

Article

Four-Band Thermal Mosaicking: A New Method to Process Infrared Thermal Imagery of Urban Landscapes from UAV Flights

Yichen Yang ^{1,2}  and Xuhui Lee ^{2,1,*}

¹ Yale-NUIST Center on Atmospheric Environment, Nanjing University of Information Science and Technology, Nanjing 210044, China; yichen.yang@yale.edu

² Yale School of Forestry & Environmental Studies, Yale University, 195 Prospect Street, New Haven, CT 06511, USA

* Correspondence: xuhui.lee@yale.edu; Tel.: +1-203-432-6271

Received: 21 April 2019; Accepted: 4 June 2019; Published: 6 June 2019



Abstract: Unmanned aerial vehicles (UAVs) support a large array of technological applications and scientific studies due to their ability to collect high-resolution image data. The processing of UAV data requires the use of mosaicking technology, such as structure-from-motion, which combines multiple photos to form a single image mosaic and to construct a 3-D digital model of the measurement target. However, the mosaicking of thermal images is challenging due to low lens resolution and weak contrast in the single thermal band. In this study, a novel method, referred to as four-band thermal mosaicking (FTM), was developed in order to process thermal images. The method stacks the thermal band obtained by a thermal camera onto the RGB bands acquired on the same flight by an RGB camera and mosaics the four bands simultaneously. An object-based calibration method is then used to eliminate inter-band positional errors. A UAV flight over a natural park was carried out in order to test the method. The results demonstrated that with the assistance of the high-resolution RGB bands, the method enabled successful and efficient thermal mosaicking. Transect analysis revealed an inter-band accuracy of 0.39 m or 0.68 times the ground pixel size of the thermal camera. A cluster analysis validated that the thermal mosaic captured the expected contrast of thermal properties between different surfaces within the scene.

Keywords: unmanned aerial vehicle; structure-from-motion; four-band thermal mosaicking; positional error; object-based calibration; transect analysis; cluster analysis

1. Introduction

Unmanned aerial vehicles (UAVs) have become a reliable observation platform for environmental remote sensing applications, including wildfire mapping [1,2], atmospheric studies [3–5], precision agriculture [6–8] and plant identification [9–11]. Compared to the traditional remote sensing platforms such as satellites and manned aircrafts, which have low spatiotemporal resolutions and high operational costs [12], light-weight UAVs are less costly and more flexible. In spite of some limitations related to weather tolerance, maximum payload, civil aviation regulation, and traveling distance [13], UAVs can collect high temporal resolution data at ultra-high spatial resolutions [14] and are versatile due to their ability to change speed, direction and elevation as required [15]. Benefitting from the advantages of the UAV platform, UAV photogrammetry and remote sensing have opened new opportunities for landscape visualization and analysis. Lucieer et al. mapped the landslide displacement using high-resolution and multi-temporal UAV photos [16]. Techniques of 3-D modelling, like structure-from-motion (SfM) [17], were combined to process the acquired UAV photos and to construct the digital elevation model (DEM). The landscape of an urban area can be

investigated in detail with the acquisition of multi-band UAV imagery [18], which can yield more valuable information than satellite imagery.

The individual photos collected by a UAV camera should be mosaicked for remote sensing applications. Image mosaicking is a technique that combines multiple image patches to form a larger image [19]. To date, it has been integrated with other sophisticated spatial analysis and detection technologies, such as undersea exploration [20,21], high-resolution satellite imaging [22], fingerprint registration [23], 3-D visualization [19,24], and microscopic observation [25]. The past two decades have witnessed the rapid growth of ‘structure-from-motion’ (SfM), a revolutionary, low-cost, user-friendly photogrammetric technique for image mosaicking [26]. SfM extracts common features in different photos using the scale invariant feature transform (SIFT) algorithm. Then, with the help of a bundler algorithm [27], a sparse point cloud is generated, which is subsequently densified by the clustering view for multi-view stereo (CMVS) algorithm [28]. The patch-based multi-view stereo (PMVS2) algorithm [28] is finally used to reconstruct a 3-D model of the measurement target. One important advantage of SfM over other methods is its high automation. The whole process of SfM can be completed without prior knowledge of camera pose and its 3-D location [26]. SfM has become increasingly user-friendly since the release of SfM-based commercial software, such as Agisoft Photoscan (<https://www.agisoft.com>) and Pix4D (<https://www.pix4d.com>). As part of the workflow in these software products, multiple photos are stitched together to form an orthographically projected map called orthomosaic.

The quality of an orthomosaic can be improved by other effective algorithms, such as speeded up robust features (SURF), oriented FAST and rotated BRIEF (ORB), and robust features along the edge (RFAE) [29]. Meanwhile, as narrow-band multispectral cameras (e.g., mini-MCA6) and thermal cameras (e.g., FLIR Vue Pro) are increasingly used on board UAVs, image mosaicking is no longer limited to RGB photos. Thermal maps have important merits in different fields of science. For example, thermal maps can be combined with geomorphologic data to monitor potential volcanic unrest in a 3-D view [30], or used to reconstruct the thermal property of plant canopies [31]. Other applications include assisting the model calculation of surface energy balance [32] and surveying land use and local climate [33–35].

So far, we lack efficient procedures for processing thermal images. Often, thermal images are used as individual frames and without the clear characterization of positional accuracy. Mosaicking UAV thermal infrared (TIR) imagery has usually been not as successful as RGB and multispectral imagery for the following reasons:

1. As opposed to RGB cameras and multispectral cameras, thermal cameras available on the market usually have low-resolution lenses [36]. The resulting coarse-grid raster imagery is poor in texture and poses a challenge for key-point identification. The chance for SfM workflow to fail at sparse point generation, the first step of the workflow, is high [37].
2. RGB and multispectral cameras both provide multi-band images. With the assistance of multi-band contrast, there is a high tolerance of failed point match. However, most of the thermal data are single-band images, which lowers the probability of finding matched points correctly.
3. To compensate for the downsides mentioned above, UAV flights to collect thermal images are usually carried out separately at a lower elevation and at a slower speed compared to the flight for other bands [36,38]. The separated flight enables a closer watch on the objects, increases photo overlaps, decreases the pixel size and improves image resolution. In addition, auxiliary data such as camera global positioning system (GPS) data and ground control points (GCPs) are usually more indispensable during thermal mosaicking than RGB mosaicking [18,39,40]. However, planning a separated flight mission and setting GCP boards on the ground [31,38] beforehand are both labor intensive, not to mention that the thermal band will have a smaller sampled area compared to the RGB band. Moreover, not all types of thermal camera have built-in GPS data, and some camera types are incompatible with the UAV to synchronize its GPS data. Even though the thermal mosaicking without geotag can be successful, it is still challenging to

register, with sufficient positional accuracy, the thermal mosaic to the RGB imagery of the same target for proper interpretation of the former [29].

Motivated by these challenges and the importance of thermal mosaicking, this work aims to design a novel method, referred to as the four-band thermal mosaicking (FTM), in order to process thermal images acquired by a UAV, with three goals in mind:

1. The method overcomes the difficulty of mosaicking low-resolution, single-band thermal imagery. The flight preparation and the mosaicking process in the SfM-based applications are no more complicated than those used for RGB photos. The method does not require either onboard GPS or GCP data.
2. The final product, the thermal orthomosaic, can be easily registered to the RGB orthomosaic of the same target. There is no loss of the sampled area in the thermal orthomosaic, and the method allows pixel-by-pixel analyses between the thermal and the RGB bands.
3. The method provides a simple and robust way to establish relative positional errors and to validate the temperature map.

