GROWPRO: A FLEXIBLE AND HIGH-RESOLUTION IMAGING TOOL FOR HIGH-THROUGHPUT PLANT PHENOTYPING IN THE FIELD

A Thesis Submitted to the College of Graduate and Postdoctoral Studies in Partial Fulfillment of the Requirements for the degree of Master of Science in the Department of Computer Science University of Saskatchewan Saskatoon

By
Ilya Vladimirovich Ovsyannikov

©Ilya Vladimirovich Ovsyannikov, June 11th, 2018. All rights reserved.
PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada
ABSTRACT

Various remote sensing platforms are being used in agriculture research, but their capability and flexibility may not be enough for carrying out efficient computer analysis. Low-quality data can affect the accuracy of methods and techniques used for high-throughput phenotyping in plant breeding. Despite the fact that there are many available remote sensing tools, researchers are seeking new platforms with superior features and performance over already proven systems.

In this thesis, we describe a novel approach for remote sensing for field-based phenotyping called the GrowPro. This thesis presents an overview of related work and commonly available approaches to remote sensing, description of the designed system, data gathering procedure, post-processing pipeline and best practice for capturing plants.

We show the ease of use of the GrowPro, simplicity of data gathering and qualify the accuracy of the stitching process. We proved that by using this novel approach, high-resolution RGB stitched images of regions of interests (e.g., an individual plot or range of plots) can be obtained. This method appeared to be stable over time, different trials and weather conditions. Examples of RGB stitched images of a variety of crops at various stages of growth have been described.
ACKNOWLEDGEMENTS

In the conduct of experiments we gratefully acknowledge: Summer students of University of Saskatchewan for field filming: Matthew Galbraith, Eileen van Heerde, Jaden Ball, James Scarrow, Michael van der Kamp. Field access and plot labeling: Steve Shirtliffe, Steve Ryu, Sally Vail, Christina Eynck. Agisoft Photoscan tutorials: William Van der Kamp. Field crews at AAFC and UofS for plot seeding, maintenance and ground truthing.
# CONTENTS

<table>
<thead>
<tr>
<th>Permission to Use</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Contents</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 Motivation

1.1.1 Technical point of view

1.1.2 Software point of view

1.2 Problem Statement

1.3 Proposed Solution

1.3.1 Similar approaches

1.4 Contributions

## 2 Background Related Work

2.1 Airborne Systems

2.2 Fixed In-field Systems

2.3 Mobile Vehicular Systems

2.4 Image collection management

2.4.1 Image clustering and classification

2.5 Orthophotos

2.6 Discussion

## 3 GrowPro Filming Tool and Data Acquisition

3.1 Functional design

3.2 Equipment research

3.3 Camera rig description

3.3.1 Assembling of the camera rig

3.3.2 Best practices for filming

3.3.3 Walking patterns

3.4 Data collection procedure

3.5 Amount of data collected

3.6 Data Overview

3.7 Field layout description

3.8 Discussion

## 4 Data Pre-processing Pipeline

4.1 Pre-processing pipeline description

4.1.1 Range Segmentation

4.1.2 Prior image adjustments and corrections

4.1.3 Data verification and range labeling

4.1.4 Stitching and Orthomosaic Generation

4.1.5 Data Management
4.2 Implementation of pipeline for 2-pass data
4.2.1 GPS-based data verification and range labeling
4.2.2 GPS-based range segmentation
4.2.3 GPS pre-processing for 3D model generation
4.3 Discussion

5 Results
5.1 Metrics of success
5.1.1 Visual approach
5.1.2 Automated approach
5.2 Ground resolution comparison
5.3 Results over different crops
5.4 Results over different stages of growth
5.5 Results over different lighting conditions
5.6 Discussion

6 Conclusion
6.1 Contributions
6.2 Limitations
6.3 Recommendations and future work
6.3.1 Tool and filming procedure recommendations and potential improvements
6.3.2 Data analyses recommendations and potential improvements
6.3.3 Straightening GPS paths
6.4 Conclusion

References

Appendix A GrowPro user guide
A.1 Before filming on the field
A.2 Prepare the camera rig for filming
A.3 Filming procedure
A.4 Always remember!
A.5 Remote clicker technique
A.6 Additional notes
A.7 Gathering data procedure

Appendix B Data management
B.1 Orthomosaic generation and exporting
B.2 Data preparation for Plot Segmentation tool
# List of Tables

1.1 Attributes of flexibility and capability in field-based high-throughput phenotyping systems ...................................................... 2

2.1 Available types of remote sensing platforms ........................................................................................................................................................................ 11

3.1 Camera comparison ................................................................................................................................................................................................. 16

3.2 Part list of the proposed camera system .............................................................................................................................................................. 17

3.3 Amount of data collected during the 2017 growing season .............................................................................................................................. 23

4.1 Performance of the algorithm for images of Canola DFL captured on June 5th, 2017, $k = 30$ .............................................................. 48

4.2 Performance of the algorithm for images of Canola DFL captured on June 13th, 2017, $k = 30$ .............................................................. 48

4.3 Evaluation of flattening technique ................................................................................................................................................................. 61

5.1 Stitching results over different stages of growth regarding different artifacts visually detected in the orthomosaic images ......................................................................................................................... 68

5.2 Stitching performance over different lightning conditions for Canola and Wheat images of one- and two-pass walking pattern ......................................................................................................................... 70
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The spectrum of field-based phenotyping systems in terms of flexibility versus capability. Our proposed GrowPro system is highly flexible, with higher capability than aerial drones in terms of ground resolution.</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>Orthophoto sample</td>
<td>13</td>
</tr>
<tr>
<td>3.1</td>
<td>FeiyuTech G5 Triaxial Handheld Gimbal [1]</td>
<td>17</td>
</tr>
<tr>
<td>3.2</td>
<td>Camera rig</td>
<td>18</td>
</tr>
<tr>
<td>3.3</td>
<td>Parts of the camera rig</td>
<td>18</td>
</tr>
<tr>
<td>3.4</td>
<td>Filming procedure</td>
<td>19</td>
</tr>
<tr>
<td>3.5</td>
<td>One-pass walking pattern</td>
<td>20</td>
</tr>
<tr>
<td>3.6</td>
<td>Two-pass walking pattern</td>
<td>20</td>
</tr>
<tr>
<td>3.7</td>
<td>Data collection workflow. After images have taken by the GrowPro, they should be transferred from the camera by the desktop application “GoPro Quik”. This application generates the required file structure and group images per range.</td>
<td>22</td>
</tr>
<tr>
<td>3.8</td>
<td>r02 range of Canola, captured on June 28th</td>
<td>24</td>
</tr>
<tr>
<td>3.9</td>
<td>r136 range of Carinata, captured on June 28th</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>Visual representation of the proposed pre-processing pipeline. The entire process is divided into six major steps.</td>
<td>26</td>
</tr>
<tr>
<td>4.2</td>
<td>Range Segmentation step in pre-processing pipeline</td>
<td>27</td>
</tr>
<tr>
<td>4.3</td>
<td>Example of colour instability caused by white balance difference</td>
<td>28</td>
</tr>
<tr>
<td>4.4</td>
<td>Sample of image with operator’s feet and drop shadow from the camera rig</td>
<td>28</td>
</tr>
<tr>
<td>4.5</td>
<td>Stitching performance before and after cropping out of destructive subjects (shadows) from the images.</td>
<td>29</td>
</tr>
<tr>
<td>4.6</td>
<td>Top down view of GPS coordinates of images from Carinata trial. The figure shows only a cropped sector of the Carinata field. The dataset consists of the top block with two ranges and the bottom block with 16 ranges. Ranges with high level of nosiness are circled red.</td>
<td>30</td>
</tr>
<tr>
<td>4.7</td>
<td>Top down view of GPS coordinates of images from Canola DFL trial. From right to left: the 5th and 11th ranges have missing data (circled red).</td>
<td>29</td>
</tr>
<tr>
<td>4.8</td>
<td>Stitching and Orthomosaic generation step in pre-processing pipeline</td>
<td>30</td>
</tr>
<tr>
<td>4.9</td>
<td>Data management step in pre-processing pipeline</td>
<td>31</td>
</tr>
<tr>
<td>4.10</td>
<td>Required file structure for the designed software tool. Root directory here represents a required directory for the program and should only contains one level of sub-directories.</td>
<td>34</td>
</tr>
<tr>
<td>4.11</td>
<td>Principle of GrowPro Viewer software</td>
<td>34</td>
</tr>
<tr>
<td>4.12</td>
<td>Graphical User Interface of the GrowPro Viewer</td>
<td>35</td>
</tr>
<tr>
<td>4.13</td>
<td>Field layout information represented in .JSON file</td>
<td>35</td>
</tr>
<tr>
<td>4.14</td>
<td>GPS coordinates of images of Canola DFL, taken on July 7th, 2017. The distinct empty areas on the left and right side of the trial can be observed.</td>
<td>36</td>
</tr>
<tr>
<td>4.15</td>
<td>GPS coordinates of images of Carinata, taken on June 13th, 2017. The distinct empty areas in the middle of the trial can be observed.</td>
<td>36</td>
</tr>
<tr>
<td>4.16</td>
<td>Field layout of the AAFC field</td>
<td>37</td>
</tr>
<tr>
<td>4.17</td>
<td>Plotted GPS coordinates of the third block of Carinata (blue points) and the field layout represented by bounding boxes (red lines).</td>
<td>38</td>
</tr>
<tr>
<td>4.18</td>
<td>Example of a GPS track with low level of noisiness.</td>
<td>39</td>
</tr>
<tr>
<td>4.19</td>
<td>Example of a GPS track with high level of noisiness.</td>
<td>39</td>
</tr>
<tr>
<td>4.20</td>
<td>Axis of the range</td>
<td>39</td>
</tr>
<tr>
<td>4.21</td>
<td>Finding an axis of the range using LSE approximation technique. In the experiment, two sets of points approximated by the lines $l_{s1}$ and $l_{s2}$ were used to find a range axis $l_{axis}$.</td>
<td>40</td>
</tr>
<tr>
<td>4.22</td>
<td>A vector defined by a pair of points in the point set</td>
<td>41</td>
</tr>
</tbody>
</table>
4.23 Definition of local noise for the selected pair of points ............................................. 42
4.24 Estimation of noisiness for images of Canola DFL taken on June 28th, 2017, window size - 15 points ................................................................. 43
4.25 Classification results for images of Canola DFL (12 ranges, captured on June 5th, 2017), k = 30 ................................................................. 47
4.26 Classification results for images of Canola DFL (12 ranges, captured on June 13th, 2017), k = 30 ................................................................. 47
4.27 Absolute differences between successfully classified images and the actual number of images over different numbers of the nearest neighbours (from 1 to 40), for images of Canola DFL, June 5th. Smallest absolute difference is highlighted green for each range. ................................................................. 49
4.28 Absolute differences between successfully classified images and the actual number of images over different numbers of the nearest neighbours (from 1 to 40), for images of Canola DFL, June 13th. Smallest absolute difference is highlighted green for each range. ................................................................. 49
4.29 Visualization of the convolutional algorithm (ten points window size). Algorithm calculates the angle between each vector and the North vector ................................................................. 50
4.30 CCA for images of Canola DFL, 12 ranges, June 5th, window size = 50 points ......................... 51
4.31 GPS tracks of images of Canola DFL, June 05th, and the number of turning points ......................... 51
4.32 CCA for images of Canola DFL, 12 ranges, June 13th. 23 distinctive peaks have been identified, window size = 50 points ................................................................. 52
4.33 GPS tracks of images of Canola DFL, June 13th, and the number of turning points ......................... 52
4.34 Turning areas (highlighted red), identified by the algorithm within images of Canola DFL, June 5th. Threshold value = 70deg, window size = 50 points ................................................................. 53
4.35 Turning areas (highlighted red), identified by the algorithm within images of Canola DFL, June 13th. Threshold value = 70deg, window size = 50 points ................................................................. 54
4.36 GPS coordinates of images 4th range of Canola, taken on 9 different filming dates. Red and green dots represent starting and finishing point respectively ................................................................. 55
4.37 Representation of a mesh model in Agisoft Photoscan GUI with different positions ......................... 56
4.38 Altitude nosiness within images of 25 ranges of Carinata (lateral view) .................................................. 57
4.39 All altitudes are assigned to the same value and belong to the same plane .................................................. 57
4.40 GPS points of images for Canola DFL 2017-06-21 dataset .................................................. 58
4.41 3D mesh models before GPS adjustments .................................................. 58
4.42 3D mesh models after GPS adjustments .................................................. 59
4.43 GPS points of images for Canola DFL 2017-06-26 dataset .................................................. 59
4.44 3D mesh models before GPS adjustments .................................................. 60
4.45 3D mesh models after GPS adjustments .................................................. 60
4.46 Example of GPS track of the range that potentially can be correctly positioned parallel to XY coordinate plane .................................................. 61
4.47 Evaluation of flattening technique for different filming dates .................................................. 62
5.1 Examples of artifacts found in the generated orthomosaic images .................................................. 64
5.2 An example of a .csv file report generated for the range images of Carinata. Instances in the table are sorted by the total number of non-white pixels .................................................. 65
5.3 Ground resolution of images taken from UAVs and GrowPro .................................................. 66
5.4 Samples of orthomosaic images of different crops .................................................. 67
5.5 Canola plants at different stages of growth. Labels of the images refer to a number of weeks from seeding .................................................. 68
5.6 Filming results on cloudy days .................................................. 69
5.7 Filming results on sunny days .................................................. 69
6.1 An example of perpendicular axis to a plane, parallel to a 3D model mesh .................................................. 76
6.2 Example of a GPS track with low level of noisiness .................................................. 76
Chapter 1  
Introduction

High throughput phenotyping (HTP) is an emerging area in plant breeding aimed at increasing the quality and quantity of phenotypic information in plant breeding experiments in order to overcome the so-called phenotyping bottleneck [20]. A number of systems have been proposed for HTP in controlled growing environments, such as conveyor belt [26] and robotic systems [25] to automatically photograph plants grown in pots or trays and image analysis software to extract phenotypic measurements from the resulting pictures [17].

HTP in the field presents a number of practical and important challenges beyond controlled environments, such as densely grown plants, uncontrolled image backgrounds, variable lighting conditions, and inclement weather.

1.1 Motivation

Despite these challenges, field phenotyping is an important pursuit in plant breeding due to the noted difficulties in translating greenhouse results to the field. A variety of field-based HTP systems have been proposed, including aerial drones, motorized vehicles, push-carts and fixed in-field imaging systems [18]. Each type of system provides a different trade-off in terms of ground resolution, coverage, ease-of-operation, cost, etc. However, to date, a low-cost phenotyping system that can be used with any type of breeder field and capture images that can resolve fine plant features required to measure phenotypes such as emergence counts, has yet to be realized.

1.1.1 Technical point of view

The spectrum of field-based phenotyping systems can be described by a trade-off between imaging flexibility and imaging capability (Table 1.1). By flexibility, we refer to the ability for a given phenotyping system to be easily operated across a range of crop types and field locations. Higher flexibility permits use in a wider range of phenotyping experiments and allow for a greater range of environmental conditions to be assessed as the system can be transported to (or replicated at, in a cost-effective manner) different trials across a particular growing region or across different growing regions around the world. Flexibility also includes the ease at which a particular system can be operated and the degree of specialized training required to operate it. Finally, we also include the quantity of images collected as an aspect of flexibility, whereby a
more flexible phenotyping system can gather a higher quantity of images, but perhaps at a lower quality per image. By imaging capability, on the other hand, we refer to the variety and quality of image data produced by a phenotyping system. For example, a system that can support a heavier payload can be simultaneously outfitted with multiple sensors and afford larger and heavier sensors such as hyperspectral cameras [15]. Imaging capability also includes the final ground resolution of the imaging system as high-resolution images better capture the fine-scale features of a particular plant. High capability includes imaging systems that capture data other than spectral data, including 3D shape and structure of a plant and/or crop canopy. Finally, the capability also refers to the ability for an imaging system to precisely and accurately place sensors at specific locations within the field, e.g. at the precise height above the canopy, or at the same location within a plot on subsequent days.

