

Crop Row Detection and Waypoint Definition Using Aerial Images of Fields

Alireza Amiri

Deptartment of Mechatronics Engineering
Khaje Nasir Toosi University of Technology (KNTU)
Tehran, Iran
ali.amiri@email.kntu.ac.ir

Saeed Khankalantary

Deptartment of Mechatronics Engineering
Khaje Nasir Toosi University of Technology (KNTU)
Tehran, Iran
s.kalantary@kntu.ac.ir

Abstract—As autonomous agriculture evolves, using wheeled mobile robots for various tasks necessitates precise waypoint generation to define the robots' paths accurately. This paper introduces a method that leverages aerial imagery to detect crop positions and determine waypoints. A specialized hardware setup, consisting of a high-resolution camera, wireless transmitter, and receiver, is developed to capture and transmit live images of the agricultural field. In the image processing stage, crops are identified through two parallel techniques—Unet and K-means clustering. Subsequently, the Hough Transform is applied to detect crop row lines, refined through filtering to ensure that a single, accurate line represents each row. Finally, by selecting specific points on the paths between these rows and converting them into global coordinates, the system facilitates real-time crop detection and precise waypoint generation, supporting autonomous navigation for agricultural robots.

Index Terms—Crop Row Detection, Hough Transform, Image Processing

I. INTRODUCTION

The integration of advanced technologies, such as artificial intelligence, sensing systems, and autonomous robots, is crucial in addressing global food challenges by enhancing agricultural productivity and sustainability [2], [3]. Precision Agriculture (PA) has emerged as a smart management system that optimizes input distribution, such as water and fertilizers, based on site-specific needs, thereby improving crop yield and resource-use efficiency while minimizing environmental impact [5], [6]. Mobile robots in precision agriculture have been utilized in various field tasks, such as fertilization, irrigation, weeding, harvesting and crop picking [2], [3].

In precision agriculture, traditional GPS-based path planning for agricultural machinery remains common, yet it presents challenges, such as the risk of seedling injury due to deviations between the ideal path and the actual crop rows [1]. To address these issues, the use of machine vision-based crop row detection on unmanned agricultural machinery has gained attention, allowing for real-time, precise path planning that minimizes crop damage [1], [8]. However, the unstructured agricultural environment complicates accurate navigation and autonomous operations, necessitating the integration of onboard sensors, such as scanning lasers and machine vision cameras, to enhance the robot's ability to sense and interact with its surroundings [2], [3]. Despite these advancements,

ground-based platforms face challenges, including soil compaction and vibrations from uneven terrain, which can be mitigated by utilizing UAVs for high-resolution, low-altitude aerial sensing [10].

Unmanned Aerial Vehicles (UAVs) have increasingly become a vital tool in precision agriculture, offering a flexible, cost-effective platform for high-resolution remote sensing [9], [12]. UAVs can capture detailed imagery under different conditions, providing fine spatial resolutions and covering significant areas, which is essential for monitoring crop variability and supporting temporal analysis [10], [12]. These platforms excel in applications such as vegetation segmentation, weed management, and crop row detection, filling the gap between terrestrial and satellite-based remote sensing [7], [13]. Despite certain limitations, such as flight endurance, the ability to conduct self-automated flights and provide timely data collection makes UAVs an indispensable tool in modern agriculture [11], [13]. However, aerial image post-processing is necessary to differentiate crop rows from soil and weeds, highlighting the complexity of their integration in precision farming [6]. The initial step in the procedure of crop row detection is semantic segmentation of aerial images, in order to determine the positions where vegetation exists. Different approaches were experimented in previous works in order to achieve a trusted and accurate model for segmentation.

The literature on crop row detection encompasses a range of approaches, each employing distinct methodologies to enhance precision in agricultural applications. Vegetation indices (VIs) like NDVI, ExG, and SAVI are frequently utilized as inputs for various detection methods, including thresholding algorithms, K-means clustering, and the Minimum Distance to the Mean (MDM) classifier. These methods are instrumental in segmenting vegetation from soil backgrounds and effectively identifying crop rows [1], [6], [13]. Fusion approaches, which combine RGB and NDVI data, are also explored to improve segmentation processes, particularly for autonomous robotic navigation in agricultural fields [5].

