

THERMAL 3D MODELS ENHANCEMENT BASED ON INTEGRATION WITH VISIBLE IMAGERY

F. Dadras Javan *, M. Savadkouhi

Dept. of Photogrammetry, School of Surveying and Geospatial Engineering, University College of Engineering, University of Tehran, Tehran, Iran - (Fdadrasjavan, M.savadkouhi)@ut.ac.ir

KEY WORDS: UAV Photogrammetry, Thermal Imagery, Point Cloud Integration, Key frame Selection, Thermal Camera Calibration, Textured Mesh

ABSTRACT:

In the last few years, Unmanned Aerial Vehicles (UAVs) are being frequently used to acquire high resolution photogrammetric images and consequently producing Digital Surface Models (DSMs) and orthophotos in a photogrammetric procedure for topography and surface processing applications. Thermal imaging sensors are mostly used for interpretation and monitoring purposes because of lower geometric resolution. But yet, thermal mapping is getting more important in civil applications, as thermal sensors can be used in condition that visible sensors cannot, such as foggy weather and night times which is not possible for visible cameras. But, low geometric quality and resolution of thermal images is a main drawback that 3D thermal modelling are encountered with. This study aims to offer a solution for to fixing mentioned problem and generating a thermal 3D model with higher spatial resolution based on thermal and visible point clouds integration. This integration leads to generate a more accurate thermal point cloud and DEM with more density and resolution which is appropriate for 3D thermal modelling. The main steps of this study are: generating thermal and RGB point clouds separately, registration of them in two course and fine level and finally adding thermal information to RGB high resolution point cloud by interpolation concept. Experimental results are presented in a mesh that has more faces (With a factor of 23) which leads to a higher resolution textured mesh with thermal information.

1. INTRODUCTION

In the last few years Unmanned Aerial Vehicles (UAVs) are being frequently used to acquire high detailed and resolution images and consequently producing Digital Surface Models (DSMs) and orthophotos in a photogrammetric procedure for topography and surface process research. Various applications such as monitoring (Cryderman et al., 2014; Siebert and Tiezer, 2014; Lelong et al., 2008; Berni et al., 2009), providing volume estimates in earthworks (Zhou, 2009; Zhou, 2010), Change detection (Qin, 2014) and particularly low altitude and high resolution images have great potential on geomatics application such as 3D city modelling, large scale mapping, archaeology, true orthophoto generation and etc. (Zhang et al., 2011 and Reemondino et al., 2011). In addition, UAVs are flying due to collecting different remote sensing information and are using in several applications such as forest fire monitoring, agriculture field survey, search and rescue and etc. (Berni et al., 2009).

Recently, thermal imaging systems are used in various and wide fields of applications such as archaeological heritage documentation (Brumana et al., 2013), landslide hazard assessment (Teza et al., 2015), crop health monitoring (Mangus et al., 2016) and photovoltaic plants inspection (Tsanakas et al., 2017).

Thermal mapping is getting more important in civil applications, as thermal sensors may be applied in condition that visible sensors cannot, such as foggy weather and night times which is not possible for visible camera. Thermal cameras usually detect radiation in the long-infrared range of the electromagnetic spectrum (roughly 9,000–14,000 nanometers or 9–14 μm) and produce images of that radiation, called

Thermograms. Since infrared radiation is emitted by all objects with a temperature above absolute zero according to the black body radiation law, thermography makes it possible to see one's environment with or without visible illumination.

However, low geometric quality and resolution of thermal images is a main drawback that 3D thermal modelling encounters. In another word, 3D models reconstructed from thermal images have low spatial resolution (e.g. buildings with soft edges) due to relatively large pixel size of thermal sensors and their small focal length. Thermal data integration with visible 3D models simplifies its perception and will produce appropriate high spatial resolution thermal 3D model. In this regard there are various research with different methods, So that different 3D geometries can be integrated with thermal data as Building Information Models (BIM) (Mikeleit and Kraneis, 2010), 3D building models via texture mapping (Hoegner et al., 2007; Iwaszczuk et al., 2011), 3D point clouds via assignment and interpolation of the measured temperature to the points (Cabrelles et al., 2009; Borrmann et al., 2012; Vidas et al., 2013) or a point cloud and aerial photographs at the same time (Boyd, 2013). Using point cloud as spatial reference enables fast generation of results with a high level of detail and is appropriate for visual interpretation (Iwaszczuk and Stilla, 2017).