<table>
<thead>
<tr>
<th>Imaging Flexibility</th>
<th>Imaging Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can be operated in a range of weather conditions</td>
<td>Can capture with high spatial resolution</td>
</tr>
<tr>
<td>Can be transported or replicated at different locations</td>
<td>Can capture with high temporal resolution</td>
</tr>
<tr>
<td>Can be operated with minimal training</td>
<td>Can capture with multiple sensor types</td>
</tr>
<tr>
<td>Can be replicated at low cost and without specialized engineering requirements</td>
<td>Can capture information other than reflectance, e.g. 3D shape / structure</td>
</tr>
<tr>
<td>Can capture a large quantity of data, but at a lower quality per image</td>
<td>Can precisely, accurately and repeatably position sensors at known locations within a field</td>
</tr>
</tbody>
</table>

Table 1.1: Attributes of flexibility and capability in field-based high-throughput phenotyping systems

Figure 1.1: The spectrum of field-based phenotyping systems in terms of flexibility versus capability. Our proposed GrowPro system is highly flexible, with higher capability than aerial drones in terms of ground resolution.

Using this characterization of flexibility versus capability, we can differentiate the main classes of field phenotyping systems as shown in Figure 1.1. On one end of the HTP flexibility/capability spectrum, aerial drones are a highly flexible phenotyping platform. Drones can be used to quickly capture images of an entire field trial, e.g. flying at 15m altitude a 1-2 acre field can be imaged with sufficient image-to-image overlap to create a stitched orthomosaic image of the entire field. Drones can also be easily transported between field sites and drone hardware is now commercially available around the world. Costs for high-end imaging drones
remain relatively high ($10,000-$50,000), but substantially less than typical fixed field phenotyping systems. The main inflexible aspects of aerial drones relates to the inability to fly in windy conditions, and a high degree of specialized training required to operate. In particular, drone flying is heavily regulated in many jurisdictions for non-recreational use, i.e. for research or commercial use, even in isolated regions such as agricultural fields. Drones have a substantially lower imaging capability, as compared to other phenotyping systems, primarily due to the severely restricted payload weight and battery life of commercial drone aircraft. Despite the proliferation of high resolution RGB sensors (e.g., 4K video, > 20 MP still images) aerial drones typically have a minimum flight altitude, either due to design or due to wind instability close to the ground, which limits overall ground resolution for the images. Drone positioning while capturing images is also limited by GPS resolution and instability due to wind, therefore aerial images are typically stitched into an orthomosaic image of the entire field from which individual plots can be segmented and analyzed.

On the opposite end of the spectrum, fixed in-field imaging systems have a tremendous capability for imaging crops. For example, a gantry crane installed within a field can deploy a large and heavy array of sensing equipment at precise locations directly above the crop canopy [55], [39]. Therefore, fixed systems can generate very high-resolution images with precise plot localization. A large imaging payload permits a variety of data to be captured simultaneously, including high-resolution visible images, canopy height (ultrasound), 3D shape (laser scanning), temperature (thermal imaging), hyperspectral imaging, fluorescence imaging, etc. However, as their name suggests, the capability of fixed field systems comes at the cost of flexibility. In particular, spatial field flexibility suffers as these systems are most often permanently installed at one field site. Therefore, while able to characterize a single one-acre field in great detail, such characterization is associated with a single growing environment. Further, the high cost associated with fixed field installations (+$1M for a commercial gantry system [Rothamsted Lemntec] prevents them from being practically replicated at multiple growing sites or in regions of the developing world that do not have sufficient research infrastructure. Another barrier to performing experiments on a particular crop of interest with fixed-field systems is the requirement for seasonal crop rotation, which reduces the size of each experiment (so that three rotating crops can be planted within the workspace) or limits a particular crop to be imaged only every three years (due to the typical three-crop rotation in many parts of the world).

Close to the fixed field systems on the HTP spectrum, but more flexible, are vehicular HTP systems. These systems blend the larger payload capacities of fixed field systems with mobility afforded by aerial drones, though the capability nor the mobility matches those of these counterparts. Motorized vehicles can support moderately heavy payloads, but cannot localize and stabilize sensors as precisely as fixed field systems [35]. Manually-powered vehicles, such as push carts, require a lighter payload (so that they can be pushed), but can be more flexibly adapted to a wider variety of breeder plot layouts because the wheels are smaller and wheelbase more easily configurable [15]. The use of either type vehicle is restricted in wet muddy field conditions. Further, aerial drones (and gantry cranes) can operate over any field layout; however, vehicles require pathways for their operation. The flexibility required here should not be understated. Each crop and
breeding experiment can have very different structure in terms of the size, spacing, and layout of individual plots. For example, early generation nursery breeding trials are often limited to single row plots (due to the small amount of germplasm available) while late generation yield trials typically have larger multi-row plots. The spacing between blocks of single-row plots or between yield plots can also vary from large alleys, to single wheel tracks, to no separation at all. Due to the high costs of phenotyping vehicles ($50,000-$250,000) it is common for one vehicle to be used across crops and experiments, which therefore requires vehicle flexibility. While experimental layouts can sometimes be adapted to accommodate a particular phenotyping vehicle, the range of possible plot sizes and layouts can be restricted by the particular seeding equipment used. Moreover, until the utility of HTP systems is proven within a breeding program, there can be significant resistance from breeders and field technicians to changing the status quo.

**Flexibility as multi-dimensional feature space**

For a better understanding of the place of the GrowPro system in the HTP spectrum, we define flexibility in terms of task and conditions. UAV systems may have higher flexibility in terms of the task compared to the GrowPro, because well-developed UAVs can automatically perform preliminarily written tasks, such as covering a specified area on the field by continuously capturing images while following a defined path in space. Such flying option allows automatically covering substantial areas with the required image-to-image overlap. However, UAVs are less flexible in terms of conditions, because the performance of these systems strongly depends on weather conditions. Besides, to operate UAV systems, an operator needs specialized training and/or licensing, which makes these systems more difficult for scaling and also limits overall ease of use. In this perspective, UAV systems have low flexibility in terms of conditions.

On the other hand, the GrowPro system has lower flexibility in terms of the task, because the GrowPro system requires more human-machine interaction (an operator is supposed to be walking and carrying the camera rig, taking control over the filming procedure and camera position in relation to plants of the field). The more manual filming procedure also leads to more human errors, whereas UAV filming procedure can be completely automated. However, the GrowPro system is more usable across different weather conditions. Also, there is no need for specialized training to operate the GrowPro system which turns this system into an easy-to-scale solution for field phenotyping and remote sensing.

**1.1.2 Software point of view**

HTP systems with more flexibility implies different scenarios for remote sensing and imaging. Such cheap, flexible and replicable systems often produce substantial amount of data which subsequently requires more extensive data processing and well-developed methods for data collection and management, data verification etc. Thus, a reliable, stable and easy-to-use software platforms are very often necessary for effective data processing. In other words, there is more burden on the software analysis of data, acquired with more flexible remote sensing tool. Hence, there is a need for more extensive pre-processing pipeline.
1.2 Problem Statement

We have identified a gap in the HTP spectrum [1] for a system with higher image resolution capability than aerial drones, but with greater flexibility than push-carts so that it can be easily used and replicated for breeding trials with any field layout. Moreover, for HTP systems that can potentially fill the gap in the defined HTP spectrum, there is a high demand for an automated or semi-automated data pre-processing tool to provide overall flexibility and ease of use of the entire system and turn such system into an end-to-end solution. By end-to-end solution we refer to a complete pipeline that would consists of the two mains steps: data gathering step and the data pre-processing step.

Because the current UAV approaches cannot produce images of the entire field with resolution, comparable with closer-view images of plants (e.g. taken at 1-2 meters above the canopy), the only way to obtain a high resolution image of the entire area or field is to build an orthomosaic image of the field by stitching many single closer-view overlapped images. Higher up imaging from the drone though produces less amount of data, whereas closer-view imaging with 1s or 0.5s capturing step (so-called time-lapse mode) would generate more than 250MB of data in less than a minute. Thus, closer-view imaging of plants will produce a significant amount of high-resolution detailed images that will require stitching into a single orthomosaic image of an area of interest (in our experiment, a range is a minimal unit of the field).

From this perspective, data pre-processing pipeline can be considered as one of the most important process within an HTP system (following after data collection by a remote sensing tool) and often consists of multiple sub-processes, such as data verification and labelling using image processing and/or machine learning techniques. Hence, a software package that can handle a substantial amount of data with the ability to effectively classify, label and verify a data stream is in high demand.
1.3 Proposed Solution

We propose a novel field HTP system, called the GrowPro, consisting of a downward-facing camera that is held out over the crop canopy on a pole, stabilized with a motorized gimbal, and set to continuously capturing images while the operator walks along a row of plots. The GrowPro can support the same imaging payload as an aerial drone, but placed at a height much closer to the ground, which generates images with comparably higher ground resolution. Coverage of an entire field can be accomplished more quickly with a drone, given the larger field-of-view per image, and the GrowPro does require the manual effort of walking through the field. However, the GrowPro is ten times less expensive than the least expensive aerial drone and is more flexible in terms of requiring little operator training and being able to operate in windy weather.

To support and increase flexibility and capacity of the GrowPro system, we also propose a software solution called GrowPro Viewer that will allow users to semi-automatically process substantial amount of data. By data processing we refer to a sequence of steps necessary to generate high resolution orthomosaic images from captured RAW images of plants. Such pipeline would include image classification and labeling step, data verification step, preliminary evaluation of data accuracy and data completeness step, preparing images for stitching and forwarding processed data for orthomosaic generation. We also propose preferred practice for filming procedure, data gathering procedure, ways of effective GrowPro camera rig and GrowPro Viewer application usage and data management.

1.3.1 Similar approaches

Researchers can use cheap digital action cameras in their research projects. Due to the availability of a substantial amount of action cameras of the market, small-size digital RGB sensors become good solutions for remote sensing in field phenotyping. For example, low-cost action cameras have been used in phenotyping applications, such as mounted to the side of trucks within a vineyard to capture continuous images during routine field maintenance [14]. In this experiment an action camera was mounted horizontally, and was capturing video at 30 frames per second. Video recording was based on per row basis.

Focused on developing a new simple chip methods for HTP that could potentially increase how widely HTP is adopted, researchers used an inexpensive system of hemispherical imaging to identify a set of key Quantitative trait loci (QLTs) for biomass production [8]. Hemispherical images were taken with a GoPro Hero 3+ digital action camera with a fully hemispherical lens with a self-leveling gimbal.

An image acquisition system composed of a high-resolution DSLR camera was used for vineyard yield estimation by [19]. This system was mounted on an utility vehicle driven by an expert human operator at a fixed pace and at an approximate distance of two meters from vines. In this experiment images were captured at a fixed interval with no overlapping. However, we report here the first purposeful use of action cameras for HTP to capture full coverage of breeder crops in a similar manner to aerial drone and vehicular phenotyping systems and a complete end-to-end software solution for image pre-processing. Considering the
GrowPro system, we see the following advantages over similar approaches:

- The ability to generate high-resolution orthomosaic images of an area of interest.
- The GrowPro system can be used by a human only, whereas similar approaches implied using of a car, a walk platform or a stationary system to carry necessary sensors.
- The action camera used in the GrowPro system is adaptive to lightning which increases usability of the whole system and allows imaging without extra auxiliary artificial light. In contrast to similar systems where extra lighting is used which reduces overall flexibility of the system.
- The proposed GrowPro system comes with a software application that allows users to check and verify data completeness, label images based on a field layout, run image segmentation using Machine Learning techniques and generate orthomosaic images for specific areas of interest or for the entire trial (field). The proposed software application should accommodate all necessary steps for generating orthomosaic images out of preliminary acquired RAW images and represents a fully completed solution for data pre-processing.
- The use of preliminary GPS noise estimation to predict overall stitching accuracy for each area of interest. Such noise estimation is especially useful before the actual stitching process because orthomosaic generation is the most time-consuming step within the entire pre-processing pipeline.

Thus, the GrowPro system can potentially become a cheap, reliable and competitive solution for field phenotyping.

### 1.4 Contributions

Our intended application leads to a number of methodological questions, including:

- Is it possible to stitch together a sequence of top-down images to reconstruct a single image of a large breeder plot (e.g. 6m x 2m) when the images are captured at a low height and therefore have significant parallax differences between the plants from one image to the next?
- Is it feasible to generate orthomosaic stitched images of an entire field consisting of a large quantity of small close-up images as compared to a smaller quantity of larger images as is captured by an aerial drone?
- Will GrowPro imaging work sufficiently well for different crop types and different field layouts?
- Will variable lighting and weather conditions in the field affect the quality of images captured with the GrowPro?
While trying to answer these methodological questions, we have defined the following contributions of our research work:

- Development and deployment of the GrowPro for weekly imaging of seven different breeding trials across four different crop types (b. Napus, b. Carinata, Lentil, Wheat).

- Data acquisition and training others to use the GrowPro. We formulated a procedure of initial data collection, GPS-based data verification and range labelling procedure, implemented a solution for GPS adjustments for increasing stitching accuracy. We compared different operation techniques that can be applied for moving the camera over plants (different walking patterns, a speed of walking and height of the camera above the canopy) to determine which setup is most likely to produce satisfactory orthomosaics.

- Development of a semi-automated system, called GrowPro Viewer for data pre-processing and handling that generates stitched orthomosaics on a per plot bases and localizes and labels each plot. This software solution provides tools for data verification and labelling based on a field layout, GPS data corrections and automatic image stitching and orthomosaics generation. Machine Learning solutions for GPS-based range segmentation have been tested and evaluated.

- Development and implementation of an effective programming solution for GPS noise evaluation over a substantial number of image sets, as well as a set of additional software solutions for data management, such as synchronization of the output with a data server and linking output results with a database.

- Evaluation of stitching results across different conditions (type of crop, weather, etc.).

In this project, we have shown that the GrowPro hand-held remote sensing platform allows getting ground resolution superior to a commonly-used drone, as well as a closer view of plants compared to UAV approaches. We have shown good orthomosaic generation results across different types of a crop (Wheat, Canola, Carinata, Lentil), different stages of growth, different weather conditions (windy/sunny/cloudy days).

Here in this research work, we introduce an imaging system with greater flexibility, but that requires more complicated software for managing and processing images after they are captured. Ease of use, high-resolution sensor and good scalability of the system are the primary reasons for acquiring a vast amount of data that subsequently requires effective and complicated data processing and analysis techniques. In other words, high flexibility of the remote sensing system leads to more degrees of freedom in terms of pre-processing data analyses and the number of methods and algorithms that could potentially be used to implement these analyses.
CHAPTER 2

BACKGROUND RELATED WORK

The spectrum of field-based phenotyping system described in [1] is presented with many different remote sensing tools. Each tool takes specific place on the spectrum, where a position on the axis “Flexibility - Capability” is based on their features and characteristics. We grouped currently used HTP systems into three main groups: Airborne systems, Fixed In-field Systems and Mobile Vehicular Systems. The following sections look at each group more closely.

2.1 Airborne Systems

Implementation of UAVs for field phenotyping depends on characteristics of the UAV (stability, control, positioning, sensor mount, etc.) and sensor characteristics (specific spectral wavelengths, resolution, weight, calibration, the field of view, etc.) [37], [15]. One of the most commonly used types of UAV is Rotorcopters (DJI Inspire, Mikrocopter ARK, OktoXL 6S12, Yamaha RMAX). These tools can be used with waypoint navigation, have good hovering capabilities and can different sensors from thermal, multispectral to hyperspectral. However, a good payload may decrease battery usage and flight time [47].

Helicopters allow users to do rapid time-sensitive measurements of a large number of plots. A customized robotic helicopter was used for aerial imaging and field phenotyping by [13]. Unmanned helicopters are used for estimating yield and total biomass [49], thermal and narrowband multispectral remote sensing [11], determination of seasonal changes and variety differences in the brightness of the leaf color [40].