AI-based approaches have gained prominence, particularly deep learning models, which have significantly improved detection accuracy. For instance, CRowNet, which combines a convolutional neural network (CNN) with the Hough transform, demonstrates robust detection capabilities across various

crop types and field conditions, achieving high detection rates even in complex scenarios like curved or intersecting rows [8], [14]. Additionally, networks like U-Net, SegNet, and ModSegNet are used in conjunction with both traditional and deep learning-based semantic segmentation methods. These networks have shown varying degrees of effectiveness, with deep learning models generally outperforming classical approaches, particularly in challenging conditions where traditional methods may falter [5], [13].

Traditional computer vision techniques are widely employed in crop row detection, with the Hough Transform being a prevalent method for line detection [2], [15]. Despite its extensive use, the Hough Transform has inherent limitations such as high computational complexity and sensitivity to noise, which compromise its suitability for real-time applications [2]. To address these issues, various adaptations, including the Probabilistic and Multi-scale Hough Transforms, have been developed [2]. Moreover, combining the Hough Transform with deep learning has shown to improve detection accuracy [8]. Linear Regression is another commonly used technique, valued for its simplicity and computational efficiency, especially when integrated with preprocessing steps like image segmentation and feature extraction [2], [3]. Other methods such as the Horizontal Strips Method and Blob Analysis are also utilized in crop row detection. The Horizontal Strips Method enhances computational efficiency by bypassing additional segmentation steps but may suffer from accuracy issues due to factors like camera angle and missing rows [2]. Blob Analysis, which groups connected pixels into blobs to generate crop rows, can struggle in environments with high weed density [2]. The Random Sample Consensus algorithm is another approach, providing robust row detection by estimating mathematical models from data with outliers, though its effectiveness depends on factors such as the quality of extracted feature points [2]. Machine learning methods, including clustering techniques like K-means and deep learning models such as Faster R-CNN, YOLOv3, and SegNet, are increasingly being applied in this domain [2], [5]. These methods offer significant advantages, although challenges remain, particularly in handling varying field conditions and limited annotated data [2], [5].

In precision agriculture, extracted crop row and inter-row information serves as a foundation for various tasks, such as guiding autonomous ground vehicles by defining navigation paths between rows, which are crucial for tasks like automatic path computation and the development of vigor maps for field partitioning [11]. For aerial images, the navigation path is identified as the line between two crop rows, while for unmanned agricultural vehicles, it is defined by the central angle between two rows [1].

This paper is structured as follows. In Section A, the hardware setup for aerial imagery data acquisition is detailed. Section B explains the image processing procedure, including image splitting and the application of two methods—K-means clustering and the U-Net model—for detecting plants from aerial images. Section C describes the crop row detection pro-

cess, utilizing both linear regression and the Hough transform. Section D outlines the path detection process. In Section E, the reconstruction of the split images is addressed. Finally, Section F explains the conversion of local waypoint positions in the images to global coordinates.

II. METHODOLOGY

A. Data Acquisition

The initial step in this research involves acquiring essential environmental data, including aerial images and the global coordinates of the camera at the time of image capture. This data is crucial for converting the local pixel positions in the images into global coordinates. Various cameras, such as Sequoia, RedEdge, Micasense, and DJI Zenmuse X7, were considered based on their technical properties and available resources. The GoPro Hero 4 camera was selected for this project because it captures high-resolution RGB images at a high frame rate. Although it lacks NIR and thermal capabilities, RGB images are sufficient for the image processing algorithms used in this study. In addition to aerial images, global coordinates are obtained using a Ublox-Neo-6m GPS module, which transmits the data from the camera's location to the processing unit. The camera's height is predefined rather than measured and transmitted in real-time to simplify the setup.

B. Image Processing

This section focuses on the crucial task of crop detection within the captured images, which is essential for identifying crop rows and determining the path for subsequent operations. The process begins with a preprocessing step that converts the raw images into a format compatible with the models used for detection. Two crop detection methods, K-means and Unet, are then implemented, with a detailed explanation of each approach. The performance of these models is compared to select the most effective one for this application. Additionally, insights from previous studies using these algorithms inform the analysis.



