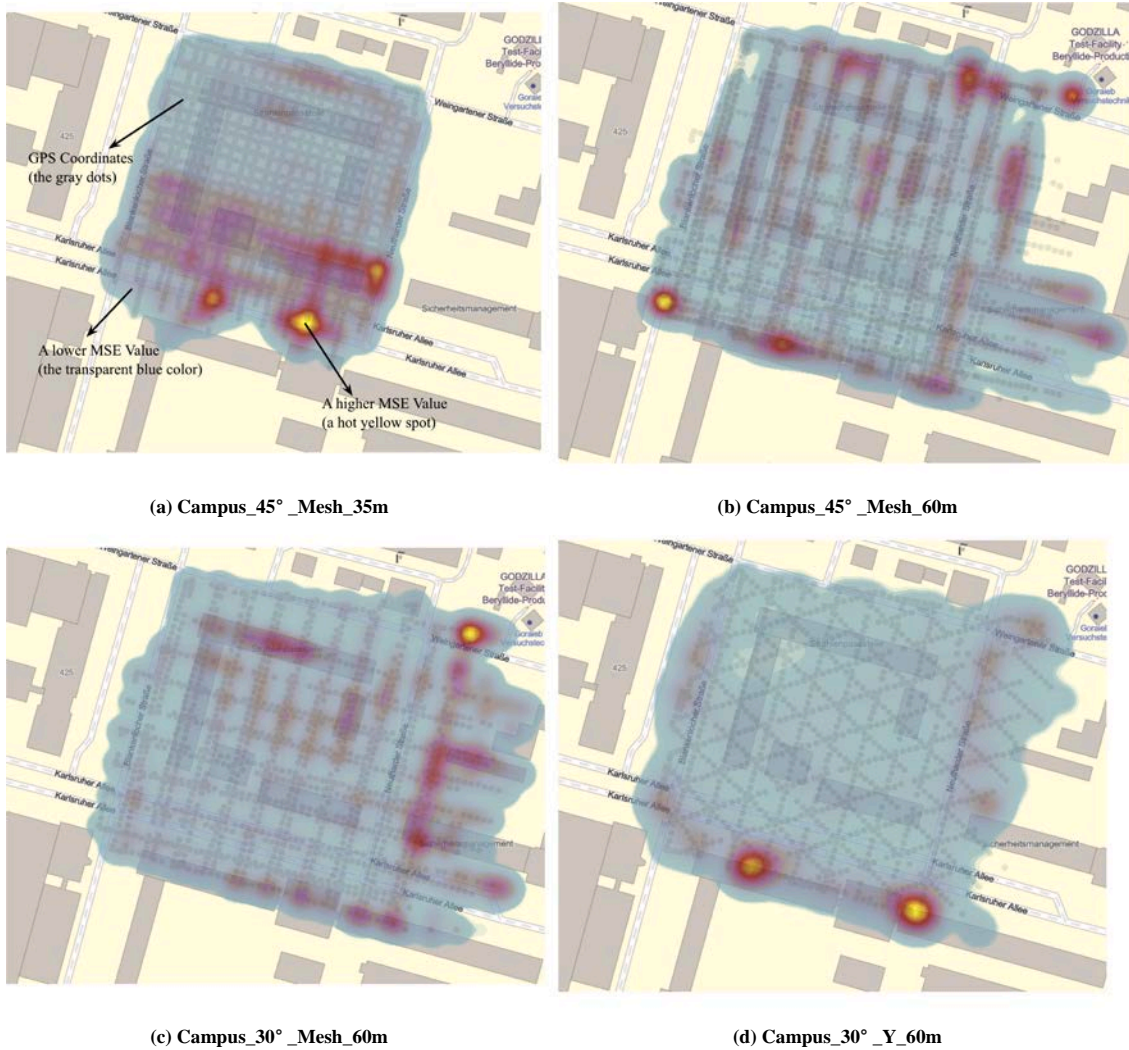


Fig. 7. Heatmaps of MSE values.

“Campus_30°_Mesh_60m,” “Campus_30°_Y_60m,” and “City_45°_Mesh_60m.” As can be seen in Fig. 7, all cases have a common trait in that high MSE areas are at the edges of areas of interest where the images were taken.

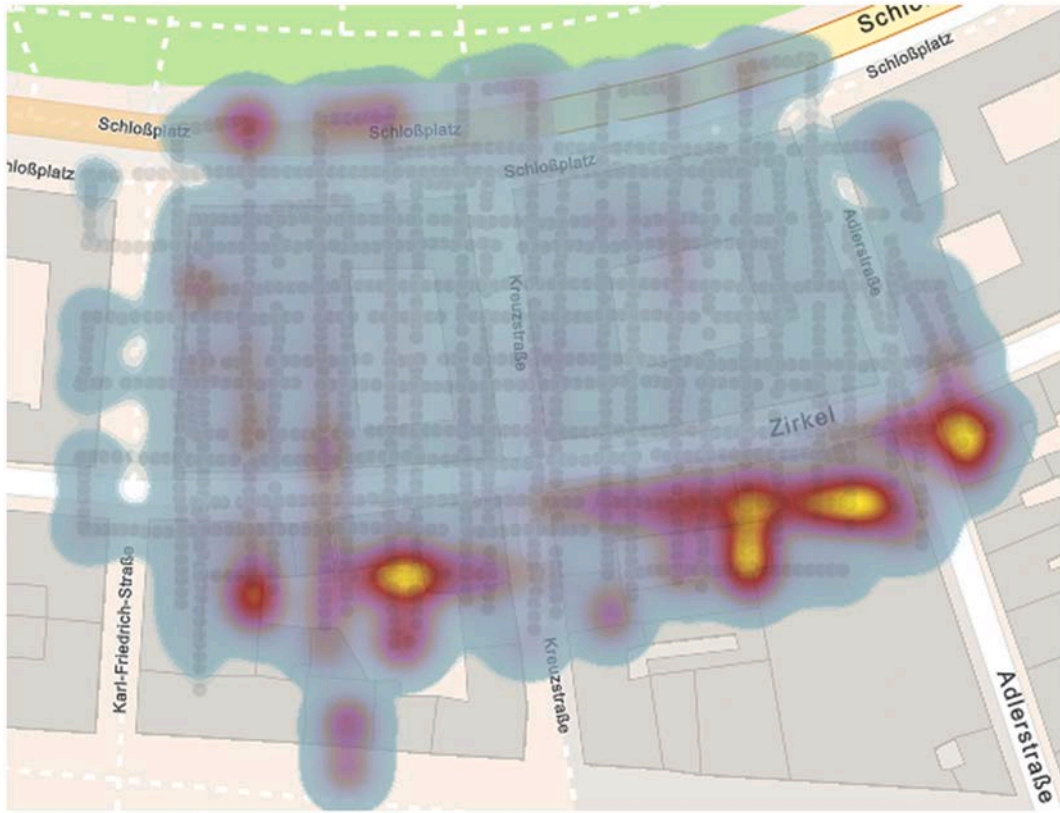
5. Discussion

5.1. SAV and EMD value analysis

Considering the thermal image quality, Rahaghi et al. 2019 argued that the accuracy of stitched-image edges was important [85]. Hoegner et al. 2009 found that features seldom matched at the edges of the images in a thermal mapping model [63]. It has been confirmed that the accuracy of recorded temperature values might vary depending on pixel location in thermal images [86]. The pixels in the center of a thermal image usually yield a more accurate temperature record than pixels in corners or edges. Therefore, thermal information from corners and edges of images that are projected onto a 3D point cloud might result in more errors. The thermal images $T^{(m)}$ shown in the upper right corner of Fig. 3 illustrate the concept of the corners and edges of thermal images. In $T^{(m)}$, 70% of the thermal image’s pixels were in the center, and the rest were in corners or edges.

