PFG 2015 / 5, 0381 – 0393
Stuttgart, October 2015

Report

Generation of TIR-attributed 3D Point Clouds from UAV-based Thermal Imagery

PATRICK WESTFELD, DAVID MADER & HANS-GERD MAAS, Dresden

Keywords: Unmanned aerial vehicle, thermal infrared imaging, calibration, structure from motion, 3D point cloud

Summary: This contribution investigates the possibility to efficientl reconstruct 3D scene geometry purely from thermal image data by using structure from motion techniques. The images are acquired by a low-cost, lightweight thermal camera mounted on an octocopter platform. The result is a 3D point cloud containing geometric information plus thermal attributes. Such a dataset may be valuable for building energy evaluation, where existing geospatial data may thus be enriched with thermal information.
The paper gives an overview of the automatic data processing chain and also considers aspects of geometric thermal camera calibration. For an extensive practical test, a raster pattern was fl wn, capturing 15,000 thermal images of the façades of a large building complex. To validate the geometric quality of the results, the point cloud was aligned to a terrestrial laser scanning 3D model of the scene. With a precision in the order of 25 mm, related to an object dimension of 50 m × 50 m × 20 m, the analysis of the model differences substantiated the geometric potential of thermal cameras in SfM tools.

Zusammenfassung: *Generierung von thermalen 3D-Punktwolken aus UAV-gestützten Wärmebildkameradaten*. Diese Arbeit untersucht das Potential von einzig auf Wärmebilddaten basierenden Structure-from-Motion Ansätzen zur effiziente 3D-Objektrekonstruktion. Getragen von einer Oktokopter-Flugroboterplattform wurden die Bilddaten mit einem preiswerten, kompakten thermischen Sensor aufgenommen. Zusätzlich zur Raumkoordinate enthält die resultierende 3D-Punktwolke eine thermische Information für jeden Objektpunkt. Vor allem im Bereich der energetischen Bewertung von Gebäuden kann diese Kombination von Geometrie und Thermografi von Interesse sein. Im Beitrag behandelt werden neben der Prozesskette zur automatischen Datenverarbeitung auch Aspekte der geometrischen Kalibrierung einer Wärmebildkamera. Im Rahmen eines umfangreichen praktischen Experiments wurden die Fassaden eines weitläufige Gebäudekomplexes rasterförmig abgefloge und ca. 15.000 Wärmbilder aufgenommen. Die Registrierung der erzeugten 3D-Punktwolke zu einem terrestrischen Laserscannermodell ermöglichte die Validierung der geometrischen Qualität der Ergebnisse. Mit 3D-Objektkoordinatenabweichungen in der Größenordnung von 25 mm, bezogen auf eine Objektgröße von 50 m × 50 m × 20 m, konnte das geometrische Potential in der Auswertung von Wärmebildkameradaten mit SfM-Werkzeugen belegt werden.

1 Introduction

Efficien energy handling is playing an increasingly important role in economy. Energy efficien y is clean, particularly favourable and immediately available. Taking diminishing resources and increasing prices for raw materials into consideration, energy efficien y forms one pillar of the energy revolution (KOHLER et al. 2013). European legislation requires an energy efficien y increase of 20 % by 2020 (2012/27/EU). In order to realize the directive, the German federal government offers incentives for individual energetic redevelopment of residential property, such as building envelope insulation or window replacements.

www.schweizerbart.de
DOI: 10.1127/1432-8364/2015/0274
1432-8364/15 /0274 $ 3.50

Heat loss from buildings can easily be visualized by thermograms. The common way in practice is that energy consultants acquire thermal infrared (TIR) images from individual buildings using hand-held thermal cameras. The TIR data are interpreted visually during on-site inspections, leading to recommendations for restructuring measures. This procedure is rather time-consuming and cost-intensive. Furthermore, a sharing and pooling of knowledge and data is impeded. However, it would be desirable – especially for the survey and assessment of larger areas like streets of houses, districts or even complete cities – to enrich spatial information stored in a geographic information system (GIS) with thermal attributes.

Several studies deal with the combination of thermal measurements and geometric data. The contribution of HOEGNER et al. (2007) focuses on the extraction and integration of façade textures from low resolution infrared image sequences. LAGÜELA et al. (2011b) and GONZÁLEZ-AGUILERA et al. (2012) describe approaches to multi-sensor registration of TIR images and terrestrial laser scanner (TLS) data. In these approaches, common features are extracted in both datasets to combine thermographic and geometric data in a joint thermographic 3D model. In BANNEHR et al. (2013), data from an airborne laser scanner, a hyperspectral and a thermal camera were fused in order to extract urban parameters like roof materials or geometric information. HOEGNER et al. (2013) present a method for the co-registration of time-of-fligh (TOF) camera generated 3D point clouds and TIR images. In HOEGNER & STILLA (2015), the relative orientation of terrestrial TIR image sequences from different views is estimated and in the fina analysis used to detect 3D building façade objects. In RUDOL & DOHERTY (2008), a technique is introduced which allows the detection and geolocalization of human bodies in thermal and colour imagery, acquired during several flight with an unmanned aerial vehicle (UAV). In PECH et al. (2013), urban heat islands are detected, monitored and analyzed. PECH et al. (2013) use RGB and thermal images, simultaneously captured by two cameras mounted on an UAV, to generate multi-temporal thermal orthophotos. IWASZCZUK et al. (2012) use aerial TIR image sequences for the detection of weak spots in building insulation by findin the best fi between a given 3D building model and the acquired TIR images. MAURIELLO & FROEHLICH (2014) give initial thoughts on how to automate the 3D reconstruction process of building façades from UAV-based collected thermal images.