Fig. 1. An example of aerial images

1) Preprocessing: High-resolution aerial images covering large areas are typically too large to use in machine learning algorithms directly. To address this, the images are initially split into equal-sized sections, making them manageable for algorithmic processing. The size of these sections is determined by the specific requirements of the algorithms used, ensuring compatibility and optimized performance. Maintaining uniform image sizes across all sections enhances processing speed and efficiency.

2) Crop Detection: This project's most computationally intensive aspect is crop detection, which is crucial in ensuring accurate outcomes for subsequent procedures. A well-defined layer mask for field images is essential, as the crop detection system must effectively distinguish plants and crops from other objects, including background soil and unwanted vegetation such as grass and weeds.

a) Color Filtering Based Image Segmentation: By analyzing aerial images of crop fields, the idea emerged of applying color filtering to isolate areas within a specific range of green, corresponding to vegetation. Given that the input images are in RGB format, the initial approach involved using only the green channel to filter out pixels based on their green intensity. However, this method proved unsatisfactory for several reasons. First, white pixels containing high green intensity could not be effectively filtered using just the green channel. Additionally, areas with a more yellowish hue were incorrectly excluded by this filter. To address these issues, the RGB images were converted to the HSV (Hue, Saturation, and Value) format, which allows for more precise color filtering based on hue. The vegetation color range was then defined and applied to the images, resulting in improved segmentation of crops from the soil. The output of this process was a binary black-and-white mask, where white pixels represented areas containing plants.

A binary K-means clustering algorithm was developed and applied to group the white pixels to refine the layer masks further, effectively clustering vegetation areas. Through an iterative process, the algorithm parameters were tuned to achieve optimal results.

b) Unet-Based Image Segmentation: Although the color filtering method demonstrated satisfactory performance, it struggled to distinguish crops from other vegetation, such as grass or weeds, present in the field. This limitation arises because the algorithm passes all vegetation through the color filter, disrupting crop detection. These challenges highlight the inadequacy of color filtering for this task and underscore the need to utilize machine-learning models that intelligently differentiate crops from other vegetation.

Access to a suitable dataset is essential to train a machine learning model. For this purpose, the vineyard aerial images dataset [REF] was employed, which includes annotated layer masks indicating the positions of plants within the images. The images were initially split into equal-sized segments of 128 by 128 pixels to ensure consistency in the input to the model, eliminating the need for resizing during processing. Subsequently, a Unet model was defined and compiled based

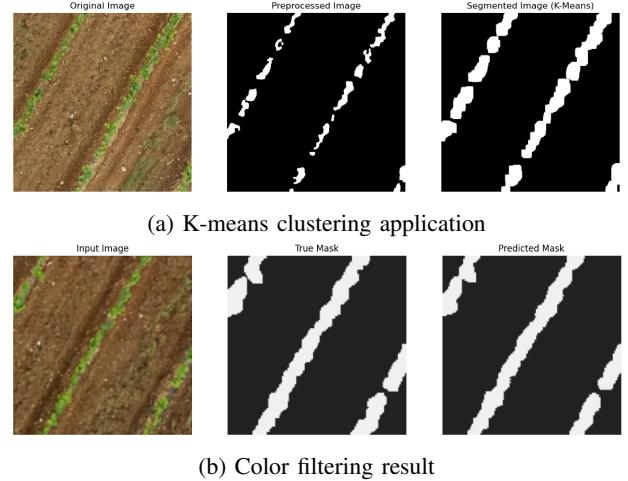


Fig. 2. Crop Detection methods: a) Using Color filtering and Kmeans clustering. b) Using Unet model

on the architecture outlined in [REF]. The model achieved a final accuracy of 96% which was obtained after seven epochs, with early stopping criteria applied. This result confirms that the Unet model can effectively capture image features and accurately segment the crop areas.

C. Crop Row Detection

Crop row detection involves assigning a line with a defined slope and intercept to each crop row. Based on the binary mask of segmented crops obtained in the previous step, lines must be fitted to the pixels to minimize the least square error from the line. This study analyzed and implemented two approaches—linear regression and the Hough Transform—for this purpose.