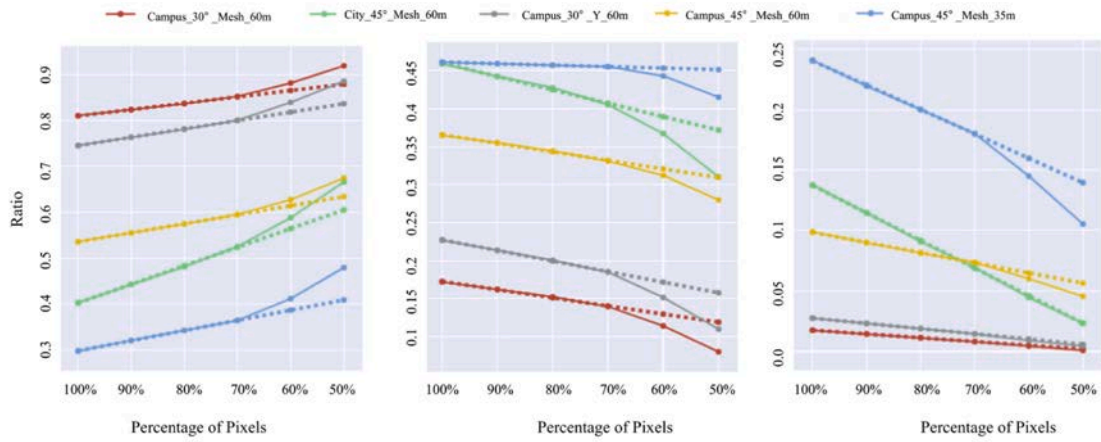
Eliminating a proper number of corner or edge pixels could reduce errors, but if too many pixels were eliminated, accuracy would be

difficult to guarantee. For an extreme example, tie points could only receive one thermal value from one thermal image. In order to determine a proper pixel-maintenance percentage, the ratios of a certain number of tie points to the total number of tie points in all experiments under varying conditions conducted in the study are plotted in Fig. 8. The number of tie points was decided by the numbers of thermal values that tie points can receive. The numbers of received thermal values were 0–5, 5–10, and 10–more. The experiments under varying conditions were: (1) Campus_45°_Mesh_35m, (2) Campus_45°_Mesh_60m, (3) Campus_30°_Mesh_60m, (4) Campus_30°_Y_60m, and (5) City_45°_Mesh_60m. The horizontal axes represent the percentage of pixels in the thermal images used in the different experiments. Six cases were tested: 100%, 90%, 80%, 70%, 60%, and 50%. In the same experiment, models and tie points were created and detected using the same RGB images, and the percentage of pixels used for the tie point data fusion only referred to the thermal images. For example, a 100% case meant that all the thermal images were used for the data fusion approach and a 70% case meant that 70% of the pixels in the thermal images were used for the data fusion approach. The vertical axes represent the ratio of the number of tie points to the total number of tie points as previously discussed. In Fig. 8 (a), (b), and (c), the numbers of received thermal values were 0–5, 5–10, and 10–more, respectively. As shown in Fig. 8, if the scatter dots were connected with solid lines and the dashed regression lines were drawn to fit dots between the 100% and the 70% case,



(e) City_45°_Mesh_60m

Fig. 7. (continued).



(a) The number of received thermal values: 0–5

(b) The number of received thermal values: 5–10

(c) The number of received thermal values: 10–more

Fig. 8. Changes of three ratios with different maintenance pixel percentages.

the slopes of the lines after the 70% case then suddenly changed. To be precise, the slopes increase in Fig. 8 (a) and decrease in Fig. 8 (b) and (c).

We also plotted the evaluation of the tie point data fusion approach in different experiments using various pixel percentages in the thermal images (see Fig. 9). The horizontal axes represent the pixel percentage in the thermal images used in different experiments. A vertical axis in Fig. 9 (a) represents the SAV evaluation, and in Fig. 9 (b) it represents the EMD evaluation. These two figures show a similar trend to the three figures in

Fig. 8, indicating that 70% of the pixels mark a watershed. To be precise, reducing pixels in the corner and using no fewer than 70% of the pixels in the center of the thermal images to conduct the tie data fusion approach could increase accuracy, because the SAV and the EMD values are decreasing. However, if the SAV and EMD values started to plunge when continuously reducing pixels, it meant that the tie points received far fewer thermal values, and the results could be inaccurate. Based on these discussions, 70% is the proper pixel-maintenance percentage. .

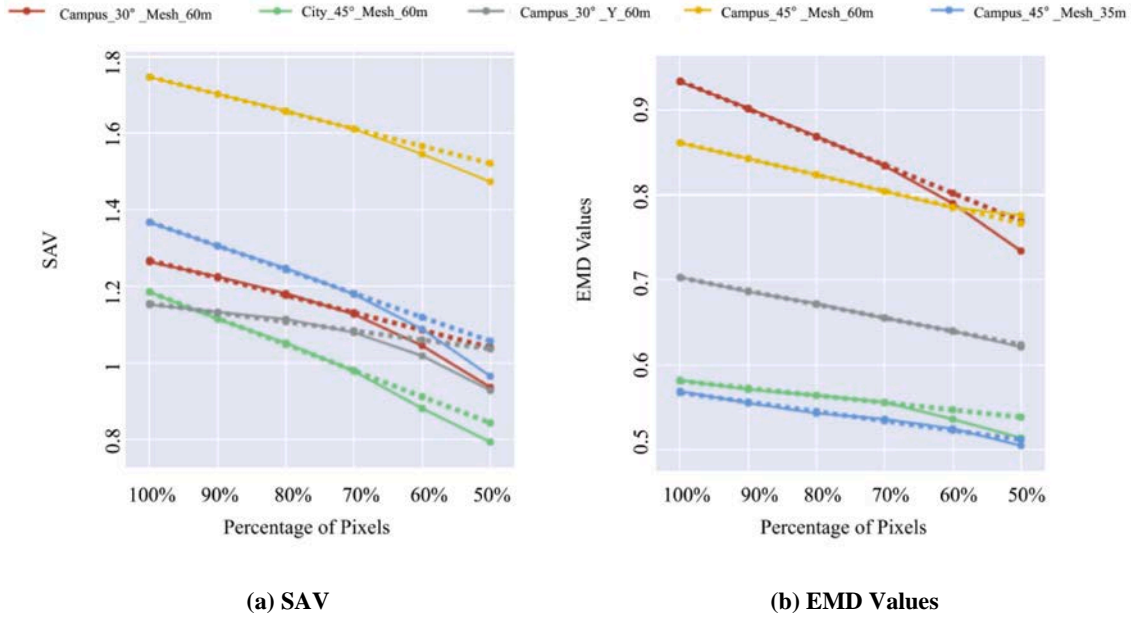


Fig. 9. Changes in SAV and EMD values with a different pixel-maintenance percentage.