The data for this study was captured by a thermal imaging camera onboard a UAV. This platform is part of a flee of unmanned aerial vehicles which are equipped with different sensors like laser scanner, RGB camera, near infrared camera and thermal camera. The fly ing robots have been designed and built in the course of the project ADFEX (Adaptive Federative 3D Exploration of Multi Robot Systems). Information on the entire project can be found on http://www.adfex.eu/.

The paper investigates the possibility to efficientl reconstruct 3D scene geometry purely from thermal image data by using structure from motion techniques. The paper gives an overview of the automatic data processing chain and also considers aspects of geometric thermal camera calibration. The results of the work presented here are TIR-attributed dense 3D point clouds of building façades. The paper is also focusing on the accuracy potential of thermal cameras in 3D point cloud generation through SfM.

The UAV sensor system is presented in section 6.2. The thermal sensor has to be calibrated firstl for accurate temperature measurements and secondly to correct for geometric errors affecting the optical imaging process (section 3). The data acquisition process to capture the façades of the courtyard of a building complex is described in section 4. Similar to the work of MAURIELLO & FROEHLICH (2014), this publication investigates the possibility to efficientl reconstruct camera poses and 3D scene geometry purely from overlapping thermal image sequences by using structure from motion (SfM) techniques. The necessary processing steps are treated in section 5. A critical analysis of the accuracies in 3D point cloud and single 3D points is given in 6. Finally, section 7 summarizes the work.

Fig. 1: (a) CAD model of an octocopter UAV sensor platform equipped with a thermal imaging camera FLIR A65. (b) Colour-coded thermal image of 640 pixel 512 pixel.

2 Sensor Systems

Any object surface having a temperature above absolute zero emits, reflect or transmits thermal radiation. The energy emitted is a measure for surface temperature and may be detected by suitable sensors such as cooled photon detectors and uncooled thermal detectors, DERENIAK & BOREMAN (1996).

In this study, we used a low-cost thermal imaging camera FLIR A65 which is based on an uncooled microbolometer array. The thermal sensor is sensitive in the infrared range between 7.5 µm and 13 µm. It produces thermal images of 640 pixel × 512 pixel with 17 µm pixel size (Fig. 1b) at 30 fps with a thermal sensitivity of less than 50 mK in a range of −40 °C to +160 °C. A fi ed 13 mm wide-angle lens (45° × 37°) provides a suitable fiel of view for the application at hand. The compact and lightweight camera body (106 mm×40 mm×43 mm; 200 g) enables an integration on a payload-restricted UAV.

The FLIR A65 is mounted on a multi-rotor flyin platform Cadmic Goliath Coax 8 (Fig. 1a). This 4-arm copter with eight rotors is part of an UAV fleet designed and used for 3D exploration tasks in urban environments. More detailed specification on the UAV platform are given in KLIX et al. (2014).

3 Thermal Camera Calibration

In order to achieve accurate and reliable measurements, a radiometric and geometric calibration of the thermal imaging sensor is necessary. Therefore, the model parameters for the temperature calculation as well as the imaging process have to be determined.

3.1 *Radiometric Calibration*

The formula to convert a 14 bit object signal S, acquired during frame grabbing, to a temperature value T (in K) is (FLIR 2012):

$$T = \frac{B}{\log(R \cdot (S - O)^{-1} + F)} \qquad (1)$$

The conversion parameters R, B, F and O are determined by pointing the camera at objects with known temperature. The black body infrared calibration source PYROTHERM CS 110 is used as a reference. During calibration, thermal image data are acquired in a range of −10 °C to +50 °C in steps of 10 K, and the four unknown radiometric calibration parameters are estimated by regression analysis. The a-posteriori standard deviation for measuring the absolute temperature of a black body within the specifie range of temperature can be stated with 70 mK (FLIR 2012). However, this level of accuracy can usually not be achieved in fiel applications due to differing emissivities and reflections

3.2 *Geometric Calibration*

The geometric camera model used here is slightly different from the one commonly used in photogrammetry (LUHMANN et al. 2006). It refers to a concept taken from computer vision (HARTLEY & ZISSERMAN 2004), to which the SfM tools used in section 5 are related.

Deviations between the ideal geometric projection model and the actual camera geometry have to be modeled to exhaust the accuracy potential of the image data. The imaging geometry of a thermal camera is similar to the pinhole camera model of a conventional camera. The transformation from an object point $\mathbf{X}(X, Y, Z)$ to a local camera coordinate $\mathbf{x}(x, y)$ is:

$$x = \frac{\begin{array}{l} r_{11} \cdot (X - X_0) \\ +r_{21} \cdot (Y - Y_0) + r_{31} \cdot (Z - Z_0) \end{array}}{\begin{array}{l} r_{13} \cdot (X - X_0) \\ +r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0) \end{array}}$$

$$y = \frac{\begin{array}{l} r_{12} \cdot (X - X_0) \\ +r_{22} \cdot (Y - Y_0) + r_{32} \cdot (Z - Z_0) \end{array}}{\begin{array}{l} r_{13} \cdot (X - X_0) \\ +r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0) \end{array}} \quad (2)$$

where

$\mathbf{X_0}$	Projection centre
$r_{r,c}$	Elements of a rotation matrix $\mathbf{R}$

The coordinates projected into camera's coordinate system (origin at the camera projection centre; z-axis points toward the viewing direction; x-axis points to the right; y-axis points down) are further corrected for imaging errors caused by radial and tangential distortion effects. Lens distortion is modeled by a frequently used polysonomial model (BROWN 1971):

$$\begin{aligned} x' &= x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \\ &\quad + p_2(r^2 + 2x^2) + 2p_1 xy) \\ y' &= y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \\ &\quad + p_1(r^2 + 2y^2) + 2p_2 xy) \end{aligned} \quad (3)$$

where

$\mathbf{x}'$	Image coordinate free of distortions
$k_{[1;3]}$	Radial distortion coefficient
$p_{1,2}$	Decentering distortion coefficient
$r = \sqrt{x^2 + y^2}$	Radial distance