2. Materials and Methods

2.1. An Overview of Four-band Thermal Mosaicking (FTM)

Inspired by [14] which shows how multi-band images are mosaicked simultaneously, we seek to combine each thermal photo with its synchronized RGB photo in order to form a four-band image. The four-band images are then imported to the Pix4D software for processing. In the SfM algorithm, the “tie points” are identified using only the RGB bands; no weight is given to the thermal band. To generate the thermal orthomosaic, the software simply combines and projects the thermal band in the same way as the RGB orthomosaic is projected from the tie points. The final four-band large orthomosaic is saved in the TIFF format. The band sequence is red, green, blue and thermal infrared. All the bands have the same pixel numbers and the same pseudo resolution.

2.2. Study Area and Instruments

2.2.1. Study Area

Our study area is Beaver Pond Park (Figure 1; center coordinate 41.33 N, 72.94 W), in the city of New Haven, Connecticut, United States. The study area covered by the UAV observation has multiple surface types, including trees, grassland, bare soil, water body, and asphalt road. The contrasting surface characteristics are helpful to recognize prominent features on the temperature map due to large temperature differences.



Figure 1. The satellite map of Beaver Pond Park, New Haven, Connecticut, United States, with an inset map indicating its location. The green box indicates the approximate area covered by the UAV observation.

2.2.2. Instruments

We used a DJI Phantom 4 Pro Quadcopter (Figure 2a, Table 1) as our UAV platform. Its built-in RGB camera was not used; instead a FLIR DUO R dual-sensor RGB/thermal camera (Figure 2b, Table 2) was attached to the UAV to acquire photos of the landscape. During the flight over the target area (the green box in Figure 1), the FLIR camera took photos in the nadir direction, in its burst mode allowing consecutive capture of synchronized RGB and thermal images at one pair of photos per second. The RGB photos were saved in the 8-bit JPG format and the thermal images in the 16-bit TIFF format. The resolution of the visible lens (1920 by 1080 pixels) was much higher than that of the thermal lens (160 by 120 pixels). The FLIR camera was not embedded with a GPS sensor and no GCPs were set in advance in the study area.



(a)



(b)

Figure 2. Photos of instruments: (a) is DJI Phantom 4 Quadcopter and (b) is the FLIR DUO R dual-sensor thermal camera.

Table 1. Technical specifications of DJI Phantom 4 Pro Quadcopter.

Property	Parameters
Weight	1388 g
Diagonal Size	350 mm
Max Speed	45 mph (Sport mode); 36 mph (Altitude mode); 31 mph (GPS mode)
Max Flight Time	Approx. 30 minutes
Max Service Ceiling Above Sea Level	6000 m
Operating Temperature Range	0 to 40 °C
Satellite Positioning Systems	GPS/GLONASS
Hover Accuracy Range (with GPS Positioning)	Vertical: ± 0.5 m, Horizontal: ± 1.5 m

Table 2. Technical specifications of FLIR DUO R dual-sensor imager.

Property	Parameters
Dimensions	41 × 59 × 29.6 mm
Weight	84 grams
Spectral Band (thermal)	7.5–13.5 μm
Thermal Frame Rate	7.5 Hz (NTSC); 8.3 Hz (PAL)
Thermal Imager	Uncooled Vox Microbolometer
Thermal Measurement Accuracy	+/-5 °C
Thermal Sensor Resolution	160 × 120
Visible Camera Resolution	1920 × 1080

2.2.3. Image Acquisition

The flight route over Beaver Pond Park was set up using DJI GS PRO, the DJI App for mission planning. The flight mission was carried out at about 10:00 AM on 31 October 2018, a clear day. Air temperature and relative humidity at the time of the flight were about 8 °C and 60%, respectively. These atmospheric parameters and the sky condition were set to the camera prior to the flight in order to correct the effect of atmospheric absorption. The flying height was 85 m and the whole flight took about 10 minutes. Four swaths of consecutive photos were taken in the NW-SE direction in parallel to the long dimension of the rectangular target area shown in Figure 1, with the front and side overlaps of 85% and 75% for RGB photos and of 85% and 50% for thermal photos, respectively. A total of 440 pairs of RGB and thermal photos were selected after manually filtering out some blurry RGB photos.

2.3. The Workflow of Four-band Thermal Mosaicking

2.3.1. Pre-processing: Creating Four-band Images

Figure 3 includes the workflow of image pre-processing. The thermal lens has a narrower Field of View (FOV = 57°) than the visible lens (FOV = 72°) in the horizontal directions, and has exactly the same FOV in vertical direction (44°). Because the RGB lens has a vertical resolution (1080) that is a multiple of that the thermal does (120), it is possible to up-sample the coarser thermal image by duplicating each of the original pixel to a 9 by 9 matrix of smaller pixels with identical Digital Number (DN) value. After the up-sampling, the thermal photo has the same number of rows as the RGB photo. The next step is cropping the visible photo to the same size as the thermal photo by removing an edge of 240 by 1080 pixels from each side. The side overlap was decreased to 50% because of cropping. After resizing, the thermal band can be stacked onto the visible band to create a four-band TIFF image with the sequence of red, green, blue, and thermal infrared. This batch process can be done in MATLAB (MATLAB and Statistics Toolbox Release 2017b, The MathWorks, Inc., Natick, Massachusetts, United States).

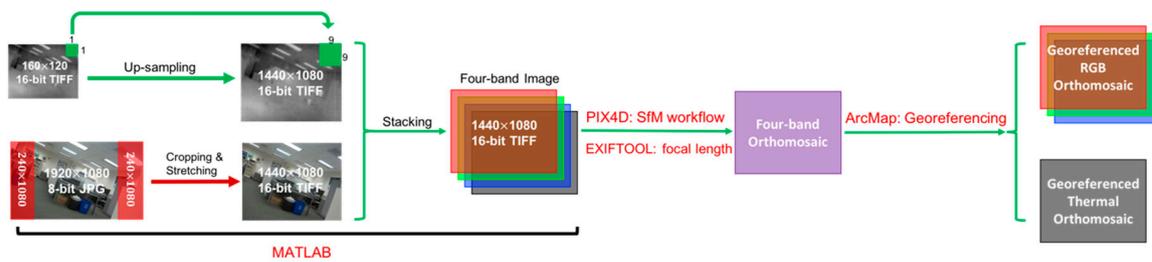


Figure 3. The workflow of pre-processing the UAV images, mosaicking, and post-processing the orthomosaics. A sample of office scene is used to demonstrate how the thermal image (in gray scale) is up-sampled and then stacked onto its corresponding RGB image, the latter of which have been cropped and radiometrically stretched. The four-band images are mosaicked by SfM workflow in PIX4D. The generated four-band orthomosaic is georeferenced by Esri’s Arcmap. The labels in red indicate the software required in the workflow.

Before stacking, the RGB bands in the 8-bit JPG format are converted to the 16-bit “TIFF” format. This conversion is necessary because the DN value in the 8-bit range (0–255) is too small compared with the DN value in the 16-bit range (0–65535). However, a complete linear stretching from 255 to 65535 would sometimes lead to failure of tie point generation in the SfM workflow. Even if the workflow was able to complete, the final RGB orthomosaic would be blurry. One possible reason for this is that linear stretching increases the minimum possible difference between DN values, and by doing so the contrast between similar features is overly magnified, making key-point detection harder. The end result is that more images are incorrectly registered.