For the flight altitude, as seen in Fig. 8, we concluded that if the drone flew higher, such as, the red, gray, and yellow lines with altitudes of 60 m, the tie point percentages with five or more related numbers of thermal images would be lower when compared to the other lines. This observation confirms that using high flight altitude could reduce the ability of photogrammetry algorithm to detect more tie points derived from the RGB images. Therefore, fewer thermal images were linked to the tie points. As seen in Fig. 9, red, gray, and yellow lines generally have higher SAV and EMD values than the other lines. This observation demonstrates that a higher flight altitude could introduce more errors than a lower flight altitude. The potential explanation was that if the drone flew higher, the thermal sensors would be further from the objects, which could then reduce the thermal sensor performances. Rakha et al. 2018 summarized the current research and showed that it was important to investigate the altitude at which the UAS flew [87]. Our study shows the relation between flight altitude and data collection accuracy.

For camera angles, as seen in Fig. 8, blue lines and green lines are shown for the 45-degree experiments that had more tie points with five or more related numbers of thermal images than the red lines and gray lines for the 30-degree experiments. This observation indicates that 45 degrees could enhance the possibilities for tie points to find more thermal image sources. We speculate that a 45-degree camera angle can capture more façade detail areas than a 30-degree angle. Features on façades are easier to detect because shapes and characteristics are more diverse than those found on roofs. For example, there are many components on façades, such as a variety of shapes of windows and doors. The study by Hoegner et al. 2009 had a similar observation that most feature points in their model were detected at the window edges [63]. When tie points are easier to detect, more thermal image sources could be linked to the tie points. As demonstrated in Fig. 9, both the blue and green lines generally have lower EMD values. The green line also has the lowest SAV, but the blue line does not correspondingly have a lower SAV. Therefore, these observations support the idea that although a 45-degree camera angle might introduce fewer errors, it would not necessarily perform better than a 30-degree camera angle. More evaluation criteria are needed.

As for flight paths, as seen in Fig. 8, the Y path (gray line) received more thermal values from thermal images than the mesh grid path (red line). According to Fig. 9, the Y path experiment had both a lower SAV

and a lower EMD value compared to the mesh grid experiment, and thus the Y path performed better than the mesh grid path in the proposed tie point data fusion approach.

For a comparison of various building styles in urban areas, as shown in Fig. 8, the tie points in the city model (green line) received more thermal values from images than those in the campus model (yellow line). According to Fig. 9, the tie-point SD values' distribution for the city model had a lower SAV and a lower EMD value than the distribution for the campus model. A potential explanation for this observation could be the discrepancy of architectural styles. Modern buildings on campus are simple with flat roofs, while European traditional urban buildings are more complex with many designs on the façades and roofs. Therefore, features on traditional buildings' façades and roofs were easier to detect with the algorithms, and since the features were precisely captured, the tie points were more accurate, and fewer errors were introduced in this experiment.

5.2. MSE analysis

MSE is a widely used approach to estimate the changes between a test case and a target case. Lin et al. 2019 [59] fused thermal imagery with point cloud models and MSE was used to evaluate the differences between the models and the images. Wiens et al. 2019 used MSE to evaluate corresponding points in two models [72]. Hoegner et al. 2016 [88] projected a 3D point cloud back to the thermal images using recorded camera settings. The quality was measured by comparing the differences of the back projection in different cases using MSE. The MSE evaluation method was used in our study to calculate the errors between tie points in the models and pixels in thermal images.

We determined 70% as the proper percentage of pixels to be maintained as discussed in Section 5.1. Therefore, the MSE evaluation was repeated for all the experiments using just 70% of pixels in the thermal images. First, Table 5 summarizes the highest and lowest MSE values among all thermal images and the average \overline{MSE} of all thermal images used for each experiment. Table 5 presents a comparison of two cases: the whole images' pixels used for the experiments and 70% thermal images' pixels used in the experiments. The table demonstrates that in the 70% case, the average \overline{MSE} values were reduced in all experiments. Second, each model and tie points in each experiment were created and detected by m pair of thermal and RGB images. For example, $\{MSE_{100}^{(1)}$,

Table 5

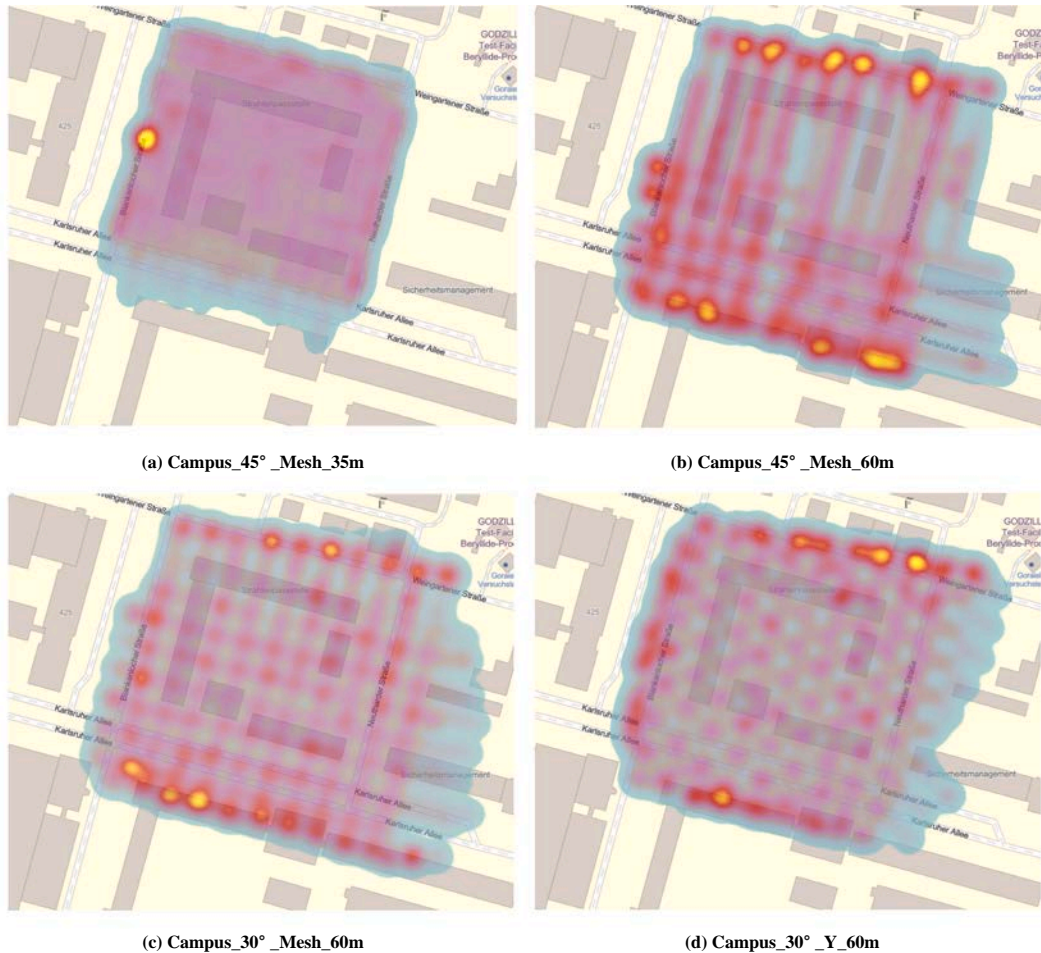
Comparison between using whole thermal images and using 70% of pixels in thermal images for MES evaluation.

