Development of an interactive software tool for the handling of scientific data collected by the UoM Airborne Remote Sensing Facility

1. Introduction

Remote sensing is the monitoring process of the characteristic of land by measuring emitted and reflected radiation. [1] The remotely sensed data has been utilised for various scientific research, including precision agriculture, bio-security research, ecology, forestry, pollution monitoring, emergency disaster and fire monitoring. [4] Those digital images used for research are acquired with sensors aboard satellites, aircraft, and drones. Especially, aerial photography is the cost-effective and widely used method in remote sensing [2]. Although the advent of drones with light-weighted sensors allowed low-cost alternatives to obtain high-resolution data for remote sensing research [3], manned aircraft is still a commonly used and suitable data collection method for heavy sensors and faster data collection in the wider geographic area. [5]

In the Hyperspectral Remote Sensing and Precision Agriculture Laboratory at the University of Melbourne, manned aircraft are frequently utilised for data collection on the agricultural field for research. Hyperspectral imageries are acquired by the high-resolution hyperspectral and thermal sensors installed in tandem on the manned aircraft. [6-8] The trajectory lines at the study site were obtained with several technical settings, including the specific field of view, which is the maximum angle of view effectively detected by a sensor [9], and a targeted altitude. The acquired linearly scanned files for each flight line are calibrated and geo-corrected, followed by a semi-automatic mosaic generation that requires the addition of ground control points. This process of mosaic generation will take weeks or months in general. Despite this long processing time, if the resulting mosaic was not successfully visualised, for example, contains large gaps, the data collection on the study site must be conducted again which will require extra cost and time for aircraft operation. Thus, it is highly beneficial to have a software tool to assess the flight quality at the study site immediately after the data collection.

This report will discuss the development of an interactive software tool that allows a quick assessment of the flight performance for the image collection that will support the decision making on the recollection of data. The data loss of the specific area is often due to the inconsistent conditions during the flight which can be represented by the gaps between trajectory lines. The software is expected to visualise the potential gaps on the image of the study site from the coverage of the adjacent swaths. To assess the quality of the flight, the two main functionalities are introduced to the software tool. The first is to filter based on the auxiliary data obtained during flight, and the second is to visualise the trajectory of the aircraft and the risks of gaps due to errors in the executed flight plan. Further, this report will discuss the extension of the current software tool to create a mosaic of the selected area of interest, based on the synchronised data and files stored in the system during the flight.

2. Methods

2.1 Data files

Three data files are required for the software tool to assess the quality of the flight, including a GPS file, a polygon file and an image file of the study site.

2.1.1 GPS file

GPS file is the data obtained during the whole flight with 7 attribute columns separated by spaces in the text format. The attributes include Roll, Pitch, Yaw, Lat, Lon, Alt and Time. The time attribute is in the form of HH:MM:SS SSSSSS. The first 2 digits indicate hours, the second two indicate minutes, the third two indicate seconds and the last 6 digits indicate milliseconds.

2.1.2 Polygon file

Polygon file is the polygon of the study site in KMZ format projected with a coordinate system EPSG: 4326. This software assumes that the polygon is a single square shape. This means that the file should be processed per single study site rather than multiples. The coordinates must be indicated in the following format when converted to the KML file: <coordinates>lon1,lat1 lon2,lat2 ... lonx,latx lon1,lat1</coordinates>. Each coordinate is separated by a space and latitude and longitude values are separated by a comma. The first and last coordinates must be the same.

2.1.3 Image file of the study site

The image file is in the TIF format and used for the background image of the software. The image should contain GPS information with EPSG: 4326. The trajectory lines of the flight will be visualised on the image.

2.1.4 Hyperspectral and thermal image datasets

The flight data utilised for this project was collected on 26th November 2020 with the hyperspectral (Hyperspec VNIR E-Series model) and the thermal (FLIR Systems SC655c) sensors installed in the manned aircraft. The field of view parameters was 45 degrees for the thermal sensor and 66 degrees for the thermal and hyperspectral sensors. A custom flight planner, which considers the FOV, swath, altitude of the flight, percentage of overlap and the altitude of the terrain, was utilised for the flight plan. The software was tested on the flight data with the field of view parameter of 30, 45 and 60 degrees and the coverage threshold of 0.2, 0.45, 0.5 and 0.8 %.