1) Linear Regression: This experiment used the binary layer mask of segmented crops to assign each pixel to its corresponding crop row. The first step involved determining the optimal angle of the crop rows through an iterative algorithm, assuming that all rows are parallel and share the same slope. Once the slope was defined, the x and y coordinates of the white pixels were rotated to align the crop rows vertically, simplifying the clustering process.

At this stage, the K-means algorithm was applied with a customized loss function designed to minimize the least square error of the distances between the white pixels and a vertical line. Unlike standard loss functions that measure the distance from a centroid point, this approach considered each cluster centered around a line, not a point.

Given the variability in the number of clusters across different images, the clustering algorithm is initiated by grouping nearby pixels. When a pixel was too distant from the existing group, it was treated as the center of a new cluster. However, this approach led to the formation of numerous unwanted clusters. To address this, adjacent clusters were merged in a subsequent step. Finally, the initial rotation of the pixel coordinates was reversed to visualize the resulting clusters.

As illustrated in the results, even after merging close clusters, the method failed to accurately identify crop rows, leading to complications in the later stages of the project. To conclude the experiment, linear regression was applied to fit a line to the pixels within each cluster. However, the results indicated that this clustering and linear regression approach is ineffective for defining crop rows, necessitating exploring alternative methods.

2) Hough Transform: Another widely used solution for crop row detection is the Hough Transform. This computer vision technique is primarily designed to identify geometric shapes in images, making it particularly effective for detecting lines and curves. In this project, the Hough Transform was employed as an alternative to the linear regression model for predicting crop rows. Implementing this method requires significantly less preprocessing and image manipulation, mainly due to the availability of related packages in OpenCV. When the Hough Transform was applied to the images, multiple lines were detected for each crop row, as illustrated in the results. These lines successfully covered the crop rows, indicating the method's effectiveness in line detection. To refine the results and assign only one line per crop row, the detected lines were merged, and a single line was calculated using the average slope and intercept of the detected lines. The final results are presented in FIGURE.

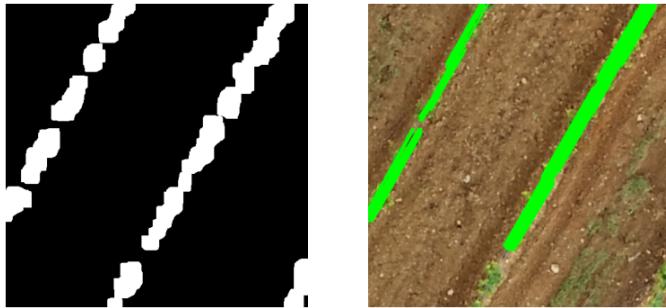


Fig. 3. An example of aerial images

D. Path Waypoint Detection

Defining waypoints becomes straightforward once the crop rows have been established. Initially, a path is defined as a line parallel to the crop rows and equidistant from two neighboring crop rows. A predefined number of equally spaced points are selected along each path line. The coordinates of these points are recorded for use in subsequent steps. FIGURE

E. Reconstruction of Aerial Image

Following the image splitting described in section B.1, the image processing tasks outlined earlier will be applied to each segment of the leading aerial image. The goal is to determine and record the waypoint coordinates within each segment. Once the waypoint coordinates are computed, the initial image must be reconstructed and assembled. Subsequently, the waypoint coordinates need to be converted to align with the coordinates of the newly assembled image.

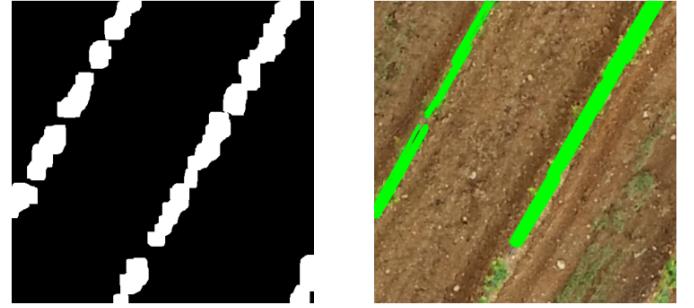


Fig. 4. An example of aerial images

F. Image Coordination to Global Coordination Conversion

An analytical approach and mathematical algorithm are required to obtain the corresponding global coordinates of the points identified in the images. This process involves using the global coordinates of the camera at the time of image capture, along with the camera's height. The conversion process is broken down into two phases: first, converting pixel coordinates into meter distances, and second, converting these meter distances into the global coordinate system.