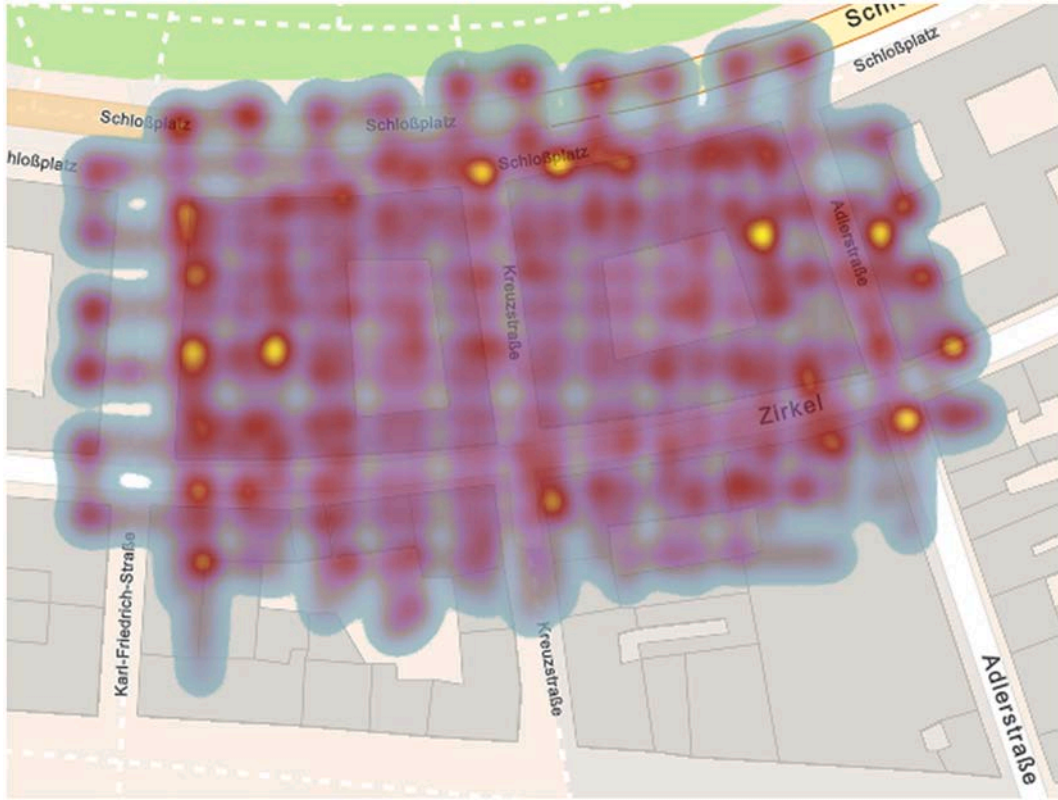
Abbreviation of each experiment	Campus_45°_Mesh_35m	Campus_45°_Mesh_60m	Campus_30°_Mesh_60m	Campus_30°_Y_60m	City_45°_Mesh_60m
Whole Thermal Images					
MSE_{max}	221.31286	214.53633	60.29066	632.37457	188.73010
MSE_{min}	0.00754	0.01038	0.02433	0.00779	0.00000
\overline{MSE}	7.73743	13.19546	3.73976	6.32774	6.69587
70% of Pixels in Thermal Images					
MSE_{max}	436.84775	283.00235	230.32526	632.37457	177.23818
MSE_{min}	0.00000	0.00000	0.02887	0.00000	0.00000
\overline{MSE}	5.66219	10.78289	2.92376	5.79497	5.03659
Ratio of Number of positive $MSE_{diff}^{(i)}$ to Number of all $MSE_{diff}^{(i)}$	80.6%	78.9%	82.3%	74.8%	78.8%

$MSE_{100}^{(2)}, \dots, MSE_{100}^{(m)}$ was used to denote the MSE values for each thermal image in the experiment in which whole pixels in thermal images were used to fuse thermal information with the tie points' RGB information, and $\{MSE_{70}^{(1)}, MSE_{70}^{(2)}, \dots, MSE_{70}^{(m)}\}$ was used to denote MSE values for each thermal image in the experiment in which 70% of the pixels were used to conduct the tie point data fusion. The differences were then calculated between the two cases and $\{MSE_{diff}^{(1)}, MSE_{diff}^{(2)}, \dots, MSE_{diff}^{(m)}\}$ was used to denote the differences of the thermal images' MSE values. If $MSE_{diff}^{(i)}$ was positive, it meant that the 70% case reduced the error. As shown in Table 5, the ratios of the number of positive $MSE_{diff}^{(i)}$ values to the number of all $MSE_{diff}^{(i)}$ values are around 78%, which meant that many thermal images' MSE values were reduced by using 70% of the pixels to conduct the data fusion approach. Additionally, a higher $MSE_{diff}^{(i)}$ signifies that 70% of cases performed better. To visualize the $MSE_{diff}^{(i)}$ values,

heatmaps of $MSE_{diff}^{(i)}$ values were also drawn for different experiments. As shown in Fig. 10, the brighter color, hot yellow spots represent higher $MSE_{diff}^{(i)}$ values while the transparent blue color represents lower $MSE_{diff}^{(i)}$ values. The GPS coordinate information of the flight path is the same as that in Fig. 7. Therefore, to make the figures easier to read, the flight paths are not shown in Fig. 10. In general, the contours of the mesh grid and Y flight path are obvious in each figure, since almost 78% of the images in each experiment have positive $MSE_{diff}^{(i)}$ values. In particular, the thermal images' MSE values at the survey-area edges covered by the drone are reduced more than other MSE values of thermal images, because the MSE values' differences are larger.

In summary, it became clear that larger data collection areas, as well as the use of 70% thermal image pixels would yield higher accuracy results in this study.

**Fig. 10.** Heatmaps of differences of every MSE value (MSE_{diff}) between 100% of cases and 70% of cases.



(e) City_45°_Mesh_60m

Fig. 10. (continued).

5.3. Error analysis and limitations

There are some errors that might influence the fusion approach accuracy. First, errors might be introduced by the thermal camera, which detects different levels of infrared light. The accuracy of temperature values captured by thermal cameras might be influenced by materials with different emissivity rather than their thermal behavior. These errors are not controllable. Second, errors might have occurred during data collection. Our accuracy is related to camera and angle distances between the normal ray of the point cloud surface and the normal ray of the cameras, which can result in different thermal records for the same target. As our data fusion approach presents, the various relevant thermal values for one tie point are supposed to be identical. However, because of this type of error, there is still variance between the thermal values.