Finally, the projected point coordinates (u, v) in the image coordinate system (origin at top left image pixel; u-axis points to the right; v-axis points down) are calculated by shifting the coordinate system from the principal point to the origin of the image coordinate system, scaling in both image coordinate directions by focal length and correcting for skewness of the image axes:

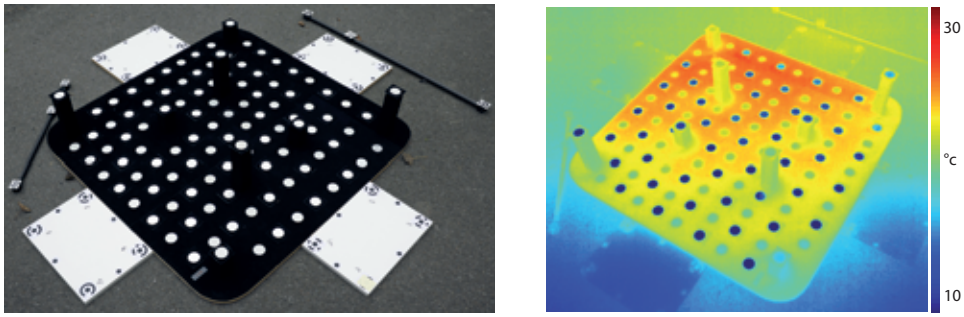
$$\begin{aligned} u &= u_0 + x' c_u + y' s_1 \\ v &= v_0 + y' c_v \end{aligned} \quad (4)$$

where

$c_{u,v}$	Focal length in u and v image coordinate direction
u_0, v_0	Principal point
s_1	Coefficien describing the skewness of the two image axes

To reduce efforts in time and instrumental resources, photogrammetric self-calibration is favoured. In our application, geometric stability was assured in the self-calibrating bundle adjustment approach by jointly adjusting the low-resolution thermal image coordinate measurements with observations made in much higher resolution images captured by a DSLR Nikon D700. This requires the design of an adequate reference fiel with targets sufficientl visible in both the RGB and the thermal imagery.

Several strategies are reported in literature to deal with the problem of making reference points visible in thermal images: BISON et al. (2012) drilled holes in an aluminum surface. YASTIKLI & GULER (2013) designed a complex iron 3D test object, equipped with plastic targets. In LAGÜELA et al. (2011a), small lamps were arranged in a grid on a wooden plank. LUHMANN et al. (2013) applied this approach and utilized target lamps, too. They



(a) (b)

Fig. 2: (a) True colour image and (b) colour-coded thermal image of a 3D target field.

Tab. 1: Thermal camera calibration parameters $\hat{x}_i$ with their standard deviations $\hat{s}_{\hat{x}_i}$. The image coordinate system is defined in section 3. The camera's pixel size is 17 µm, the camera's sensor size is 640 pixel × 512 pixel.

	c_u (px)	c_v (px)	u_0 (px)	v_0 (px)	k_1	k_2	k_3	p_1	p_2	s_1(px)
$\hat{x}_i$	774.65	774.46	315.75	259.57	-4.51e^{-2}	0.36	0	7.82e^{-4}	-1.13e^{-3}	6.37e^{-2}
$\hat{s}_{\hat{x}_i}$	0.197	0.202	0.267	0.254	1.36e^{-3}	9.80e^{-3}	fi	1.04e^{-4}	1.28e^{-4}	3.52e^{-2}

also reported about good experience with a metal reference fiel plate which reflect the cold background radiation well. The actual targets were made of self-adhesive foil which only emits radiation relating to its own temperature.

The 3D calibration plate designed for this work is also based on two material components: Silver heat protection foil with excellent reflectin properties for circular measuring marks is combined with black velour foil with good absorbing characteristics for target edging. In combination, both components provide a strong contrast in visible and infrared spectrum (Fig. 2). The centre coordinates of the circular targets (⌀40 mm) can be measured by an ellipse fit The mean standard deviation of the image point coordinate measurements was 1/40 of a pixel. These observations as well as additional scale information (reference bars) and approximate values for the unknown interior orientation parameters are introduced into a bundle adjustment procedure. Parameters were excluded from the bundle adjustment if they turned out to be insignifican in the significanc test. Outliers in the results were removed in an outlier detection procedure following 3-sigma rule.

Tab. 1 lists the interior orientation parameters and the corresponding a-posteriori standard deviations. All parameters could be determined significantl , except for the radial distortion parameter k_3. Regarding the precision of image point measurement stated above as well as the precision of the estimated geometric calibration parameters show that the calibration plate designed in combination with standard convergent imaging configuration is particularly suitable for thermographic camera calibration. The interior orientation thus estimated is used for 3D scene triangulation and dense 3D reconstruction (section 5).

4 Data Acquisition

The courtyard of an old urban building complex served as the test object in our study (Fig. 3a). The UAV introduced in section 6.2 was used to fl a raster pattern to capture thermal images of the entire surrounding façades (Fig. 3b). The façades are approximately 20 m in height and up to 50 m in length. Overall, more than 15,000 thermal images were acquired at 3 Hz with 90 % forward and 70 % side lap (Fig. 3c).

(a) (b) (c)

Fig. 3: (a) Aerial image of the subject of investigation. (b) Flying robot collecting data in front of a building façade. (c) Some of the thermal images.

