

## 3D THERMAL MAPPING OF BUILDING ROOFS BASED ON FUSION OF THERMAL AND VISIBLE POINT CLOUDS IN UAV IMAGERY

Mitra Dahaghin<sup>1</sup>, Farhad Samadzadegan<sup>1,\*</sup>, Farzaneh Dadras Javan<sup>1</sup>

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran - (samadz, fdadrasjavan, mitra.dahaghin)@ut.ac.ir

### KEY WORDS:

Thermal Infrared Imaging, Unmanned Aerial Vehicles, Building Roof, Point Cloud Generation, Fusion

### ABSTRACT:

Thermography is a robust method for detecting thermal irregularities on the roof of the buildings as one of the main energy dissipation parts. Recently, UAVs are presented to be useful in gathering 3D thermal data of the building roofs. In this topic, the low spatial resolution of thermal imagery is a challenge which leads to a sparse resolution in point clouds. This paper suggests the fusion of visible and thermal point clouds to generate a high-resolution thermal point cloud of the building roofs. For the purpose, camera calibration is performed to obtain internal orientation parameters, and then thermal point clouds and visible point clouds are generated. In the next step, both two point clouds are geo-referenced by control points. To extract building roofs from the visible point cloud, CSF ground filtering is applied, and the vegetation layer is removed by RGBVI index. Afterward, a predefined threshold is applied to the normal vectors in the z-direction in order to separate facets of roofs from the walls. Finally, the visible point cloud of the building roofs and registered thermal point cloud are combined and generate a fused dense point cloud. Results show mean re-projection error of 0.31 pixels for thermal camera calibration and mean absolute distance of 0.2 m for point clouds registration. The final product is a fused point cloud, which its density improves up to twice of the initial thermal point cloud density and it has the spatial accuracy of visible point cloud along with thermal information of the building roofs.

### 1. INTRODUCTION

Considering the growing importance of the optimal use of energy and the allocation of one-third of total energy consumption to the building sector, there is a need to model the current status of energy performance in the buildings (González et al., 2012). Consequently, modifications can be made in the building to prevent the waste of energy. Thermography is a robust tool for representing environmental heat while Thermal cameras capture the temperature as an image and could be applied for thermal inspection of different parts of buildings to detect thermal bridges, thermal irregularities and defects such as damages, cracks, air leakage sources and insulation problems (Al Lafi, 2017).

The thermal inspection of the building roof is of particular importance because it is one of the main sources of energy loss. Detecting and fixing defects will increase the life of the roof and, will save energy and costs. Unmanned Aerial Vehicles (UAVs) are presented to be useful and cost-effective platforms for data gathering over the urban and residential regions (Mader et al., 2016). To visualize energy performance and simultaneously measure and interpret it, there is a need for 3D thermal modeling (Borrman et al., 2013). Related studies in 3D thermal modeling are divided into three main categories:

- i. Mapping thermal images to an existing 3D model. Schreyer and Hoque (2009) used SketchUp software to create a CAD model of the buildings based on thermal images and then mapped thermal images to the 3D

model. Iwaszczuk and Stilla (2017) textured 3D building model in CityGML format with thermal images. For this purpose, a line-based model-to-image matching was used.

- ii. Fusion of visible images and thermal images. Lagüela et al. (2012) fused visible images and thermal images at the pixel level by using the intensity-hue-saturation (HIS) method and generated a point cloud based on fused images. Ham and Golparvar-Fard (2012) captured thermal and visible images simultaneously with a handheld infrared sensing camera and used extrinsic parameters of visible images to generate a thermal point cloud. Hoegner et al. (2016) evaluated different strategies for coregistration and fusion of UAV-based 3D point clouds and thermal infrared images. Javadnejad (2018) fused 3D point cloud with R, G, B, and thermal attributes by using boresight separation parameters between multi-cameras to project 3D coordinates from the point cloud onto the 2D thermal image space.
- iii. Thermal texture mapping to the point clouds. Borrman et al. (2013) used a robot and, multiple scans from different positions were registered by using 6D SLAM technique; then the point cloud was colored with solving the relation between scanner and cameras through calibration steps. Al Lafi (2017) created a panoramic thermal image from overlapping thermal images and mapped it to the point cloud model using reference points. Lin et al. (2019) applied a coarse registration between thermal and RGB

---

\* Corresponding author

point clouds. Reliable matching features on thermal-RGB image pairs were extracted then followed by mono-plotting of the RGB image and image resection of the thermal image. At the end an image pose refinement was used to decrease blur effects.