There are also some limitations in this study. First, the general performance evaluation criteria were not robust enough to examine different flight angles' influence on the performance of the data fusion approach, which is due to inconsistent results when such performances were examined. Additionally, although we defined a mean value from relevant thermal values for one tie point as its final thermal value, it is better to consider other calculations rather than simple average values. Second, we used images taken from 45-degree and 30-degree camera angles separately; as such, the varieties in performance by combining these different camera angles should be investigated. Also, these experiments were complex and conducted with limited combinations of test factors; therefore, other camera angles in addition to 45 and 30 degrees should be examined. Third, the flight altitude factors tested in this study were discrete rather than continuous. Other flight altitudes should also be examined.

6. Conclusions and future work

In this study, we proposed the first use of the tie point data fusion method. We tested the method under varying conditions to consider four test factors including altitudes, camera angles, flight paths, and comparisons of building styles in urban areas. This focus was not only to understand the performance of the proposed approach under varying conditions, but also to provide suggestions for future data collection.

Several important conclusions can be drawn from this study. (1) This research demonstrated that using high flight altitudes could reduce the ability of tie points to link to more thermal images. (2) The number of tie points, which had the number of k (number of associated thermal images) greater than 5 in the 45-degree experiment was greater than the number of tie points in the 30-degree experiment. The 30-degree experiment also diminished the possibilities for tie points to find more thermal image sources. (3) A Y flight path that was not used in previous research performed better than the mesh grid flight path for the tie point data fusion approach. (4) This research demonstrated that the data fusion approach performed better when surveying the traditional urban European buildings than when surveying modern buildings on a campus. (5) The studies showed that images taken from the edge of the survey areas often introduced more errors as our study illustrated in the heatmaps. (6) Eliminating the proper number of pixels in the corners can reduce data fusion errors, and the pixels in the center of a thermal image usually yield a more accurate record. Considering the accuracy after testing different cases, maintaining 70% of pixels in the central thermal-image area is recommended. The research also for the first time validated that 70% of pixels in the thermal images can reduce MSE values of those images located at the edge of the data collection area, as discussed in Section 5.2.

Based on these findings and discussions in this paper, we have some suggestions for researchers who plan to use the proposed tie point data

fusion approach. (1) Choose a proper flight altitude that is higher than the highest building in the mapping area, but not higher than 1.5 times the highest building. In our experiment, 1.5 times the highest building is an altitude of 35 m and introduced more errors, thus flight altitudes should be lower than 1.5 times the highest building. (2) Selection of camera angles should be based on survey requirements. If the users' interests are roofs, 30-degree camera angles are recommended, but if the users' interests are façades, 45-degree camera angles are better. (3) The Y flight path is recommended for tie point data fusion. Our study for the first time used Y flight path, and it proved to perform better than the mesh grid flight path. (4) Broadening the data collection area to avoid inaccuracies is recommended, as well as using only 70% of pixels in the thermal images and eliminating the pixels at the edges of thermal images as we proved in the discussion.

In the future, several research topics need to be further investigated. First, it will be necessary to refine our evaluations based on the tie points of captured objects such as building roofs, façades, trees, and the ground, because distinct sub-groups of the point clouds have different distances from the cameras, and various angles exist between the normal ray of the point cloud surface and the normal ray of the cameras. The different distances and angles might influence the results as we discussed in the error analysis. Based on these facts, the SD values' distribution of tie points that belong to distinct sub-groups are expected to be distinct. Thus, a detailed performance analysis of the data fusion approach is required. Further, this study only focused on tie point data fusion because of its importance for reconstructing a mapping model. Data fusion for all points in a point cloud model should also be studied.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the CSC (China Scholarship Council), the KIT International Department, and the DAAD (Deutscher Akademischer Austauschdienst, German Academic Exchange Service) for their support. Specifically, the DAAD provided travel support. Furthermore, the authors thank Marinus Vogl, the remote pilot, and his drone company, Air Bavarian GmbH, for their support with equipment and service.