The coordinates of 20 signalised circular targets, evenly distributed across one façade, were measured with a geodetic total station theodolite (TST) in order to provide a ground truth for evaluating the results (section 6). The design of those reference points is similar to the design shown in section 3, but the targets are larger, having a diameter of 140 mm. Following the law of error propagation for TST measurements, the RMSE of the resulting 3D coordinates can be estimated to be 4 mm. If TST coordinate differences are used as additional scaling distance information, the relative accuracy can be stated to be 2 mm. The corresponding image point coordinates of the centres are measured in thermal images by ellipse fi with a mean standard deviation of 1/30 of a pixel.

A terrestrial laser scanning point cloud acquired with a Riegl LMS-Z420i (3 million points, ca. 800 pts/m^2) is also used as a reference to analyze the quality of the point cloud obtained from the thermal images (section 6). The RMSE of a 3D point in TLS data depends on several influencin factors such as incidence angle, surface material or inaccuracies in registration and is given with about 10 mm in the dataset at hand (SCHNEIDER & MAAS 2007).

5 Data Processing and Results

The main goal of the work presented here is to reconstruct camera poses and 3D scene geometry purely from overlapping thermal imagery by using structure from motion (SfM) processing methods.

SfM techniques (e. g. FURUKAWA & PONCE 2007, MAYER 2008) generate 3D representations from 2D image sequences without initial information. Feature points are extracted from the images and matched, employing robust estimation techniques such as RANSAC (random sample consensus; FISCHLER & BOLLES 1981). This set of homologous points is then used to orient the images in a sequential bundle adjustment. Once interior and exterior orientations are calculated, dense image matching method such as semi global matching (SGM; HIRSCHMÜLLER 2005) can be employed.

The 14 bit signal of the thermal imaging sensor encodes temperatures in a range of –40 °C to +160 °C (sections 6.2 and 3). The environmental temperature during the measurement phase was +10 °C to +30 °C. In order to improve contrast for a reliable feature matching, the range of each image was expanded in a preprocessing step by histogram normalization. At the same time, the image depth was reduced to 8 bit in order to comply the image input formats required by the SfM tools used.

Image orientation was determined by VisualSFM, an integrated, open source SfM software package (WU 2011). Firstly, the SIFT operator detects and describes local feature points in the images (scale-invariant feature transform; LOWE 1999, WU 2007). The SIFT keypoints are, secondly, matched to identify homologous points between two images in a pair-wise image matching procedure. These correspondences are the basis for a bundle adjustment with robust outlier detection. It iteratively builds a sparse 3D model of the scene and simultane-

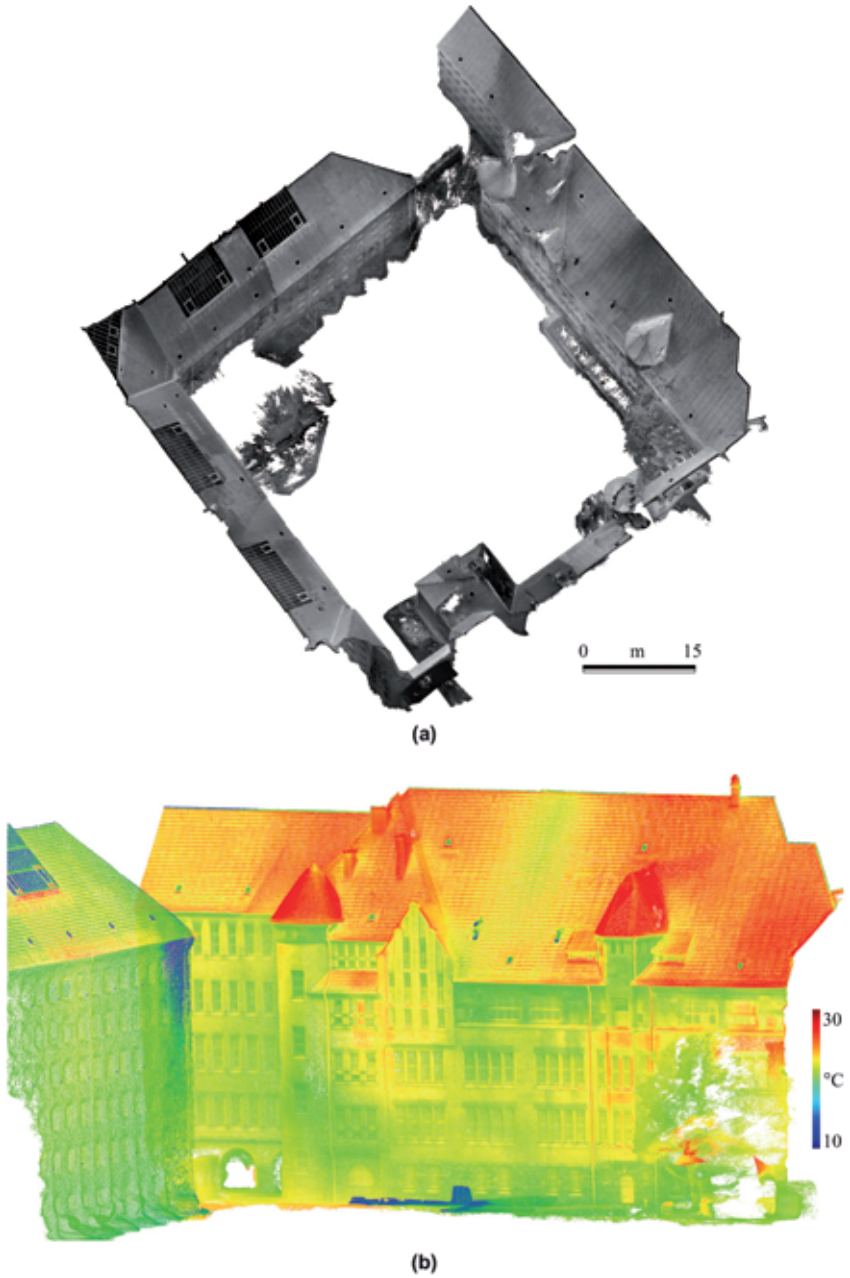


Fig. 4: 3D point cloud reconstructed from thermal images only: (a) Top view of the entire 3D model. (b) Temperature coded submodel of one façade.