2.2 Software implementations

Figure 2.2 depicts the overview of the implementation of the interactive software tool. The required input parameters and the functionality for processing data and the output are indicated in the flow chart from left to right. The software tool consists of three parts: (1) the data upload, (2) the application of filtering parameters and coverage threshold, and (3) the visualisation of trajectories and coverages. One extra functionality to select a region of interest is included which contributes to the more credible flight assessment method with further implementation on a mosaic generation.

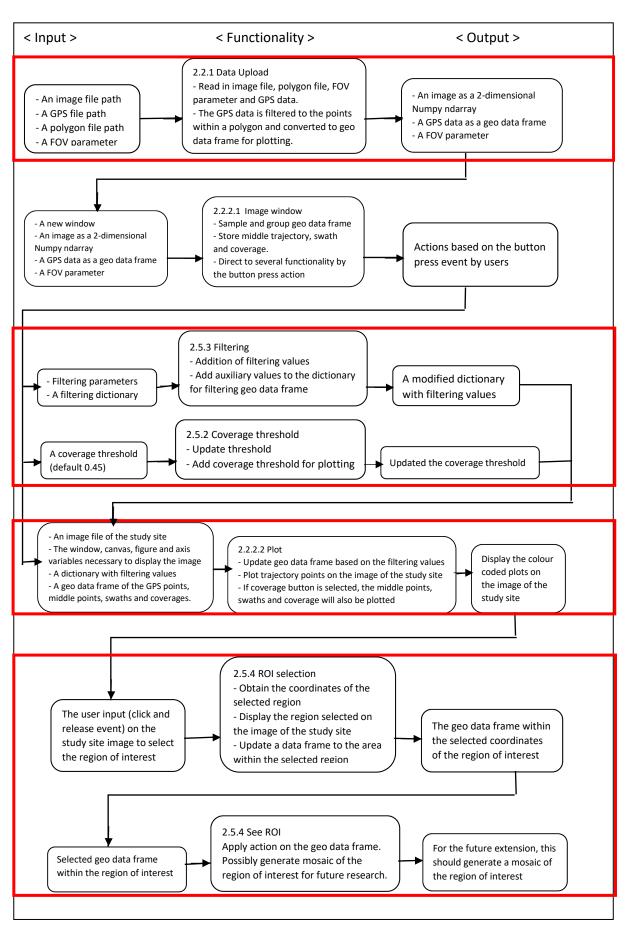


Figure 2.2 Implementation of the interactive software tool for flight assessment

2.2.1 Data upload

After the data collection by a manned aircraft, a polygon file and an image file of the study site and the collected GPS file of the flight are uploaded to the interactive software tools by users. Each file should be in the form as described in section 2.1. The image file is read into the system using rasterio module in python. The field of view parameter value is entered into the software tool at the time of file upload.

The system will extract the minimum and maximum latitude and longitude from the polygon file. These coordinates are applied to filter the GPS data frame which is created from the GPS file. The resulting data frame is converted to a geo data frame that includes the GPS coordinated of each point. The geo data frame is further filtered to extract the area within the polygon, which allows faster computational time to process the flight data. Since the geo data frame contains geographical information, this data format enables the visualisation of coordinates on the raster image file uploaded to the software tool. After the upload is completed, the software will prompt the user to proceed to the next window. The image of the study site in 2-dimensional Numpy ndarray, GPS data in geo data frame and a field of view parameter is passed to the next window.

2.2.2 Visualisation of the trajectory lines of the flight

In the new window, the user can check the trajectory lines of the flight displayed on the orthophoto of the study site. The GPS points can be filtered by the auxiliary data, including roll, pitch, yaw and altitude from the input of the graphical user interface (Refer to section 2.5.3).

2.2.2.1 Data sampling and grouping

The trajectory lines of the flight can be inspected using the software tool. Firstly, the GPS points of the flight in the geo data frame are sampled by every 2 degrees difference in latitude.