This study aims to offer a solution in order to fix mentioned problem and generate a thermal 3D model with high spatial resolution by thermal and visible point clouds integration. This integration leads to generate a more accurate thermal point cloud and DEM with reduced spacing distance which is appropriate for 3D thermal modelling. The main steps of this study are composed of generating thermal and RGB point

* Corresponding author

clouds separately, registration of them in two course and fine level and finally adding thermal information to RGB high resolution point cloud based on interpolation concept.

2. LITERATURE REVIEW

Integration of thermal data with high resolution data such as laser scanner point cloud or visible RGB image triangulation result is a common way to increase thermal point cloud density. In the last decade, several studies have addressed this issue; Coiras et al. 2000 proposed an edge base method in order to register thermal and visible images which consists of three basic steps; extraction of features from the images (persistent edges), matching of the extracted features and determination of a warping function from the determined images (Coiras et al., 2000). In order to detect foreground objects in tele-immersive spaces, (Johnson and Bajcsy, 2008) offered a robust method by integration of thermal and visible imagery, this fusion-based foreground detection is done in object level by fusing objects that were detected from visible and thermal images separately. Integration of thermal-visible imagery with TLS (Terrain Laser Scanner) data is proposed by (Cabrelles et al., 2009) in order to record an architectural monument 3D model, which is based on 3D reconstruction and mesh generation by TLS and generating a photo-realist and thermo-realistic data. The former is done by visible images bundle adjustment and adding visible texture to TLS mesh, and the latter is produced by single thermal image resection. In 2012 Borrmann (Borrmann et al., 2012) registered thermal-visible imagery and TLS point cloud in order to produce an accurate thermal 3D model of building facets, in order to prevent the waste of energy. He used visible images for heat sources and thermal bridges. With the same purpose, Vidas et al. 2013 offered an indoor mapping method by simultaneous exploiting a thermal and a RGB-D camera which yields a precise 3D model of a building indoor by combining thermal-visible images with depth map which is achieved from RGB-D camera. Hoegner et al. 2014 has considered the use of thermal imagery inappropriate for people detection because of low geometric resolution of that and absence of strong features. In order to compensate this weakness, he suggested a hybrid system containing a thermal and TOF (Time of Flight) camera. The NIR image of TOF camera is used due to its similar behaviour to the visible images and the depth map of that was used in order to depth finding and thermal images were exploited for people detection.

Some research has done in order to evaluate the potential of thermal imagery for 3D model reconstruction. By this goal, Khodaei et al. 2015 evaluated 3D model generated from thermal video frame sequences, and as a result has considered that the accuracy of thermal DSM is comparable with a visible one. In 2016 Hoegner extracted building textures with three methods, where one of them was integration of RGB and thermal images because of low geometric resolution of thermal images. uses the image sequence and camera calibration information only to reconstruct the scene in model coordinates and coregisters this model to a given 3D building model to derive optimized orientation parameters (Hoegner et al. 2016 a). With the same goal, (Hoegner et al. 2016 b) proposed a method in order to automatic registration and fusing thermal and visible point cloud which is done in five different methods; first is based on 2D line for both thermal and visible images, second is registering both point clouds using 2D image projection registration. Third, is based on registering 2D lines in thermal images and 3D lines in RGB point cloud, fourth is done by ICP (Iterative Closest Point) method for 3D point clouds and plane registering and the last method is based on point features in

both thermal and visible images. As a result, the RMS value achieved at its minimum for ICP method. Tsanakas (Tsankas et al., 2017) offered a new method for advance inspection of photovoltaic installation in large scale by aerial triangulation and terrestrial georeferencing of thermal and visible imagery. Iwaszczuk and Stillas, integrated thermal and visible images for camera pose refinement by matching uncertain 3D building models with thermal infrared image sequences for high quality texture extraction (Iwaszczuk and Stillas, 2017). Therefore, a methodology for co-registration of uncertain 3D building models with airborne oblique view images is presented. For this purpose, a line-based model for image matching is developed, in which the uncertainties of the 3D building model, as well as of the image features are considered. Matched linear features are used for the refinement of the exterior orientation parameters of the camera in order to ensure optimal co-registration. Maset et al 2017 in his research first evaluated the potential of thermal data for automatic 3D modelling and second assessed the ICP model ability to register thermal and visible point clouds (Maset et al., 2017). He generated thermal and visible point clouds separately and after registering them, thermal texture is added to the mesh generated from visible point cloud.

In this paper a method is proposed to produce a DSM with resolution of visible images and containing thermal data, in another word, a thermal 3D model is generated with high spatial resolution by thermal and visible point clouds integration. In the following section the methodology is described in more details.

3. PROPOSED METHOD

As shown in figure 1, proposed method consists of four main parts as described in the following sections.