A good solution is to stretch the DN value of the RGB bands by multiplying with a specific integer. After several trial-and-error attempts, we have chosen a multiplier of 10, bringing the range of the stretched DN values to a similar magnitude of the 16-bit range. We have found that if a larger multiplier is used, the processing time will become exceedingly long, but without obvious quality improvement.

2.3.2. Mosaicking and Post-processing

After importing the four-band images into Pix4Dmapper (version 4.3.31; <https://www.pix4d.com>), we need to allocate the band weight when editing the camera mode in the image properties editor (Table 3). The thermal band (IR) should have no weight, and via tests we determined that the green band should have the most weight. It is also important that the focal length be set to 34.8 mm, the focal length of the thermal lens. Since the four-band image has been cropped to the same FOV as the thermal image, the choice of this focal length is appropriate. This focal length has been confirmed via the EXIFTOOL [41].

Table 3. The band weight allocation in PIX4D Image Properties Editor.

Band Name	Central Wave Length [nm]	Weight
Red	660.0	0.2126
Green	550.0	0.7152
Blue	470.0	0.0722
IR	1000.0	0.0000

What follow next are the normal steps of the SfM algorithm: finding sparse point clouds, generating dense point clouds, building mesh and digital surface model (DSM), and producing orthomosaic. The whole SfM workflow in Pix4D took about 48 minutes. The four-band orthomosaic is later split to an RGB orthomosaic and a thermal orthomosaic via MATLAB. After splitting, the RGB band DN values are converted back to the 8-bit radiometric resolution, by dividing with the factor 10. Because of the rich texture of the RGB orthomosaic, distinct features are easily detected. These features are used as control points to register the UAV orthomosaics on the World Imagery base map in Esri’s ArcMap.

By selecting control points (a total of 8 in the present case) in the RGB orthomosaic, the thermal orthomosaic is georeferenced along the way (Figure 3).

The pixel DN value of the thermal band is proportional to the flux of infrared radiation received by the thermal lens. The DN value can be converted to temperature on the principle of the Stefan-Boltzmann Law [42].

3. Results

3.1. Image Orthomosaics

The resulting orthomosaics show that the surveyed area covers 12.18 hectares. The pixel size is 6.36 cm. Comparison of the RGB and the thermal orthomosaic in Figure 4 reveals strong contrasts between difference surfaces. For example, the pond water was relatively warm because the large specific heat capacity of liquid water helped it maintain high temperature in the morning. A garage with metal roof, identifiable on the RGB orthomosaic, showed an extremely low thermal signature (DN value of 2250) due to the low-emissivity of the roof. Trees were cooler than the water surface because of insufficient heating by solar radiation in the morning. A more detailed comparison of thermal properties is given in Section 3.5.

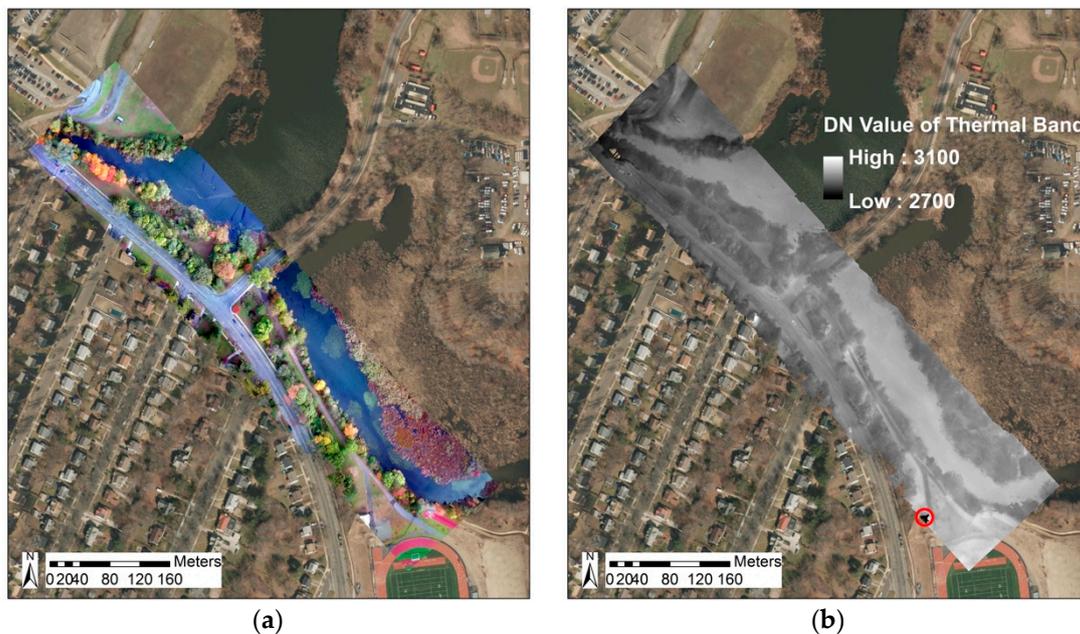


Figure 4. The georeferenced (a) RGB and (b) thermal orthomosaics of Beaver Pond Park. In the right panel, the legend is the main range of DN values of the thermal band. The red circle on the thermal orthomosaic indicates a black spot with an extremely low DN value (2250) due to the low-emissivity of a metal roof on a garage (identifiable on the RGB).

This preliminary result has already achieved our first goal: the FTM process is no more complicated than the standard RGB mosaicking in the SfM workflow. In the FTM process, the thermal mosaicking is simultaneously accomplished when the corresponding RGB images are stitched the classic way. No geographic information is used in this process. Instead, the orthomosaic is georeferenced using a simple ArcMap georeferencing tool as a post-mosaic step. As for our second goal, the thermal orthomosaic is generated from the same point cloud of the RGB orthomosaic, therefore having an identical sampled area.

3.2. Positional Error Assessment

Inter-band registration is a process that aligns the same features in difference bands accurately. Before the registration, it is important to understand the reasons for misalignment, or in other words, the sources of positional error.

One error type, the lens spacing error, is related to the separation between the visible lens and the thermal lens on the camera body (Figure 2b). This type of error can be determined accurately. Figure 5a explains how the objects at different heights in the camera FOV are projected to the background, which is the ground surface in the present study, to form a 2-D image. As for the object right on the ground surface (the green circle), the sight lines (green arrows) of the visible and the thermal lenses intersect on the ground, so the pixels of the object are projected to have the same geolocation in the RGB and the thermal images. However, the sight lines (red arrows) of an object above the ground (the red circle) intersect aloft, and the object projected in the thermal image occurs to the right of its projection in the RGB image. This misalignment will increase as the object being observed becomes higher above the ground. In Figure 5b, we have:

$$d_1 = \frac{H}{D} \times d, \quad (1)$$

where d is the lens spacing and d_1 is the misalignment between the two projections, D is the vertical distance between the on-board camera and the object being observed, and H is the height of the object above the ground.