References

- [1] Commission of the European Communities, Action Plan to Improve Energy Efficiency in the European Community, 2000, p. 2000. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0247:FIN:EN:PDF> (accessed Jun. 13, 2020).
- [2] European Commission, Energy Efficiency and Its Contribution to Energy Security and the 2030 Framework for Climate and Energy Policy, 2014, p. 2014. <https://www.consilium.europa.eu/en/policies/climate-change/2030-climate-and-energy-framework/> (accessed Jun. 13, 2020).
- [3] European Commission, 2030 Climate & Energy Framework. https://ec.europa.eu/clima/policies/strategies/2030_en, 2020 (accessed Jun. 13, 2020).
- [4] Legislation of UK, Climate Change Act 2008, 2008, p. 2008. <http://www.legislation.gov.uk/ukpga/2008/27/contents> (accessed Jun. 13, 2020).
- [5] J. Fankhauser, S. Averkhenkova, A. Finnegan, 10 years of the UK Climate Change Act, in: The London School of Economics and Political Science, 2018, pp. 1–43. <https://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2018/03/10-Years-of-the-UK-Climate-Change-Act-Fankhauser-et-al.pdf>.
- [6] K. Appunn, J. Wettengel, Germany's climate action law, *J. Energy Trans.* 2019 (2019). <https://www.cleanenergywire.org/factsheets/germanys-climate-action-law-begins-take-shape> (accessed Jun. 13, 2020).
- [7] Federal Ministry for the Environment, Climate Action Plan 2050 - Principles and Goals of the German Government's Climate Policy, 2016. https://www.bmu.de/fileadmin/Daten_BMU/Pool/Broschueren/klimaschutzplan_2050_en_bf.pdf, 2016.
- [8] The German Energy Agency, Designing energy-efficient buildings. <https://www.dena.de/en/topics-projects/energy-efficiency/buildings/#:~:text=Buildings,2020> account for around 35, residential and non-residential buildings. (accessed Jun. 13, 2020).
- [9] C. Cleveland, C. Morris, Dictionary of Energy, 2014. ISBN: 9780080968124.
- [10] A. Syed, *Advanced Building Technologies for Sustainability*, 2012. ISBN: 9780470546031.
- [11] I. El-Darwish, M. Gomaa, Retrofitting strategy for building envelopes to achieve energy efficiency, *Alexandria Eng. J.* 56 (4) (2017) 579–589, <https://doi.org/10.1016/j.aej.2017.05.011>.
- [12] A. Taibeb, H. Dekkiche, Infrared imaging as a means of analyzing and improving energy efficiency of building envelopes: the case of a LEED gold building, *Procedia Engineering* 118 (2015) 639–646, <https://doi.org/10.1016/j.proeng.2015.08.497>.
- [13] FLIR Systems, An informative guide for the use of thermal imaging cameras for inspecting buildings, solar panels and windmills, in: *Thermal Image Guidebook for Building and Renewable Energy Applications Content 2011*, 2011. http://www.flirmedia.com/MMC/THG/Brochures/T820325/T820325_EN.pdf (accessed Jun. 13, 2020).
- [14] Irish Standard, Thermal performance of buildings - Determination of air permeability of buildings - Fan pressurization method (ISO 9972:2015), 2015. <https://www.iso.org/standard/55718.html>, 2015 (accessed Jun. 13, 2020).
- [15] S. Dudić, I. Ignjatović, D. Sešlija, V. Blagojević, M. Stojiljković, Leakage quantification of compressed air using ultrasound and infrared thermography, *Meas. J. Int. Meas. Confed.* 45 (7) (2012) 1689–1694, <https://doi.org/10.1016/j.measurement.2012.04.019>.
- [16] J.M.A. Márquez, M.Á.M. Bohórquez, S.G. Melgar, A new metre for cheap, quick, reliable and simple thermal transmittance (U-value) measurements in buildings, *Sensors (Switzerland)* 17 (9) (2017), <https://doi.org/10.3390/s17092017>.
- [17] I. Nardi, D. Paoletti, D. Ambrosini, T. De Rubeis, S. Sfarra, U-value assessment by infrared thermography: a comparison of different calculation methods in a guarded hot box, *Energy Build.* 122 (2016) 211–221, <https://doi.org/10.1016/j.enbuild.2016.04.017>.
- [18] Y. Ham, M. Golparvar-Fard, Automated cost analysis of energy loss in existing buildings through, in: *ISARC 2013 - 30th International Symposium on Automation and Robotics in Construction and Mining, Held in Conjunction with the 23rd World Mining Congress*, 2013, pp. 1065–1073, <https://doi.org/10.22260/ISARC2013/0117>.
- [19] Y.K. Cho, Y. Ham, M. Golparvar-fard, Advanced engineering informatics 3D as-is building energy modeling and diagnostics: a review of the state-of-the-art, *Adv. Eng. Inform.* 29 (2) (2015) 184–195, <https://doi.org/10.1016/j.aei.2015.03.004>.
- [20] G. Baldinelli, F. Bianchi, Windows thermal resistance: infrared thermography aided comparative analysis among finite volumes simulations and experimental methods, *Appl. Energy* 136 (2014) 250–258, <https://doi.org/10.1016/j.apenergy.2014.09.021>.
- [21] T. Taylor, J. Counsell, S. Gill, Energy efficiency is more than skin deep: improving construction quality control in new-build housing using thermography, *Energy Build.* 66 (2013) 222–231, <https://doi.org/10.1016/j.enbuild.2013.07.051>.
- [22] Y. Hou, L. Soibelman, R. Volk, M. Chen, Factors affecting the performance of 3D thermal mapping for energy audits in a district by using infrared thermography (IRT) mounted on unmanned aircraft systems (UAS), in: *2019 International Symposium on Automation and Robotics in Construction (ISARC)*, 2019, pp. 266–273, <https://doi.org/10.22260/ISARC2019/0036>.
- [23] Residential Energy Services Network, RESNET Interim Guidelines for Thermographic Inspections of Buildings. <https://www.resnet.us/articles/resnet-revises-interim-guidelines-for-thermographic-inspections-of-buildings/> (accessed Jun. 13, 2020).
- [24] E. Lucchi, Applications of the infrared thermography in the energy audit of buildings: a review, *Renew. Sust. Energy Rev.* 82 (May 2017) (2018) 3077–3090, <https://doi.org/10.1016/j.rser.2017.10.031>.
- [25] A. Kyllili, P.A. Fokaides, P. Christou, S.A. Kalogirou, Infrared thermography (IRT) applications for building diagnostics: a review, *Appl. Energy* 134 (2014) 531–549, <https://doi.org/10.1016/j.apenergy.2014.08.005>.
- [26] Pix4D Mapper Support Page, Pix4D. <https://support.pix4d.com/hc/en-us/articles/202557459-Step-1-Before-Starting-a-Project-1-Designing-the-Image-Acquisition-Plan-a-Selecting-the-Image-Acquisition-Plan-Type#label5>, 2020 (accessed Jun. 13, 2020).
- [27] Y. Ham, M. Golparvar-Fard, An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery, *Adv. Eng. Inform.* 27 (3) (2013) 395–409, <https://doi.org/10.1016/j.aei.2013.03.005>.
- [28] R. Albatici, A.M. Tonelli, M. Chiogna, A comprehensive experimental approach for the validation of quantitative infrared thermography in the evaluation of building thermal transmittance, *Appl. Energy* 141 (2015) 218–228, <https://doi.org/10.1016/j.apenergy.2014.12.035>.
- [29] L. Hoegner, U. Stilla, Mobile thermal mapping for matching of infrared images with 3D building models and 3D point clouds, *Quantif. InfraRed. Thermogr. J.* 6733 (2018) 1–19, <https://doi.org/10.1080/17686733.2018.1455129>.
- [30] J.R. Martínez De Dios, A. Ollero, J. Ferruz, Infrared inspection of buildings using autonomous helicopters, in: *IFAC Proceedings Volumes (IFAC-PapersOnline)* 4, 2006, pp. 602–607, <https://doi.org/10.3182/20060912-3-DE-2911.00105>, no. PART 1.
- [31] P. Kim, J. Park, Y.K. Cho, J. Kang, UAV-assisted autonomous mobile robot navigation for as-is 3D data collection and registration in cluttered environments, *Autom. Constr.* 106 (July) (2019) 102918, <https://doi.org/10.1016/j.autcon.2019.102918>.
- [32] H. Freimuth, M. König, A framework for automated acquisition and processing of as-built data with autonomous unmanned aerial vehicles, *Sensors (Switzerland)* 19 (20) (2019) 1–22, <https://doi.org/10.3390/s19204513>.
- [33] H. Hamledari, B. McCabe, S. Davari, A. Shahi, E. Rezazadeh Azar, F. Flager, Evaluation of computer vision- and 4D BIM-based construction progress tracking on a UAV platform, in: *6th CSCE-CRC International Construction Specialty*