According to the mentioned methods, the first category includes some limitations such as the difficulty of aligning IR images to the model that needs some manual work, the 3D model does not represent as-is building condition, and the final model shows relative temperature differences. In the texture mapping method, there are issues with shooting direction of IR camera that should be perpendicular to the facades, and sequential thermal images have to overlap 50% for image registration (Wang et al., 2012). In this research, the advantages of thermal and RGB image fusion method are used. The purpose of this paper is to generate a high-resolution thermal point cloud of the building roofs. For this purpose, thermal and visible images are acquired using a small UAV system. The lower spatial resolution of thermal images causes sparse thermal point cloud, and the edges of the buildings are distorted and not accurate enough for 3D mapping applications. So, it is fused with a higher resolution visible point cloud in order to yield combined point cloud with the spatial accuracy of visible point cloud and thermal information content of the building roofs.

## 2. PROPOSED METHOD

The proposed methodology consists of 5 steps: (1) camera calibration, (2) point clouds generation, (3) point clouds registration, (4) building roofs extraction, and (5) fusion of visible and thermal point clouds. Fig.1 shows the workflow of the proposed strategy.

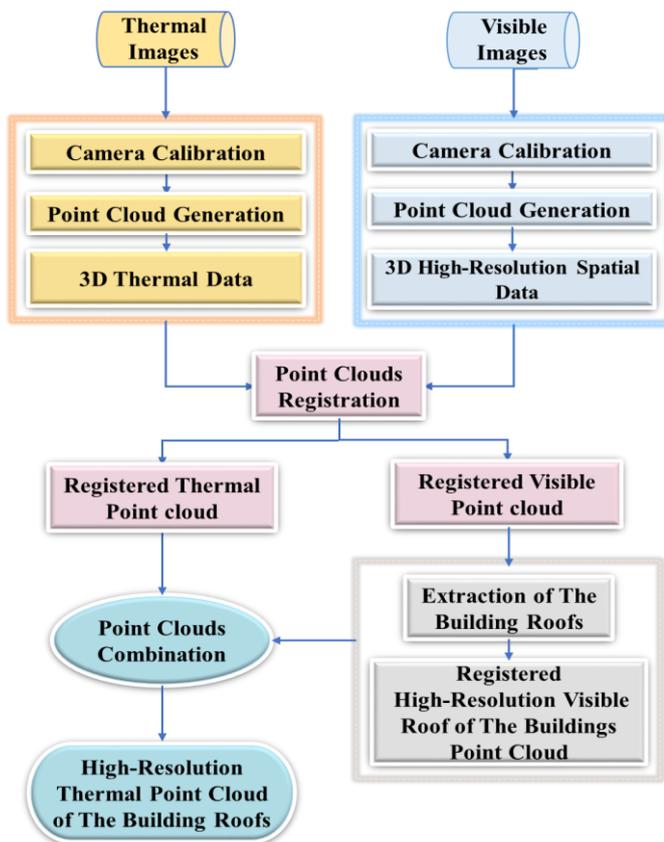


Figure.1 Workflow of the proposed strategy.

## 2.1 Camera Calibration

In the first processing step, camera calibration is performed to obtain interior orientation parameters of both visible and thermal sensors. Self-calibration is available for visible images but, because of the low spatial resolution of thermal images, thermal camera calibration is applied. For the purpose, Zhang's method is considered in which a calibration target with a specific pattern is captured in multiple views from different convergent directions. Circular pattern and chessboard are two common patterns used for calibration targets. Datta et al. (2009) and Usamentiaga et al. (2017) compared the camera calibration with two circular and chessboard patterns, and results in higher accuracy were obtained for the circular pattern. Thus, in this paper, circular targets are chosen for thermal camera calibration. In each image, the center of circles are needed to be detected that requires an ellipse extraction algorithm. Ouellet and Hébert (2008) algorithm is executed to the estimation of ellipse parameters by exploiting tangent lines. This algorithm used the advantage of dual conic model and constrained the dual conic to a dual ellipse. After locating the patterns and forming the object coordinates, the camera calibration parameters are calculated. The Homography matrix  $H$  is used to relate the object space and the image space as Eq.1.

$$x = HX, \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

In equation 1,  $[x \ y \ z]$  represents the coordinates of a point in the image space and  $[X \ Y \ Z]$  represents the coordinates of a point in the object space. The Levenberg–Marquardt method is then used to optimize the Homography by minimizing the re-projection error (Zhang, 2000). Given  $N$  pairs of calibration images with  $m$  dots in each one, the cost function is defined as:

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \hat{m}(A, R_i, t_i, M_j)\|^2 \quad (2)$$

where  $m_{ij}$  is the image coordinate and  $\hat{m}$  is the projection coordinate. By the accurate estimation of the Homography matrix, the internal camera parameters are calculated.