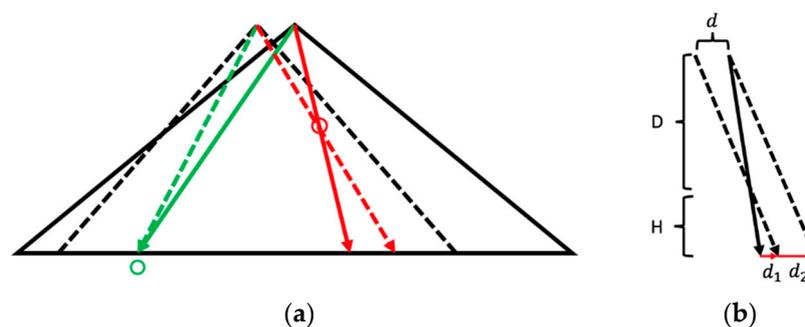


Figure 5. The diagram showing the lens spacing error. In (a), the solid triangle and the dashed triangle represent the FOV of the visible and the thermal lens, respectively. The thermal lens FOV is narrower and is to the left of the visible lens FOV when the camera is in forward motion. For the object located on the ground (green circle), the sight lines of the visible lens (solid green arrow) and the thermal lens (dashed green arrow) both point to the object on the ground (green circle), and for the object above the ground (red circle), the intersection point of the sight lines (solid red and dashed red arrows) is elevated. In (b), the solid and the dash arrows are the sight lines of the visible and the thermal lenses, respectively, where d is lens spacing, D is distance between the on-board camera and the object being observed, H is the height of the object above the ground, and d_1 and d_2 refer to two different contributors of Lens Spacing Error.

During the band stacking, we align the RGB and thermal images along their central lines instead of offsetting the thermal images appropriately by the distance of lens spacing. Therefore, the lens spacing d itself also becomes part of the lens spacing error (Equation (2)), marked as d_2 in Figure 5b, where

$$d_2 = d, \quad (2)$$

and the total lens spacing error E_d is given by Equation (3):

$$E_d = d_1 + d_2, \quad (3)$$

For our flight over the Beaver Pond Park, the tallest objects are the trees, which are no more than 15 m high. Given the H of 15 m, the flying height $H + D = 85$ m, and the lens separation d of 2.8 cm, the lens spacing error E_d is about 3.4 cm, which is smaller than the size of one RGB pixel (6.36 cm). Therefore, no correction is made for this type of positional error.

Another type of positional error is caused by the slight shutter delay between the visible and the thermal lenses (Supplementary Figure S2). This error depends on the moving direction and speed of the UAV, and in theory can be corrected for each swath and flight level. But such a swath-by-swath procedure is labor intensive. The better way of correction is to remove the error for the whole orthomosaic, as discussed below.

3.3. Object-Based Calibration

3.3.1. Quantifying and Correcting the Misalignment

We designed an object-based method to remove the positional error between the RGB and the thermal bands. Unlike other object-based algorithms [14,29], which register each of the individual photos prior to mosaicking, this method uses objects classified from the orthomosaic (Figure 6c). In total, 20 objects were analyzed. Two of the objects are shown in Figure 6a,b. Each object is characterized by a misalignment vector between the thermal band and the RGB bands. These 20 vectors are plotted in Figure 6d. In each case, the vector base is the object identified in the RGB mosaic and is anchored at the origin, and the vector tip is the object identified in the thermal mosaic. The vector plot helps us to determine the total misalignment, i.e., the length of the vector, as well as orientation of the misalignment. By using objects scattered across the whole region, we can determine whether different parts of the thermal orthomosaic have a generally consistent misalignment. Since these vectors have a similar length and orientation, it is acceptable to average them in order to obtain a “mean vector”, whose two orthogonal components, in pixel number, are $x = 28.20$ and $y = -5.55$. The magnitude of the mean vector is 28.74, representing the mean positional error in pixel number or an actual distance of 182.79 cm. We then corrected the misalignment by offsetting the thermal orthomosaic 28 pixels leftward and six pixels upward. The corrected thermal orthomosaic should have a much reduced positional error relative to the RGB orthomosaic.

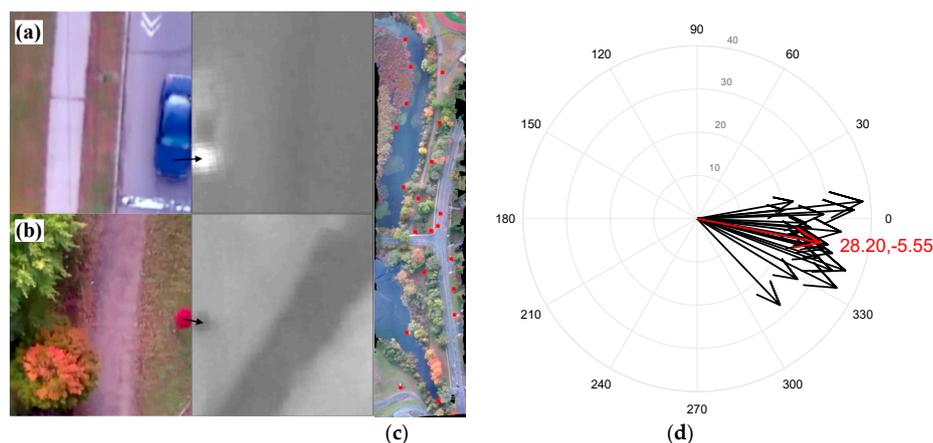


Figure 6. Removal of misalignment between the RGB and the thermal orthomosaics. (a,b): two typical views clipped from the orthomosaics, with two recognizable objects [a hot car engine in (a) and a cold red tree in (b)]. The left panel shows the RGB view and the right panel shows the thermal view. The black arrow is the vector of misalignment. (d): 20 misalignment vectors (black arrows) sampled from (c) and the mean misalignment vector (red arrow). The two red numbers are components of the mean vector.

3.3.2. Validating the Object-Based Calibration

To validate our calibrating method, both visual comparison and quantitative analysis are necessary. Figure 7 is an example showing an area in the RGB mosaic with a portal in the center showing the linked view of the same area in the thermal mosaic, before (panel a) and after calibration (panel b). A segment of gravel road is clearly recognizable in the thermal mosaic because it had a relatively low temperature. Comparing the road boundary inside and outside of the portal, it is clear that the thermal mosaic is registered much better after calibration. The same conclusion is reached using transect analysis. Figure 7c,d are variations of the DN value in thermal, green and blue bands along the horizontal transects marked in Figure 7a,b, respectively. The alignment in Figure 7d is better than in Figure 7c, because two black lines marking the two edges of the road in the Blue and Thermal bands becomes more vertical. (The road edges are not recognizable in the Green band.)