- Conference 2017 1, 2017, pp. 621–630, no. March, <https://purl.stanford.edu/wh873cw2351>.
- [34] D. Iwaszczuk, U. Stilla, Quality assessment of building textures extracted from oblique airborne thermal imagery, *ISPRS Annals of the Photogrammetry, Remote Sens. Spatial Inform. Sci.* 3 (1) (2016) 3–8, <https://doi.org/10.5194/isprs-annals-III-1-3-2016>.
- [35] P. Westfeld, D. Mader, H.G. Maas, Generation of TIR-attributed 3D point clouds from UAV-based thermal imagery, *Photogram. Fernerkund. Geoinform.* 2015 (5) (2015) 381–393, <https://doi.org/10.1127/1432-8364/2015/0274>.
- [36] J. Park, P. Kim, Y.K. Cho, J. Kang, Framework for automated registration of UAV and UGV point clouds using local features in images, *Autom. Constr.* 98 (November 2018) (2019) 175–182, <https://doi.org/10.1016/j.autcon.2018.11.024>.
- [37] T.J.B. Dewez, S. Yart, Y. Thuon, P. Pannet, E. Plat, Towards cavity-collapse hazard maps with Zeb-Revo handheld laser scanner point clouds, *Photogramm. Rec.* 32 (160) (2017) 354–376, <https://doi.org/10.1111/phor.12223>.
- [38] M. Rodriguez-Martin, P. Rodriguez-Gonzalez, D. Gonzalez-Aguilera, E. Nocerino, Novel approach for three-dimensional integral documentation of machine rooms in hospitals using portable Mobile mapping system, *IEEE Access* 6 (2018) 79200–79210, <https://doi.org/10.1109/ACCESS.2018.2884922>.
- [39] Y. Chang, D.J. Lee, Y. Hong, J. Archibald, Unsupervised video shot detection using clustering ensemble with a color global scale-invariant feature transform descriptor, *EURASIP J. Image Video Process.* 2008 (1) (2007), 860743, <https://doi.org/10.1155/2008/860743>.
- [40] A. Alahi, R. Ortiz, P. Vanderghelyst, FREAK: fast retina Keypoint, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, 2012, pp. 510–517, <https://doi.org/10.1109/CVPR.2012.6247715>.
- [41] D. Mistry, A. Banerjee, Comparison of feature detection and matching approaches: SIFT and SURF, *GRD J. Global Res. Dev. J. Eng.* 2 (4) (2017) 7–13, <https://www.grdjournal.com/uploads/article/GRDJE/V02/I04/0013/GRDJEV02I040013.pdf>.
- [42] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. Assoc. Comput. Machin.* 24 (6) (1981) 381–395, <https://doi.org/10.1145/358669.358692>.
- [43] J.L. Schonberger, J.-M. Frahm, Structure-From-Motion Revisited, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, <https://doi.org/10.1109/CVPR.2016.445>.
- [44] D. Forsyth, J. Ponce, *Computer Vision: A Modern Approach*, 2011. ISBN: 9780136085928.
- [45] L. Acampora, I. National, F. De Filippis, A. Martucci, L. Sorgi, 3D reconstruction of thermal images 3D reconstruction of thermal images, in: 26th Aerospace Testing Seminar, 2011, pp. 263–276, https://www.researchgate.net/publication/233407848_3D_Reconstruction_of_Thermal_Images.
- [46] F. Remondino, From point cloud to surface: the modeling and visualization problem, *Int. Archiv. Photogramm. Remote Sens. Spatial Inform. Sci.* 34 (2003), XXXIV/5/W10, <https://doi.org/10.3929/ethz-a-004655782>.
- [47] S.L. López, J.C. García, J.M. Sánchez, D.R. Bernárdez, H.L. Cimadevila, Thermographic mobile mapping of Urban environment for lighting and energy studies, *J. Daylight.* 1 (1) (2014) 8–15, <https://doi.org/10.15627/jd.2014.2>.
- [48] G. Clarkson, S. Luo, R. Fuentes, Thermal 3D modelling, in: 34th International Symposium in Automation and Robotics in Construction, 2017, <https://doi.org/10.22260/ISARC2017/0068>.
- [49] M. Chen, A. Feng, K. McCullough, B. Prasad, R. Mcalinden, L. Soibelman, Semantic Segmentation and Data Fusion of Microsoft Bing 3D Cities and Small UAV-based Photogrammetric Data 20220, 2020, pp. 1–12, doi: arXiv:2008.09648v1.
- [50] M. Chen, et al., 3D Photogrammetry Point Cloud Segmentation Using a Model Ensembling, *Framework* 34 (6) (2020) 1–20, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000929](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000929).
- [51] E. Maset, A. Fusiello, F. Crosilla, R. Toldo, D. Zorzetto, Photogrammetric 3D building reconstruction from thermal images, *ISPRS Ann. Photogram. Remote Sens. Spat. Inform. Sci.* 4 (2W3) (2017) 25–32, <https://doi.org/10.5194/isprs-annals-IV-2-W3-25-2017>.
- [52] N. Bayomi, S. Nagpal, T. Rakha, C. Reinhart, J.E. Fernandez, Aerial thermography as a tool to inform building envelope simulation models, *SimAud 2019* (2019) 45–48, <https://doi.org/10.5555/3390098.3390121>.
- [53] A.M. Kim, R.C. Olsen, F.A. Kruse, Methods for LiDAR point cloud classification using local neighborhood statistics 8731, 2013, pp. 1–10, <https://doi.org/10.1117/12.2015709>.
- [53] Y. Hou, R. Volk, L. Soibelman, A novel building temperature simulation approach driven by expanding semantic segmentation training datasets with synthetic aerial thermal images, *Energies* 14 (2) (2021), <https://doi.org/10.3390/en14020353>.
- [54] H. Xu, L. Cheng, M. Li, Y. Chen, L. Zhong, Using ocretes to detect changes to buildings and trees in the urban environment from airborne LiDAR data, *Remote Sens.* 7 (8) (2015) 9682–9704, <https://doi.org/10.3390/rs70809682>.
- [55] M. Campos-Taberner, et al., Processing of extremely high-resolution LiDAR and RGB data: outcome of the 2015 IEEE GRSS data fusion contest-part a: 2-D contest, *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 9 (12) (2016) 5547–5559, <https://doi.org/10.1109/JSTARS.2016.2569162>.
- [56] J. Castagno, E. Atkins, Roof shape classification from LiDAR and satellite image data fusion using supervised learning, *Sensors (Switzerland)* 18 (11) (2018), <https://doi.org/10.3390/s18113960>.
- [57] R. Iehag, D. Bulatov, P. Helmholz, D. Belton, Classification and representation of commonly used roofing material using multisensorial aerial data, *Int. Archiv. Photogram. Remote Sens. Spatial Inform. Sci. ISPRS Archiv.* 42 (1) (2018) 217–224, <https://doi.org/10.5194/isprs-archives-XLII-1-217-2018>.
- [58] M. Yamaguchi, T.P. Truong, S. Mori, Superimposing Thermal-Infrared Data on 3D Structure Reconstructed by RGB Visual Odometry 5, 2018, pp. 1296–1307, <https://doi.org/10.1587/transinf.2017MVP0023>.
- [59] D. Lin, M. Jarzabek-rychard, X. Tong, H. Maas, Fusion of thermal imagery with point clouds for building façade thermal attribute mapping, *ISPRS J. Photogramm. Remote Sens.* 151 (December 2018) (2019) 162–175, <https://doi.org/10.1016/j.isprsjprs.2019.03.010>.
- [60] M. Chen, E. Koc, Z. Shi, L. Soibelman, Proactive 2D model-based scan planning for existing buildings automation in construction, *Autom. Constr.* 93 (September) (2018) 165–177, <https://doi.org/10.1016/j.autcon.2018.05.010>.
- [61] S. Lagüela, J. Martínez, J. Armesto, P. Arias, Energy Efficiency Studies Through 3D Laser Scanning and Thermographic Technologies 43, 2011, pp. 1216–1221, <https://doi.org/10.1016/j.enbuild.2010.12.031>.
- [62] D. Borrmann, H. Afzal, J. Elseberg, N. Andreas, Mutual Calibration for 3D Thermal Mapping, 2012, pp. 3–8, <https://doi.org/10.3182/20120905-3-HR-2030.00073>.
- [63] L. Hoegner, U. Stilla, Thermal leakage detection on building facades using infrared textures generated by mobile mapping, in: 2009 Joint Urban Remote Sensing Event, 2009, <https://doi.org/10.1109/URS.2009.5137681>.
- [64] S. Lagüela, J. Armesto, P. Arias, J. Herráez, Automation of thermographic 3D modelling through image fusion and image matching techniques, *Autom. Constr.* 27 (2012) 24–31, <https://doi.org/10.1016/j.autcon.2012.05.011>.
- [65] P.A. Fokaides, S.A. Kalogirou, Application of infrared thermography for the determination of the overall heat transfer coefficient (U-value) in building envelopes, *Appl. Energy* 88 (12) (2011) 4358–4365, <https://doi.org/10.1016/j.apenergy.2011.05.014>.
- [66] E. Barreira, R.M.S.F. Almeida, J.M.P.Q. Delgado, Infrared thermography for assessing moisture related phenomena in building components, *Constr. Build. Mater.* 110 (2016) 251–269, <https://doi.org/10.1016/j.conbuildmat.2016.02.026>.
- [67] M. Vollmer, K.P. Mollmann, *Infrared Thermal Imaging: Fundamentals, Research and Applications*, 2010. ISBN: 9783527413515.
- [68] M. Weimann, J. Leitloff, L. Hoegner, B. Jutzi, U. Stilla, S. Hinz, Thermal 3D Mapping for Object Detection in Dynamic Scenes, vol. II, no. November, 2014, pp. 17–20, <https://doi.org/10.5194/isprsannals-II-1-53-2014>.
- [69] D. González-Aguilera, P. Rodríguez-González, J. Armesto, S. Lagüela, Novel approach to 3D thermography and energy efficiency evaluation, *Energy Build.* 54 (2012) 436–443, <https://doi.org/10.1016/j.enbuild.2012.07.023>.
- [70] D. Lin, Thermal Texture Selection and Correction for Building Facade Inspection Based on Thermal Radiant Characteristics, no. May, 2018, <https://doi.org/10.5194/isprs-archives-xlii-2-585-2018>.
- [71] S. Lagüela, L. Díaz-Vilarino, J. Martínez, J. Armesto, Automatic thermographic and RGB texture of as-built BIM for energy rehabilitation purposes, *Autom. Constr.* 31 (2013) 230–240, <https://doi.org/10.1016/j.autcon.2012.12.013>.
- [72] A. Wiens, W. Kleiber, K.R. Barnhart, D. Sain, Surface estimation for multiple misaligned point sets, *Math. Geosci.* (2019), <https://doi.org/10.1007/s11004-019-09802-y>.
- [73] L. Hoegner, S. Tutas, U. Stilla, 3D building reconstruction and construction site monitoring from RGB and TIR image sets, in: 2016 12th IEEE International Symposium on Electronics and Telecommunications (ISETC), 2016, 2016, pp. 305–308, <https://doi.org/10.1109/ISETC.2016.7781118>.
- [74] T.P. Truong, M. Yamaguchi, S. Mori, V. Nozick, H. Saito, Registration of RGB and thermal point clouds generated by structure from motion, in: 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017, 2020, pp. 419–427, <https://doi.org/10.1109/ICCVW.2017.57>.
- [75] B.G. Rayson, N.T.P. Enna, J.P.M. Ills, GPS Precise Point Positioning for UAV Photogrammetry, vol. 33, no. December, 2018, pp. 427–447, <https://doi.org/10.1111/phor.12259>.
- [76] M.M.R. Mostafa, K. Schwarz, Digital image georeferencing from a multiple camera system by GPS/INS, *ISPRS J. Photogramm. Remote Sens.* 56 (1) (2001) 1–12, [https://doi.org/10.1016/S0924-2716\(01\)00030-2](https://doi.org/10.1016/S0924-2716(01)00030-2).
- [77] F. Sheeba, R. Thamburaj, J.J. Mammen, M. Kumar, V. Rangslang, Convex hull based detection of overlapping red blood cells in peripheral blood smear images, *IFMBE Proc.* 52 (2015) 51–53, https://doi.org/10.1007/978-3-319-19452-3_14.
- [78] R. Bai, J. Zhao, D. Li, W. Meng, A new technique of camera calibration based on X-target, in: *Communications in Computer and Information Science* 97, 2010, pp. 39–46, https://doi.org/10.1007/978-3-642-15853-7_6. CCIS, no. PART 1.
- [79] Bentley Systems, ContextCapture - 3D Reality Modeling Software - Bentley Official Website. <https://www.bentley.com/en/products/brands/contextcapture> (accessed Jun. 13, 2020).
- [80] F. Nielsen, On the Jensen-Shannon symmetrization of distances relying on abstract means, *Entropy* 21 (5) (2019) 1–23, <https://doi.org/10.3390/e21050485>.
- [81] F. Pérez-Cruz, Kullback-leibler divergence estimation of continuous distributions, in: *IEEE International Symposium on Information Theory - Proceedings*, 2008, pp. 1666–1670, <https://doi.org/10.1109/ISIT.2008.4595271>.
- [82] D. Jurafsky, J.H. Martin, *Speech and Language Processing*, 2008. ISBN: 9780131873216.
- [83] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, 2010. ISBN: 9780136042594.
- [84] Y. Rubner, C. Tomasi, L.J. Guibas, The earth Mover’s distance as a metric for image retrieval, *Int. J. Comput. Vis.* 40 (2000) 99–121, <https://doi.org/10.1023/A:1026543900054>.
- [85] A. Irani Rahaghi, U. Lemmin, D. Sage, D.A. Barry, Achieving high-resolution thermal imagery in low-contrast lake surface waters by aerial remote sensing and image registration, *Remote Sens. Environ.* 221 (March 2018) (2019) 773–783, <https://doi.org/10.1016/j.rse.2018.12.018>.