Core advantages of fixed-wing UAVs (Landcaster Precision Hawk, senseFly eBee) are waypoint navigation, better flight time and ability to operate with multiple sensors mounted. However, it has the limited hovering capacity and high speed of flight (lower speeds are required for image overlapping and stitching). Fixed wing UAVs are used for remote sensing of crop biomass and nitrogen status [27], acquisition of bidirectional reflectance factor dataset [23]. Powered paragliders are used in remote sensing and phenotyping mainly because they can carry substantial loads, use proven technology and low-cost [50].

A research group IMAP3D use a low-cost blimp for aerial photogrammetric applications. This device is based on helium zeppelin and mounted digital camera with video and radio control [38]. Blimps are also used for monitoring the growth characteristics of cotton in an irrigation study [44], detecting the attributes of a wheat crop [30] and field photography. Blimps allow carrying multiple sensors on board for field photography
Researchers in Belgium developed a new form of close-range aerial photography based on a camera suspended from a compilation of both a helium balloon and kite wings [54].

2.2 Fixed In-field Systems

Fixed systems include those where an array of sensors can be automatically transported to different areas of the field. The advantages of Fixed in-field imaging platforms include an excellent capability for imaging crops, ability to obtain high-resolution images with precise plot localization over large areas. Some examples of such devices for field phenotyping include the system designed for high-throughput monitoring of crop performance. A fully automated robotic field phenotyping platform, mounted on fixed rails was designed by [55]. Sensors of this platform can be positioned in three-dimensions and can produce a detailed description of canopy development with high accuracy and reproducibility.

Similarly, common irrigation systems can be used as a positioning tool for transporting an array of sensors to a specific place in the field. One of the examples of such devices is the platform with a two-span linear-move irrigation system for moving a set of reflectance and infrared sensors. [33]. Another application of the irrigation system as a transporting tool is the system used to simultaneously monitor water status, nitrogen status, and crop growth at 1m spatial resolution [22].

Thermal camera and digital RGB camera were mounted on a truck-crane about 15 m above the canopy by [41] for field phenotyping. A fork-lift at about 8 m height was used for thermal imaging of Solanum tuberosum by [43].

2.3 Mobile Vehicular Systems

Mobile vehicular HTP systems afford better mobility and precise positioning of the sensors. Motorized tractors have been used in agriculture for many years and have become efficient platforms for carrying various phenotyping systems. Tractors can carry a set of sensors for simultaneous measurements over large areas and the number of plots [7].

Mobile phenotyping system varies from small pushcarts to driver-operated phenomobiles. The most straightforward system is flexible lowcost hand-controlled push cart with an array of sensors mounted on wheeled frame [58]. The next level of complexity is autonomous field robot platform for individual plat phenotyping [46] and robotic carrier for precision agriculture research [29]. Another example of the middle-class mobile system is a cable robot with the 3D sensor and a set of phenotyping sensors on board for automatically distinguishing between plant species, designed by [56]. On the next stage of development are motorized vehicles represented by manned buggies and phenomobiles. Examples include a high-clearance tractor for Geo-referenced crop canopy spectral measurements [9], remote sensing buggies for large-scale field phenotyping with LiDAR, RGB camera, thermal infra-red camera and imaging spectroradiometer sensor.
on board [15]. A high-clearance tractor was used for canopy height, temperature, and spectral reflectance measurements [57].

<table>
<thead>
<tr>
<th>System type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed wing UAV</td>
<td>Waypoint navigation, better flight time and ability to operate with multiple crop sensors</td>
<td>Limited hovering capacity and high speed of flight.</td>
</tr>
<tr>
<td>Powered paraglider</td>
<td>Can carry substantial loads, use proven technology, and low-cost.</td>
<td>Needs a professional operator, weather dependent, limited hovering capacity, limited hover altitude.</td>
</tr>
<tr>
<td>Blimps, Balloons, Kites</td>
<td>Relatively cheap compared with other aerial platforms, good hovering ability, long flight time</td>
<td>Limited to low wind speed, low localization precision, limited payload.</td>
</tr>
<tr>
<td>Fixed</td>
<td>Unmanned repetitive operations available, precise locations, large array of sensors, high resolution</td>
<td>Generally expensive; can only monitor a very limited number of plots, permanently installed, does not follow seasonal crop rotation.</td>
</tr>
<tr>
<td>A robotic platform, mounted on fixed rails</td>
<td>Capability for imaging crops, ability to obtain high-resolution images with precise plot localization.</td>
<td>Limited area of crop cannot be scaled, only one trial available for constant measurements.</td>
</tr>
<tr>
<td>Two-span linear-move irrigation system</td>
<td>Give precise, high resolution images from a fixed angle, multiple measurements available.</td>
<td>Very expensive, can be used only for a limited area.</td>
</tr>
<tr>
<td>Fixed Unmanned repetitive operations available, precise locations, large array of sensors, high resolution.</td>
<td>Good payload, weather independent.</td>
<td>Limited hovering capacity, good spatial resolution.</td>
</tr>
<tr>
<td>Autonomous field robot platform</td>
<td>Continuous operation, good payload, multiple measurements available.</td>
<td>Expensive, as commercial solutions available, take a long time to cover a large-scale experiment.</td>
</tr>
<tr>
<td>Manned buggy/Phenomobile</td>
<td>Flexibility with the design, very adaptable, can carry multiple sensors, good for large-scale experiments.</td>
<td>Expensive.</td>
</tr>
</tbody>
</table>

Table 2.1: Available types of remote sensing platforms

2.4 Image collection management

Digital photography popularization, as well as technological breakthrough in developing compact digital cameras, have led to a sharp increase in the number of images people capture in their daily life. Such a vast amount of data can no longer be sorted manually. Large image collections are difficult to navigate due to their size and difficulties for computers to understand the content.

Current approaches to the image collection management issue can be grouped into the following clusters:

- **Descriptive labelling of items in the image set** [48].
  - **Visual scanning and inspection of images and content definition** [10].
  - **Automatic image content identification techniques. This approach has become the basis for substantial research work in the area of Image Processing and Machine Learning for image recognition and content definition** [28, 53].
  - **Automatic generation of metadata with a list of categories and its values, such as authorship, timestamp, geo-location references, exposure, etc.** [21, 24].

The location information category that is embedded in metadata file of an image is now successfully used for image collection management [42, 52]. Location information obtained for each image refers to the location of the camera at the time of capture. The location is represented by three main parameters: latitude, longitude
and altitude. Such geo-referenced images can be grouped geographically or displayed on a digital map, which provides spatial context. By using this spatial distribution, a user can successfully identify data patterns or clusters (for example, images that belongs to a specific area on the field) for further data analysis.

2.4.1 Image clustering and classification

The range segmentation and image labelling problem can be solved with either supervised or unsupervised learning approaches. If considering supervised learning techniques, the range segmentation problem can be transformed into a multi-class classification problem. In this case, each range can be viewed as a single class. Because in the experiment all ranges are explicitly separate from each other, in the multi-class classification problem all class labels would be mutually exclusive.

For image classification problem, preliminary labeled and manually classified images can be used as a training set for a given classification algorithm. Image classification/clustering is supposed to be done based on GPS coordinates of images. The field layout remains the same throughout the season, hence there’s no need to generate a training set each time after the filming session.

On the other hand, if we consider unsupervised learning techniques, the range segmentation problem transforms into a problem of clustering. The filming approach is repetitive, that is why it produces distinct patterns in the GPS dataset (tracks of points, snaking patterns, changes of direction, etc.). Clusters in this perspective refer to ranges in the field. In that case, the number of clusters is known a priori, because the number of ranges is always constant.
2.5 Orthophotos

Orthophotos are one example of many applications of photogrammetry. Orthophotos (also named as orthomosaic images) combine geometry and realistic representation of subjects to give users a metric visualization of the area. Building orthophotos from remotely captured images is popular and used in many fields, such as cartography, space industry, environmental monitoring, civil engineering, city planning, etc.

An orthophoto is a remotely captured photograph or image geometrically modified and adjusted so that the scale is fixed and uniform: images that form an orthomosaic photo have the same lack of distortion as a cartographic map (Figure 2.1) [31]. An orthophotograph can be used to calculate real distances because orthophotos represent an accurate visualization of the Earth’s surface [2].

Figure 2.1: Orthophoto sample

Any aerial images can potentially be influenced by a phenomenon known as relief displacement or hidden areas. This effect represents the geometric distortion that takes place due to height irregularity in the terrain being captured. High objects, such as buildings, mountains and trees will be misplaced on the periphery from the centre of the image. The taller the object and the further the object is from the centre-of-view, the greater the radial distortion.

Ortho-rectification is the process that corrects such distortions by performing a mathematical transformation on the image. Algorithms that perform ortho-rectification consider the shape of the terrain represented by a digital terrain model (DTM). Aerial orthophotos can be generated following the typical photogrammetric workflow that consists of the following steps:

- Image orientation and alignment. This step is sometimes based on location information of geo-tagged
images [4].

- Re-projection with a digital terrain model (Ortho-rectification) - the process of transforming pixels in a scanned image to their proper location in a digital orthophoto [51].

- Image mosaicing. The most technically challenging part of the pipeline, because the rectification process moves subjects into their proper locations but will also leave holes or blind spots in the place from which they came. Hence, considerable overlap between photographs is critical for filling the blind spots and creating images with the most coverage possible [51].

The orthomosaic image quality depends on image resolution, camera colour calibration accuracy, camera orientation, and DTM accuracy [34]. Orthomosaic generation error becomes more significant when images are gathered with UAV systems mainly because of their lower sensor resolution and detailing at high altitudes [16]. Relief displacement and artifacts caused by shadows and image mosaicking are the leading causes of information loss in large orthomosaics. Both types of damage significantly reduce the overall quality of orthophotos. At the dawn of orthophoto technology, for solving the problem of information loss caused by differences in elevation a new term was introduced in [6]. True-orthophoto - an orthomosaic image that can be built by using a digital surface model (DSM). However, a true-orthophoto sometimes leads to additional artifacts, such as self-occlusions.

2.6 Discussion

Review of the use and implementation of HTP systems for field-based phenotyping, as well as deploying image processing techniques has shown that there is still room for improvements, mainly due to diversity and flexibility of the methods and tools used alongside with HTP systems (image processing, machine learning).

Such opportunities allow implementation and deploying not only high-end technologies and pieces of equipment but also easy-to-purchase publicly available solutions (cameras, accessories, integrated development environments for programming, etc.) for building a new and unique filming tool. The GrowPro project itself would represent and symbolize such flexibility and variety in the use of available equipment and well-established data analyses techniques.
CHAPTER 3

GROWPRO FILMING TOOL AND DATA ACQUISITION

In this chapter, we describe the development of the GrowPro filming tool. We define a functional design step, equipment research step and tool assembly step. We also give recommendations on best practices for filming and describe data managing and handling. Finally, we provide an overview of data collection during the 2017 growing season across Canola DFL, Canola DRIL and Carinata field trials.

3.1 Functional design

Using UAVs, fixed in-field systems or mobile vehicles for field-based remote sensing and phenotyping sometimes may have some restrictions, so we aimed to create a new remote sensing platform with technical characteristics based on recommendations by the researchers in this domain. Unfortunately, resolution and quality of images taken from UAV platforms sometimes are not enough for carrying out efficient computer analysis. Low-quality images can affect the accuracy of image analysis and machine learning algorithms used for estimating phenotypes, such as emergence vigor, biomass, etc. Therefore, a closer-view, high-resolution sensing platform is needed. Elimination of weather factors is also important because sometimes wind can adversely affect the UAV imaging. The required product is supposed to be a hand-held filming tool an operator carrying for a relatively extended period, hence the total weight of the product should be the minimal possible. Operating the system shouldn’t need specialized training and should be easy to use. An opportunity to teach a technician of how to use the device is also important, especially when it can be done in the shortest possible time.

According to Canadian Legislation, when flying a UAV in Canada for non-recreational purposes, you must follow the rules in the Canadian Aviation Regulations (section 602.41 Unmanned air vehicles) and undergo specialized training, so a new remote sensing vehicle should be designed in such a way that does not fall under these regulations.

The image acquisition apparatus should:

- Ability to acquire RGB images of a subject with highest possible spatial and temporal resolution.
- Weather independent (wind-resistant and stable performance)
- Minimal total weight possible.
Operating of the system shouldn’t need specialized training.

Long battery life

Low-cost

### 3.2 Equipment research

Choosing the right digital camera is essential for the correct running of the system and post-processing. Post-processing software has specific requirements for images that are used for building orthomosaic pictures of subjects. In the workflow that we have built, Agisoft Photoscan software is used for creating orthomosaic RGB images. This software requires that all photos have geolocation information. GPS data of each image is used for aligning images that form an orthomosaic image. Orthomosaic images can be created only when overlap between two images exceeds 60%. To obtain this overlap, a camera with an option to take still images in a continuous mode with the smallest interval between photos available, or take high-resolution video footage. In case of video footage, the resulted video file will have to be transformed into a set of still images. Based on the results of camera comparison, GoPro Hero 5 Black camera was chosen because it meets the requirements for the image acquisition as well as for the pre-processing. The following table 3.1 describes the main technical parameters of the cameras tested.

<table>
<thead>
<tr>
<th>Features</th>
<th>Xiaomi Yi</th>
<th>GoPro Hero 5 black</th>
<th>DJI Osmo</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS data</td>
<td>Not available</td>
<td>Available</td>
<td>Available</td>
</tr>
<tr>
<td>Time-lapse photo mode</td>
<td>Separate still images</td>
<td>Separate still images</td>
<td>Video file time-lapse</td>
</tr>
<tr>
<td>Image stabilizer</td>
<td>No</td>
<td>Yes</td>
<td>Built-in gimbal</td>
</tr>
<tr>
<td>Image resolution</td>
<td>12MP</td>
<td>12MP</td>
<td>12MP</td>
</tr>
<tr>
<td>Video resolution</td>
<td>4k wide</td>
<td>4k wide</td>
<td>4k linear</td>
</tr>
<tr>
<td>Battery</td>
<td>Built-in, can be charging while filming</td>
<td>Built-in, can be charging while filming</td>
<td>Built-in, a special external charger is needed</td>
</tr>
<tr>
<td>Price</td>
<td>$300</td>
<td>$400</td>
<td>$800</td>
</tr>
</tbody>
</table>

**Table 3.1:** Camera comparison
Hand-held camera systems imply a degree of instability and shaking, despite there are many digital cameras which has built-in stabilization system. For high quality, reliably captured images, an external additional stabilizing system is needed.

Electronic gimbals can be used for stabilizing digital cameras. These devices are widely used because of its usability, convenience, and relatively low price. In our experiment, we need a device that can be continuously charged during the filming process to increase the battery life of the camera rig.

FeiyuTech G5 Triaxial Handheld Gimbal (Figure 3.1) suits all our needs well. All specifications and technical characteristics can be found on the official website and or user manual.

### 3.3 Camera rig description

We suggest a camera rig, where a telescopic plastic pole (monopod) is the main frame of the rig (Figure 3.2). All electronic equipment is mounted on the monopod. The camera rig consists of the following parts (Table 3.2):

<table>
<thead>
<tr>
<th>Part</th>
<th>Model</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action camera</td>
<td>GoPro Hero 5</td>
<td>$390</td>
</tr>
<tr>
<td>Gimbal</td>
<td>FeiyuTech G5 Triaxial Handheld Gimbal</td>
<td>$290</td>
</tr>
<tr>
<td>Monopod</td>
<td>AmazonBasics 67-Inch</td>
<td>$19</td>
</tr>
<tr>
<td>Remote control</td>
<td>Smart GoPro Wi-Fi</td>
<td>$52</td>
</tr>
<tr>
<td>Ball head camera mount</td>
<td>1/4-inch universal thread mount</td>
<td>$11</td>
</tr>
<tr>
<td>20000 mAh Battery pack (two items)</td>
<td>AUKEY 20000mAh Portable External Battery Charger Power Bank</td>
<td>$40</td>
</tr>
<tr>
<td>USB extender 3ft (two items)</td>
<td>iMBAPrice 3 Feet USB 2.0 Extender Cable</td>
<td>$22</td>
</tr>
<tr>
<td>Lightweight retractable USB cable</td>
<td>Cable Matters (2-Pack) Retractable USB to Micro-USB</td>
<td>$25</td>
</tr>
<tr>
<td>USB to microUSB cable</td>
<td>CableCreation 90 Degree USB 2.0 A to Micro USB B Cable</td>
<td>$11</td>
</tr>
<tr>
<td>microUSB to USB-C adapter</td>
<td>Anker USB-C (male) to Micro USB Adapter (female)</td>
<td>$11</td>
</tr>
<tr>
<td>Battery charger hub (4 outputs)</td>
<td>PowerPort Speed 4 Ports</td>
<td>$30</td>
</tr>
<tr>
<td>Camera shoulder strap</td>
<td>LNKOOO Camera Strap</td>
<td>$20</td>
</tr>
<tr>
<td>Assembly</td>
<td>GrowPro</td>
<td>$921</td>
</tr>
</tbody>
</table>

Table 3.2: Part list of the proposed camera system
3.3.1 Assembling of the camera rig

To assemble the camera rig, GoPro camera must be fixed to the gimbal. The gimbal should also be mounted on the monopod using the ball head (Figure 3.2). It is necessary to point the gimbal down to maintain the required one degree of freedom.