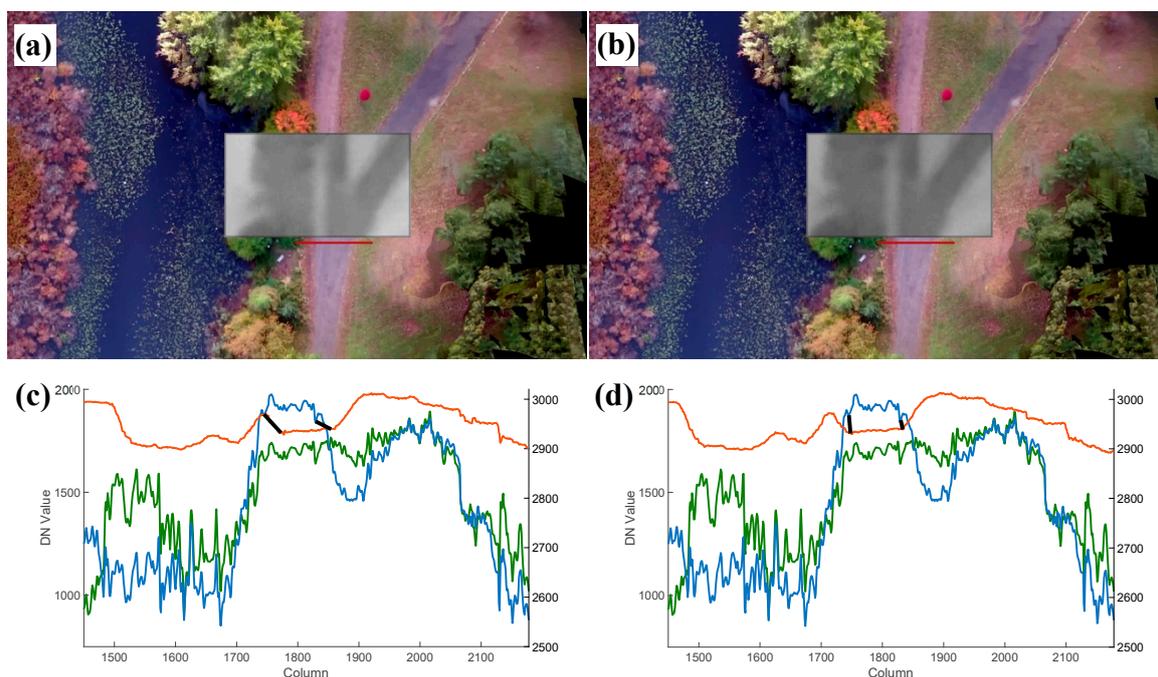


Figure 7. An example showing removal of image misalignment. (a,b): clipped views of the same area before and after calibration, respectively. (c,d): band profiles along the transects marked in (a) and (b), with the Infrared band in red corresponding with the right y-axis, the Green band in green, and the Blue band in blue, both corresponding with the left y-axis. The black lines mark the road edges.

The alignment accuracy is further quantified with additional transects. We identified 10 transects for each of the four directions, including horizontal, vertical, main-diagonal (top-left to bottom right), and anti-diagonal (top-right to bottom-left). (Figure 7 is an example of horizontal transect). Each transect passes through an object whose temperature is visibly different from its surrounding. The offset (in pixel numbers) between the edge of the object in the thermal mosaic and in the RGB mosaic is the residual positional error. The root mean square (RMS) error for each direction is given in Table 4. The overall mean error is 6.12 (38.92 cm), which is much better than the mean error before calibration (21.64, 137.63 cm). The RMS errors are in the range of 5.09 to 7.13. The largest misalignment is 7.13 pixels, or 45.35 cm in the physical dimension, which is smaller than 57.24 cm, the size of the original raw thermal pixel before up-sampling. Furthermore, the error patterns do not show any obvious directional dependence, supporting the use of a mean vector to align the two mosaics.

Table 4. Misalignment error in pixel number in four directions after calibration.

	Horizontal	Vertical	Main-diagonal	Anti-diagonal	Mean
R1	3	0	5	1	\
R2	1	4	8	-5	\
R3	4	6	6	-10	\
R4	1	1	0	-11	\
R5	7	-14	1	-2	\
R6	9	4	1	5	\
R7	4	12	10	10	\
R8	-1	0	0	0	\
R9	6	-8	-8	-7	\
R10	7	-6	5	-4	\
RMS	5.09	7.13	5.62	6.64	6.12

3.4. Radiometric Calibration

After the calibration procedure, the thermal orthomosaic has acceptable positional error. Thus we can inspect the temperature distribution of different objects directly using their positions in the RGB orthomosaic. Prior to this pixel-by-pixel analysis, all the DN values in the thermal band should be converted to actual temperature values.

The temperature conversion formula was obtained with measurements made simultaneously with an infrared thermometer (Figure 8a; working wavelengths 8–14 μm) and the thermal camera (Figure 2b; working wavelengths 7.5–13.5 μm). The emissivity of the infrared thermometer was set to 0.95. The measurements were carried out at 14:00 PM on 9 January 2019, on 40 ground objects. The air temperature and the sky condition were close to those of the UAV flight. However, the relative humidity was much lower (~40%), requiring a proper adjustment of atmospheric parameters of the camera to correct atmospheric absorption. A matrix of 20 by 20 pixels centered on the object was used to obtain the mean DN value of the object. Linear regression was performed between the sampled temperature values and the corresponding DN values (Figure 8b). The regression equation was used to convert the thermal orthomosaic to a temperature map (Supplementary Figure S1).

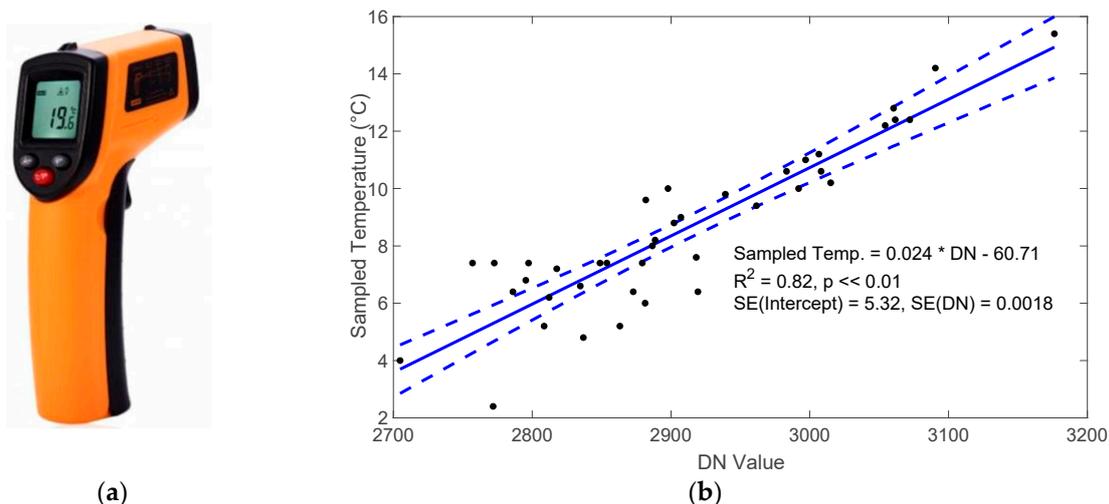


Figure 8. (a): Infrared thermometer used to measure temperature of ground objects; (b): Linear regression between the sampled temperature of 40 objects and their thermal DN values. Also shown are the regression equation, coefficient of determination (R^2), confidence level (p), the confidence interval of the predictions (blue dashed lines), and the standard errors of coefficients (SE).

3.5. Cluster Analysis

We use a cluster analysis to evaluate the temperature map. The analysis is comprised of two steps, rule-based classification and construction of temperature histograms.

A rule-based classification was applied to the RGB orthomosaic using a MATLAB program. The rules were set only based on the RGB contrast. In all, 77 % of the pixels were grouped into 9 different information classes, and 23 % remained unclassified (Figure 9). Next, a histogram of temperature distribution was constructed for each class (Figure 10). The typical temperature range of difference surfaces can be determined from the histograms. The temperature contrasts can be examined by comparing the peaks of these histograms. As expected for this time of the day (~10:00 local time), tree pixels are the coolest. Grass pixels have narrower range than tree pixels, and the majority of grass and tree pixels have similar temperatures as indicated by the very close peak values. Soil is warmer than grass because it warms up faster in the morning under sunlight. Of the three types of aquatic plants, the “cyan” plants form a thin layer floating on the water, so their temperatures are close to the temperature of the water body. The “green” and “red” plants extend above the water, so their temperatures are more different from the water temperature. The temperatures of asphalt and sandy road are intermediate. Water is the warmest because of its large heat capacity. The cluster analysis confirms that the thermal orthomosaic has captured the expected temperature distributions across the landscape.