The sampled points are assigned to a block which is the collection of data within the same latitude range. The number of blocks for the dataset is derived by the range of the latitude of the study site divided by the latitude interval. For each block, the average latitude is assigned as follows:

average latitude = minimum latitude + latitude interval * (block number counted from lowest latitude + 1/2).

The average latitude of the assigned block for each GPS point is utilised as the key to get the longitudes of the same latitude points. In the calculation, the latitude is multiplied by 10000 to avoid the floating-point underflow.

The sampled GPS points in the data frame are also grouped into each trajectory line group with similar longitude values. The grouping is based on time as there is a larger time gap between each trajectory line obtained during the flight.

2.2.2.2 Middle trajectories, swaths and coverages

The centre of the two trajectory lines, the swath of each trajectory point and the coverage which is the percentage of the overlaps of the swath are calculated as the GPS points are sampled. After the filtering values are applied, the updated data frames of trajectories are visualised on the orthophoto of the study site. If the coverage button is selected in the software, the middle trajectories, swaths and coverages are also visualised. The coverage under and over the given threshold are colour coded separately as indicated in section 2.5 of this report.

For the visualisation of the middle trajectories, swaths and coverage, several dictionaries and formulas were applied. A dictionary is created to convert the average latitude key to a list of lists with trajectory group number, longitude value and index of a data frame with all the GPS points. This dictionary is

utilised to obtain the longitude of the adjacent trajectory point to plot their middle point and calculate the swath coverage. The middle points of the two trajectory lines and coverage are stored in a separate data frame for visualisation on the orthophoto.

To calculate the swath for the corresponding altitude, the following formula was utilised for each sampled GPS point with a given field of view parameter in degrees.

Swath =
$$(2 * altitude) * tan(radians(FOV) / 2)$$

The coverage is the percentage of the overlap of two swaths from adjacent trajectory lines out of the one swath length, which is denoted by the following expression.

```
Coverage = n/2m \%
```

where n = overlap of the field of view of the trajectory lines, m = the half of the field of view

The conversion method from meters to a longitude value to plot the swath and coverage on the orthophoto was calculated as follows.

Meters to longitude conversion:

The radius of the circumference of the earth at the given latitude:

```
ri = R \cos(radians(latitude))
```

where R = 6371000 m (radius of the earth)

The meter change per longitude change:

$$ti = 2 (ri) * (pi) / 360$$

The longitude change for the given x meter change:

longitude change in degree = x / ti

Two dictionaries are created to link the index of a data frame of the sampled GPS point to the index of its corresponding middle points and swath data frames. This was used to obtain the index of the filtered middle trajectories, swaths and coverage data frame.

The index of the coverage data frame that corresponds to the swath data frame was calculated as follows:

The index of the swath data frame – (the number of swaths – the number of coverages)

The number of coverages subtracted from the number of swaths indicates the number of swath points in the firstly sampled trajectory group. These index conversion methods were utilised to filter the middle trajectories, swaths and coverage data frames when the sampled GPS points are filtered by the software.

2.3 Programming language, modules and libraries

Python was utilised as a coding language of this software tool. The rasterio module was applied to display the orthophoto of the study site. The tkinter were used for graphical user interface design. Mainly pandas and geopandas library were utilised for the data structure for processing GPS information. Other libraries are used for the implementation of this software tool including shapely, matplotlib, numpy, zipfile, lxml, math, datetime, collections, functools.

2.4 Assumptions

Under the limited time frame of creating the software visualisation tool, it has the following assumptions:

- (1) The polygon of the study site is single square-shaped. This assumption is to have a correct dictionary implementation to match the indexes of swath lines and the coverage data frame.
- (2) Trajectory lines must be in the North-South direction. This assumption is due to the grouping implementations and meter conversion calculations that require directional restrictions, as shown in section 2.2.2.2.
- (3) No error for the GUI input is assumed. The leading objective of this project was to visualise the flight information for detecting image gaps. Thus, the exception handling codes for the expected error cases were not put on focus within the given time frame. All the files must be in the format indicated in section 2.1, and input must be correctly entered into the tool as explained in section 2.5.