The battery pack can be attached to the monopod using tape. USB extenders are used to plug the camera and the gimbal to the battery pack. So, in this case, we need to build a chain of cables, which goes from both devices to the battery pack. First cable chain goes from the camera (micro USB to USB-C adapter + lightweight USB cable + USB extender), second cable chain goes from the gimbal (USB to micro USB cable (short) + USB extender). USB extenders can be fixed with tape along the monopod (Figure 3.3).

![Camera rig](image_url)

**Figure 3.2:** Camera rig

![Parts of the camera rig](image_url)

(a) Camera and gimbal  (b) Battery pack  (c) Gimbal in use

**Figure 3.3:** Parts of the camera rig
3.3.2 Best practices for filming

A hand-held filming approach inevitably involves human factors and sometimes may be inconsistent and unstable. Additionally, the camera rig consists of components that need to be adjusted before filming. All necessary equipment and components should be assembled correctly. The camera should always work in a correct mode (time-lapse images with 0.5sec interval); a stabilizing system should be adjusted each time before the filming starts. Moreover, an operator should move evenly and with the regular constant pace (Figure 3.4). Control of all of these factors can make filming procedure complex. To avoid issues and difficulties, and make filming easy for an operator, a step by step guide has been made, so an operator might follow the guideline each time to avoid errors (See appendix A).

![Filming procedure](image)

**Figure 3.4:** Filming procedure

3.3.3 Walking patterns

Depending on the size of a minimal unit (i.e., plot, range, row) filming should be done differently. We suggest two canonical walking patterns. If the width of a field unit is small enough to fit in the field of view of the camera, then filming can be done with a one-pass walking pattern alongside the unit (Figure 3.5). Arrows represent walking direction; green areas represent extra areas that should be fit into the frame to have an entire area of the minimal field unit covered. Grey rectangles represent a stack of images taken in series.
If a field unit is wide and cannot be covered entirely in a single pass, then a two-pass walking pattern needs to be used. An operator steps along one side of a field unit then turns around and keeps walking along the other side of the unit. This approach will allow the total width of the field unit covered.

In the case of a two-pass walking pattern, 60% transverse and 80% lengthwise overlapping is needed (Figure 3.6). The blue area on Figure 3.6 represents overlapping between two stacks of images correspond to two different passes.

While filming with the two-pass walking pattern, an operator should try to avoid shadows on the plants by choosing the walking direction depending on the position of the sun. Narrow plots can be filmed with one-pass walking pattern and at a lower level which increases overall ground resolution.
3.4 Data collection procedure

Unlike UAV images, GrowPro images sometimes cannot be stitched into orthomosaic images of the whole field, mainly because of the much closer position of the GrowPro camera to the canopy and smaller field of view. Close position to the canopy and narrow field of view of the camera may lead to insufficient covering of the entire field.

Therefore, stitching of images captured by the GrowPro can only be performed per each minimum unit of the field (i.e., range, row or plot). Since the GrowPro just focuses on ranges of plants within the field, and not on the entire area, a whole-field orthomosaic image cannot be obtained. However, larger number of images produced by the GrowPro compared to UAV tools leads to better spatial resolution (See 5.2).

In this case, plot segmentation programming tool that is typically required for images taken by UAV filming tools, cannot be used for pictures taken by the GrowPro. Instead, we first collect and group unstitched RAW photos that correspond to a single range, then stitch RAW photos of each grouping set to get orthomosaic images of all ranges, and then crop out individual plots.

To group images that represent one range, during filming procedure, we use remote control clicker and film only area of the range, pausing recording at the end of each range. After pausing, the system of the camera automatically groups all images into a separate directory. Going through the entire field and pausing recording at the end of each range, after the entire filming session (filming of each range in series) we obtain RAW images of plants that grouped according to the field layout. However, during the beginning of the experiment, for early filming dates of the 2017 growing season (May 30th - June 12th), the remote clicker wasn’t used and image data-sets for those filming dates were represented as unsorted stacks of images.

After a filming session, we use “GoPro Quik” application to upload all images from the camera to external storage. This step is critically important for transferring the file structure from the camera to external storage. As a result of using “GoPro Quik,” we have an image data set with the file structure that represents the field layout (Figure 3.7).
3.5 Amount of data collected

For the experiment of seven different breeding trials across four different crop types have been observed and tested for filming with the designed camera setup. From the end of May to the end of September 7 fields of Canola plants, one field of Carinata plants, four fields of Lentil plants and two fields of Wheat have been captured. More than 890000 still images have been collected which exceeds 1900GB of data. These images will subsequently be processing for making orthomosaic pictures of plots.

The GrowPro filming tool due to its flexibility and ease of use allowed to obtain a significant amount of data (Table 3.3). Such approach requires a stable automated solution for processing all images with minor manual interference. In this case, a semi-automated pre-processing tool is in high demand.
<table>
<thead>
<tr>
<th>Field</th>
<th>Size</th>
<th>Filming days</th>
<th>Data gathered (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carinata</td>
<td>270x775ft</td>
<td>55</td>
<td>669</td>
</tr>
<tr>
<td>Canola DFL</td>
<td>88x186ft</td>
<td>13</td>
<td>71</td>
</tr>
<tr>
<td>Canola DRIL</td>
<td>88x465ft</td>
<td>19</td>
<td>230</td>
</tr>
<tr>
<td>Canola Seed Nursery</td>
<td>88x527ft</td>
<td>7</td>
<td>60</td>
</tr>
<tr>
<td>PhenoLentil</td>
<td>183x116ft</td>
<td>10</td>
<td>35</td>
</tr>
<tr>
<td>PhenoWheat</td>
<td>183x116ft</td>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>PhenoCanola 207R1</td>
<td>88x480ft</td>
<td>11</td>
<td>84</td>
</tr>
<tr>
<td>PhenoCanola 207R2</td>
<td>88x480ft</td>
<td>11</td>
<td>109</td>
</tr>
<tr>
<td>PhenoCanola 1</td>
<td>234x204ft</td>
<td>14</td>
<td>256</td>
</tr>
<tr>
<td>PhenoCanola 2</td>
<td>153x192ft</td>
<td>8</td>
<td>87</td>
</tr>
<tr>
<td>Nasser Lentil</td>
<td>85x33ft</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Rosthern Lentil Agile</td>
<td>312x72ft</td>
<td>6</td>
<td>94</td>
</tr>
<tr>
<td>Rosthern Lentil Agile BIO</td>
<td>85x33ft</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Sutherland Lentil Agile BIO</td>
<td>85x33ft</td>
<td>7</td>
<td>31</td>
</tr>
<tr>
<td>Sutherland Lentil Agile LDP</td>
<td>312x72ft</td>
<td>8</td>
<td>122</td>
</tr>
<tr>
<td>Overall</td>
<td>2342x3784ft</td>
<td>191</td>
<td>1915</td>
</tr>
</tbody>
</table>

Table 3.3: Amount of data collected during the 2017 growing season

### 3.6 Data Overview

The entire data set of acquired RAW images can be split in the following way:

- Images taken with one pass walking pattern (432259 images of PhenoCanola, PhenoWheat, PhenoLentil and Lentils):
  
  Non-grouped (26366 images, 6%).

  Grouped (pre-sorted) images by using remote clicker that automatically split sequence of images into different folders (405893, 94%).

- Images taken with two passes walking pattern (465160 images of Canola and Carinata)

  Non-grouped (29839 images, 6.4%).

  Grouped (pre-sorted) images by using remote clicker that automatically split sequence of images into different folders (435321 images, 93.6%).

The entire pre-processing procedure implies mapping between a region of interest (row, plot, range) and corresponding pictures in the image set. The most important here is to assign each subset of images to a
relevant area on the field. Each subset of images should be labeled in the same way as the actual field was labeled.

### 3.7 Field layout description

As stated above, in this experiment we focused on taking images of Canola DFL, Canola DRIL and Carinata fields. The Carinata field consists of 3 blocks of 50 ranges, the Canola DFL field consists of 12 ranges, and Canola DRIL field consists of 27 ranges. Each range consists of rows of plants spaced 1 foot apart. Each row is 10 feet long, and a range consists of 44 rows for Canola DFL and DRIL (Figure 3.8), and 46 rows for Carinata (Figure 3.9). Thus, each range represents a rectangular area of the field, adjoining each other with larger sides and also spaced apart.

![Figure 3.8: r02 range of Canola, captured on June 28th](image)

![Figure 3.9: r136 range of Carinata, captured on June 28th](image)

### 3.8 Discussion

The GrowPro hand-held sensing tool has taken an intermediate position between less flexible pushcarts and lower image resolution aerial drones. The GrowPro had the same payload as a flying drone but placed at a height much closer to the ground, which generates images with higher ground resolution. However, the GrowPro does require the manual effort of walking through the field and takes more time to film the entire field than a drone.
CHAPTER 4

DATA PRE-PROCESSING PIPELINE

The Pre-processing pipeline plays one of the most critical roles for acquiring high-resolution stitched images of regions of interest. As shown in the Background Related Work chapter, there are many many different approaches and techniques for image clustering and classification (2.4.1), orthomosaic generation (2.5) and data collection management (2.4). In the following sections, we propose and implement ideas for building a data pre-processing pipeline that produces orthomosaic range images of high resolution.

4.1 Pre-processing pipeline description

In this project we define 6 core steps in the pre-processing pipeline:

1. GPS-based range segmentation - grouping of RAW images based on image GPS data. The following parameters should be used for image classification or clustering:
   - Filming date (or stage of growth),
   - Walking pattern (one or two passes),
   - Field layout (per plot/range),
   - Missing data.

2. Prior image adjustments and corrections:
   - Colour calibration and adjustments,
   - Cropping out regions of the images that contains objects that may affect stitching accuracy (feet, shadows, etc.),
   - GPS pre-processing for 3D model generation (zeroing altitudes for better orthomosaic generation).

3. Data verification and range labeling:
   - Identification of missing data,
   - Range labeling based on field layout,
   - Forward images to stitching.
4. Stitching and orthomosaic generation in Agisoft Photoscan [4]:
   - Photo alignment,
   - Dense point cloud generation,
   - Building a mesh,
   - Orthomosaic generation,
   - Export generated stitched image.

5. Data management:
   - Transfer local copies of generated orthomosaics to a data server (Onomi),
   - Upload stitched images to database,
   - Resize images for Plot Segmentation tool.

6. Plot Segmentation:
   - Each generated orthomosaic should be segmented per plot (two rows of plants within a range).
   - Segmentation can be done with existing software (Plot segmentation tool), or manually.

The following figure describes the pre-processing pipeline (Figure 4.1).

**Figure 4.1:** Visual representation of the proposed pre-processing pipeline. The entire process is divided into six major steps.
4.1.1 Range Segmentation

For images taken with the one-pass walking pattern, Image Processing techniques and algorithms are required to carry out range segmentation (such as histogram analysis, morphological image processing in the forms of boundary extraction, connected components extraction, image segmentation, object recognition and so on). For images taken with the two-pass walking pattern range segmentation can be executed based on GPS coordinates and or implementation of a clustering algorithm (Figure 4.2).

![Figure 4.2: Range Segmentation step in pre-processing pipeline](image)

First, the necessary file structure that follows the field layout should be obtained by using GoPro Quik application [3.4]. By transferring all images from the camera to a storage drive using GoPro Quik application, generated file structure represents the field layout (grey block on figure 4.2). Second, if the remote clicker wasn’t used during the filming, this means that the entire image set hadn’t been sorted and in this case, all images should be processed by a clustering or classification algorithm.
4.1.2 Prior image adjustments and corrections

Colour adjustments may be needed in case of different white balance from image to image (Figure 4.3). Moreover, cropping of images might also be needed to get rid of some artifacts (Figure 4.4) that might affect stitching (drop shadows of the operator and the camera rig). Preliminary tests of influence destructive subjects on the images shown in Figure 4.5. However, for two-pass 2017 dataset, no colour adjustments was required, because all images were taken with the same camera settings throughout the filming season. GoPro Hero 5 camera allowed automatically adjust white balance during the filming based on lightning conditions and colours of the scene.

![Example of colour instability caused by white balance difference](image1)

(a) ![Sample of image with operator’s feet and drop shadow from the camera rig](image2)

(b)

**Figure 4.3:** Example of colour instability caused by white balance difference

**Figure 4.4:** Sample of image with operator’s feet and drop shadow from the camera rig

According to preliminary tests, destructive subjects (operator’s feet and drop shadow from the camera rig) may affect stitching results in a way that the final orthomosaic image has different scale within the image area (Figure 4.5a) which leads to a trapezoidal form of the image. In the case of cropped images, the results orthomosaic looks rectangle-like (Figure 4.5b). However, for the two-pass 2017 dataset, image cropping was not required, because destructive subjects on images were not the primary causes of bad orthomosaic generation.
Figure 4.5: Stitching performance before and after cropping out of destructive subjects (shadows) from the images

4.1.3 Data verification and range labeling

All images contain EXIF information, hence plotting GPS coordinates would be helpful for evaluating the completeness of data set, as well as for tracking any inconsistencies and operational errors. For example, visual inspection of GPS tracks displayed as a point sequence on the map can help to localize gaps in the sequence of images. If some images are missing, there will be a distinctive empty area on the map where there are no GPS coordinates displayed (Figure 4.6). Moreover, all images are captured from the same height above the ground, so we can consider only two-dimensional lat-long GPS representation.

Figure 4.6: Top down view of GPS coordinates of images from Canola DFL trial (12 ranges). From right to left: the 5th and 11th ranges have missing data (circled red).

Displaying GPS coordinates can also be useful to preliminary estimate nosiness of the data. Inaccurate GPS coordinates of images may lead to subsequent errors in the stitching process. By knowing the level of nosiness, a user can conclude on possible stitching accuracy (Figure 4.7). Range labeling procedure should be based on the information about missing data and field layout (number of ranges, the name of the first range, etc.).
Figure 4.7: Top down view of GPS coordinates of images from Carinata trial. The figure shows only a cropped sector of the Carinata field. The dataset consists of the top block with two ranges and the bottom block with 16 ranges. Ranges with high level of nosiness are circled red.

4.1.4 Stitching and Orthomosaic Generation

Stitching step can be treated as the most important within the designed pipeline (Figure 4.8). This work is concentrated on developing, testing and implementation of a fully automated stitching solution using the existing Python API in Agisoft Photoscan.

Agisoft PhotoScan is an advanced image-based 3D modeling solution aimed at creating professional quality 3D content from still images [4]. Agisoft Photoscan works with arbitrary images. Images can be taken from any position, providing that the object to be reconstructed is visible on at least two photos. Both image
alignment and 3D model reconstruction are fully automated. The procedure of photographs processing and 3D model construction comprises five main stages.

1. The first stage is camera alignment.

2. The next stage is building dense point cloud.

3. The third stage is building 3D mesh.

4. The fourth stage is orthomosaic generation.

5. The last step is exporting the data.

Note, that orthomosaic generation is a multi-step process. First, previously generated meshes should be positioned parallel to one of the planes in the coordinate system. Then, a bash process should be run for all meshes to create orthomosaics.