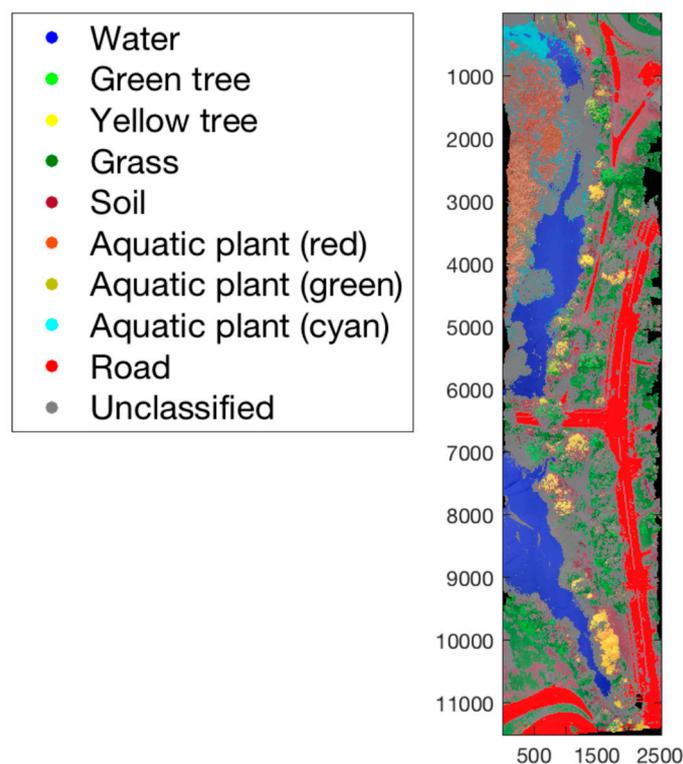


Figure 9. Rule-based classification of the RGB orthomosaic. The coordinate axes are number of the pixels.

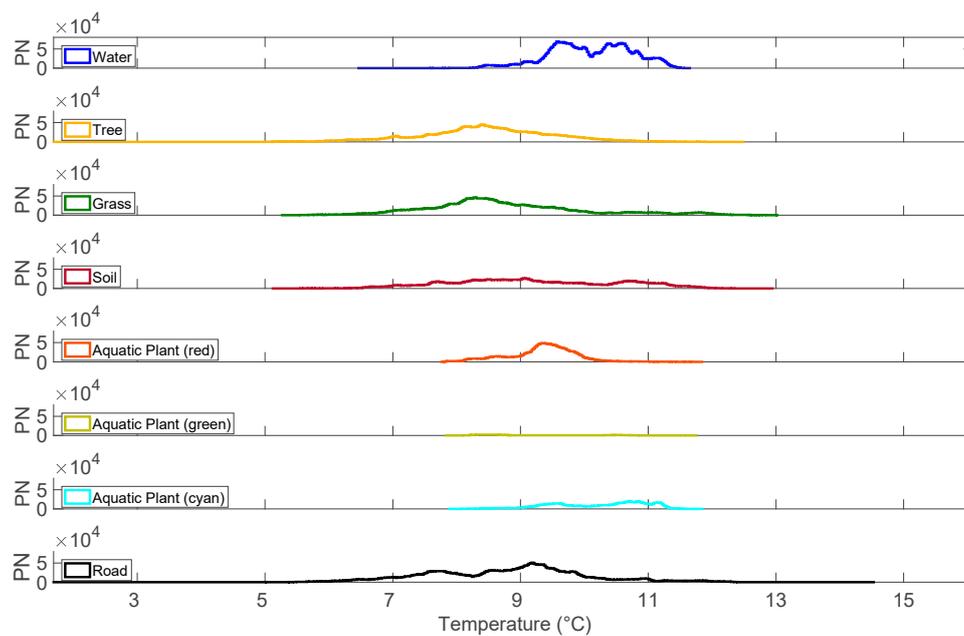


Figure 10. The temperature histograms of the clusters showing in Figure 9. PN is the pixel number. Here, clusters “Green Tree” and “Yellow Tree” have been merged into one cluster called “Tree”.

4. Discussion

This work aims to design a novel method, four-band thermal mosaicking (FTM), for mosaicking thermal imagery. The key to overcoming the difficulty associated with the coarse resolution and lack of single-band contrast of thermal images lies in processing all the images carefully prior to the mosaicking. Once the four-band images are generated, the workflow in Pix4D, an SfM-based application, is similar to mosaicking RGB images. No geographic information is needed in advance. The up-sampling process should be done with care. For the FLIR DUO PRO camera, the visible and thermal lenses have the same vertical FOV. That the visible band resolution is a multiple of the thermal-band resolution greatly simplifies the up-sampling process. For other dual cameras types without consistent FOVs in any direction or multiple relationship between resolutions, both RGB and thermal images need to be cropped to have matched FOVs and then up-sampled to the resolution of their minimum common multiple. Toward that end, more tests are needed for other camera types (including self-designed dual-lens cameras). Because cropping will decrease image size, feature matching may not be as precise as using the uncropped RGB images.

In the present study, the thermal and RGB orthomosaics have identical sampled areas. Two types of positional error are analyzed. The first type, the lens spacing error, is negligible (less than 6.36 cm). Therefore, we only need to calibrate the whole orthomosaic in order to remove the second type of error—misalignment caused by shutter delay between the two lenses. It is worth noting that object-based calibration is only feasible if the vectors of misalignment of the selected calibration objects are similar, or if the thermal orthomosaic deviates from the RGB orthomosaic by a uniform offset. Erratic error vectors are likely if the UVA flight pattern has no dominant direction of motion. Under this circumstance, the misalignment needs to be fixed for each single swath. In the present case, a simple offset correction to the thermal orthomosaic has produced acceptable results, as shown by the transect analysis. Compared to some other studies [36] (p. 4016), our positional error is higher in actual distance (39 cm), but is within one original thermal pixel size (57 cm).

The error analysis is based on the ground features found in the final orthomosaics. The overall performance is essentially limited by the resolution of the thermal lens: features smaller than one original thermal pixel are not identifiable in the thermal orthomosaic. The primary goal of the object-based calibration in this work is just to reduce the positional error to less than the original

thermal pixel size (57 cm), which we have achieved. Further improvement of the performance requires the use of higher-resolution thermal cameras with a smaller raw pixel size.

In theory, UAV images can also be processed in a continuous domain like HSV (hue-saturation-value) since the HSV domain separates the information of color (hue) and intensity (value) from the RGB composite. It is also possible that the contrast between adjacent objects becomes sharper after the conversion from RGB to HSV. We tested this idea by converting the RGB photos to HSV images and then scaling the HSV values to a 16-bit format. We then processed the images three times using Pix4D with three different band allocations: (1) main band-weight on hue, (2) main band-weight on saturation, and (3) main band-weight on value. We found that the orthomosaic is more successful with the main band-weight allocated to value. This HSV-based orthomosaic is overall consistent to the original RGB-based orthomosaic (Supplementary Figure S3). However, the two orthomosaics differ in detail. In many places, the HSV-based orthomosaic shows serious errors in mapping the true colors (Supplementary Figure S4). The erratic color distribution would lead to a poor registration of the thermal band at these places. For this reason, we have decided to use the original RGB-mosaicking method.

Our temperature map has been evaluated by cluster analysis. However, our temperature measurement assumed a fixed value of emissivity. This may explain why some low-emissivity objects, such as the metal roof of a garage (Figure 4) has a very low temperature ($-6.7\text{ }^{\circ}\text{C}$, Supplementary Figure S1). In the work of Turner et al. [36], no unexpected temperature values are evident even though a similar measurement method was used. This is because the surface of our study area is more heterogeneous as opposed to the Moss Bed in their study. Nevertheless, the primary purpose of radiometric calibration is to convert pixel DN values to values on an actual temperature scale. The cluster analysis is not meant to indicate that the actual temperature had been correctly retrieved for all the objects in the scene. Instead, these temperature values are used to determine if the thermal mosaic has captured the expected contrast of thermal properties across the landscape.