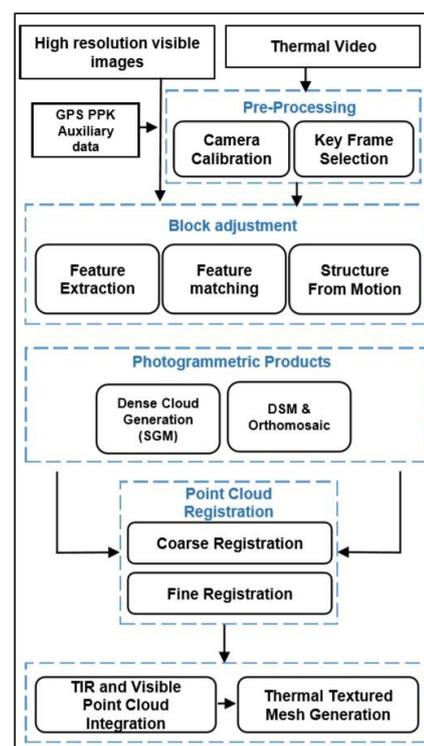


Figure1. Diagram of the proposed method.

3.1 Pre-Processing

Camera Calibration: for calibration of thermal camera, it is necessary to select or design an appropriate test field which is suitable for thermal imagery. There are several effective factors in designing this test field like material, shape, targets dimension and so on. In this study a new test field is designed, where a special plane filled with ellipses is designed so that a cool plane with holes such as ellipse were put in front of a warm monitor. These holes become lighter in thermal images while the cool plane will be dark, these lighter ellipses will act such a target in images. (See fig. 2) Although it is better to use non-planar test field, we used a planar test field due to problems such as construction and the importance of homogenous distributing of heat all over the test field (Zhang 1999 and S. Lagüela 2011). After capturing convergent images from this plane, an ellipse fitting algorithm proposed by Ouellet Jean-Nicolas is used. (Ouellet and Hebert, 2009) In this algorithm, first a gradient operator is used to extract the edge of ellipses then the geometric centers of these ellipses are calculated.

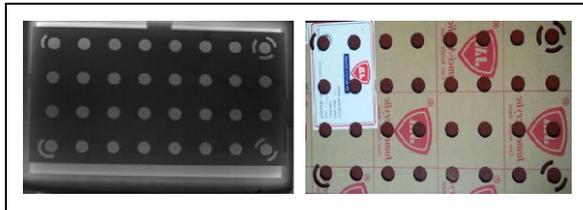


Figure 2. Calibration checkerboard image by a thermal camera (Left) and Calibration checkerboard view by an optical camera (Right)

This study considers the pinhole camera geometry for thermal camera in order to calculate the IO parameters. Equation (1) and (2) transform a point from 3D object space to 2D image space which is called collinearity equation:

$$x = -c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \quad (1)$$

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

where c = focal length
 x, y = image coordinates
 X_0, Y_0, Z_0 = coordinates of projection center
 X, Y, Z = object coordinates
 r = elements of rotation matrix.

Lens distortion parameters are obtained by frequently used Brown model eq. (3) and (4): (Brown, 1971)

$$x' = x(1 + k_1r^2 + k_2r^4 + k_3r^6 + p_2(r^2 + 2x^2) + 2p_1xy) \quad (3)$$

$$y' = y(1 + k_1r^2 + k_2r^4 + k_3r^6 + p_1(r^2 + 2y^2) + 2p_2xy) \quad (4)$$

where x' = image coordinate which is free of distortion
 k = radial distortion coefficient
 p = decentring distortion coefficient
 r = radial distance from projection of perspective point on image plane.

3.1.1 Key Frame Selection: Extracting key frames of a video sequence is a useful way for video abstracting. Key frames are the outstanding frames of a video sequence which are the whole frame representation. So they are also called representative frames (Truong and Venkatesh, 2007). Key frames can be selected in radiometric and geometric aspects. In radiometric key frame selection, the quality of the frames is checked and the blur frames are removed. In geometric key frame selection, the frames are checked for geometry of frames. Several studies have been proposed in order to control and quantify the emergence of the blur effect (Boult et al., 1997; Marziliano et al., 2002; Ong et al., 2003). One of these metrics with an acceptable operation is the BlUM metric, which is based on the human blur perception (Crete et al. 2007) The key idea of this blur estimation is to blur the initial image and to analyse the behaviour of the neighbouring pixels' variation. The blurring step should be done with a strong low-pass filter in order to be sure to compare the initial image with an image which seems blurred for the human perception. The choice of the type of filter is not exhaustive if it is a strong filter. The output of this algorithm is a score between 0 and 1. The results close to 0 represent the best sharpness and results close to 1 represent the worst blur of image, respectively. Geometric key frame selection is the procedure of selecting key frames due to the geometrical aspects of frames such as number of frames, position of selected frames, stability of them and etc. There is several research which studied about geometric key frame selection methods, three of these methods are based on Geometric Robust Information Criterion (GRIC) factor (Torr, 1998). (Pollefeys et al, 2004) and (Knorr, 2006) modified this metric by using different weights for it and (Ahmed, 2009) integrated that with another factor named PELC. (Gibson et al., 2002) and (Seo et al., 2008) had their exclusive scores for key frame selecting.