2.5 Graphical User Interface

2.5.1 Pan and zoom

Users can pan and zoom the image after selecting the pan/zoom button of a toolbar. The image can be panned in any direction by dragging the image with a mouse pad. If you press the zoom button with the picture of the magnifier, users can select the region on the image to zoom in. Users can go back to the previous or the next view with the back or forward arrow button and go back to the original image composition with the home button. Users can also save the image with the save button on the right. The configuration of the image can be modified with the subplot configuration button.



Figure 2.5.1.1 Tool bar: from the left, the home, back, forward, pan/zoom, zoom, subplot-configuration and save button [10].

2.5.2 Visualisation of the trajectory lines and coverage

The "Display trajectories" button will plot the trajectory line of the aircraft in red (Figure 2.5.2.2). The "Display coverage" button will visualise middle lines between trajectory in yellow, a swath of each GPS point in blue, overlap of the swath above the coverage threshold in green and below the threshold in red (Table 2.5.2.1 & Figure 2.5.2.3). Thus, users can identify the potential gap of the image by looking at the image on the screen. The user can input the coverage threshold from the entry box displayed on the left side of the screen and press "Apply" (Figure 2.5.3.1). The default coverage threshold is set as 0.45.

Button name	Visualised data	Colour in the image
Display	Trajectory lines	Red
trajectories		
Display	Middle lines between two adjacent trajectory lines	Yellow
coverage	swath	Blue
	Swath overlap above a threshold	Green
	Swath overlap below a threshold	Red

Table 2.5.2.1 Description of button and colour coded points

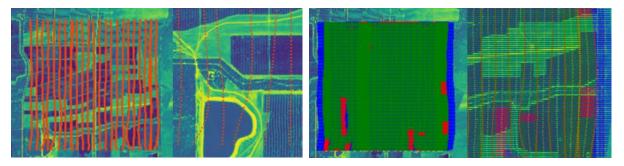


Figure 2.5.2.2 Trajectory lines

Figure 2.5.2.3 Coverage

2.5.3 Filtering GPS points by auxiliary data

Users can filter the GPS data by entering the attribute values of interest. Users can enter a minimum value and a maximum value in the entry box, press "Apply", and press either "Display trajectories" or "Display coverage" to visualise the filtered values (Figure 2.5.3.1). The values outside of the entered minimum and maximum interval are obtained when filtered by the roll, pitch and yaw. The values within the minimum and maximum interval are obtained when filtered by the altitude. If no value is entered, the maximum or minimum parameter value is applied for maximum and minimum entries by default.

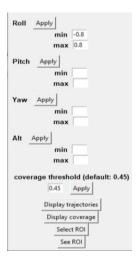


Figure 2.5.3.1 Filtering entries

2.5.4 Region of interest

Users can select a region of interest after pressing the "Select ROI" button (Figure 2.5.4.1). The selector is activated by pressing the key "s" and deactivated by pressing the key "q". Users can also apply an action to the selected region of interest by pressing the "See ROI" button. Although no action is applied for the current software version, the output for pressing "see ROI" can be extended for a more accurate assessment method with further implementation. The future extension of this functionality is discussed as the "suggestion for improvement for future research" in section 4.4 of this paper.

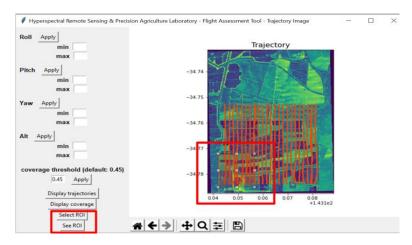


Figure 2.5.4.1 Region of interest selections

3. Results

3.1 Software upload functionality

The software tool first prompted the user to upload three files necessary for the flight assessment (Figure 3.1). The field of view parameter was entered from the entry box followed by pressing the enter button. Then, an image file, GPS file and a polygon file were selected from the local folder. The upload status was displayed on the right side of the screen. Once the upload was completed, the "proceed" button appeared on the left bottom side of the screen to move to the next window for the image visualisation.