Although not helpful to positional accuracy, the correction of atmospheric interference may bring improvements to surface temperature estimation. MODTRAN (radiative transfer modeling for atmospheric correction) is regarded as a reliable model to correct the impact of atmospheric absorption of radiation [43]. Berni et al. [44] applied MODTRAN to their UAV photogrammetry and found that the error due to atmospheric absorption is dependent on surface temperature and is lower for cooler objects. At the maximum temperature value of $15\text{ }^{\circ}\text{C}$ in our landscape, we infer from their data that the error in our study should be no more than $1\text{ }^{\circ}\text{C}$. The actual error should be lower than this because our UAV flight was carried out in late fall at a lower atmospheric moisture and at a lower flight height than theirs.

5. Conclusions

In this study, a novel method, termed four-band thermal mosaicking (FTM), was developed to process thermal images from unmanned aerial vehicles (UAVs). The method requires the synchronized capture of RGB and thermal images using a dual-lens camera. The thermal images are subsequently resized and stacked onto their corresponding RGB photos in order to create four-band images. Afterwards, the four-band images are mosaicked via an SfM workflow using only the RGB bands. The method overcomes the problem of coarse resolution and low contrast in the thermal band in thermal mosaicking. The four-band orthomosaic can then be separated into an RGB and a thermal mosaic for further analysis.

The FTM was applied successfully to images acquired in a flight mission over an urban landscape. To improve the result, we developed an object-based calibration method to correct the positional error between the RGB orthomosaic and the thermal orthomosaic. In the current study, the error arose mainly from shutter delay between the lenses. The calibration reduced the mean positional error from 138 cm to 39 cm, which is less than one thermal pixel size (57 cm).

The FTM workflow opens opportunities for low-cost but high-resolution UAV observation of the thermal environment. It has scientific merits in many fields of application such as wildlife survey [45,46], wildfire detection [1,2], and urban climate studies [47], all of which require detailed thermal data. Future work is needed to test the method with different camera configurations. With the help of higher-resolution thermal cameras, further reduction of inter-band positional error may be possible. Additional improvements may also result from more rigorous correction for atmospheric interference and the use of ground targets of known temperatures in order to generate the actual temperature map.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/11/11/1365/s1>, Figure S1: The temperature map of Beaver Pond Park, Figure S2: The misalignment caused by shutter delay, Figure S3: Overall comparison between the HSV-based and the original RGB-based orthomosaic, Figure S4: Detailed comparison between the HSV-based and the original RGB-based orthomosaic.

Author Contributions: conceptualization, X.L. and Y.Y.; methodology, X.L. and Y.Y.; formal analysis, Y.Y.; investigation, Y.Y.; writing—original draft preparation, Y.Y.; writing—review and editing, X.L.; visualization, Y.Y.; supervision, X.L.; project administration, X.L.; funding acquisition, X.L.

Funding: This research was partially funded by the NUIST-Wuxi Research Institute.

Acknowledgments: All the authors gratefully acknowledge the technical support of Yale Center for Earth Observation (YCEO). The first author, Yichen Yang, acknowledges the guidance of Dr. Natalie Schultz on software operation, and appreciates her useful comments on the experiment design.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ambrosia, V.G.; Wegener, S.; Zajkowski, T.; Sullivan, D.V.; Buechel, S.; Enomoto, F.; Lobitz, B.; Johan, S.; Brass, J.; Hinkley, E. The Ikhana unmanned airborne system (UAS) western states fire imaging missions: From concept to reality (2006–2010). *Geocart. Int.* **2011**, *26*, 85–101. [[CrossRef](#)]
2. Phan, C.; Liu, H.H.T. A cooperative UAV/UGV platform for wildfire detection and fighting. In Proceedings of the 2008 Asia Simulation Conference-7th International Conference on System Simulation and Scientific Computing, Beijing, China, 10–12 October 2008; pp. 494–498.
3. Andersen, T.; Scheeren, B.; Peters, W.; Chen, H. A UAV-based active AirCore system for measurements of greenhouse gases. *Atmos. Meas. Tech.* **2018**, *11*, 2683–2699. [[CrossRef](#)]
4. Fladeland, M.; Sumich, M.; Lobitz, B.; Kolyer, R.; Herlth, D.; Berthold, R.; McKinnon, D.; Monforton, L.; Brass, J.; Bland, G. The NASA SIERRA science demonstration programme and the role of small–medium unmanned aircraft for earth science investigations. *Geocart. Int.* **2011**, *26*, 157–163. [[CrossRef](#)]
5. Ryan, J.C.; Hubbard, A.; Box, J.E.; Brough, S.; Cameron, K.; Cook, J.M.; Cooper, M.; Doyle, S.H.; Edwards, A.; Holt, T.; et al. Derivation of High Spatial Resolution Albedo from UAV Digital Imagery: Application over the Greenland Ice Sheet. *Front. Earth Sci.* **2017**, *5*, 40. [[CrossRef](#)]
6. Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. *Remote Sens.* **2014**, *6*, 10395–10412. [[CrossRef](#)]
7. Candiago, S.; Remondino, F.; De Giglio, M.; Dubbini, M.; Gattelli, M. Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sens.* **2015**, *7*, 4026–4047. [[CrossRef](#)]
8. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E.; Walthall, C.L. Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
9. Baena, S.; Moat, J.; Whaley, O.; Boyd, D.S. Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. *PLoS ONE* **2017**, *12*, e0188714. [[CrossRef](#)] [[PubMed](#)]
10. Mohan, M.; Silva, A.C.; Klauberg, C.; Jat, P.; Catts, G.; Cardil, A.; Hudak, T.A.; Dia, M. Individual Tree Detection from Unmanned Aerial Vehicle (UAV) Derived Canopy Height Model in an Open Canopy Mixed Conifer Forest. *Forests* **2017**, *8*, 340. [[CrossRef](#)]
11. Nevalainen, O.; Honkavaara, E.; Tuominen, S.; Viljanen, N.; Hakala, T.; Yu, X.; Hyypä, J.; Saari, H.; Pölonen, I.; Imai, N.N.; et al. Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. *Remote Sens.* **2017**, *9*, 185. [[CrossRef](#)]