According to (Javidi et al., 2017) experiments, we used BlUM as a metric for radiometric aspect and the method proposed by Seo as a function for geometric key frame selection from our thermal aerial video frame sequences. Seo (Seo et al., 2004), considered three measures. First one is the ratio of the number of corresponding points about feature points, second is distribution of corresponding points about the frame and the last measure is the homography error. Equation (5) is the combination of these three measures which is presented by (Seo et al., 2008):

$$S = w_1(1 - \frac{N_c}{N_f}) + w_2\sigma_c + w_3H_{err} \quad (5)$$

where S = Score to select the key-frame
 N_c, N_f = Number of corresponding points and that of feature points
 σ_c = Standard deviation of the point density
 H_{err} = Homography error
 w = Weight used to alter the relative significance of each score.

By dividing a frame to sub-regions equation (6) calculates standard deviation of the point density.

$$\sigma_c = \sqrt{\frac{1}{N_s} \sum_{i=1}^{N_s} (N_{ci} - \frac{N_c}{N_s})^2} \quad (6)$$

where N_s = Number of sub-regions

N_e, N_{ci} = Number of inliers and that in the i 'th sub-region.

3.2 Block Adjustment

The process of block adjustment starts with feature extraction and matching, then Structure from Motion (SfM) algorithm is used in order to compute camera alignment. Detailed descriptions of block adjustment workflow will be discussed in the following subsections. Mentioned workflow is used for both thermal extracted key frames and high resolution visible images.

3.2.1 Feature Extraction and Matching: Feature points are extracted from the images as salient points. The SIFT operator (Scale Invariant Feature Transform; (Lowe, 2004)) is used to detect and describe local feature points in the image. The number of matched key-points is dependent on the image texture and resolution. This algorithm consists of six consecutive part. First, generating a "scale space" of original images to ensure scale invariance, second, using Laplacian of Gaussian (LoG) operator for finding key points. Third part is finding key points which are maxima and minima in the Difference of Gaussian (DoG) obtained in previous part. Forth is related to eliminating bad key points (Edges and low contrast regions), then an orientation is assigned to each key point, all further usage is related to these orientations. Finally SIFT features are calculated, this features are scale and orientation invariant and are appropriate for image matching.

After feature extraction, key points are matched by orientation information assigned to them before, in this level image matching is done. The presence of similar areas will produce false matches (outliers). Validation of matches is verified by outlier detection RANSAC (Random Aample Consensus; (Fischler and Bolles, 1981)) algorithm. It attempts to fit the data into a mathematical model by iteration, the more the iterations the higher the probability that the fitted data contains only inliers.

3.2.2 Structure from Motion: Structure from Motion or SfM is a photogrammetric algorithm which calculate camera poses (Motion) and 3D scene structure by finding correspondent features between a couple of image or multiple images captured from same scene (Ulmann, 1979) particularly producing 3D point cloud in case of aerial imagery due to producing accurate 3D model and Ortho-rectified images. The process of reconstructing sparse 3D point cloud of the scene from series of overlapping images can be done using structure from motion algorithm (Snavely et al., 2006). This study has used this algorithm in order to generate sparse 3D point cloud. As mentioned before, Feature points are extracted from the images as key-points. Then sequential bundle adjustment is used to orient the images using this set feature points. It iteratively builds a sparse 3D model of the scene and simultaneously estimates the exterior orientations of all images.

3.3 Dense Point Cloud Generation

After block adjustment, Semi Global Matching (SGM) algorithm is used to generate dense point cloud and DEM for both thermal and visible images, separately. Semi global-based matching use pixel-wise, Mutual Information (MI) matching for compensating the radiometric differences of input images instead of intensity based matching, because it is robust against many complex intensity transformations and even reflections (Hirschmueller, 2005). This method supported by a smoothness constraint expressed as a global cost function, but take local

disparities into consideration and performs a fast approximation by path-wise optimizations from all directions. (Hirschmueller, 2011).