## 2.2 Point Clouds Generation

To create the point cloud, from un-ordered, un-calibrated RGB images, state-of-the-art Structure from Motion (SfM) and Multi-View-Stereo (MVS) photogrammetry and computer vision techniques are applied. After acquisition of block of RGB images with at least overlap of 80%, a scale-invariant feature transform (SIFT) then identifies common feature points across the image set (Lowe, 2004). SIFT feature descriptors are invariant to scale, orientation, affine distortion and, partial illumination changes. The fundamental matrix (F-matrix) can be calculated using common keypoints of image pairs (Longuet-Higgins, 1979). The F-matrix establishes the relationship between the two images and preserving collinearity constrain. Candidate F-matrices are evaluated using the Random Sample Consensus (RANSAC) method (Fischer and Bolles, 1981). After further refinement, all 'outlier' matches are removed (Smith et al., 2016). From the correct feature correspondences, SfM uses sparse bundle adjustment (Snavely et al., 2008), to transform measured image coordinates into 3-D points covering the area of interest (Micheletti et al., 2015). The sparse point cloud is then intensified using MVS techniques (Furukawa and Ponce, 2010). The same operation is repeated with a block of thermal images, but due to the limitation of thermal images, the internal

orientation parameters obtained in the previous step will be used instead of a self-calibration procedure.

### 2.3 Point Clouds Registration

In order to register the two generated visible and thermal point clouds, at first, the visible point cloud is geo-referenced via the control points that are extracted using Google Earth. A geo-referenced point cloud is given in an Earth-fixed coordinate system, that in this study, WGS84 is used as an earth-centered system. To specify the coordinate of the point cloud in a real-world coordinate system a 3D Helmert transformation (with seven parameters: three translations, three rotations and, one scale) can be used by considering common GCPs and solve the absolute orientation (Turner et al., 2012). It is necessary to have at least three control points to calculate the transformation, though more is needed to assess the quality of the transformation, independently. The control points should have a good dispersion and be distributed throughout the model (James and Robson, 2012). Control points are measured manually in the images, and their 3D locations are determined in the model through the intersection. After adding the control points, the position of the cameras is optimized, and a dense visible point cloud is generated which the ortho-image can be obtained based on it. Then, by applying control points captured from visible ortho-image to the thermal point cloud, two point clouds are registered.

### 2.4 Building Roofs Extraction

To detect the buildings and extract their roofs, in the first step, filtering of ground points is required. For the purpose, cloth simulation filtering (CSF) algorithm is applied (Zhang et al., 2016). This method is based on a simulation that if the point cloud is turned upside down and then a piece of cloth is dropped on the inverted surface because of the gravity and the cloth is defined with rigidity, then the final shape of the cloth is the DTM. Fig. 2 illustrates the overview of the concept of this algorithm.

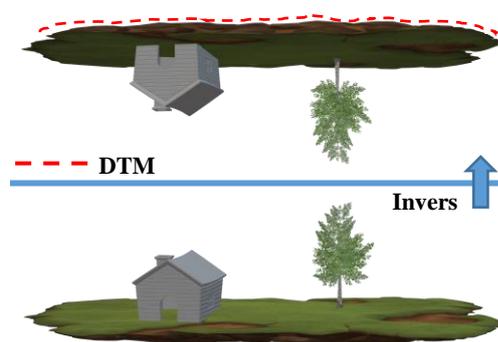


Figure 2. Overview of the basic concept of CSF algorithm.

By filtering the DTM, ground points are separated from non-ground points. Then, to remove vegetation and trees the RGBVI index that defined as the normalized difference of the squared green reflectance and the product of blue × red reflectance is used (Bendig et al., 2015) as follows:

$$RGBVI = (R_G * R_G) - (R_R * R_B) / (R_G * R_G) + (R_R * R_B) \quad (3)$$

where  $R_R$  = red,  $R_G$  = green,  $R_B$  = blue. A height threshold is applied to remove points with elevations below the specified value with the concern that a building cannot be that low.