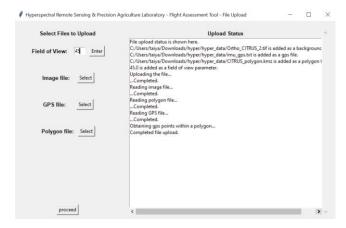


Figure 3.1.1. Upload window

3.2 Trajectory and coverage visualisation

In the next window, users were able to display the trajectory lines of the flight data on the orthophoto of the study site by selecting the "Display trajectories" button, as shown in figure 3.2.1.

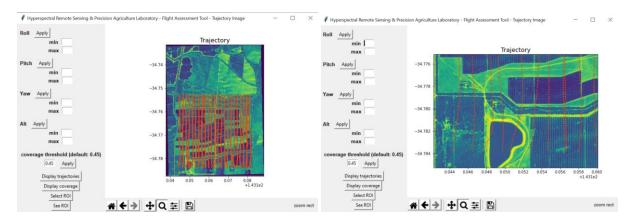


Figure 3.2.1. Trajectory lines (left), Figure 3.2.2. Zoomed trajectory lines (right).

The coverage of the trajectory lines was displayed by selecting the "Display coverage button" (Figure 3.2.3). The trajectory points were plotted with red dots, the middle points of trajectory lines were plotted with yellow, swaths were shown in the blue line, coverages over the threshold were in green and under the threshold were in the red.

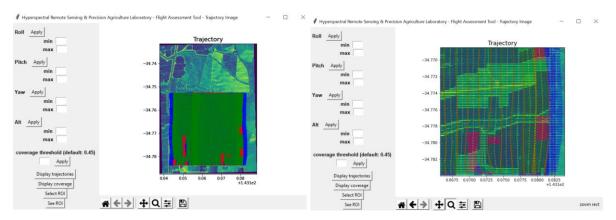


Figure 3.2.3. Coverage (left), Figure 3.2.4. Zoomed coverage (right).

Users were able to pan and zoom the images displayed on the right side of the screen as shown in figures 3.2.2 and 3.2.4.

The filtering values of the auxiliary data were applied from the entry boxes on the left side of the screen. For this test case, the roll of minimum -0.8 and maximum 0.8 was applied, and both the trajectory lines and the coverage were successfully filtered by the auxiliary values (Figure 3.2.5, 3.2.6, 3.2.7 & 3.2.8).

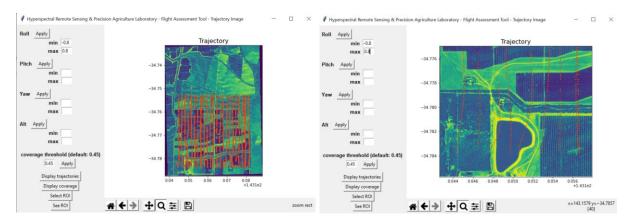


Figure 3.2.5. Filtered trajectory lines (left), Figure 3.2.6. Zoomed filtered trajectory lines (right).

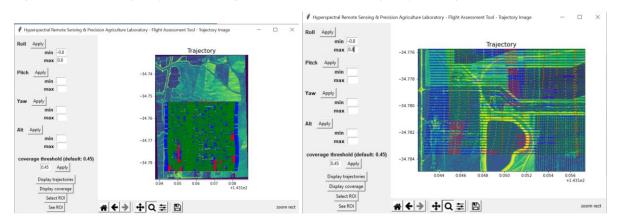


Figure 3.2.7. Filtered coverage (left), Figure 3.2.6. Zoomed filtered coverage (right).

3.3 Region of Interest

Users can select the region of interest on the orthophoto of the study site after clicking the "select ROI" button. Users can use a mouse pad to select the region which is indicated in a red square (Figure 3.3.1). The coordinates of the selected region were read into the software tool, as shown in figure 3.3.2.

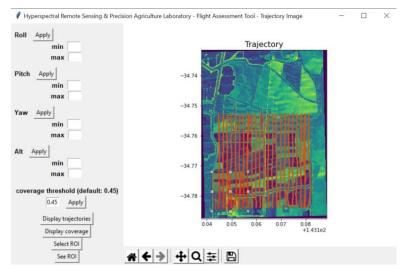


Figure 3.3.1 Region of interest selection.

```
roi dict
{'Lon': [143.1443375522871, 143.15655863683466], 'Lat': [-34.78371442232143, -34.7718919235831]}
(143.141863228, -34.784929440) --> (143.156558637, -34.771891924)
```

Figure 3.3.2 Coordinates of the region of interest read into the system.