- [86] FLIR, FLIR Technical Note, 2010, p. 2010. <https://www.flir.com/discover/suas/uas-radiometric-temperature-measurements/> (accessed Jun. 13, 2020).
- [87] T. Rakha, A. Gorodetsky, Review of unmanned aerial system (UAS) applications in the built environment: towards automated building inspection procedures using drones, *Autom. Constr.* 93 (January) (2018) 252–264, <https://doi.org/10.1016/j.autcon.2018.05.002>.
- [88] L. Hoegner, S. Tuttas, Y. Xu, K. Eder, U. Stilla, Evaluation of methods for coregistration and fusion of RPAS-based 3D point clouds and thermal infrared images, *Int. Archiv. Photogram. Remote Sens. Spatial Inform. Sci. ISPRS Archiv.* 41 (July) (2016) 241–246, <https://doi.org/10.5194/isprsarchives-XLI-B3-241-2016>.

Repository KITopen

Dies ist ein Postprint/begutachtetes Manuskript.

Empfohlene Zitierung:

Hou, Y.; Volk, R.; Chen, M.; Soibelman, L.

[Fusing tie points' RGB and thermal information for mapping large areas based on aerial images: A study of fusion performance under different flight configurations and experimental conditions.](#)

2021. Automation in construction, 124.

doi: [10.5445/IR/1000129014](https://doi.org/10.5445/IR/1000129014)

Zitierung der Originalveröffentlichung:

Hou, Y.; Volk, R.; Chen, M.; Soibelman, L.

[Fusing tie points' RGB and thermal information for mapping large areas based on aerial images: A study of fusion performance under different flight configurations and experimental conditions.](#)

2021. Automation in construction, 124, Art.-Nr.: 103554.

doi: [10.1016/j.autcon.2021.103554](https://doi.org/10.1016/j.autcon.2021.103554)