12. Berni, J.A.J.; Fereres Castiel, E.; Su-rez Barranco, M.D.; Zarco-Tejada, P.J. Thermal and Narrow-band Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 722–738. [[CrossRef](#)]
13. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [[CrossRef](#)]
14. Laliberte, A.S.; Goforth, M.A.; Steele, C.M.; Rango, A. Multispectral Remote Sensing from Unmanned Aircraft: Image Processing Workflows and Applications for Rangeland Environments. *Remote Sens.* **2011**, *3*, 2529–2551. [[CrossRef](#)]
15. Villa, T.F.; Jayaratne, E.R.; Gonzalez, L.F.; Morawska, L. Determination of the vertical profile of particle number concentration adjacent to a motorway using an unmanned aerial vehicle. *Environ. Pollut.* **2017**, *230*, 134–142. [[CrossRef](#)] [[PubMed](#)]
16. Lucieer, A.; Jong, S.M.D.; Turner, D. Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Prog. Phys. Geogr. Earth Environ.* **2013**, *38*, 97–116. [[CrossRef](#)]
17. Niethammer, U.; James, M.R.; Rothmund, S.; Travelletti, J.; Joswig, M. UAV-based remote sensing of the Super-Sauze landslide: Evaluation and results. *Eng. Geol.* **2012**, *128*, 2–11. [[CrossRef](#)]
18. Burud, I.; Vukovic, M.; Thiis, T.; Gaitani, N. Urban surfaces studied by VIS/NIR imaging from UAV: Possibilities and limitations. In Proceedings of the Sixth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2018), Paphos, Cyprus, 26 March 2018. [[CrossRef](#)]
19. Jaechoon Chon, T. Three-Dimensional Image Mosaicking Using Multiple Projection Planes for 3-D Visualization of Roadside Standing Buildings. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2007**, *37*, 771–783. [[CrossRef](#)]
20. Garcia, R.; Battle, J.; Cufi, X.; Amat, J. Positioning an underwater vehicle through image mosaicking. In Proceedings of the Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), Seoul, Korea, 21–26 May 2001; Volume 2773, pp. 2779–2784.
21. Marks, R.L.; Rock, S.M.; Lee, M.J. Real-time video mosaicking of the ocean floor. *IEEE J. Ocean. Eng.* **1995**, *20*, 229–241. [[CrossRef](#)]
22. Kerschner, M. Seamline detection in colour orthoimage mosaicking by use of twin snakes. *ISPRS J. Photogramm. Remote Sens.* **2001**, *56*, 53–64. [[CrossRef](#)]
23. Moon, Y.S.; Yeung, H.W.; Chan, K.C.; Chan, S.O. Template synthesis and image mosaicking for fingerprint registration: An experimental study. In Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 17–21 May 2004; p. 409.
24. Szeliski, R. Video mosaics for virtual environments. *IEEE Comput. Gr. Appl.* **1996**, *16*, 22–30. [[CrossRef](#)]
25. Ott, S.R. Acquisition of high-resolution digital images in video microscopy: Automated image mosaicking on a desktop microcomputer. *Microsc. Res. Tech.* **1997**, *38*, 335–339. [[CrossRef](#)]
26. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314. [[CrossRef](#)]
27. Lourakis, M.I.A.; Argyros, A.A. SBA: A software package for generic sparse bundle adjustment. *ACM Trans. Math. Softw.* **2009**, *36*, 1–30. [[CrossRef](#)]
28. Furukawa, Y.; Ponce, J. Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1362–1376. [[CrossRef](#)] [[PubMed](#)]
29. Yahyanejad, S.; Rinner, B. A fast and mobile system for registration of low-altitude visual and thermal aerial images using multiple small-scale UAVs. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 189–202. [[CrossRef](#)]
30. Lewis, A.; Hilley, G.E.; Lewicki, J.L. Integrated thermal infrared imaging and structure-from-motion photogrammetry to map apparent temperature and radiant hydrothermal heat flux at Mammoth Mountain, CA, USA. *J. Volcanol. Geotherm. Res.* **2015**, *303*, 16–24. [[CrossRef](#)]
31. Webster, C.; Westoby, M.; Rutter, N.; Jonas, T. Three-dimensional thermal characterization of forest canopies using UAV photogrammetry. *Remote Sens. Environ.* **2018**, *209*, 835–847. [[CrossRef](#)]
32. Hoffmann, H.; Nieto, H.; Jensen, R.; Guzinski, R.; Zarco-Tejada, P.; Friborg, T. Estimating evaporation with thermal UAV data and two-source energy balance models. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 697–713. [[CrossRef](#)]
33. Coutts, A.M.; Harris, R.J.; Phan, T.; Livesley, S.J.; Williams, N.S.G.; Tapper, N.J. Thermal infrared remote sensing of urban heat: Hotspots, vegetation, and an assessment of techniques for use in urban planning. *Remote Sens. Environ.* **2016**, *186*, 637–651. [[CrossRef](#)]

34. Honjo, T.; Tsunematsu, N.; Yokoyama, H.; Yamasaki, Y.; Umeki, K. Analysis of urban surface temperature change using structure-from-motion thermal mosaicing. *Urban Clim.* **2017**, *20*, 135–147. [[CrossRef](#)]
35. Tsunematsu, N.; Yokoyama, H.; Honjo, T.; Ichihashi, A.; Ando, H.; Shigyo, N. Relationship between land use variations and spatiotemporal changes in amounts of thermal infrared energy emitted from urban surfaces in downtown Tokyo on hot summer days. *Urban Clim.* **2016**, *17*, 67–79. [[CrossRef](#)]
36. Turner, D.; Lucieer, A.; Malenovsky, Z.; King, H.D.; Robinson, A.S. Spatial Co-Registration of Ultra-High Resolution Visible, Multispectral and Thermal Images Acquired with a Micro-UAV over Antarctic Moss Beds. *Remote Sens.* **2014**, *6*, 4003–4024. [[CrossRef](#)]
37. Pech, K. Generation of multitemporal thermal orthophotos from UAV data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *1*, 305–310.
38. Nishar, A.; Richards, S.; Breen, D.; Robertson, J.; Breen, B. Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): A case study of the Wairakei–Tauhara geothermal field, Taupo, New Zealand. *Renew. Energy* **2016**, *86*, 1256–1264. [[CrossRef](#)]
39. Burud, I.; Thiis, T.; Gaitani, N. Reflectance and thermal properties of the urban fabric studied with aerial spectral imaging. In Proceedings of the Fifth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2017), Paphos, Cyprus, 20–23 March 2017. [[CrossRef](#)]
40. Harvey, M.C.; Rowland, J.V.; Luketina, K.M. Drone with thermal infrared camera provides high resolution georeferenced imagery of the Waikite geothermal area, New Zealand. *J. Volcanol. Geotherm. Res.* **2016**, *325*, 61–69. [[CrossRef](#)]
41. Harvey, P. ExifTool by Phil Harvey: Read, Write and Edit Meta Information! Available online: <https://www.sno.phy.queensu.ca/~phil/exiftool> (accessed on 6 October 2018).
42. Kuenzer, C.; Dech, S. *Thermal Infrared Remote Sensing*; Springer: Dordrecht, The Netherlands; New York, NY, USA, 2013.
43. Berk, A.; Anderson, G.P.; Bernstein, L.S.; Acharya, P.K.; Dothe, H.; Matthew, M.W.; Adler-Golden, S.M.; Chetwynd, J.H., Jr.; Richtmeier, S.C.; Pukall, B. MODTRAN4 radiative transfer modeling for atmospheric correction. In Proceedings of the Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, Denver, CO, USA, 19–21 July 1999. [[CrossRef](#)]
44. Berni, J.A.J.; Zarco-Tejada, P.J.; Suárez, L.; González-Dugo, V.; Fereres, E. Remote sensing of vegetation from UAV platforms using lightweight multispectral and thermal imaging sensors. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2009**, *38*, 6. [[CrossRef](#)]
45. Christiansen, P.; Steen, A.K.; Jørgensen, N.R.; Karstoft, H. Automated Detection and Recognition of Wildlife Using Thermal Cameras. *Sensors* **2014**, *14*, 13778–13793. [[CrossRef](#)] [[PubMed](#)]
46. Focardi, S.; De Marinis, A.M.; Rizzotto, M.; Pucci, A. Comparative Evaluation of Thermal Infrared Imaging and Spotlighting to Survey Wildlife. *Wildl. Soc. Bull.* **2001**, *29*, 133–139.
47. Voogt, J.A.; Oke, T.R. Thermal remote sensing of urban climates. *Remote Sens. Environ.* **2003**, *86*, 370–384. [[CrossRef](#)]