3.4 Thresholds

Various coverage threshold values were applied to experiment with the software functionality of assessing areas under a coverage threshold. With a 0.2 % coverage threshold, no coverage was identified as under the threshold as shown in figure 3.4.1. With a 0.45 % and 0.5 % coverage threshold, the software tool outputted areas under the threshold as shown in figures 3.4.2 and 3.4.3. The threshold value of 0.8 % gave most of the coverage as under the threshold (Figure 3.4.4). These results have shown that with a higher threshold value, more areas are classified as under the threshold where the potential gaps exist. This indicates that the threshold value is successfully applied to the system and the overlap of the swath is visualised correspondingly for the assessment of the potential gaps.

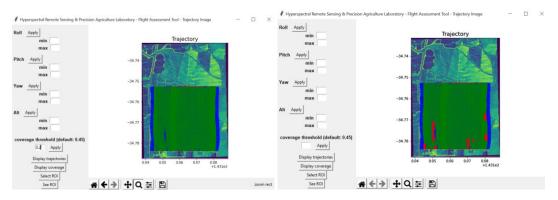


Figure 3.4.1 Coverage with a 0.2 % threshold, Figure 3.4.2 Coverage with a 0.45 % threshold.

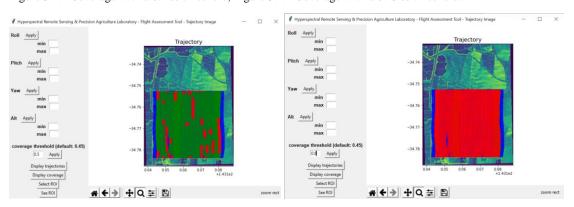


Figure 3.4.3 Coverage with a 0.5 % threshold, Figure 3.4.4 Coverage with a 0.8 % threshold.

3.5 Field of view

Various field of view parameters including 30, 45 and 60 degrees were tested on this software to assess the visualisation of the swath based on the fields of view. Figures 3.5.1, 3.5.2 and 3.5.3 show the coverage for the field of view parameters 30, 45 and 60 degrees respectively for a 45 % coverage threshold. As the field of view increases, the width of the swath became wider and fewer overlaps under the threshold was observed. This indicates that the software implements the swath width on the image successfully corresponding to the field of view parameter.

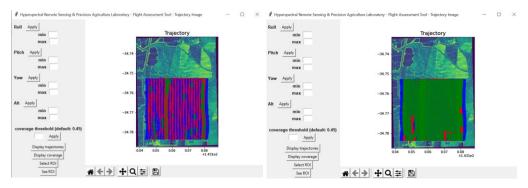


Figure 3.5.1 Coverage with FOV 30°, Figure 3.5.2 Coverage with FOV 45°.

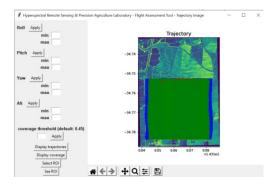
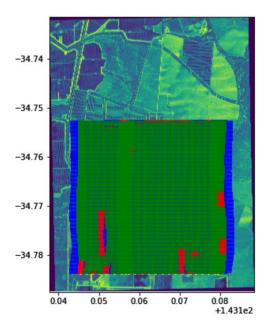


Figure 3.5.3 Coverage with FOV 60°.

3.6 Accuracy

The software output with a field of view parameter of 45 degrees and 45 % coverage threshold was compared with the actual image obtained by a thermal sensor with a field of view parameter of 45 degrees. Although they are not the complete match, the software tool gave a rough area under the coverage threshold as shown in figure 3.6.1 that is similar to the region of the gap observed in the thermal image indicated in white in figure 3.6.2. On top of the coverage value indicated with colour, zooming into the image and observing the trajectory lines could also help identify the potential gap created through flight.