Because of the different orientation of normal vectors in z-direction for roofs and walls, it is possible to extract the facet of roofs and isolate them from walls by computing the normal vectors in the z-direction and applying a predefined threshold. To compute normal vectors, Boulch and Marlet (2016) method is used which map a Hough accumulator for normal estimation into an image and use it as an input of a convolutional neural network (CNN) to learn to estimate a normal from such an image-accumulator. The framework of this method is pictured in Fig. 3. This method is robust to noise, outliers, and density variation, especially in the presence of sharp edges.

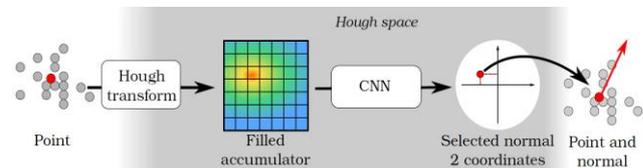


Figure 3. Framework of CNN-based normal estimation (Boulch and Marlet, 2016).

After this step remaining points are belongs to the roof of the buildings and some others are excess points. The use of Gaussian filter to remove noise makes the edges soften, so 3D Connected Components Labeling is used to separate building roof segments from the rest of the points that is done by a connectivity analysis in TIN (Triangulated Irregular Network), which was modified based on an algorithm suggested by (Lumia et al., 1983). Finally building roofs are extracted and isolated from ground points, vegetation points, walls and excess points.

### 2.5 Combination of Visible and Thermal Point Clouds

During previous steps, points of the building roofs that are extracted from the visible point cloud are registered to the corresponding points in the thermal point cloud. So, for the combination of visible and thermal data, each visible point is referred to its corresponding point in the thermal point cloud and its radiometric information is extracted and assigned to the visible point as stated in Eq.3 and 4:

$$(X', Y', Z')_T = FindNearestNeighbor(X, Y, Z)_V \quad (3)$$

$$F(X, Y, Z)_{fusion} = G(X', Y', Z')_{Thermal} \quad (4)$$

where  $(X', Y', Z')$  and  $(X, Y, Z)$  are the spatial coordinate in the thermal point cloud and the visible point cloud,  $F_{fusion}$  is the radiometric information of the new generated point cloud and  $G_{Thermal}$  is the radiometric information of the thermal point cloud. In this way, the process of resampling is carried out based on the nearest neighbor interpolation strategy. Finally, a fused point cloud is generated with the spatial accuracy of the visible point cloud and the radiometric information of the thermal point cloud that represents the heat map for each building roof.

## 3. RESULTS

In this research, for thermal camera calibration, a rectangular calibration board with 13 \* 17 hollow circles, 12 mm in diameter is used where the distances between centers of the circles are 24 mm. Six coded targets are considered on the calibration board to specify target positions in different images (Fig. 4(a)). To make the targets more distinguishable and make a better contrast in the images, the calibration board heated, then imaged by a thermal camera in multiple views (Fig. 4(b)).

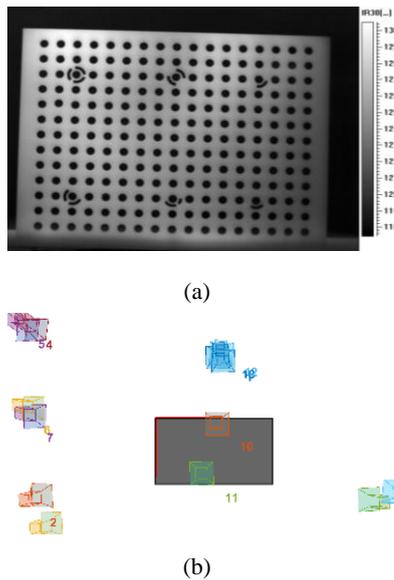


Figure 4. Thermal camera calibration: (a) circular calibration target in the thermal camera, (b) thermal camera positions.

Calibration board images are changed to binary images using the adaptive thresholding (Prakash and Karam, 2012), and then the centers of the circles are identified through the algorithm expressed in (Ouellet and Hebert, 2008) that is based on the intersection of lines perpendicular to the gradient (Fig. 5). The circles in each image are numbered so that the same circles in the different images have the same number.