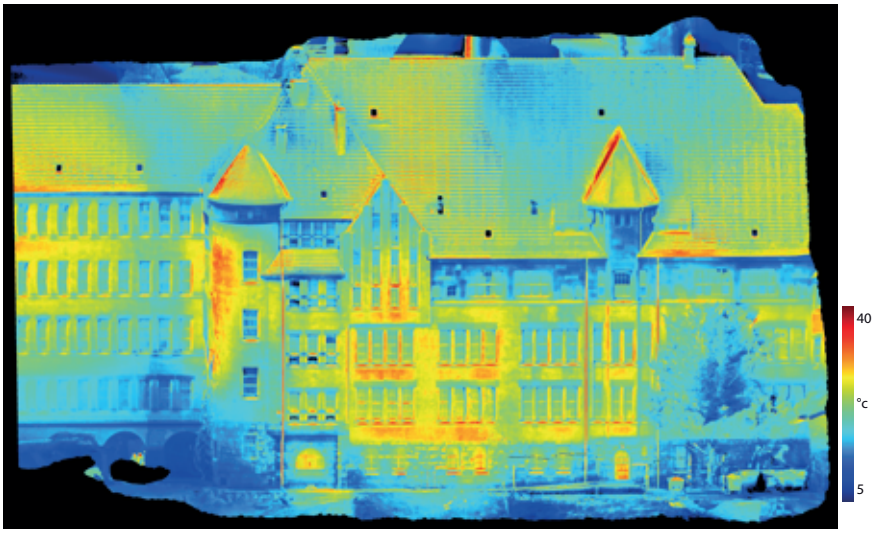


Fig. 5: TIR-attributed ortho-image of a part of one façade.

ously estimates the exterior orientations of all images; the interior orientation were kept constant, using the parameters derived as described in section 3. VisualSFM allows to produce multiple 3D models to reconstruct the entire scene, which is a particularly fl xible functionality for complex objects like the courtyard observed in this study. The internal camera model of VisualSFM includes only one radial distortion parameter. In order to consider the fully parametrized camera calibration model as described in section 3, the image coordinates of all homologous points as well as the camera orientations were imported into the commercial software package Agisoft PhotoScan. The camera calibration was fi ed to the parameters listed in Tab. 1, and another bundle adjustment was performed. In the next processing step, dense point clouds were derived from the given set of images and their orientation parameters by the dense image matching technique implemented in Agisoft PhotoScan.

The histogram stretching applied above must finall be reversed in order to provide thermal attributes. Each 3D point is re-projected into all corresponding image spaces where its temperature value is interpolated in a resampling process, weighted in accordance to the distance to the principal point.

The resulting 3D models can be scaled and geo-referenced either by using the tachymetric observations of the signalised control points or by aligning them to the previously acquired TLS dataset (section 4). In this study, an iterative closest point algorithm (ICP) minimizes the distance from the thermal point cloud calculated by structure from motion (as source) to the TLS point cloud (as reference) by iteratively applying a 3D Helmert transformation. The fina result consists of about 12 million TIR-attributed 3D points and is shown in Fig. 4.

Finally, ortho-rectifie façade images can be generated (Fig. 5). Orthographic projections of the building façades enable thermographic professionals to measure true distances directly in the image.

6 Accuracy Analyses

As the focus of the work presented here is on the accuracy potential of thermal cameras in 3D point cloud generation through SfM, several tests were performed to assess the accuracy. These tests include internal precision measures for image point measurements, camera poses and 3D object points (section 6.1) as well as external accuracy measures to validate the geo-

metric quality of the data using TST and TLS data as reference (sections 6.2 and 6.3).

The coordinate system of each submodel is oriented in a way that the façade forms the YZ-plane and the X-axis is oriented in camera's direction of view.

6.1 *SfM Internal Precision*

The SfM software tools used in this study (VisualSFM and Agisoft PhotoScan) deliver only very sparse information on the internal precision. Therefore, the SfM project data obtained after sparse reconstruction were fed into our own bundle adjustment software for a thorough accuracy assessment. The bundle adjustment delivers information on the precisions of the measured observations as well as the unknown camera poses and 3D object coordinates. In order to evaluate the pure results obtained from SfM tools, no further blunder checks were implemented. In the following, those results are presented.

The 2D coordinates of valid SIFT keypoints were introduced as a first group of observations. A second group contained the image point measurements of the signalised reference points. A variance component estimation (VCE; KOCH 2004) was used to ensure the adequate weighting of the two types of observations. The distances between the tachymetric observations of the signalised control points served as reference to scale the scene. The resulting additional constraint equations were weighted by a static variance component in accordance with the relative accuracy specification of a TST measurement given in section 4. The interior camera geometry was fixed to the parameters estimated in section 3. The exterior camera orientation parameters as well as the 3D coordinates of the object points calculated during SfM sparse reconstruction were treated as unknowns.

Tab. 2: A-posteriori standard deviations $\hat{s}_{\hat{x}_i}$ and mean point error $\hat{s}_{\hat{X}\hat{Y}\hat{Z}}$ of a 3D object point coordinates from a sparse point cloud in (mm). Row 1 and 3 are the mean, row 2 and 4 the median.