4.1.5 Data Management

Next step is to export and save generated orthomosaics. After, exported orthomosaics should be synced with a database for further segmentation with plot segmentation tool. Figure 4.9 shows the principle of data management process.

The GrowPro filming tool does not allow generating orthomosaic images of the whole field, because it only can produce detailed images of ranges. However, range images also have to be sorted, mainly because each range has more than 35 rows of plants, where each pair of rows would refer to different plant phenotype. Such areas have to be segmented out for the range image for further image processing analysis.
Problem

As it was described earlier, the GrowPro Viewer program exports orthomosaic range images, which are stored and labelled according to the field layout as .jpeg or .tif files. Before exported orthomosaics can be used for further analysis with Plot Segmentation Tool, they should be transferred to a server and imported into a database.

Proposed solution

For these purposes, a set of python scripts has been designed to manage orthomosaic images. Data management can be described as a sequence of required steps. Each step refers to the execution of a python script. A detailed step-by-step guide for data management using the designed python scripts is described in Appendix B.2.

4.2 Implementation of pipeline for 2-pass data

Implementation of the described GrowPro pre-processing pipeline required writing a set of python scripts. It is important to adapt the proposed programming solution for processing of a substantial amount of data. More specifically, a user should provide to the program only a path to the input dataset, make some data adjustments (GPS altitude corrections, data verification and labelling, data clustering or classification if necessary) using a graphical user interface, and avoid manual, repetitive interaction with the program during the actual data processing.

After all necessary scripts and programs have been written, evaluation of stability and reliability of the designed software has been carried out. It has been shown that:

- The programs allow users to automatically generate orthomosaic images from the RAW unstitched images (taken with the two-pass walking pattern)
- The programs should work as a stand-alone solution without user interaction during the data processing.
- Final generated orthomosaic images should be labelled and stored in the database, following the required field layout and ready for plot segmentation tool.

We have split the entire image set into two different groups (1- and 2-passes walking patterns), hence here in this project we target images taken with the 2-passes walking pattern. The main reason for this is that over 60% of images were taken using 2-passes walking pattern. Moreover, for images taken with the 1-pass pattern, a different approach for pre-processing and segmentation needs to be used. More specifically, image processing techniques should be used to segment region of interests from the entire orthomosaics based on images content (for example, greenness, exposure, texture or shape). As stated in 4.1.3 for 2-passes part of the dataset verification and processing is based on GPS coordinates of images and visualization of the entire GPS tracks.
While building the pre-processing pipeline, we solved three major problems. First, we implemented data verification and range labelling procedures which are based on GPS coordinates of RAW images. Then, we tested two algorithms for automatic range segmentation process that is also based on GPS coordinates of plant images. To increase the accuracy of orthomosaic generation, we developed and tested a GPS pre-processing technique. As an additional tool, we also suggested an approach for evaluation of GPS data nosiness. Results of using the designed pre-processing pipeline, stitching results assessment and overall stability and reliability have been shown in the following sections.

4.2.1 GPS-based data verification and range labeling

Problem

As it was described in 4.1.3, visual inspection of plotted GPS coordinates of images can help to localize data losses. These GPS points are displayed on the map as a point sequence and form point tracks that refer to the actual camera trajectory during the filming procedure.

For missing data, there will be a distinctive empty area with no GPS coordinates displayed. By missing data, we refer to any missing images during the filming, caused by remote clicker malfunctioning, operator’s errors, or sudden termination of the filming caused by the weather.

Plotting GPS coordinates will also be useful for estimating data nosiness. By knowing the level of nosiness, a user can conclude on possible stitching accuracy (Figure 4.7). GPS nosiness may lead to poor stitching results (Table 5.2).

Proposed Solution

For data verification and labelling, a software tool called GrowPro Viewer has been designed. This software has the graphical user interface with multiple options for data verification and labelling. The toolset consists of the following options:

- Plot GPS coordinates of each area of interests (per range) or of the entire trial.
- Visually detect missing gaps/images by plotting GPS coordinates of images.
- Calculate the level of GPS nosiness for each range of the selected filming date.
- Automatically adjust GPS coordinates of images for better stitching by assigning GPS altitude to the same value, so all points belong to the same plane.
- Automatically label data set according to the field layout and information about missing data.
- Automatically run stitching for the entire chosen filming day.

At the first step, the program requires a path to a directory that contains images for one filming date. This directory must contain only one level of sub-directories (4.10). Then, a user should choose the type of
data set (pre-sorted or non-sorted image set) and define, whether the image set had preliminary been sorted with the remote clicker, or it was just an unsorted stack of images.

**Figure 4.10:** Required file structure for the designed software tool. Root directory here represents a required directory for the program and should only contains one level of sub-directories.

For a sorted set of images, a corresponded GUI activates that allows users to plot GPS coordinates, automatically label the data, evaluate GPS nosiness, display information about the number of images, Adjust GPS coordinates for better stitching and forward prepared image set to stitching. Information about stitching process and its stages is displayed in real time.

If the image set wasn’t sorted, then the program allows users to either sort selected image set using built-in algorithms (4.2.2) or plot GPS coordinates of this set.

The workflow for the GrowPro Viewer is depicted in Figure **4.11**.
The graphical user interface is shown in Figure 4.12.

**Figure 4.12:** Graphical User Interface of the GrowPro Viewer

**Implementation of automated data labeling**

Before using the designed GUI for data verification and labelling, a separate JSON file should be created with the information about all fields and all trials (Figure 4.13).

```
"trials": [  
  
  ]
```

**Figure 4.13:** Field layout information represented in .JSON file

Description of each trial consists of the name of the field, name of the trial, season, number of ranges, filming pattern and range labels. This file allows the program to automatically pick up all necessary information about field layout and label the selected dataset. If there are missing ranges in the dataset, the user should first identify these ranges by the visual inspection of the plotted GPS coordinates, and then list index numbers in the corresponded input box of the GUI.

**Data verification technique**

To correctly identify the labels of missing ranges, it is important to locate plotted GPS points relative to the field layout and the edges of the field. To do so, bounding boxes that represent the field layout are plotted on top of the GPS points of the data. Red lines in Figure 4.16 and 4.14 show trials' boundaries for verifying the dataset location within the trial.
The position of bounding boxes has been calculated by a python script `generate_layout_points.py` based on the GPS coordinates of the complete image sets (image sets with the full number of filmed ranges). We couldn’t get the ground truth GPS data of the location of each trial, that is why the location of the boxes is defined as min-max latitudes and longitudes of the selected data set. Two middle lines that delimitate the blocks of Carinata are located at the same distances from the top and the bottom boundaries of the whole Carinata trial.

Red and green dots represent starting and finishing points of the filming session respectively. Identifying the starting point allows users to label the dataset correctly.

If some of the first or the last ranges are missing (Figure 4.14), there will be a shift between GPS points and the boundary of the bounding box. With this shift, it will be easy to define the label of the first and last filmed ranges.

**Figure 4.14:** GPS coordinates of images of Canola DFL, taken on July 7th, 2017. The distinct empty areas on the left and right side of the trial can be observed.

The above strategy can also be used for identifying data losses in the middle of the trial (Figure 4.15).

**Figure 4.15:** GPS coordinates of images of Carinata, taken on June 13th, 2017. The distinct empty areas in the middle of the trial can be observed.
Layout of the Agriculture and Agri-Food Canada (AAFC) farm is depicted in Figure 4.16.

**Figure 4.16**: Field layout of the AAFC field

GPS coordinates of images. Red dot - start, Green dot - finish
Figure 4.17: Plotted GPS coordinates of the third block of Carinata (blue points) and the field layout represented by bounding boxes (red lines).
GPS noisiness calculation

In the 2-pass walking pattern (3.3.3), the operator captures images by walking up one side of a range, turning around and then captures images walking down the opposite side of the same range. Therefore the ideal pattern of GPS coordinates for each range would be a U-shape with long straight sides (Figure 4.18). For noisy GPS coordinates the pattern would be a non-consistent sequence of points (Figure 4.19).

![Figure 4.18: Example of a GPS track with low level of noisiness.](image)

![Figure 4.19: Example of a GPS track with high level of noisiness.](image)

For the 2017 growing season dataset, all ranges captured with 2-pass walking pattern have the same shape and size, representing as a rectangle areas that form a trial 3.7. Plants for 2017 season were seeded in a way that all ranges adjoin each other. Each range has a rectangular shape, so a larger axis which divides the range into two halves can be defined for each range. Definition of the range axis will be described in the following section.

Finding axis of range

To calculate the range nosiness, we need to identify a line that would refer to a larger axis of the range i.e. a vector parallel to the range axis. Geometrical interpretation of the range axis is shown in Figure 4.20.

![Figure 4.20: Axis of the range](image)
To find a vector collinear to the range axis, we deployed the least square error (LSE) approximation technique. A set of discrete GPS points is approximated with a linear function \( p = p(t) = \alpha + \beta t \) represented by the solid line. We determined two constants \( \alpha \) and \( \beta \) such that fitted the data as good as possible in the sense of least squares. We wanted the solution to minimize the function

\[
F(\alpha, \beta) = \sum_{i=1}^{n} (\alpha + \beta t_i - y_i)^2,
\]

where \((t_i, y_i)\) is a GPS point in lat-long space. In order to minimize \( F \) with respect to \( \alpha \) and \( \beta \), we solved

\[
\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial \beta} = 0
\]

Approximation here is based on a set of GPS points along the larger side of the range. We define this set of points as a point sequence with a ten images offset from the starting point of each range. Number of points \( s \) that form lines \( l_{s_1} \) and \( l_{s_2} \) (Figure 4.21) equals to 70% out of the mean number of one half of the size of a range image set:

\[
s = 0.7 \sum_{i=0}^{r} \frac{L(r)}{2r},
\]

where \( r \) - number of ranges, \( L(r) \) - number of images for each range. Selection of \( s \) is based on the intention to pick only points that lay on long straight sides (first and second pass), but not from the turning area. This can potentially increase accuracy of detection range axis \( \vec{l}_{axis} \). Specifically, this set of points defines the position of the whole range (Figure 4.21). For better estimation, two sets of points from both sides can be used for finding the axis of the range. \( \vec{l}_{axis} \) is defined as a vector equidistant from vectors \( \vec{l}_{s_1} \) and \( \vec{l}_{s_2} \).

**Figure 4.21:** Finding an axis of the range using LSE approximation technique. In the experiment, two sets of points approximated by the lines \( l_{s_1} \) and \( l_{s_2} \) were used to find a range axis \( \vec{l}_{axis} \).
Definition of Noise

As described above, image location information is indicated in Degrees, Minutes and Seconds (DMS) format. For calculations in Euclidean metric, DMS coordinates have been translated to Decimal Degrees (DD) and then to UTM format of UTM zone 13 (Saskatchewan province of Canada). Considering Cartesian coordinate system, GPS coordinates of images would represent a set of 2D points, where each image/point defines as \( p(x_1, x_2) \) (here we do not take into account elevation in the form of altitude). Thus, two adjacent points form a vector \( \vec{a} = (p_i, p_{i+w}) \), where \( i \) is a point in the set, and \( w \) is a window size (Figure 4.22).

![Diagram](image)

**Figure 4.22:** A vector defined by a pair of points in the point set

Note that the vector can be defined not only between two adjacent points, but between any two points in the set: \( p_i \) and \( p_{i+w} \), where \( w \) - is *moving window size*.

Selection of moving window size is based on overall nosiness of the data set (which is determined visually) and the shape of the data. If image GPS coordinates look noisy (Figure 4.19), then larger \( w \) will be required to reduce the effect of the noise. Moreover, selection of \( w \) may also depend on the number of images per range (different speed of walking during the filming process may lead to a different number of RAW range images captured).
We define local noise for the selected pair of points through the dot product of the two vectors: $\vec{a} = (p_i, p_{i+w})$ and the main axis $\vec{b} = l_{axis} = (b_1, b_2)$ (Figure 4.23). The dot product of these two vectors is defined by the next expression:

$$<\vec{a}, \vec{b}> = |\vec{a}||\vec{b}|\cos(\phi_i)$$

, where $\phi_i$ is the angle between the given vectors. GPS data nosiness in this case is an angle $\phi_i$ between a pair of vectors and is defined by the next expression:

$$\cos(\phi_i) = \frac{p_i b_1 + p_{i+w} b_2}{|\vec{a}||\vec{b}|}$$

![Figure 4.23: Definition of local noise for the selected pair of points](image)

To evaluate the noise of the entire range, we sum nosiness of all sequentially selected pairs of points and divide this sum by the number of pairs to get the mean nosiness $N_m$ for the range:

$$N_m = \frac{\sum_{i=0}^{n-w} \cos(\phi_i)}{(n-w)}$$

where $n$ - is the number of points in the range and $w$ - moving window size, $i$ - point in the pointset.
Results of GPS noise estimation can be represented as a bar chart (Figure 4.24) for each filming date, or exported as .csv file.

**Figure 4.24:** Estimation of noisiness for images of Canola DFL taken on June 28th, 2017, window size - 15 points
4.2.2 GPS-based range segmentation

Range segmentation is a required step for building stitched orthomosaic images. Since with the GrowPro filming tool, it is not feasible to generate a stitched image of the whole field at once, compare to UAV approaches, there is a need to film each range of the trial individually and then carry out data segmentation per range. For 2-pass data, we stitch together images at the range level, and the crop out individual plots from the range orthomosaic. Therefore, we need to group images per range, a process we call range segmentation.

Problem

As described in 3.6, 6.4% of images taken with the two-pass walking pattern during the 2017 growing season are unsorted (i.e., start-pause recording with the remote control clicker wasn’t made) and stored in a single directory. Even though the unsorted part of the two-pass dataset is relatively small, these images represent plants on their early stages of growth and might be important for further image analysis.

If the manual range segmentation technique (i.e., start-pause recording with the remote control clicker) was not used during the filming, then there is a need to cluster and group unsorted RAW images after the filming, i.e. carry out image segmentation per range.

Proposed solution

Here in this project, we define the following segmentation approaches:

- Image-based range segmentation, when a user should visually inspect all images and group them according to the field layout.

- GPS-based range segmentation. In this case, a range segmentation problem turns into a classification or clustering problem of two-dimensional spatial data, where well-established machine learning methods can be used.

Instead of using subjects on RAW images as a base for segmentation, each RAW image can be considered as a GPS point in space. In this perspective, classification or clustering can be carried out on GPS points in space. For this case, each range within the field would represent a unique class or cluster.

Data characterization

Selection of an appropriate clustering or classification method depends on data characteristics, as well as the ground truth availability. The following list describes the key features of the data:

- Dataset represents spatial information about each image so that Euclidean metric can be used. Originally, image location information is indicated in Degrees, Minutes and Seconds (DMS) format. For calculations in Euclidean metric, DMS coordinates have been translated to Decimal Degrees (DD) and then to UTM format of UTM zone 13 (Saskatchewan province of Canada).
The data set is of the two-dimensional type.

Field layout implies non-overlapping ranges, so range segmentation can be transformed into a multi-class classification problem because each range can be viewed as a class. In the experiment all ranges are explicitly separate from each other, that is why all class labels are mutually exclusive.

Number of classes remains constant and known a priori because the number of ranges is fixed.

Ground truth data is available and fixed for the entire test set.

There are numerous classification algorithms available. For initial testing of a supervised learning approach, we suggest implementing one of the most popular algorithms - K-nearest neighbours (KNN). As for a unsupervised learning approach, we propose initial implementation and testing of a convolutional clustering method.

K-nearest neighbors algorithm

The traditional KNN classifier finds the k nearest neighbours based on some distance metric by detecting the distance of the target data point from the training dataset, then finding the class from those nearest neighbours by some voting mechanism. The efficient implementation of KNN algorithm is of particular interest in Geographic Information Systems [45]. For instance, a user can request a system to define N closest points of interest to his or her location on the map. Moreover, classification of spatial data can also be carried out when the training dataset changes often [32]. Particular interest is focused on solving the clustering problem in high-dimensional spaces [12].

As stated above, majority of images of 2017 growing season were taken and sorted using start-pause recording technique. However, for images of Canola DFL, 37% of images (June 5th and June 13th filming dates) are unsorted and 63% of images (June 21th through August 21th filming dates) are sorted ones.