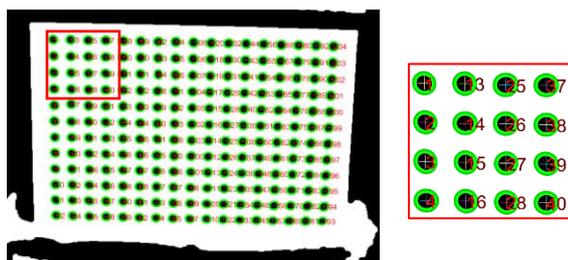


Figure 5. Detection of centers of circles based on Ouellet method.

Center of the circles are considered as image coordinate and by forming object coordinate, camera calibration is done based on Zhang's method. In order to evaluate the accuracy of thermal camera calibration mean re-projection error is calculated through mapping the calibration point positions from the object coordinate system to the image plane, using the obtained set of calibration parameters.

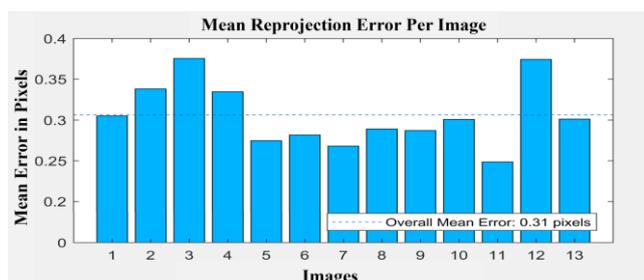


Figure 6. Mean re-projection error chart for thermal camera calibration.

Fig. 6 illustrates the mean re-projection error per image with an average of 0.31 pixels over all 13 views and 221 calibration points. In the next step, the dense visible and thermal point cloud are generated (Fig. 7).

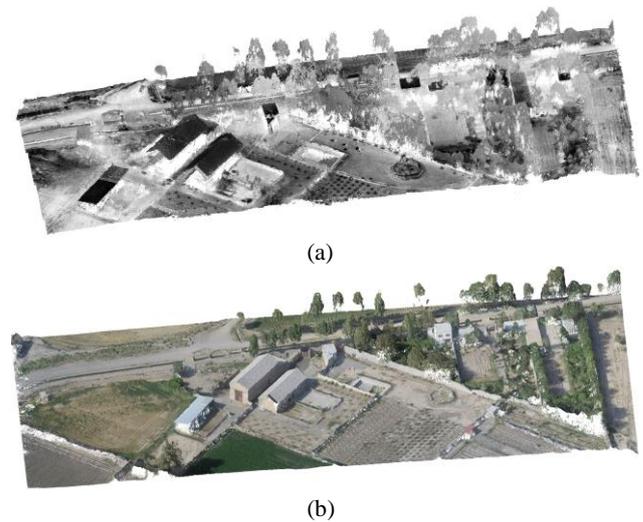


Figure 7. Point cloud generation: (a) thermal point cloud, (b) visible point cloud.

712 thermal frames that they are 640 x 480 in dimension with pixel size of 17µm and ground sampling distance (GSD) of 4.45 cm/pix were extracted from recorded video and 696 RGB images that they are 3000 x 2250 in dimension with pixel size of 1.73µm and GSD of 2.22 cm/pix captured via UAV. Thermal and visible point clouds are generated with a point density of 106 points/m<sup>2</sup> and 227 points/m<sup>2</sup> respectively. The point clouds are then geo-referenced based on the control points that are depicted in Fig.8.