	$\hat{s}_{\hat{X}}$	$\hat{s}_{\hat{Y}}$	$\hat{s}_{\hat{Z}}$	$\hat{s}_{\hat{X}\hat{Y}\hat{Z}}$
SIFT	204.75	53.84	45.64	216.58
	160.40	34.19	32.22	167.14
Ref.	22.87	7.30	8.77	25.56
	23.12	6.19	8.82	25.50

The a-posteriori standard deviation $\hat{s}_{xy}$ of the original SIFT keypoints automatically adjusted by VCE was 1.1 pixel. This lack of subpixel accuracy corresponds to investigations on SIFT descriptors given in literature: REMONDINO (2006) states that with region detectors like SIFT the number of matched correspondences is quite high but the accuracy is poor, probably caused by perspective effects. The a-posteriori standard deviation $\hat{s}_{xy}$ of the signalised targets used as reference points is 1/3 of a pixel, and is thus significantly worse than the precision of the best-fit ellipses stated above. Reasons for that could be a too optimistic internal precision measure of ellipse-fitting. Furthermore, the size of the individual VCE observation groups is not well-balanced, to the disadvantage of the reference measurements (1:60). Self-calibration did not lead to any improvements in the estimated precision.

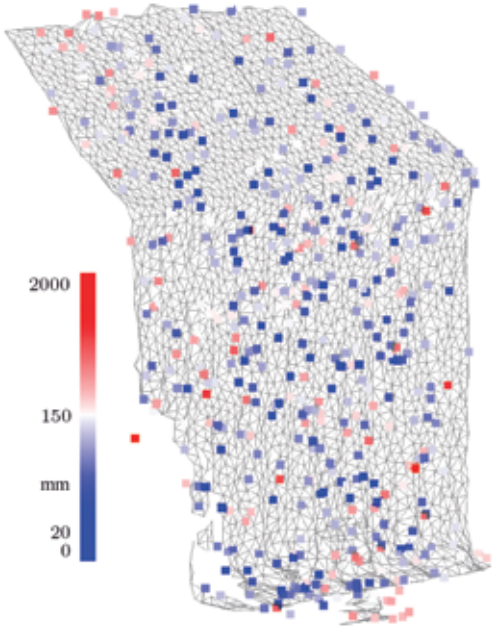


Fig. 6: 3D mesh of a part of the TLS reference façade, overlaid with colour-coded, logarithmically scaled mean point error of all 3D object points from sparse point cloud ranging from 20 mm to 1.8 m

Tab. 3: Mean a-posteriori standard deviations $\hat{s}_{\hat{x}_i}$ of camera's projection centres in (mm) and rotation angles in (°).

$\hat{s}_{\hat{X}_0}$	$\hat{s}_{\hat{Y}_0}$	$\hat{s}_{\hat{Z}_0}$	$\hat{s}_{\hat{X}_0\hat{Y}_0\hat{Z}_0}$
46.25	132.04	139.12	197.30

$\hat{s}_{\hat{\omega}}$	$\hat{s}_{\hat{\varphi}}$	$\hat{s}_{\hat{\kappa}}$
1.495	0.151	1.501

Tab. 4: Single 3D point precision in mm.

	$RMSE_X$	$RMSE_Y$	$RMSE_Z$	$RMSE_{XYZ}$
Intersection	18.28	4.45	4.92	19.45
Point picking	17.13	15.59	17.97	29.32

On an average, the a-posteriori standard deviations of the unknown 3D object coordinates $\mathbf{X}$ is 50 mm in planimetry and 200 mm in depth for all object points originating from SIFT keypoint observations (Tab. 2, row 1). The average a-posteriori standard deviation of the 3D object coordinates of the reference targets is 25 mm (Tab. 2, row 3). The accuracy for 3D SIFT points looks very high at a firs glance; however, it should be noted that this represents an average for all 3D points, including problematic regions such as edges or vegetation as well as points affected by a poor intersection geometry. As one can also see from Fig. 6, the depth precision of most of the façade points is between 20 mm to 100 mm. A more detailed evaluation of the single 3D point accuracy potential achieved after dense reconstructionis given in sections 6.2 and 6.3.

The a-posteriori standard deviations $\hat{s}_i$ of the unknown exterior orientation parameters are in average 200 mm for camera positions $\mathbf{X_0}$ and 1.0° for camera rotations $\omega\varphi\kappa$ (Tab. 3). Due to correlations between several camera orientation parameters, the internal precision figure of the projection centres are worse than the 3D point precision achieved above, an effect which is also known from aerial photogrammetry. The lower precision in roll (ω) and yaw (κ) can be explained by a partly fragmentary cross overlap in image configuratio for this part of the court yard.

6.2 *External 3D Point Accuracy*

For an external accuracy test, the 3D coordinates of the 20 reference target points were calculated by spatial intersection, using the measured image coordinates and the camera orientation parameters as determined in section 5. These 3D object coordinates were compared with the TST reference coordinates. The residual discrepancies after a 3D Helmert transformation between the two datasets are listed in Tab. 4, row 1. The RMSE of a 3D object coordinate $\mathbf{X}$ is about 20 mm.

The 3D coordinates of the reference points can also be measured directly in the dense point cloud. Due to sampling errors, they actually do not match exactly with the true centres of the circular reference targets. This is reflecte in a too pessimistic RMSE in lateral direction, but the RMSE in depth direction (here X) is comparable to its equivalent determined by spatial intersection for points on a planar façade (Tab. 4, row 2).

The façades were captured with a ground sampling distance (GSD) of 26 mm for an average measurement distance of 20 m. Thus, the RMSE in depth direction is 2/3 of the GSD, which is a very promising ratio for 3D model reconstruction from low resolution thermal images and well comparable to the accuracy obtained for digital surface models automatically created from aerial images (HIRSCHMÜLLER & BUCHER 2010).