As defined, 63% of the dataset is sorted, we can treat it as a training set for the KNN classification algorithm, because in this case, all images can be classified into 12 non-overlapping classes (from 1 to 12, which is the number of ranges in Canola DFL trial). Thus, for images of Canola DFL of 2017 growing season there are 7385 unsorted images that will represent a test set and 19800 sorted images that will represent a training set.

The neighbours are taken from a set of GPS points of sorted images for which the class (i.e range) is known and stored in a text file with three columns, where the frist two columns represent latitude and longitude information, and the third columns contains class labels. This can be thought of as the training set for the algorithm, though there are no explicit training step. The training examples are GPS points in a two-dimensional feature space, each with a class label. Features here are latitude and longitude, a class would represent a range (for Canola DFL there are 12 ranges i.e 12 non-overlapping classes). Test set for the algorithm is a text file with GPS data of unsorted images and contains only latitude and longitude information. No boot strapping needed because we have got a lot of training data, so classification here is
based on preliminary labeled training data. We chose to use the pre-segmented data as a training dataset, rather than for bootstrapping, because we have a large amount of pre-segmented data (22000 images), as compared to the amount of unsegmented images that required labeling (7300 images).
Visualization of performance of the KNN algorithm have shown in Figure 4.25 and 4.26. The Figures show classification of images of Canola DFL captured on June 5th and 13th, 2017 growing season with the number of nearest neighbours equals to 30.

**Figure 4.25:** Classification results for images of Canola DFL (12 ranges, captured on June 5th, 2017), k = 30.

**Figure 4.26:** Classification results for images of Canola DFL (12 ranges, captured on June 13th, 2017), k = 30.
To evaluate performance of the KNN algorithm for GPS-based image segmentation, confusion matrices have been calculated for two unsorted filming dates (images of Cabola DFL, June 5th and June 13th) with 30 nearest neighbours for the KNN algorithm. To generate confusion matrices for unsorted part of the dataset, we manually sorted (June 5th and 13th) to have ground truth classification of these filming dates. Then we run KNN classification for unsorted version of Canola DFL June 5th and 13th and subsequently generated the confusion matrices (Tables 4.1 and 4.2). When manual segmentation over June 5th and 13th filming dates has been done, we counted the number of images for each range (column “all images” in figures 4.27 and 4.28). Note that the use of sorted dataset as training set is limited because they are not representative of the continuous path taking in the unsorted dates, because of the gap in GPS coordinates when the remote was used to turn the camera off at the end of one range and then back on at the start of the next range.

<table>
<thead>
<tr>
<th>481</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>359</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>353</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>326</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>330</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>348</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>360</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>357</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>325</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>341</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>333</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>369</td>
</tr>
</tbody>
</table>

**Table 4.1:** Performance of the algorithm for images of Canola DFL captured on June 5th, 2017, k = 30.

<table>
<thead>
<tr>
<th>157</th>
<th>52</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>200</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>221</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>190</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>212</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>256</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>237</td>
<td>29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>68</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>165</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>220</td>
<td>71</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>242</td>
</tr>
</tbody>
</table>

**Table 4.2:** Performance of the algorithm for images of Canola DFL captured on June 13th, 2017, k = 30.
To see the impact of the number of nearest neighbours (k) on the algorithm’s performance we run image classification of Canola DFL (June 5th and 13th) with different k parameters to see how the number of correctly sorted images changes with the number of nearest neighbours used by the algorithm. KNN classification performance varies with different k and from range to range, mainly because of the GPS nosiness (Figure 4.36). To better visualize classification performance over different k, we calculated the absolute differences between the actual number of images and the number of successfully classified images, taken from confusion matrices (the main diagonal of a matrix) for each k. Smallest absolute difference row-wise is highlighted green for each range (Figures 4.27 and 4.28).

<table>
<thead>
<tr>
<th></th>
<th>k = 1</th>
<th>k = 5</th>
<th>k = 10</th>
<th>k = 15</th>
<th>k = 20</th>
<th>k = 25</th>
<th>k = 30</th>
<th>k = 35</th>
<th>k = 40</th>
<th>all images</th>
</tr>
</thead>
<tbody>
<tr>
<td>r01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>481</td>
</tr>
<tr>
<td>r02</td>
<td>18</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>367</td>
</tr>
<tr>
<td>r03</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>15</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>15</td>
<td>364</td>
</tr>
<tr>
<td>r04</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>326</td>
</tr>
<tr>
<td>r05</td>
<td>18</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>9</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>338</td>
</tr>
<tr>
<td>r06</td>
<td>43</td>
<td>27</td>
<td>20</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>366</td>
</tr>
<tr>
<td>r07</td>
<td>25</td>
<td>24</td>
<td>24</td>
<td>23</td>
<td>21</td>
<td>20</td>
<td>18</td>
<td>16</td>
<td>14</td>
<td>378</td>
</tr>
<tr>
<td>r08</td>
<td>30</td>
<td>23</td>
<td>17</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>358</td>
</tr>
<tr>
<td>r09</td>
<td>35</td>
<td>38</td>
<td>44</td>
<td>46</td>
<td>42</td>
<td>45</td>
<td>42</td>
<td>43</td>
<td>40</td>
<td>367</td>
</tr>
<tr>
<td>r10</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>18</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>358</td>
</tr>
<tr>
<td>r11</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>341</td>
</tr>
<tr>
<td>r12</td>
<td>34</td>
<td>36</td>
<td>35</td>
<td>30</td>
<td>52</td>
<td>58</td>
<td>55</td>
<td>47</td>
<td>37</td>
<td>424</td>
</tr>
</tbody>
</table>

Figure 4.27: Absolute differences between successfully classified images and the actual number of images over different numbers of the nearest neighbours (from 1 to 40), for images of Canola DFL, June 5th. Smallest absolute difference is highlighted green for each range.

<table>
<thead>
<tr>
<th></th>
<th>k = 1</th>
<th>k = 5</th>
<th>k = 10</th>
<th>k = 15</th>
<th>k = 20</th>
<th>k = 25</th>
<th>k = 30</th>
<th>k = 35</th>
<th>k = 40</th>
<th>all images</th>
</tr>
</thead>
<tbody>
<tr>
<td>r01</td>
<td>44</td>
<td>39</td>
<td>40</td>
<td>56</td>
<td>54</td>
<td>57</td>
<td>52</td>
<td>53</td>
<td>50</td>
<td>209</td>
</tr>
<tr>
<td>r02</td>
<td>55</td>
<td>47</td>
<td>34</td>
<td>29</td>
<td>27</td>
<td>27</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>225</td>
</tr>
<tr>
<td>r03</td>
<td>22</td>
<td>21</td>
<td>18</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>217</td>
</tr>
<tr>
<td>r04</td>
<td>21</td>
<td>21</td>
<td>18</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>23</td>
<td>242</td>
</tr>
<tr>
<td>r05</td>
<td>60</td>
<td>63</td>
<td>64</td>
<td>63</td>
<td>62</td>
<td>62</td>
<td>61</td>
<td>62</td>
<td>61</td>
<td>251</td>
</tr>
<tr>
<td>r06</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>224</td>
</tr>
<tr>
<td>r07</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>259</td>
</tr>
<tr>
<td>r08</td>
<td>32</td>
<td>33</td>
<td>30</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td>266</td>
</tr>
<tr>
<td>r09</td>
<td>72</td>
<td>71</td>
<td>74</td>
<td>73</td>
<td>69</td>
<td>69</td>
<td>68</td>
<td>68</td>
<td>68</td>
<td>243</td>
</tr>
<tr>
<td>r10</td>
<td>82</td>
<td>82</td>
<td>82</td>
<td>82</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>248</td>
</tr>
<tr>
<td>r11</td>
<td>73</td>
<td>74</td>
<td>73</td>
<td>80</td>
<td>71</td>
<td>75</td>
<td>71</td>
<td>88</td>
<td>87</td>
<td>291</td>
</tr>
<tr>
<td>r12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>242</td>
</tr>
</tbody>
</table>

Figure 4.28: Absolute differences between successfully classified images and the actual number of images over different numbers of the nearest neighbours (from 1 to 40), for images of Canola DFL, June 13th. Smallest absolute difference is highlighted green for each range.

The best choice of k also depends on the data. Larger values of k reduces data noisiness on the classification, but make boundaries between classes less distinct. Figures 4.27 and 4.28 show that finding the best k...
for accurate classification might become problematic. However, it can be seen from the Figures that \( k \) within the range from 20 to 40 neighbours works better than within the range from 1 to 15 neighbours.

**Convolutional clustering algorithm**

Another promising solution within unsupervised learning approach is convolutional clustering algorithm (CCA). The two-pass walking pattern implies a turning point in the middle of the way where an operator gradually changes his moving direction on the opposite. This turning point can be seen as the significant change of direction of the velocity vector relative to the vertical axis (i.e. North direction). This algorithm approximates the walkway and allows detecting the turning point for the two-pass walking pattern (Figure 4.29). By tracking the angle between North direction and the velocity vector, the turning point can be defined as the point, where the angle between the velocity vector and the vertical line is maximum.

![Figure 4.29: Visualization of the convolutional algorithm (ten points window size). Algorithm calculates the angle between each vector and the North vector.](image)

Convolutional clustering is considered as a more generalized approach, so we used this algorithm for solving clustering problem. The convolutional algorithm has been run for two unsorted image sets of Canola DFL (June 5th and 13th). Results of preliminary implementation of convolutional clustering algorithm have shown in Figures 4.30 and 4.32. For both image sets window size equals to 50 points. The graph represents the distribution of the direction of the velocity vector relative to the vertical axis (i.e. North direction) over time. For the data set that contains 12 ranges, there are 23 or 24 turning points which correspond to 23 or 24 peaks on the graph, depending on the starting point and/or field layout.
To compare Convolutional clustering results, we counted the number of turns during the filming of both, June 5th and 13th filming dates. Figure 4.31 shows GPS tracks of images of Canola DFL, June 5th. 24 clearly visible turns have need visually detected. Note, that GPS noisiness at the beginning of the filming session (red dot in figure 4.31) matches nosiness on the graph in Figure 4.30.
For images of Canola DFL June 13th convolutional clustering results are shown in Figure 4.32. GPS tracks represent 23 distinctive turns during the filming session (Figure 4.33).

For convolutional clustering, we consider segmentation process as successful if there are distinctly visible peaks on the graph (Figures 4.30 and 4.32) and the number of these peaks equals the number of turns. Because the results obtained by the convolutional algorithm cannot be considered as the final image clustering and only shows preliminary grouping based on velocity, identifying peaks that represent the turns during the filming procedure can be treated as a separate research task and problem. Note that for this experiment we used the image set of Canola DFL that contains images with both precise and noisy GPS information.
Identifying peaks for Convolutional clustering algorithm

To quantitatively analyze results of Convolutional clustering algorithm we thresholded images with an angle between North direction and the velocity vector greater than 70° for both June 5th and 13th Canola DFL. After applying the same threshold value for both image sets, convolutional clustering algorithm identified 126 images of Canola DFL June 5th and 88 images of Canola DFL June 13th with the angle between North direction and the velocity vector greater than 70°. Then, both subsets of images (i.e. 126 images of June 5th and 88 images of June 13th) have been inspected for identifying possible groups that would belong to each peak of the graphs (Figures 4.30 and 4.32). Images in each subset have been manually clustered into 26 and 23 groups respectively, based on the sequential image order and corresponding gaps. These image groups/clusters refer to the turning points and highlighted as red dots in Figures 4.34 and 4.35.

**Figure 4.34:** Turning areas (highlighted red), identified by the algorithm within images of Canola DFL, June 5th. Threshold value = 70deg, window size = 50 points.
Discussion

When using KNN algorithm for classification, the principal limitations we faced was the low accuracy of GPS image coordinates. Initially there were two potential ways for generating ground truth GPS data. The first option was to get GPS coordinates of the trial (GPS coordinates of each range within the trial) with an external GPS device that calculates high accuracy GPS data, so this GPS data could subsequently be used as training dataset for the KNN classification algorithm. The second option was to use GPS coordinates of already manually clustered range images as a training set for the algorithm.

We could not get precise GPS data before filming and implemented the second approach, so overall accuracy of generated ground truth GPS strongly depends on GPS accuracy of manually sorted plant images. In addition to this, even with the good relative accuracy of GPS (between two adjacent images), the absolute accuracy of GPS coordinates of each image is quite low. Each time during filming GPS data may potentially be corrupted by many factors, such as atmospheric conditions, satellite geometry and signal blockage. This is why for different filming sessions carried out on different days GPS location of a range calculated at the same location with the same receiver may differ from day to day. As a result, absolute inaccuracy leads to a shift in GPS coordinates of different filming sessions, which makes using generated ground truth impossible. As a result, GPS inaccuracy and shifting phenomenon may lead to misclassification (Figure 4.36).
However, GPS-based image classification with KNN classification algorithm might still be a reasonable and practical approach, especially if ground-truth GPS coordinates of ranges are acquired by a high-accuracy GPS receiver. Besides, selection of $k$ value for KNN should be based on data nosiness and the number of images in a range. For the current field layout (Canola DFL, Canola DRIL and Carinata) we recommend $k$ greater than 20 neighbours.

It can be seen from tables 4.1 and 4.2 that all classification errors affect only two adjacent ranges. This phenomenon can be treated as a systematic bias. Such systematic error can be caused by GPS inaccuracy that changes from day to day and depends on signal arrival time measurements, numerical calculations of the receiver, atmospheric effects (ionospheric/tropospheric delays), ephemeris and clock data of the receiver, multipath signals effect, natural and artificial interference and so on. This is another example of the GPS data shift that was described above 4.36.

Since convolutional clustering method is a preliminary and possible way of image clustering, we considered the problem of identifying peaks as extensive mathematical analysis task for future work. In our work, we only showed CCA as a “possible perspective approach.” However, we suggested and implemented a thresholding technique that has demonstrated prospective results. Note, that GPS nosiness sometimes leads to a split of peaks into multiple local extrema (the first peak in Figure 4.30) which subsequently might be misunderstood as various turning points. Moreover, a threshold value should be chosen accordingly so that each peak would be identified correctly. In Figure 4.35 there is one missing turning point, and the main reason for this is that the threshold value was fixed for both filming dates (June 5th and June 13th) and one peak in Figure 4.32 had been eventually missed. In the experiment, we see that algorithm performance strongly depends on the window size used in CCA, a threshold value used for identifying peaks and overall GPS data nosiness. All these parameters may vary depending on the dataset.
4.2.3 GPS pre-processing for 3D model generation

Problem

To fully automate stitching process of two-pass range images it is necessary to switch from using Agisoft Photoscan graphical user interface (GUI) to execute a script that uses Agisoft Photoscan Application Programming Interface (API) and can automate all required stitching steps that would be done manually in the Agisoft Photoscan GUI.

The critical limitation here is that orthomosaic generation step requires manual re-positioning of a mesh model in such a way that the surface of a model is parallel XY coordinate planes of the planar coordinate system.

![Wrong position of the model](image1.png)  ![Correct position of the model](image2.png)

**Figure 4.37:** Representation of a mesh model in Agisoft Photoscan GUI with different positions

Proposed solution

While UAV approaches can be programmed to fly at the same height above the ground, which subsequently leads to uniform GPS altitudes of all images that form UAV orthomosaics, for the GrowPro tool an operator may unintentionally move the camera up and down while walking, which produces altitude irregularity over image GPS data. To correctly position 3D mesh models we should assign altitudes of all GPS coordinates to the same constant value. In this case, all GPS points will belong to the same plane which will flatten all models and position them in a way that the 3D mesh models are parallel to one of the coordinate planes of the planar coordinate system (Figure 4.38 and 4.39).
Figure 4.38: Altitude nosiness within images of 25 ranges of Carinata (lateral view)

Figure 4.39: All altitudes are assigned to the same value and belong to the same plane

Figure 4.40 shows plotted GPS points of images of Canola DFL taken on June 21st, 2017. This dataset contains images that form relatively even GPS tracks for the first four ranges (from right to left); however, GPS of the second half of the dataset looks somewhat noisy and irregular.
Figure 4.40: GPS points of images for Canola DFL 2017-06-21 dataset

3D mesh models generated without using the flattening are positioned in 3-dimensional space in the way that not all of them are parallel to the XY plane of the plane coordinate system and do not face the same direction (figure 4.41). As it was stated, facing all 3D meshes the same Z direction is critical for automatic orthomosaic generation and export.