Mutual Information (MI) is the most applied measure to calculate the matching cost for a pair of pixels which measures the dissimilarity of corresponding pixels and handles radiometric differences. Mutual information is the difference between the entropy (probability that a system will take a certain form) of each image and the joint entropy which is used to calculate the matching cost for each pixel and its disparity (The pixel value of the matched pixel in the second image) (Hirschmueller, 2008).

3.4 Point Cloud Registration

Thermal and visible point cloud registration procedure is implemented in two levels, course and fine registration. In the first level, the rigid body transform with GCPs is implemented, after first level, RGB and TIR point clouds would be close together in the 3D space and have approximately the same scale. Therefore, the TIR point cloud can be registered to the one generated from RGB images using Iterative Closest Point (ICP) algorithm. This part is called fine registration. ICP method may summarized in the following steps for given an input point cloud and target point cloud: 1) Determining pairs of corresponding points, 2) Estimate a transformation that minimize the distances between the correspondences, 3) Apply the transformation to align input and target point cloud.

3.5 Thermal and Visible Point Cloud Integration

After complete registration of thermal point cloud with visible point cloud, it is possible to make a higher resolution thermal point cloud which density is equal to the visible one. This procedure is done by simple linear weighted interpolation between point clouds. See equation (6):

$$T_{new} = \sum_{i=1}^N \frac{d_i t_i}{\sum_{i=1}^N d_i} \quad (6)$$

where T = Thermal intensity for visible point
 d = Distance between visible point and near thermal points (we have considered four nearest points)
 t = Thermal intensity of thermal points
 N = Number of nearest points (e.g. $N=4$)

4. RESULTS

4.1 Sensors and Dataset

In this study a thermal camera and a high resolution visible camera mounted on a light weight multi-rotor with roll and pitch axis stabilization have been used up to collect thermal and visible datasets. The high resolution visible camera is a Sony a6000 24 MP with a 35 mm lens and thermal camera is an uncooled focal plane array camera 640S which used to collect thermal video.

Generally, thermal imaging cameras are divided to short-wave, middle-wave and long-wave Infrared (IR) based on their electro-magnetic spectrum range sensitivity. The thermal camera used in this study detects TIR region of IR which contains middle and long wave. More detailed specifications are presented in table (1).

Flight is planned and implemented in UAV GCS software for a part of an industrial city in south of Tehran, more detailed information is presented in Table (2).

Visible camera		
Sensor	CMOS	
Pixel size	4.04 μm	
Image Size	6000*4000	
Focal length	35mm	
Thermal camera		
Sensor	Uncooled FPA	
Pixel size	17 μm	
Video Size	640*480	
Focal length	25 mm	

Table 1. Applied sensors technical specification.

Camera	Visible	TIR	
Fight height (m)	200	200	
Overlap (percent)	70	70	
Sidelap (percent)	70	70	
Images	205	6000 frame	

Table 2. Flight planning parameters.

4.2 Pre-Process results

As previously stated, two-dimensional target plane is used for geometric calibration of camera. In the case of geometric calibration, in addition to the interior orientation parameters, the mathematical model between image and target coordinate system are also calculated. Now, having these parameters and coordinates of the control points, we can calculate the accuracy of the proposed calibration algorithm. (Fig. 3)

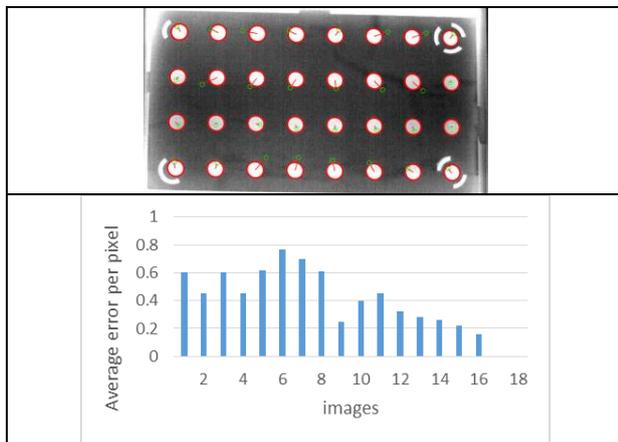


Figure 3. Points Error with a magnification of 20x (Top) and Re-projection errors of images (Bottom)

In key frame selection step, frames with low quality are eliminated from video sequence by applying the BluM metric threshold. In order to select key frame geometrically from thermal video sequence, Seo key frame selection algorithm is used. The result of key frame selection steps is shown in Table (3).