6.3 *Comparison of SfM and TLS Point Clouds*

The deviations between the SfM dense point cloud and the TLS reference point cloud are shown in Fig. 7, using a colour code for visualisation. For each point of the SfM point cloud,

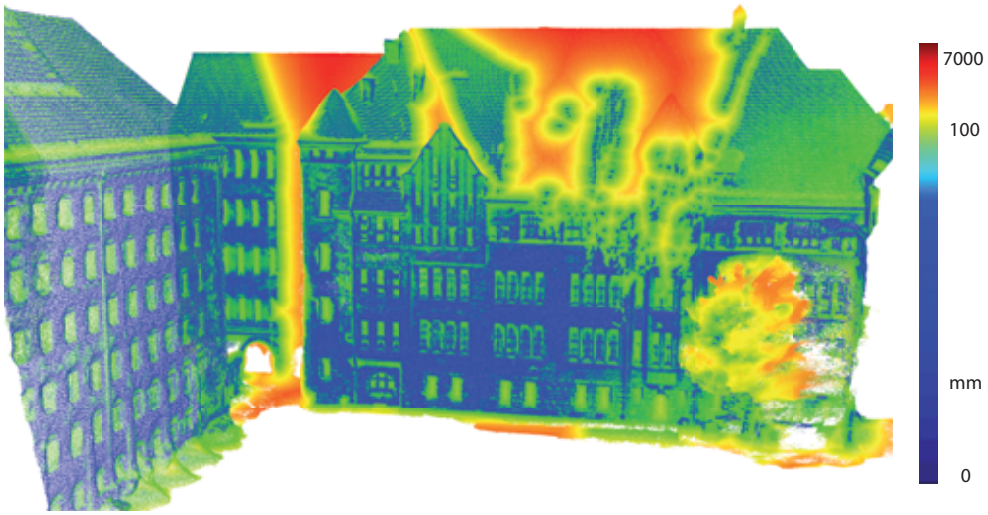


Fig. 7: Colour-coded, logarithmically scaled differences between SfM and TLS point clouds.

the Euclidean distance to the nearest point in the TLS point cloud was computed (nearest neighbour distance). Ignoring TLS blind spots caused by occlusions in the TLS dataset as well as vegetation (yellow, orange and red areas), the deviations are less than 25 mm in average for façade structures (roof areas as well as window areas reflectin TIR-radiation were excluded), obtained from a total of 12 million 3D points. The accuracy value contains the actual measurement error of SfM coordinates as well as the representation error from the sampling process.

6.4 *Discussion*

These external precision figure obtained from a comparison of the 3D point cloud as a result from dense matching with data acquired by terrestrial laser scanning constitute a rather rigorous accuracy check – and they are surprisingly good, considering the fact that it is purely based on data of a thermal camera with small sensor format as well as specifi contrast and texture conditions. The results obtained from signalised reference points (section 6.2) confir these figures Surprisingly, the internal precision figure obtained from processing the "black box" SfM project data with our own bundle adjustment are much worse. This can be explained by the fact that the SfM project data were not filtere and that the dataset also contained points in occluded areas and on vegetation, which were excluded in the analysis in section 6.3. Moreover, it should be noted that the matching tools used to produce the sparse data in section 6.1 (SIFT) and the dense representation tested in section 6.2 (dense image matching) were different, with the SIFT operator delivering a high robustness, but a lower image measurement precision.

7 Conclusion

The work presented here deals with reconstructing TIR-attributed 3D building façades purely from UAV-based thermal images. Structure from motion techniques are used to estimate camera orientations and 3D object point coordinates without any initial information. The results are very promising: The façades of a courtyard of a building complex could completely be represented by a dense 3D point cloud consisting of about 12 million points. The RMSE of the TIR-attributed 3D points was found to be less than 20 mm in a comparison with terrestrial laser scanner data, the depth accuracy is about 2/3 of the ground sampling distance.

Future work should concentrate on the radiometric resampling process. The data quality may be improved by considering the 3D object geometry in the resampling and correction of thermal attributes. The radiometric accuracy potential should also be investigated.

Acknowledgements

The research work presented in this paper has been funded by the European Social Fund (ESF) via Sächsische Aufbaubank (SAB). We would also like to thank our partners in the ADFEX project for their support and the great collaboration.

References

BANNEHR, L., SCHMIDT, A., PIECHEL, J. & LUHMANN, T., 2013: Extracting Urban Parameters of the City of Oldenburg from Hyperspectral, Thermal, and Airborne Laser Scanning Data. – PFG - Photogrammetrie, Fernerkundung, Geoinformation **2013** (4): 367–379.

BISON, P., BORTOLIN, A., CADELANO, G., FERRARINI, G., FURLAN, K. & GRINZATO, E., 2012: Geometrical correction and photogrammetric approach in thermographic inspection of buildings. – 11th International Conference on Quantitative InfraRed Thermography: http://www.ndt.net/article/qirt2012/papers/QIRT-2012-285.pdf (3.8.2015).

BROWN, D.C., 1971: Close-range camera calibration. – Photogrammetric Engineering **37** (8): 855–866.

DERENIAK, E.L. & BOREMAN, G.D., 1996: Infrared detectors and systems. – 1. edn., Wiley-Interscience: ISBN-13 978-0471122098.

FISCHLER, M.A. & BOLLES, R.C., 1981: Random sample consensus: A paradigm for model fittin with applications to image analysis and automated cartography. – Communications of the ACM **24** (6): 381–395.

FLIR, 2012: User documentation, FLIR A6xx series. – FLIR Commercial Systems. Meer, Belgium.