However, for the ranges that had even GPS tracks, the flattening algorithm has changed their positions in the way that all ranges faced the same direction and looked flat (Figure 4.42). As it can be seen, 3D meshes that relate to noisy GPS tracks have been aligned even more random, whereas 3D meshes that refer to the first four ranges have been positioned appropriately.
GPS data adjustments with the flattening techniques were also applied for images of Canola DFL taken on June 26th, 2017. GPS data of this dataset is even (Figure 4.43), doesn't have noise throughout the entire area of the trial and is the most favourable for generating high-quality orthomosaic images. Figure 4.44 shows 3D meshes that have been created without the flattening process. Figure 4.45 shows the range 3D meshes that have been generated with preliminary flattened GPS coordinates of all images.

It is important to notice that changing GPS coordinates of images does not affect dense point cloud generation, as well as mesh generation accuracy, because GPS data is used only for preliminary photo alignment and positioning of the 3D meshes and dense point cloud in the 3D space, but not for stitching.
Method evaluation

Test results of implementation the flattening technique for GPS data have shown that this method better contributes to correct positioning of 3D mesh models when GPS data nosiness (4.19) of a range is relatively low (Figure 4.18).

To evaluate the flattening technique that has been described above, we have run tests on images of Canola DFL taken with the two-pass walking pattern. First, we have generated stitched images of 108 ranges of Canola DFL without applying the GPS flattening adjustments. Then, we have produced the same set of stitched images, but with preliminary applied GPS flattening adjustments.

Noise evaluation for both data sets has been carried out using the designed technique (4.2.1). It appeared that the flattening technique does not affect nosiness, so for the comparison and evaluation, we clustered ranges into five different groups (though, the number of groups can be chosen arbitrarily depending on overall noise variance within the dataset) based on the level of noise. Group 1 is for ranges with noise between 4.1° and 10.0°, group 2 - between 10.0° and 20.0°, group 3 - between 20.0° and 30.0°, group 4 - between 30.0°
Figure 4.46: Example of GPS track of the range that potentially can be correctly positioned parallel to XY coordinate plane

and 40.0°, group 5 - between 40.0° and 80.46° where 4.1° and 80.46° are extreme values of nosiness in the entire Canola DFL dataset. Then, the number of ranges with correct positioning has been calculated for both cases - with the application of the flattening technique and without. Results have been shown in table 4.3.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of ranges tested</th>
<th>Number of ranges positioned parallel to XY coordinate plane before flattening</th>
<th>Number of ranges positioned parallel to XY coordinate plane after flattening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 (4.1 - 10.0)</td>
<td>57</td>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td>Group 2 (10.0 - 20.0)</td>
<td>34</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Group 3 (20.0 - 30.0)</td>
<td>11</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Group 4 (30.0 - 40.0)</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Group 5 (40.0 - 80.46)</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>3</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 4.3: Evaluation of flattening technique

As it can be seen from the Table 4.3, GPS flattening technique correctly positioned 70% of ranges, however, if the GPS flattening method isn’t used, then only less than 3% of ranges have been positioned correctly. 84% of ranges belong to group 1 and 2 and have GPS noise between 4.1° and 20.0°. For ranges in group 1 with the low level of nosiness, 77% of ranges have been positioned correctly after using the flattening method, however, for ranges in group 2, only 62% of them have been positioned correctly. Figure 4.47 shows the number of correctly positioned ranges for different filming dates. As it can be seen, there is no decline in the number of correctly positioned ranges for the late filming days.
4.3 Discussion

Implementation of GPS-based algorithms for image pre-processing has shown that the performance of pre-processing pipeline strongly depends on the accuracy of geo-location data of each image and GPS nosiness within a set of images. The inaccuracy of GPS sensor in the camera leads to stitching errors and affects the overall stability of the pipeline. As it was shown, GPS coordinates should be flattened to reduce stitching errors. But flattening approach works only when GPS points form a plane. Hence, for images taken with the one-pass walking pattern, GPS inaccuracy may become critical. In this perspective, the developed method for GPS noise evaluation may be beneficial for initial data verification. This method works fast and allows users to quickly verify data right after each filming session on the filed.

The accuracy of GPS-based image segmentation also depends on the accuracy of geo-location data. In this case, better image GPS data that can potentially be acquired by an external GPS sensor would positively affect image segmentation performance.

Given all advantages and disadvantages of selected algorithms and methods for building a pre-processing pipeline, these methods have to be tested on a substantial amount of images, so a better understanding of the performance can be obtained. The following chapter describes the performance of the pre-processing pipeline and shows orthomosaic generation results over different crops and filming conditions.
CHAPTER 5

RESULTS

Testing and evaluation of overall performance of the pre-processing pipeline required stitching of the entire 2017 image dataset taken with the two-pass filming technique. We have verified, labelled and stitched images of Canola DFL, Canola DRIL and Carinata to build orthomosaic range images. From over 450000 RAW unstitched images of plants, we have generated 1373 orthomosaic images of plants.

To quantitatively analyze output orthomosaics we focused on certain types of artifacts that might appear in the orthos and make it hard to use them for phenotypic analysis. These common artifacts include: holes that might happen because of the plant movements, distortion that might occur because of the incorrect position of 3D meshes during stitching process. All of these factors may decrease overall stability and reliability of the pre-processing pipeline. Moreover, by analyzing the stitching results and identifying the most common problems we can suggest possible improvements for the pre-processing pipeline, as well as for the filming procedure and/or data gathering.

To analyze the performance we took a closer view on plant images taken on three different times: early stage of growth (3-5th week from seeding, during emergence stage), middle stage of growth (the 6th week from seeding, during early vigor stage) and late stage of growth (the 7th week from seeding, during flowering stage).

5.1 Metrics of success

In our project, we identified two main approaches for quantitative assessment of stitching results. The first approach is a manual vision inspection of output orthomosaic images. Such inspection takes longer time but gives the best understanding of the results. The second approach refers to accessing to image information, such as the dimensions of the generated orthos and its total number of pixels.
5.1.1 Visual approach

To evaluate pre-processing pipeline performance, we visually inspected range images of Canola DFL, Canola DRIL and Carinata, focusing on pictures taken between 3rd and 7th weeks from seeding. We considered orthomosaic generation as “successful” or “good” if an orthomosaic image represents a rectangular, flat and non-distorted picture of a range without holes (Figure 5.1a). Images with zippering artifacts (Figure 5.1b) usually have two pieces of orthomosaics split because of the insufficient overlapping between images from each pass. Images with holes (Figure 5.1c) have “missing” or “blank” spaces in the middle and/or at the border of the canvas. This happens when tall plants move on the wind or because of the Parallax effect. Image distortion is usually caused by the incorrect orientation of 3D meshes (Figure 5.1d).

![Figure 5.1: Examples of artifacts found in the generated orthomosaic images](image)

(a) “Good” image  (b) Zippering  (c) Holes in images  (d) Distorted image
5.1.2 Automated approach

On the other hand, to automatically process and analyze a substantial amount of orthomosaic images there is a need for clear quantitative metrics. For assessment of stitching results in the GrowPro Viewer, we implemented generation of a report file in .csv format (Figure 5.2) that describes stitching results for each filming date. This file shows the following information:

- The label of the range.
- Canvas dimensions. This parameter shows the size of the whole area of the generated image including the background.
- Total number of non-white, i.e. non-background pixels. This parameter gives information about the size of an image. If an image is corrupted by artifacts (zippering, holes, distortion), then the total number of image pixels would be small. If an image is not corrupted, then the total number of image pixels is big. Agisoft Photoscan exports orthomosaics on white background, so the actual plant image can be considered as all non-background pixels.
- Number of white (i.e. background) pixels, can be found as \( \text{height} \times \text{width} - \text{number of non-white pixels} \).

<table>
<thead>
<tr>
<th>Range</th>
<th>Dimensions</th>
<th>Total non-white pixels</th>
<th>Total white pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-07-17_r29</td>
<td>(7069, 28930)</td>
<td>131272457</td>
<td>73233713</td>
</tr>
<tr>
<td>2017-07-17_r21</td>
<td>(9653, 28038)</td>
<td>117538158</td>
<td>153112656</td>
</tr>
<tr>
<td>2017-07-17_r28</td>
<td>(13645, 35994)</td>
<td>101278955</td>
<td>389859175</td>
</tr>
<tr>
<td>2017-07-17_r23</td>
<td>(9120, 28687)</td>
<td>93503530</td>
<td>168121910</td>
</tr>
<tr>
<td>2017-07-17_r01</td>
<td>(7352, 27906)</td>
<td>90524430</td>
<td>114640482</td>
</tr>
<tr>
<td>2017-07-17_r24</td>
<td>(10251, 27691)</td>
<td>90118565</td>
<td>193741876</td>
</tr>
<tr>
<td>2017-07-17_r32</td>
<td>(7508, 29973)</td>
<td>84110693</td>
<td>140926591</td>
</tr>
<tr>
<td>2017-07-17_r16</td>
<td>(6984, 27416)</td>
<td>84015825</td>
<td>107457519</td>
</tr>
<tr>
<td>2017-07-17_r17</td>
<td>(8720, 26612)</td>
<td>81198384</td>
<td>150858256</td>
</tr>
<tr>
<td>2017-07-17_r48</td>
<td>(8754, 29774)</td>
<td>79623253</td>
<td>181018343</td>
</tr>
<tr>
<td>2017-07-17_r33</td>
<td>(6263, 26119)</td>
<td>76726439</td>
<td>86856858</td>
</tr>
<tr>
<td>2017-07-17_r15</td>
<td>(7633, 24770)</td>
<td>71151347</td>
<td>117918063</td>
</tr>
<tr>
<td>2017-07-17_r34</td>
<td>(9883, 27808)</td>
<td>68759418</td>
<td>206067046</td>
</tr>
<tr>
<td>2017-07-17_r20</td>
<td>(6424, 20576)</td>
<td>66018327</td>
<td>66161897</td>
</tr>
<tr>
<td>2017-07-17_r30</td>
<td>(8571, 26974)</td>
<td>65809423</td>
<td>165384731</td>
</tr>
</tbody>
</table>

**Figure 5.2:** An example of a .csv file report generated for the range images of Carinata. Instances in the table are sorted by the total number of non-white pixels

However, these parameters can be used only within one filming date. The main reason for this is that Agisoft Photoscan application exports orthomosaics with different resolution, which subsequently leads to different file sizes and a number of pixels in images. Moreover, the quantitative analysis would also require...
deploying additional Image Processing approaches (such as Histogram analysis, Morphological image processing in the forms of boundary extraction, connected components extraction, Image segmentation, Object recognition and so on). Hence within the scope of the project, visual inspection of the output orthomosaic images is a reasonable approach.

5.2 Ground resolution comparison

Agisoft Photoscan generates an orthomosaic image that is based on preliminary generated dense point cloud and mesh model. Depending on the size of a model, the number of pictures taken, the height that filming had been made on, the final size of the orthomosaic image may reach up to 50000 pixels along the larger side.

Ground resolution of pictures taken by the GrowPro is more detailed than conventional images taken from UAVs. Average resolution of the images taken from the GrowPro at the height of 1.5m is 2mm per pixel with 12 megapixels GoPro Hero 5 sensor, whereas the average ground resolution of pictures taken from UAV sensing platforms at the height of 30m is 20mm per pixel with 1.2 megapixels sensor in Micasense RedEdge camera (Figure 5.3).

(a) UAV at the height of 30m, 1.2MP sensor, ground resolution is 20mm per pixel
(b) GrowPro at the height of 1.5m, 12MP sensor, ground resolution is 2mm per pixel

Figure 5.3: Ground resolution of images taken from UAVs and GrowPro
5.3 Results over different crops

A universal approach to filming with the GrowPro allows users to apply the system to different types of crop. Over the entire 2017 growing season, four different crops have been filmed. We have acquired high-resolution images of Canola, Carinata, Wheat and Lentil using both, one-pass and two-pass walking patterns.

Results show that it is possible to generate high-resolution orthomosaic images of different crops. However, the height of plants, i.e. stage of growth, might seriously affect stitching performance, mainly because tall plants tend to move in the wind more than plants in their early stages of growth. In the experiment we have focused on two-pass image dataset of Canola DRIL, Canola DFL and Carinata, so the primary analysis was made on these crops. However, initial stitching tests have shown promising results for building orthomosaics of Lentils and early Wheat due to relatively small height of these crops. Figure 5.4 shows best samples of stitched orthomosaic images of the experimented crops in 2017 growing season.

(a) Canola  (b) Carinata  (c) Lentil  (d) Wheat

**Figure 5.4:** Samples of orthomosaic images of different crops
5.4 Results over different stages of growth

To evaluate stitching results and overall performance of the GrowPro system we have analyzed 489 orthomosaic images of Canola DFL, Canola DRIL and Carinata of 2017 growing season to see, which artifact is the most common. Stitching has been carried out over the entire two-pass dataset for all three crops with the GrowPro Viewer pre-processing tool.

We focused on capturing plants at different stages of development, applying the two-pass walking pattern for 2m width Canola and Carinata ranges. Figure 5.5 shows samples of the stitched images of Canola plants at various stages of growth.

![Figure 5.5: Canola plants at different stages of growth. Labels of the images refer to a number of weeks from seeding](image)

Table 5.1 shows visual data analysis for identifying stitching artifacts in orthomosaic images of Canola and Carinata of their different stages of growth. As mentioned, we focused on early stage of growth (between 3rd and 5th weeks from seeding), middle stage of growth (6th week from seeding) and late stage of growth (7th week from seeding). Here we consider an orthomosaic image as “bad” when the entire image area is wholly damaged (nothing can be visually seen) and/or critically corrupted by more than one type of artifacts.

<table>
<thead>
<tr>
<th>Stage of growth</th>
<th>All images</th>
<th>Good</th>
<th>Bad</th>
<th>Zippering</th>
<th>Holes</th>
<th>Distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early stage (weeks 3-5), emergence</td>
<td>173</td>
<td>75</td>
<td>68</td>
<td>19</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Middle stage (week 6), vigor</td>
<td>158</td>
<td>84</td>
<td>46</td>
<td>21</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Late stage (week 7), flowering</td>
<td>158</td>
<td>18</td>
<td>50</td>
<td>66</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

**Table 5.1:** Stitching results over different stages of growth regarding different artifacts visually detected in the orthomosaic images
5.5 Results over different lighting conditions

Filming was carried out at different weather conditions. Cloudy days with no strong wind appeared to be the most pleasant days for filming, because clouds diffuse direct sunlight and prevent the occurrence of shadows on images. Figure 5.6 shows stitching results of pictures taken on a cloudy day over different crops and different trials on different days. Stitched images of the same crops but filmed on a sunny day showed in Figure 5.7.

![Figure 5.6: Filming results on cloudy days](image)

![Figure 5.7: Filming results on sunny days](image)
To evaluate stitching stability over different lighting conditions we have inspected 489 stitched images of Canola (5.2). By artifacts, we refer to zippering problem, holes, distortion or an image totally failed to be stitched.