For a 49 % coverage threshold, the wider areas under the threshold are indicated in red (Figure 3.6.3). Figure 3.6.4 shows the coverage under the threshold that corresponds to observed gaps in the thermal image in yellow squares. This shows that the software visualised most of the gaps in the thermal image. However, the accuracy is not 100 % as shown in figure 3.6.5 with some gaps in the thermal image not represented in the software tool.

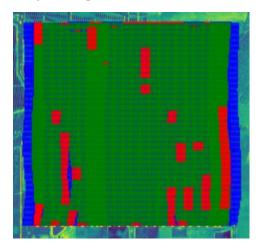


Figure 3.6.3 Coverage with FOV 45° and threshold of 49 %

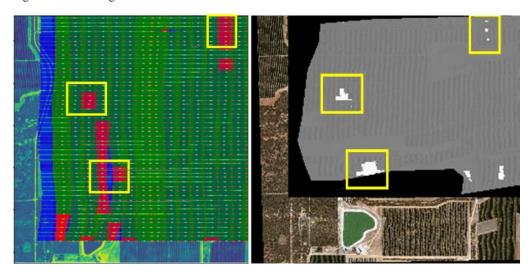


Figure 3.6.4 Comparisons of the coverage with FOV 45° and a threshold of 49% of the software (left) and thermal image gaps (right)

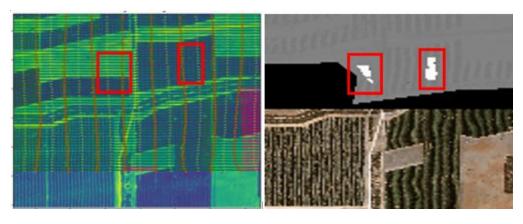


Figure 3.6.5 The gaps in a thermal image (right) not represented in the software tool with FOV 45° and a threshold of 49 % (left)

4. Discussions and future extensions

The result of the interactive software tool illustrates most of the gaps observed in the thermal image with an appropriate choice of a field of view parameter and a coverage threshold. This quick assessment of the flight is conducted from the collected GPS data and enables the onsite evaluation of potential gaps immediately after the flight. The observed visualisation functionalities of trajectory points and swath coverage based on filtering values, the field of view parameter and the coverage threshold can be utilised to target the data points with a specific condition that is more likely to cause the gaps. In terms of the accuracy of the software tool, the gap detection from the coverage threshold values did not cover some gaps observed in the thermal image. One possible reason lies in the calculation of the swath that only considers the field of view parameter and the altitude. The swath value may vary depending on the roll, pitch and yaw parameters in practice and does not always have an equal width on the right and left sides of the trajectory. To make this system more accurate, the swath calculation can be created in a more complex manner with all the required parameters. However, this tool can be utilised for the first scan to identify the potential area with gaps which can be double-checked with a closer look at the area with mosaic generations which will be discussed in section 4.4. Since the software visualisation tool was developed under a limited time frame, there are several suggestions for improvements. The following suggestions can be implemented with further tests on real-world flight data sets to adapt for unexpected error cases.

4.1 Error handling

As mentioned in the assumption, the system expects the correct input for this software. The software tool can be more useful by including error handling such as error messages for incorrect inputs. In addition, the software could be more adaptable to various inputs by automatic conversion such as the transformation of the file format available for both KML and KMZ files.

4.2 Adapting for the flexibility on data collection methods

With further implementation, the system can be extended for the data recorded in the east-west direction by storing the meter conversion calculation and grouping system by longitude value for the collected data and making this calculation available by user choice.

4.3 Usability

The software usability can be improved further by including various parameter values for filtering such as speed. The improvement of time complexity can potentially be achieved by incorporating parallel computing or time-efficient filtering methods which would be especially effective for large datasets.

4.4 Mosaic generation

The functionality to generate a mosaic for the selected region of interest will make the system more useful by checking the gap from the actual image. A quick check of the image at the data collection site will allow an efficient assessment method of flight quality. Since the selection of a region of interest is implemented in the software, the mosaic generation can be achieved by linking the flight time and GPS information to the image file collected by flight.