FURUKAWA, Y. & PONCE, J., 2007: Accurate, dense, and robust multi-view stereopsis. – IEEE Conference on Computer Vision and Pattern Recognition 2007 (CVPR '07): 1–8.

GONZÁLEZ-AGUILERA, D., RODRIGUEZ-GONZALVEZ, P., ARMESTO, J. & LAGÜELA, S., 2012: Novel approach to 3D thermography and energy efficien y evaluation. – Energy and Buildings **54**: 436–443.

HARTLEY, R.I. & ZISSERMAN, A., 2004: Multiple View Geometry in Computer Vision. – Second edn., Cambridge University Press, ISBN:0521540518, United Kingdom.

HIRSCHMÜLLER, H., 2005: Accurate and effi cient stereo processing by semi-global matching and mutual information. – IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005), Volume 2: 807–814.

HIRSCHMÜLLER, H. & BUCHER, T., 2010: Evaluation of digital surface models by semi-global matching. – DGPF Tagungsband **19** (2010): 571–580.

HOEGNER, L., KUMKE, H., MENG, L. & STILLA, U., 2007: Automatic extraction of textures from infrared image sequences and database integration for 3D building models. – PFG – Photogrammetrie-Fernerkundung-Geoinformation **2007** (6): 459–468.

HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U., 2013: Co-registration of time-of-fligh (ToF) camera generated 3D point clouds and thermal infrared images (IR). – SEYFERT, H. (ed): DGPF Tagungsband, Dreilndertagung DGPF, OVG, SGPF, Volume **22**, 481–488.

HOEGNER, L. & STILLA, U., 2015: Building facade object detection from terrestrial thermal infrared image sequences combining different views. – ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences **II-3/W4**: 55–62.

IWASZCZUK, D., HOEGNER, L., SCHMITT, M. & STILLA, U., 2012: Line based matching of uncertain 3d building models with IR image sequences for precise texture extraction. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2012** (5): 511–521.

KLIX, M., SCHNITZER, F., PFANNE, M. & JANSCHEK, K., 2014: Guidance-Navigation-Control System zur 3D-Exploration von ausgedehnten Objekten mit Multi-Roboter-Systemen. – 63. Deutscher Luft- und Raumfahrtkongress (DLRK), Augsburg: on CD-ROM.

KOCH, K.-R., 2004: Parameterschätzung und Hypothesentests in linearen Modellen. – 4. edn., Ferd. Dümmler Verlag, Bonn.

KOHLER, S., AGRICOLA, A.-C., JOEST, S., PETERS, S. & STOLTE, C., 2013: Energieeffizien als Säule der Energiewende. – Energiewirtschaftliche Tagesfragen **63**: 8ff.

LAGÜELA, S., GONZALEZ-JORGE, H., ARMESTO, J. & ARIAS, P., 2011a: Calibration and verificatio of thermographic cameras for geometric measurements. – Infrared Physics & Technology **54** (2): 92–99.

LAGÜELA, S., MARTÍNEZ, J., ARMESTO, J. & ARIAS, P., 2011b: Energy efficien y studies through 3D laser scanning and thermo-

graphic technologies. – Energy and Buildings **43** (6): 1216–1221.

LOWE, D.G., 1999: Object recognition from local scale-invariant features. – Seventh IEEE International Conference on Computer Vision, Volume **2**: 1150–1157.

LUHMANN, T., PIECHEL, J. & ROELFS, T., 2013: Geometric calibration of thermographic cameras. – Thermal Infrared Remote Sensing: 27–42, Springer.

LUHMANN, T., ROBSON, S., KYLE, S. & HARLEY, I., 2006: Close Range Photogrammetry: Principles, Methods and Applications. – Revised edn., Whittles Publishing, Dunbeath, United Kingdom.

MAURIELLO, M.L. & FROEHLICH, J.E., 2014: Towards automated thermal profilin of buildings at scale using unmannedaerial vehicles and 3D-reconstruction. – 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, UbiComp '14 Adjunct. ACM: 119–122, New York, NY, USA.

MAYER, H., 2008: Issues for image matching in structure from motion. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B3a): 21–26.

PECH, K., STELLING, N., KARRASCH, P. & MAAS, H., 2013: Generation of multitemporal thermal orthophotos from UAV data. – ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **1** (2): 305–310.

REMONDINO, F., 2006: Detectors and descriptors for photogrammetric applications. – The international archives of the photogrammetry, remote sensing and spatial information sciences **36** (3): 49–54.

RUDOL, P. & DOHERTY, P., 2008: Human body detection and geolocalization for uav search and rescue missions using color and thermal imagery. – Aerospace Conference, 2008 IEEE: 1–8.

SCHNEIDER, D. & MAAS, H.-G., 2007: Integrated bundle adjustment of terrestrial laser scanner data and image data with variance component estimation. – The Photogrammetric Journal of Finland **20**: 5–15.

WU, C., 2007: SiftGPU: A GPU implementation of Scale Invaraint Feature Transform (SIFT): http://cs.unc.edu/ccwu/siftgpu/ (3.8.2015).

WU, C., 2011: VisualSFM: A Visual Structure from Motion System: http://ccwu.me/vsfm/ (3.8.2015).

YASTIKLI, N. & GULER, E., 2013: Performance evaluation of thermographic cameras for photogrammetric documentation of historical buildings. – Boletim de Ciéncias Geodésicas **19** (4): 711–728.

Addresses of the Authors:

Dr. PATRICK WESTFELD, DAVID MADER & Prof. Dr. HANS-GERD MAAS Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing, D-01062 Dresden, e-mail: {patrick.westfeld}{david.mader} {hans-gerd.maas}@tu-dresden.de

Manuskript eingereicht: Mai 2015
Angenommen: August 2015