<table>
<thead>
<tr>
<th>Sun</th>
<th>Images inspected</th>
<th>Images with artifacts</th>
<th>Factors affected stitching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloudy</td>
<td>95</td>
<td>34</td>
<td>Parallax effect, Insufficient overlapping between adjacent images or blocks of images,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wrong orientation of 3D meshes, Movements of plants caused by the wind.</td>
</tr>
<tr>
<td>Sunny</td>
<td>394</td>
<td>278</td>
<td>Parallax effect, Insufficient overlapping between adjacent images or blocks of images,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dropped shadows of an operator or/and the camera system, Wrong orientation of 3D meshes,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Movements of plants caused by the wind.</td>
</tr>
</tbody>
</table>

**Table 5.2:** Stitching performance over different lightning conditions for Canola and Wheat images of one- and two-pass walking pattern

Although stitching accuracy for the images taken with the two-pass walking strategy is low mainly because of the zippering problem (Figure 5.1), specific areas of these images may still be used. Visual inspection has shown that areas corresponded to one of the passes (first half of the stitched image or the second half) are relatively clean and does not have artifacts, which allows using certain parts of the stitched images with the zippering problem for further required estimation or/and evaluation.
5.6 Discussion

RAW images from the action camera that form an orthomosaic can be acquired with excellent quality and high resolution, and the quality and resolution of these images remain the same and constant through the entire filming session. However, the quality of generated orthomosaic range images can be affected by many factors. Poor overlap of still RAW images that form orthomosaics often leads to the zippering problem. Bad positioning of 3D meshes may lead to image distortion or stitching failure. Plants movements caused by the high wind lead to holes in orthomosaics. Parallax effect here also plays a significant role, so taller plants tend to be stitched with more artifacts. Moreover, different lighting conditions for the adjacent images, such as shadows dropped by an operator or/and the camera system may also corrupt orthomosaics.

Besides, the zippering artifact of stitching dictates some modifications in filming procedure. More overlapping between the two passes should be obtained, so there might be a need for holding the camera higher above the canopy.

Moreover, plants in their late stages of growth are tall, and during the windy weather, they move notably. Moving of subjects during the filming, especially when multiple images are taken, often leads to stitching errors, because the same plant can appear in different positions on two adjacent images. Such numerous ghosts of the plant badly affect the stitching results.

It can be seen from the Table 5.1 that images of plants on their early and middle stages of growth form orthomosaics with a lower number of artifacts, whereas pictures of plants in their late stage of growth tend to have more artifacts. Moreover, the zippering problem occurs more often for images of plants on their late growth stage. The issue of distortion is common for all images mainly because of the suboptimal positioning of 3D meshes. Holes in late-stage images appear more often because plants at this stage are taller and tend to move more on the wind. Lightning conditions test (Table 5.2) shows that stitching works better for images taken on cloudy days.
CHAPTER 6

CONCLUSION

During the research work, all methodological questions have been answered. We have shown that the GrowPro filming system and GrowPro Viewer software solution are reliable and useful tools for high-throughput field-based phenotyping. Not only a cheap and easy-to-use novel HTP approach, but the GrowPro is also able to combine high capability and flexibility at the same time, taking a worthy place on the HTP spectrum and become a competitive tool to already well-established popular HTP systems.

Results show that it is possible to stitch together a sequence of top-down images to reconstruct a single image of a large breeder plot (e.g. 6m x 2m) when the images are captured at a low height, however, stitching accuracy will decrease as soon as plants become taller. It is not feasible to generate orthomosaic stitched images of an entire field consisting of a large quantity of small close-up images, mainly because GrowPro imaging was focused on capturing ranges within the filed. GrowPro imaging works well for plants at their early stages of growth and shows better results for images taken on cloudy days.

6.1 Contributions

In our research work we have defined the several contributions. We have shown design and development of a novel hand-held remote sensing platform called the GrowPro. This system can be built using cheap, publicly available pieces of equipment. The system is easy to assemble and easy to use in different environmental conditions. The system can be used without specialized training.

By using the GrowPro, complete coverage of the area of interest (e.g., an individual plot or range of plots) can be obtained. Taking images with the GrowPro setup allows to have a closer view of the plant, so detailed high-resolution images of the plant can be obtained. We have collected over 890000 still images of plants (Canola, Carinata, PhenoLentil, PhenoWheat, PhenoCanola, Lentils) of 2017 growing season, which exceeds 1915GB of data.

Semi-automated data pre-processing tool called the GrowPro Viewer has been developed, deployed and tested on different image sets. The GrowPro Viewer can perform data verification and labeling, GPS-based image classification or clustering and forwarding image sets for stitching and orthomosaic generation with Agisoft Photoscan. For broader data analysis, we have also developed and implemented a method for initial GPS nosiness evaluation. This approach can help to predict stitching accuracy and performance before the
actual stitching process. Moreover, a set of Python scripts has been developed for data management.

We have processed 1051 ranges of Canola DFL, Canol DRIL and Carinata of 2017 growing season and generated 578 orthomosaic range images using the GrowPro Viewer. We also showed and analyzed best samples of orthomosaic range images of a variety of crops at its different stages of growth and weather conditions.

We have suggested best practices for filming, detailed a data gathering procedure and formulated recommendations for the use of the pre-processing and stitching pipeline. Our results indicate that the GrowPro has a potential for increasing the precision of phenotype analysis by providing higher ground resolution than is possible with UAV imaging, based on two-dimensional images of plants.

The GrowPro system has already been used as a remote sensing tool for capturing images of wheat for 898 CMPT course project “Phenotypic Trait Estimation Through Image Analysis of Wheat”. A research paper “DeepWheat: Estimating Phenotypic Traits from Crop Images with Deep Learning” was accepted for publication as a paper in the proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) [5].

### 6.2 Limitations

Even though the pre-processing pipeline and the described methods and improvements contribute to overall stitching accuracy, there are still some limitations that may significantly affect the performance of data pre-processing. The following list shows all major limitations of the project.

- **GPS inaccuracy of the GPS sensor of the camera.** GoPro Hero 5 action camera that was used for the project has a built-in GPS sensor. The accuracy of this GPS sensor cannot be improved or changed, so we treated GPS information from the sensor as data with a fixed level of GPS inaccuracy. This limitation has been approached with GPS flattening technique, as well as with the method for preliminary data verification and labelling methods.

- **Windy weather.** The province of Saskatchewan has a significant amount of windy days throughout summer time and growing season. In this case, plants tend to move on the wind which subsequently affects stitching performance.

- **Tall plants.** As soon as plants become taller, the parallax phenomenon more significantly affects stitching performance. To overcome this limitation, for taking images of plants in their late stages of growth we moved the camera higher above the canopy to reduce parallax effects.

- **Human errors.** Any human interaction leads to human errors. For example, during long filming sessions (2-5 hours) an operator may lose control over the camera system and sometimes change the camera position. Such inconsistency often leads to insufficient overlapping between images and/or sets of images (for the 2-pass walking pattern).
6.3 Recommendations and future work

Even though the GrowPro system is a cheap, reliable and easy-to-use solution for field phenotyping, based on the results of the research work and user experience during 2017 growing season and data acquisition we would recommend specific improvements.

To increase overall stability of the designed remote sensing tool, the authors of this thesis suggest the following changes within the GrowPro camera system, as well as within pre-processing pipeline.

6.3.1 Tool and filming procedure recommendations and potential improvements

2017 data acquisition and filming process have shown that the major problem caused by the two-pass walking pattern is so-called zipping problem. As explained above, the zipping artifact happens when there is not enough overlapping between the two passes. To solve this problem, we suggest installation of two cameras on the horizontal bar. It will require two operators working together and holding the bar. This improvement can potentially lead to the following results:

- Potentially solve 3D-mesh positioning problem. GPS data from two different cameras would form GPS point sequence with distinct passes.
- Increase overlapping between two passes and eliminate zipping problem (two cameras can be installed in a way that guarantees two field’s of view overlapping).
- Enlarged field-of-view of the filming rig will allow users to use only one-pass walking pattern and speed up the entire process of filming.

On the other hand, the zipping problem can also potentially be solved by moving the camera higher above the canopy (app. 3m height). However, this will reduce overall resolution of the orthomosaic range images.

For the upcoming growing season and data acquisition, we recommend using the remote control unit for image segmentation and the two-pass walking pattern. The main reason for that is that GPS data inaccuracy and nosiness would remain, which may affect classification algorithm’s performance. Moreover, for orthomosaic generation procedure, there should be a distinct plane formed by the GPS points, that’s why the two-pass walking pattern is preferable.

6.3.2 Data analyses recommendations and potential improvements

The GrowPro Viewer has also been proven to be a reliable software tool for a substantial amount of data. However, there is still room for improvements within the designed data processing pipeline.

If the remote control system is not used, there is still a need for precise and effective clustering or classification algorithms. Preliminary tests and results of using K-nearest neighbours and convolutional clustering
techniques have shown their excellent performance and effectiveness for GPS-based image segmentation. However, convolutional clustering still needs more in-depth research work. Initial implementation of convolutional clustering has generated the following tasks for future research work:

- To finally cluster images, all peaks that represent turning points should be identified. The problem of finding peaks can be treated as a separate research task and implies many possible approaches.

- Windows size for the clustering algorithm is critically important and should be selected based on overall data nosiness. In this perspective, finding a connection between data nosiness and window size for the convolutional clustering algorithm would increase image segmentation accuracy.

Moreover, the flattening technique that has been implemented and used still needs some improvements. As it can be seen from the method evaluation, some 3D meshes were sometimes positioned randomly, so there is a need in new approaches and algorithms for positioning all 3D mesh models in the same way, facing the Z direction.

To re-orient a 3D model correctly, as future work we suggest fitting a plane to the set of points in the 3D model in order to automatically detect perpendicular vector to the plane (Figure 6.1). This vector can be used as a reference for a new local coordinate system of the 3D mesh model. Each 3D mesh model would subsequently be oriented in the same direction. Such transformations can be done in an external software for three-dimensional modeling (SolidWorks, Catia, NX, MeshLab, etc.).
6.3.3 Straightening GPS paths

Similar to flattening GPS coordinates of images for placing all images into a single horizontal plane [4,3] which increases overall stitching accuracy, we also recommend GPS data adjustments in two-dimensional latitude-longitude space. Figure 6.2 shows an example of a good GPS track that tends to successful stitching. For further increasing of stitching accuracy and overall pre-processing pipeline performance we suggest straightening of GPS paths so that they would represent “ideal” U-form GPS tracks with two straight lines of points that refer to larger sides of the range. These adjustments can potentially increase stitching accuracy and improve 3D mesh positions.

Figure 6.2: Example of a GPS track with low level of noisiness.
6.4 Conclusion

The GrowPro system has taken an important place within the spectrum of field-based phenotyping tools. This is a cheap, effective and easy-to-use system with flexibility comparable to UAV systems and capability comparable to pushed vehicles. Moreover, the GrowPro Viewer software, designed for data processing turned out to be a promising and, in many ways, effective solution for analyses of a substantial amount of data. Both, the GrowPro and GrowPro Viewer can be easily replicated and scaled.

We have shown the feasibility of stitching together a sequence of top-down images to reconstruct a single image of a large breeder plot when the images are captured at a low height. We generated range orthomosaic images consisting of a large quantity of small close-up images with resolution, much higher compared to images captured by an aerial drone. We identified challenges of filming large ranges with a two-pass walking pattern. Based on the analysis we recommended improvements for increasing the performance of the GrowPro system.

GrowPro imaging works well for different crop types and variable lighting and weather conditions in the field, however, early stages of growth are more preferably for imaging due to small height of the plants, which reduces parallax effect and movements of the plants caused by the wind. However, data processing procedure and pipeline still needs some improvements that would increase overall stability and reliability of the entire GrowPro HTP system.

The GrowPro system has taken a prominent place in the HTP spectrum of systems for field phenotyping. The GrowPro is more capable than UAV approaches because it can produce images of noticeably higher resolution than drone images. From this perspective, the height of the camera above the canopy can be considered as the core advantage of the GrowPro, because it is impossible to take drone images of plants from the same height because of a UAV has a minimum altitude at which it can fly. The lowest altitude UAV flights from the University of Saskatchewan drone team is 15 meters. On the other hand, the GrowPro is less flexible in terms of the task, because the GrowPro needs a significant human-machine interaction. Besides, the GrowPro is more flexible than any other pushed or motorized vehicles because the GrowPro is very easy to deploy and set up for filming, whereas more sophisticated motorized vehicles may be a lot more challenging to deploy, use or scale, and may also require more than one operator or specialized training.
REFERENCES


APPENDIX A

GROWPRO USER GUIDE

A.1 Before filming on the field

- Make sure the battery pack is charged.
- Make sure the camera is charged.
- Copy all data from the previous filming session.
- Erase the microSD card.
- Keep the lens clean and protect during transportation.

A.2 Prepare the camera rig for filming

- GoPro camera:
  - Choose time-lapse mode (0.5s, linear mode) Insert microSD card.
  - Plug in the gimbal to the battery pack.
- Place and fix the camera on the gimbal using the bracket and bolts.
- Turn on the gimbal and initialize it (see the gimbal user guide).
- Plug in the extendable cable to GoPro camera for constant charging (do it after turning the gimbal on!).

A.3 Filming procedure

- Point the camera down using the gimbal.
- Put on the camera strap.
- Press the “Start” button on the camera and begin recording.
- Use start-stop technique for automatic range segmentation (A.5).

A.4 Always remember!

- Be sure and double check if GPS is active (the GPS pictogram should be white, not grey!).
  To activate GPS, you might need to turn the camera on before going to the field, so the camera can detect satellites during your transfer to the field.
- Make sure the gimbal and the camera is connected to the battery pack during filming.
- Walk as slow as it possible.
  (60% of overlapping is required!).
- Check if the camera covers entire area of the plot.
  Adjust the level of the camera to have the range covered.
A.5 Remote clicker technique

- Use the remote control clicker and pause recording each time you finish a row/plot/range (it depends on field layout and walking pattern).

- By pausing/resuming recording you separate sets of images. These sets correspond to minimal field units (row/plot/range). In this case, each set of images will represent the entire coverage of an area of interest.

- This approach will allow to avoid manual grouping and will make data segmentation much easier for us.

A.6 Additional notes

Keep everything charged. One extra battery pack is for avoiding any troubles with battery life. Make sure that the camera, the gimbal, the remote and both battery packs are charged before going out.

Keep all your equipment clean. Keep the camera clean, in a case and apart from the rig to avoid scratches. Fix SD cards properly inside the card boxes.

Try not to take the rig apart, because this might cause troubles (not having the stuff plugged into the battery and so on). Fix all cables on the mono-pod with tape.

Make notes after each filming on paper/Slack so the data can be stored accordingly.

A.7 Gathering data procedure

Use GoPro Quick application to segment all images and collect them according to the plan of the plan. Each set of images should represent either a row, or a plot, or a range. Pipeline allows stitching only by plot/row/range.
APPENDIX B

DATA MANAGEMENT

The following guideline describes required steps for data management (3D mesh positioning, orthomosaic generation, export orthos, import files to a database), if GPS altitude zeroing hasn’t been done. A set of python scripts has been designed to process the data.

B.1 Orthomosaic generation and exporting

- First change path within the script so that the date and block match the current project you are working in. The path is organized as /ortho_view_data/block/date where the block and date are to be changed to match the projects date and block. This path is where matrix information of the current view are saved.
- Select a chunks 3D model so it is bolded, reset the view so the current chunk is centered and alone.
- Enable Show Cameras on top toolbar to see where turn is oriented. Rotate using model so it is horizontal and turn is on left hand side.
- Click Batch process and run script, change path to where the script is located and run.
- Continue with all remaining chunks.
- Change path within script to match the project you are currently working in.
- Batch process and run script to build orthomosaics for all chunks.

B.2 Data preparation for Plot Segmentation tool

Exported orthomosaic images should be synchronized with a DataBase for further segmentation.

Step 1: Sync images with data server (Onomi)

Open command prompt on your computer and type “bash” command in order to run “rsync” software that performs synchronization. This command will synchronize new exported images to a folder of the same structure on “Onomi” server. It then calls a general bulk_insert call to insert them into the database.

Step 2: Insert images into database

Run the script to insert all .tif files in the specified path to the database. It uses information from the path organization and the other tables within the database themselves to create the rows which it should insert.

Step 3: Resize images for Plot Segmentation Tool

Run the script to resize images so that they can be seen in the browser for the plot segmentation tool.

Below is the detailed step-by-step description of the data management process:

- Run insert code, it is on git in the db_connection branch in p2irc/data_server/gopro_insert.py
- Process requires a conda environment which after set up is ran in the console by: Entering source activate p2irc. Then entering python gopro_insert.py.
- Code uses config file for logging into onomi, this info should be changed to match your credentials.
• Run script to resize images so that they can be seen in the browser for the plot segmentation tool.

• Run resize_images.py with the same steps as above. It is also located in the db_connection branch in p2irc/data_server/gopro_insert.py