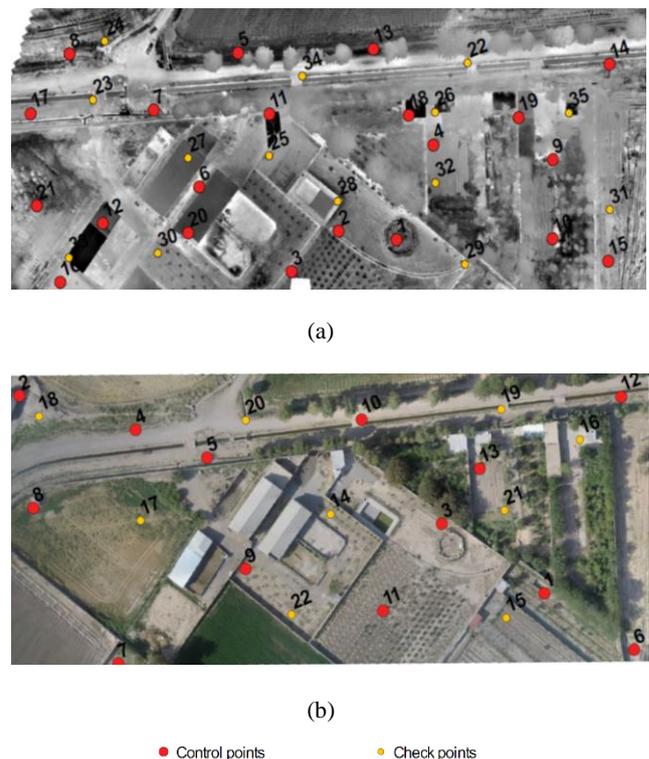


Figure 8. Geo-referenced images: (a) thermal

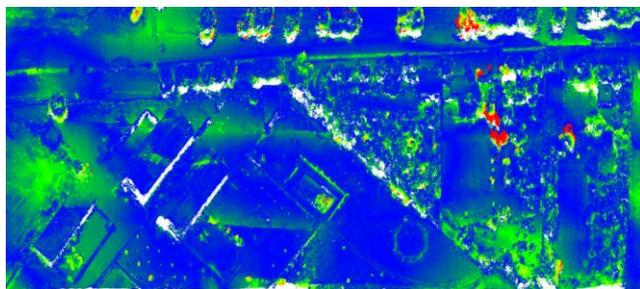
orthophoto, (b) visible orthophoto.

13 Control points and 9 check points are considered for visible point cloud and total RMSEs are 52.9628 cm and 53.5022 cm, respectively. 21 Control points and 14 check points are considered for thermal point cloud and total RMSEs are 5.03339 cm and 32.7098 cm, respectively. These errors are indicated in details in Table 1.

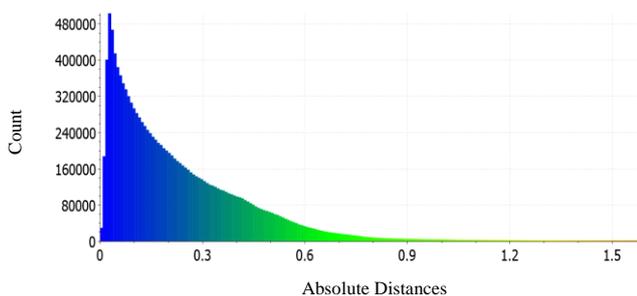
|                                  | X error<br>(cm) | Y error<br>(cm) | Z error<br>(cm) | XY error<br>(cm) |
|----------------------------------|-----------------|-----------------|-----------------|------------------|
| <b>Thermal<br/>Control Point</b> | 1.96434         | 3.43715         | 3.10844         | 3.95886          |
| <b>Thermal<br/>Check Point</b>   | 3.97794         | 3.21909         | 32.307          | 5.11728          |
| <b>Visible<br/>Control Point</b> | 11.2646         | 15.4173         | 49.4011         | 19.0942          |
| <b>Visible<br/>Check Point</b>   | 9.91278         | 10.6897         | 51.4777         | 14.5785          |

Table 1. Control points and check points RMSEs in the visible and thermal point cloud

The registration accuracy is then evaluated based on absolute distances between the visible point cloud and thermal point cloud, and it is displayed through colors where blue represents the lowest distance and red represents the highest distance (Fig.9).



(a)



(b)

Figure 9. Absolute distances between visible point cloud and thermal point cloud: (a) distances are displayed through the colors, (b) histogram of absolute distances (units are in meters).

Mean absolute distance between visible and thermal point clouds is 0.2 m. As shown, most areas are blue with the least difference. The red zone is the part of the trees where there is more difference in them. Most of the buildings are displayed in blue, which indicates the success of registration process. Results of building roof extraction stages are presented in Fig. 10-14.



Figure 10. Ground filtering of visible point cloud.

Fig. 10 illustrates visible point cloud after applying CSF algorithm with cloth resolution of 2 and classification threshold of 0.5 (the unit is same as the unit of point clouds) to separate non-ground points from ground points.



Figure 11. Applying RGBVI index.

In order to remove vegetation points, RGBVI index is calculated for all the points and then figure out an adequate threshold that removes as many vegetation points as possible without removing building-related points (Fig. 11).