4.3 Dense Cloud, DEM and Orthomosaic Generation

A dense point cloud was produced from the given set of images and their orientation parameters, using Semi-Global Matching (SGM) technique, where TIR point cloud has 28 and the visible one has 630 points per square meter. Then DSM has been generated with this source by interpolation. The resolution of DSMs is 20 cm and 4 cm for TIR and visible data respectively. Figure 4 is illustrating DSM and Orthomosaic generated from mentioned case study.

Total captured frames	6000
Removed by BluM	40
Removed by Seo	5116
Key frames	866

Table 3. Results of key frame selection.

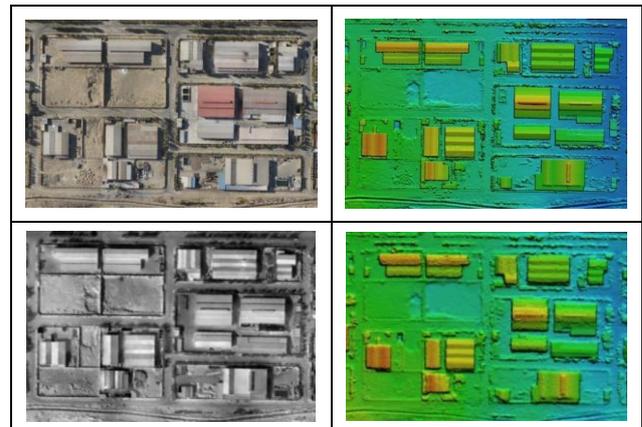


Figure 4. Orthomosaic and DEM of visible (Top) and thermal images (Bottom).

4.4 Assessment of Registration

4.4.1 Georeferencing: Thermal 3D point cloud was built in the local coordinate, for transferring into a global coordinate system and to remove scale ambiguity, the similarity transformation using three Ground Control Points (GCPs) was used. This coordinates are measured independently from visible model which its images are georeferenced by PPK GPS module.

4.4.2 Fine Registration: After georeferencing when two point clouds have the same scale, the ICP algorithm is fulfilled and region based evaluation is done after that (See Fig. 5).

4.1 Integration of TIR and Visible

The output of proposed method is a dense point cloud with both RGB and thermal information (i.e. temperature). This point cloud was produced by a linear weighted interpolation method which assign thermal information to visible points. So the geometric accuracy and point density of that is same as visible point cloud, furthermore the textured mesh generated by that is sharper and has more faces. Figure 6 shows triangulated and textured mesh produced by proposed method. This 3D model is produced by accurate registration of thermal point cloud by visible one. The accuracy of registration is validated by check points and region based differences between them.

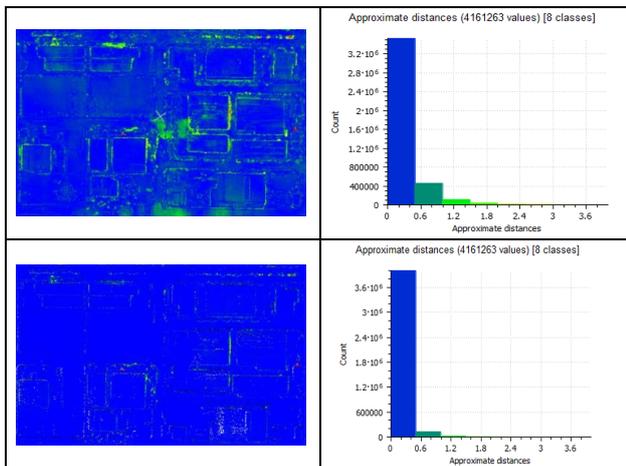


Figure 5. Error map which shows vertical distance between TIR and Visible point clouds after course (Top) and fine registration (Bottom). (Units are in meters)

The average error of difference between them is 10, 12 and 20 in three axes. And the output point cloud density is 630 points per square meter.