5. Conclusions

The interactive software tool enabled the visualisation of the potential gaps of the thermal and hyperspectral remote sensing imagery from the flight GPS data collected by the University of Melbourne Airborne Remote Sensing Facility. The trajectory lines, the middle points of the two trajectories, the swath and the coverage were displayed on the given orthophoto of the study site with respect to the field of view parameter and the coverage threshold. The software allowed the filtering of GPS points by auxiliary values to check the specific points of interest. The resulting coverage points under the threshold covered most of the observed gaps in the thermal image under the 49 % threshold and the field of view of 45 degrees. This interactive software tool can be utilised for the first quick check for the potential gaps in remote sensing imagery from the visualised colour-coded coverage analysis on an orthophoto. The system accuracy can further be improved by incorporating roll, pitch and yaw in the swath calculation and creating a mosaic image of the potential area under the threshold.

6. Acknowledgements

The author would like to thank Professor Pablo J. Zarco-Tejada and Dr Tomas Poblete from the Hyperspectral Remote Sensing and Precision Agriculture Laboratory at the University of Melbourne for supervising this research project. Although this was the first time to conduct a research project in her undergraduate degree, the author was able to complete the interactive software tool within one semester thanks to their support. The author would also like to acknowledge Associate Professor Ross Gall at the University of Melbourne for coordinating this research project.

7. References

- [1] USGS. (2021). What is remote sensing and what is it used for?. USGS Science for a changing world. https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news-science-products=0#qt-news-science-products
- [2] DAS, N. (2020). *Remote Sensing: The art behind geospatial data collection for non-experts*. AIDASH. https://www.aidash.com/resource/remote-sensing-art-behind-geospatial-data-collection-non-experts
- [3] Milas, A. S., Cracknell, A. P., & Warner, T. A. (2018). Drones the third generation source of remote sensing data. *International Journal of Remote Sensing*, 39(21), 7125-7173. https://doi.org/10.1080/01431161.2018.1523832
- [4] Toro, F. G. (2017). *Special Issue "UAV or Drones for Remote Sensing Applications"*. Sensors. https://www.mdpi.com/journal/sensors/special_issues/UAV_drones_remote_sensing#info
- [5] Singh, A. V. (2020). *Remote Sensingfor Non-Experts: How core industries can benefit from satellite tech*. AIDASH. https://www.aidash.com/resource/remote-sensing-non-experts-how-core-industries-can-benefit-satellite-tech
- [6] Camino, C., Calderon R., Parnell S., Dierkes H., Chemin Y., Roman-Ecija M., Montes-Borrego M., Landa B.B., Navas-Cortes J.A., Zarco-Tejada P.J., & Beck P.S.A. (2021). Detection of Xylella fastidiosa in almond orchards by synergic use of an epidemic spread model and remotely sensed plant traits. *Remote Sensing of Environment*, 260, 112420. https://doi.org/10.1016/j.rse.2021.112420
- [7] Hornero, A., Zarco-Tejada, P.J., Quero, J.L., North, P.R.J., Ruiz-Gomez, F.J., Sanchez-Cuesta, R., & Hernandez-Clemente, R. (2021). Modelling hyperspectral- and thermal-based plant traits for the

- early detection of Phytophthora-induced symptoms in oak decline. *Remote Sensing of Environment*, 263, 112570. https://doi.org/10.1016/j.rse.2021.112570
- [8] Suarez, L., Zhang, P., Sun, J., Wang, Y., Poblete, T., Hornero, A., & Zarco-Tejada, P.J. (2021). Assessing wine grape quality parameters using plant traits derived from physical model inversion of hyperspectral imagery, *Agricultural and Forest Meteorology*, 306, 108445. https://doi.org/10.1016/j.agrformet.2021.108445
- [9] EO-MINERS. (2021). Earth Observation in the frame of EO-MINERS Overview of remote sensing methods, sensors and applications. eo miners. http://www.eo-miners.eu/earth_observation/eo_eof_msa_techterms.htm
- [10] The Matplotlib development team. (2020). *Interactive navigation*. Matplotlib. https://matplotlib.org/3.2.2/users/navigation_toolbar.html