Figure 12. Applying the normal limit in z direction.

Calculating the normal vectors in the z-direction for the remaining points results that the normal vector for the roof planes is close to 1 or -1 and for the wall planes are close to 0 so the points related to the walls can be eliminated (Fig. 12).



Figure 13. Utilize Connected Component Labeling algorithm.

To isolate roof points from excess points Connected Component Labeling algorithm is applied. Finally, all the building roofs are completely extracted (Fig. 13). Afterwards, data resampling is carried out, and the high-resolution thermal point cloud of the building roofs is generated with the same point density of visible point cloud along with thermal information of thermal point cloud (Fig. 14(a)).

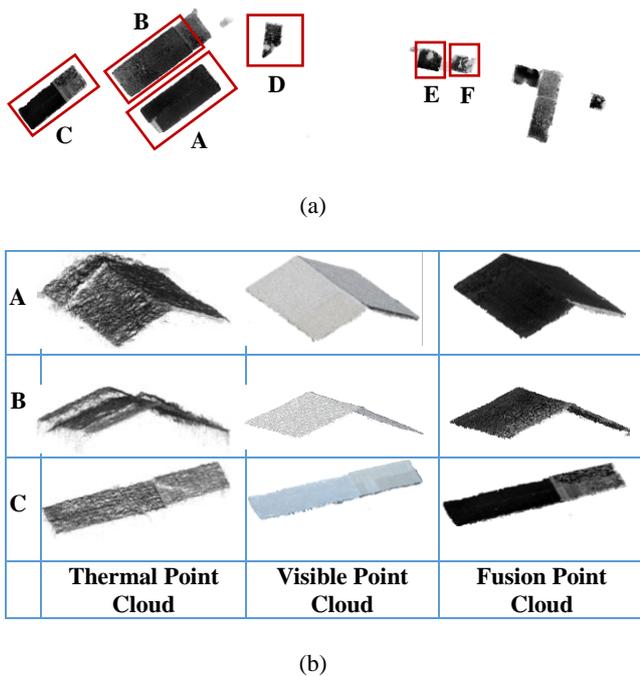


Figure 14. Final results: (a) high resolution thermal point cloud of building roof, (b) comparison of extracted roofs in thermal and visible point clouds and fused point cloud.

Sections of the buildings in the thermal point cloud, visible point cloud, and fusion point cloud are presented for comparison in (Fig. 14(b)). As it is presented, thermal point cloud distortions do not exist in the fused point cloud, and edges of the building roofs in the fused point cloud are as sharp as the visible point cloud.

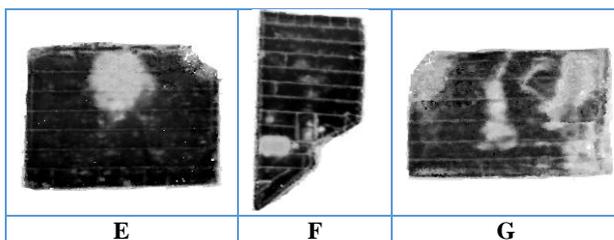


Figure 15. Thermal anomalies of building roofs.

Buildings E, F, G are shown in Fig. 15 and the thermal anomalies of building roofs are evident. Lighter radiometric data indicates more heat in the area than the area with darker radiometric information.

#### 4. Conclusions

This study has proposed a method for 3D thermal mapping of building roofs and suggested the combination of visible and thermal point cloud to address the problem of low-resolution thermal imagery. To generate a thermal point cloud, internal orientation parameters are needed, which they were obtained through thermal camera calibration. For 3D mapping of thermal information to the visible point cloud, accurate registration of two point clouds is of importance. By combining visible point cloud of the building roofs and registered thermal point cloud, a fused point cloud was generated that it has spatial accuracy of

visible point cloud along with thermal information of the building roofs. The initial density of the thermal point cloud was 106 points/m<sup>2</sup>, which eventually density of fused point cloud improved to 227 points/m<sup>2</sup>. Results showed that fusion of visible and thermal point clouds could be used to generate a high-resolution, distortion-free thermal point cloud of the building roofs. Future work will perform this algorithm on a larger dataset that includes more buildings and will investigate to segment thermal anomalies for further analysis.