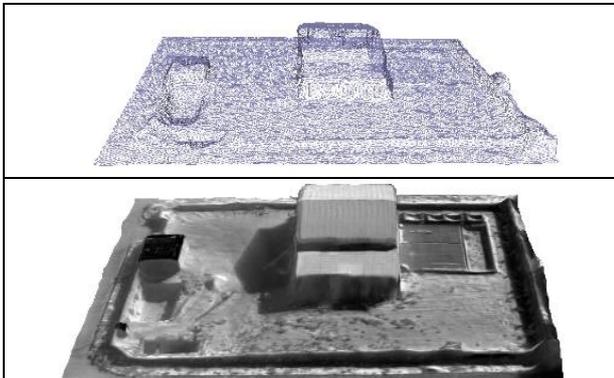


Figure 6. Triangulated and textured mesh produced by proposed method

5. CONCLUSION

The problem which is considered in this study is the low resolution of thermal images which leads to generate low density thermal point cloud. Due to literature, 3D models which produced by thermal imagery has an acceptable geometric accuracy but in this study we offered a method to enhance geometrical accuracy by registering that in object space and increasing the density of TIR point cloud by integration with a precise visible model generated by high resolution georeferenced visible images, thanks to GPS PPK technology. After precise calibration of thermal camera with 0.4 average reprojection error, and key frame selection, the registration is done by 10, 12 and 30 cm average differences between two point clouds, and finally the integration procedure was done successfully. As a result, the output of proposed method is a dense and accurate point cloud which has thermal information with minimum distortion and disparity rather to reference high accuracy visible point cloud. The number of points in this point cloud has increased by approximately 20 times and the density of that is averagely 630 points per square meter, where, the primary thermal model density was averagely 28 points per square meter. This study showed the potential of thermal

imagery in 3D modelling by integration with visible imagery without using GCPs.

REFERENCES

- Ahmed, M. (2009). Key frame extraction and browser-based visualization for 3d reconstruction from video streams. Universitat Politècnica de Catalunya (UPC) & Asian Institute of Technology (AIT).
- Borrmann, D., Elseberg, J., & Nuchter, A. (2012). Thermal 3d mapping of building facades. 8th Conference on Intelligent Autonomous Systems. Jeju Island, Korea.
- Boyd, G. (2013). Zeroing in on energy savings with thermal imaging. *Photogrammetric Engineering & Remote Sensing*, 79(4), 313-316.
- Brown, D. C. (1971). Close-Range Camera Calibration. *Photogrammetric Engineering*, 855-866.
- Brumana, R., Oreni, D., Cuca, B., Binda, L., Condoleo, P., & Triggiani, M. (2013). Strategy for integrated surveying techniques finalized to interpretive models in a byzantine church, Mesopotam, Albania. *International Journal of Architectural Heritage*.
- Cabrelles, M., Galcerá, S., Navarro, S., Lerma, J. L., Akasheh, T., & Haddad, N. (2009). INTEGRATION OF 3D LASER SCANNING, PHOTOGRAMMETRY AND THERMOGRAPHY TO RECORD ARCHITECTURAL MONUMENTS. 22nd CIPA Symposium. Kyoto, Japan.
- Cabrelles, M., Galcera, S., Navarro, S., Lerma, J. L., Akasheh, T., & N., H. (2009). Integration of 3d laser scanning, photogrammetry and thermography to record architectural monuments. 22nd CIPA Symposium. Tokyo, Japan.
- Coiras, E. (2000). Segment-based registration technique for visual-infrared images. *Optical Engineering*, 39(1), 282-289.
- Crete, F., Dolmiere, T., Ladret, P., & Nicolas, M. (2007). The blur effect: perception and estimation with a new no-reference perceptual blur metric. *Electronic Imaging*.
- Cryderman, C., Mah, S., & Shufletoski, A. (2014). Evaluation of UAV photogrammetric accuracy for mapping and earthworks computations. *Geomatica*, 68, 309-317.
- Fischler, M. A., & Bolles, R. C. (1981). random sample consensus A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Graphics and Image Processing*, 24(6), 381-395.
- Gibson, S., Cook, J., Howard, T., Hubbard, R., & Oram, D. (2002). Accurate camera calibration for off-line, video-based augmented reality. 1st International Symposium on Mixed and Augmented Reality.
- Hirschmuller, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego.
- Hirschmüller, H. (2011). Semi-Global Matching Motivation, Developments and Applications. in *Wichmann/VDE Verlag, Belin*.
- Hoegner, L., & Stilla, U. (2016). Automatic 3D reconstruction and texture extraction for 3D building models from thermal infrared image sequences. 13th International Conference on Quantitative Infrared Thermography. Munich, Germany.
- 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*, 459-468.
- Hoegner, L., Tuttas, S., Xu, Y., Eder, K., & Stilla, U. (2016). EVALUATION OF METHODS FOR COREGISTRATION AND FUSION OF RPAS-BASED 3D POINT CLOUDS AND THERMAL INFRARED IMAGES. *The International Archives*