#### REFERENCES

- Al Lafi, G. 2017. 3D Thermal Modeling of Built Environments Using Visual and Infrared Sensing. Concordia University.
- Borrmann, D., Elseberg, J. & Nüchter, A. 2013. Thermal 3D mapping of building façades. Intelligent autonomous systems 12. Springer.
- Boulch, A. & Marlet, R. Deep learning for robust normal estimation in unstructured point clouds. Computer Graphics Forum, 2016. Wiley Online Library, 281-290.
- Datta, A., Kim, J.-S. & Kanade, T. Accurate camera calibration using iterative refinement of control points. 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009. IEEE, 1201-1208.
- Fischler, M. A. & Bolles, R. C. J. C. O. T. A. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. 24, 381-395.
- Furukawa, Y., Ponce, J. J. I. T. O. P. A. & Intelligence, M. 2009. Accurate, dense, and robust multiview stereopsis. 32, 1362-1376.
- González-Aguilera, D., Rodríguez-Gonzalvez, P., Armesto, J., Lagüela, S. J. E. & Buildings 2012. Novel approach to 3D thermography and energy efficiency evaluation. 54, 436-443.
- Ham, Y. & Golparvar-Fard, M. Rapid 3D energy performance modeling of existing buildings using thermal and digital imagery. Construction Research Congress 2012: Construction Challenges in a Flat World, 2012. 991-1000.
- Hoegner, L., Tuttas, S., Xu, Y., Eder, K., Stilla, U. J. I.-I. A. O. T. P., Remote Sensing & Sciences, S. I. 2016. Evaluation of methods for coregistration and fusion of rps-based 3d point clouds and thermal infrared images. 41, 241-246.
- James, M. & Robson, S. J. J. O. G. R. E. S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. 117.
- Javadnejad, F. 2017. Small Unmanned Aircraft Systems (UAS) for Engineering Inspections and Geospatial Mapping.
- Lagüela, S., Armesto, J., Arias, P. & Herráez, J. J. A. I. C. 2012. Automation of thermographic 3D modelling through image fusion and image matching techniques. 27, 24-31.
- Longuet-Higgins, H. C. J. N. 1981. A computer algorithm for reconstructing a scene from two projections. 293, 133.
- Lowe, D. G. J. I. J. O. C. V. 2004. Distinctive image features from scale-invariant keypoints. 60, 91-110.

- Lumia, R., Shapiro, L., Zuniga, O. J. C. V., Graphics, & Processing, I. 1983. A new connected components algorithm for virtual memory computers. 22, 287-300.4
- Mader, D., Blaskow, R., Westfeld, P., Weller, C. J. I. A. O. T. P., Remote Sensing & Sciences, S. I. 2016. Potential of uav-based laser scanner and multispectral camera data in building inspection. 41.
- Micheletti, N., Chandler, J. H. & Lane, S. N. 2015. Structure from motion (SfM) photogrammetry.
- Ouellet, J.-N., Hébert, P. J. M. V. & Applications 2009. Precise ellipse estimation without contour point extraction. 21, 59.
- Prakash, C. D. & Karam, L. J. Camera calibration using adaptive segmentation and ellipse fitting for localizing control points. 2012 19th IEEE International Conference on Image Processing, 2012. IEEE, 341-344.
- Schreyer, A. C. & Hoque, S. J. P. O. I., 2009. Interactive three-dimensional visualization of building envelope systems using infrared thermography and SketchUp. 177.
- Smith, M., Carrivick, J. & Quincey, D. J. P. I. P. G. 2016. Structure from motion photogrammetry in physical geography. 40, 247-275.
- Snavely, N., Seitz, S. M. & Szeliski, R. J. I. J. O. C. V. 2008. Modeling the world from internet photo collections. 80, 189-210.
- Turner, D., Lucieer, A. & Watson, C. J. R. S. 2012. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds. 4, 1392-1410.
- Usamentiaga, R., Garcia, D., Ibarra-Castanedo, C. & Maldague, X. J. M. 2017. Highly accurate geometric calibration for infrared cameras using inexpensive calibration targets. 112, 105-116.
- Wang, C., Cho, Y. K. & Gai, M. J. J. O. C. I. C. E. 2012. As-is 3D thermal modeling for existing building envelopes using a hybrid LIDAR system. 27, 645-656.
- Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X. & Yan, G. J. R. S. 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. 8, 501.
- Zhang, Z. J. I. T. O. P. A. & Intelligence, M. 2000. A flexible new technique for camera calibration. 22.