- of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- Iwaszczuk, D., & Stilla, U. (2017). Camera pose refinement by matching uncertain 3D building models with thermal infrared image sequences for high quality texture extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 33-47.
- Iwaszczuk, D., Hoegner, L., & Stilla, U. (2011). Matching of 3d building models with ir images for texture extraction. *JURSE 2011 - Joint Urban Remote Sensing Event*, (pp. 25-28).
- Javidi, N., Arofteh, A. R., Samadzadegan, F., & Dadras Javan, F. (2017). Wide-Angle Lens Distortion Correction in 3D Key Frame Selection. *ISPRS*.
- Johnson, M., & Bajcsy, P. (2008). Integration of thermal and visible imagery for robust foreground detection in tele-immersive spaces. *11th International Conference on Information Fusion*. Cologne, Germany.
- Knorr, S., Imre, E., Alatan, A., & Sikora, T. (2006). A geometric segmentation approach for the 3d reconstruction of dynamic scenes in 2d video sequences. *Signal Processing Conference*. 14th European. Florence. Italy.
- Lagüela, S., González-Jorge, H., Armesto, J., & Arias, P. (2011). Calibration and verification of thermographic cameras for geometric measurements. *Inf. Phys. Technol.*, 54, 92-99.
- Lelong, C., Burger, P., Jubelin, G., Roux, B., Labbé, S., & Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors*, 3557-3585.
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*.
- Mangus, D. L., Sharda, A., & Zhang, N. (2016). Development and evaluation of thermal infrared imaging system for high spatial and temporal resolution crop water stress monitoring of corn within a greenhouse. *Computers and Electronics in Agriculture*, 149-159.
- Maset, E., Fusiello, A., Crosilla, F., Toldo, R., & Zorzetto, D. (2017). *PHOTOGAMMETRIC 3D BUILDING RECONSTRUCTION FROM THERMAL IMAGES*. *ISPRS*.
- Mikeleit, V., & Kraneis, D. (2010). Planen in funf dimensionen. *GIS Business*, 58-59.
- Ouellet, J.-N., & Hébert, P. (2009). Precise ellipse estimation without contour point extraction. *Machine Vision and Applications*, 21, 59-67.
- Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J., & Koch, R. (2004). Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59 (3), 207-232.
- Qin, R. (2014). An object-based hierarchical method for change detection using unmanned aerial vehicle images. *Remote Sensing*, 6, 7911-7932.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). AV photogrammetry for mapping and 3d modeling—current status and future perspectives. *International Conference on Unmanned Aerial Vehicle in Geomatics*. Zurich, Switzerland, .
- Seo, J., Hwang, Y., & Hong, H. (2004). Structure and Motion Recovery Using Two Step Sampling for 3D Match Move. *Springer-Verlag Berlin Heidelberg. MICAI*, 652-661.
- Seo, Y., Kim, S., Doo, K., & Choi, J. (2008). Optimal key frame selection algorithm for three-dimensional reconstruction in un calibrated multiple images. *Journal of the Society of Photo-Optical Instrumentation Engineers*, 47(5), 53201-53400.
- Siebert, S., & Teizer, J. (2014). Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle. *Automation in Construction*, 41, 1-14.
- Snively, S. (2006). Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics*.
- Teza, G. (2015). Integration of laser scanning and thermal imaging in monitoring optimization and assessment of rockfall hazard: a case history in the Carnic Alps (Northeastern Italy). *Natural Hazards*, 76(3), 1535-1549.
- Torr, P. H. (1998). Geometric motion segmentation and model selection. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1321-1340.
- Truong, B., & Venkatesh, S. (2007). Video abstraction: A systematic review and classification. *ACM transactions on multimedia computing, communications, and applications (TOMM)*.
- Tsanakas, J. A., Ha, L. D., & Al Shakarchi, F. (2017). Advanced inspection of photovoltaic installations by aerial triangulation and terrestrial georeferencing of thermal/visual imagery. *Renewable Energy*, 102, 224-233.
- Ullman, S. (1979). *The Interpretation of Structure from Motion*. The Royal Society, 405-426.
- Vidas, S., Moghadam, P., & Bosse, M. (2013). 3D Thermal Mapping of Building Interiors using an RGB-D and Thermal Camera. *IEEE International Conference on Robotics and Automation*.
- Zhang, Y., Xiong, J., & Hao, L. (2011). Photogrammetric processing of low-altitude images acquired by unpiloted aerial vehicles. *Photogrammetric Record*, 26, 190-211.
- Zhang, Z. (1999). Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. *IEEE*.
- Zhou, G. (2009). Near real-time orthorectification and mosaic of small UAV video flow for time-critical event response. *Remote Sensing*, 47, 739-747.
- Zhou, G. (2010). Geo-referencing of video flow from small low-cost civilian UAV. *AUTOMATION SCIENCE AND ENGINEERING* (pp. 456-466). *IEEE Transactions*.