1) Pixel to Meter Conversion: Assuming the camera is positioned at the center of the image at a height of h , and given the angle θ (defined as the angle between the standard line from the camera to the ground and the line connecting the camera to the edge of the image), the conversion from pixel position to ground distance is described by the following equations:

$$\tan \theta = \frac{x}{h} \implies x = \tan \theta \cdot h$$

where x is the distance from the camera to the edge of the image in meters.

If the image length in pixels is denoted as L , then:

$$1 \text{ pixel} = 2h \cdot \frac{\tan \theta}{L} \text{ (m)}$$

Using these equations, the position of a pixel with coordinates (x, y) in the image can be determined relative to the camera's position on the ground.

2) Meter to Global Coordination Conversion: The final step involves converting the meter distances obtained from the previous calculations into global coordinates. The relationship between meters and degrees of latitude or longitude is given by:

$$1 \text{ (m)} = 0.00001^\circ$$

This conversion factor enables the translation of local pixel positions, measured in meters, into global coordinates.

III. CONCLUSION

In response to the increasing demand for autonomous agricultural systems, there is a critical need for accurate and reliable waypoints for navigation. This paper presents a solution involving a comprehensive live image capturing, processing, and waypoint generation system. The system is divided into three main sections: data acquisition, which captures and

transmits aerial images; image processing, which identifies crops, determines crop rows, and assigns waypoints; and global coordinate conversion, which translates local waypoint coordinates into global coordinates. This approach ensures that the waypoints are precise and suitable for use by mobile robots or other devices requiring accurate navigation within agricultural fields.

REFERENCES

- [1] Y. Yang et al., "Real-time detection of crop rows in maize fields based on autonomous extraction of ROI," *Expert Systems with Applications*, vol. 213, p. 118826, 2023.
- [2] J. Shi, Y. Bai, Z. Diao, J. Zhou, X. Yao, and B. Zhang, "Row detection BASED navigation and guidance for agricultural robots and autonomous vehicles in row-crop fields: methods and applications," *Agronomy*, vol. 13, no. 7, p. 1780, 2023.
- [3] V. R. Ponnambalam, M. Bakken, R. J. Moore, J. Glenn Omholt Gjevestad, and P. Johan From, "Autonomous crop row guidance using adaptive multi-roi in strawberry fields," *Sensors*, vol. 20, no. 18, p. 5249, 2020.
- [4] N. Cunha, T. Barros, M. Reis, T. Marta, C. Premebida, and U. J. Nunes, "Multispectral image segmentation in agriculture: A comprehensive study on fusion approaches," in *Iberian Robotics conference*, 2023: Springer, pp. 311-323.
- [5] T. Barros et al., "Multispectral vineyard segmentation: A deep learning comparison study," *Computers and electronics in agriculture*, vol. 195, p. 106782, 2022.
- [6] G. Ronchetti, A. Mayer, A. Facchi, B. Ortuan, and G. Sona, "Crop row detection through UAV surveys to optimize on-farm irrigation management," *Remote Sensing*, vol. 12, no. 12, p. 1967, 2020.
- [7] M. Hassanein, M. Khedr, and N. El-Sheimy, "Crop row detection procedure using low-cost UAV imagery system," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 349-356, 2019.
- [8] M. D. Bah, A. Hafiane, and R. Canals, "CRowNet: Deep network for crop row detection in UAV images," *IEEE Access*, vol. 8, pp. 5189-5200, 2019.
- [9] I. Sa et al., "WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming," *Remote Sensing*, vol. 10, no. 9, p. 1423, 2018.
- [10] S. Sankaran et al., "Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review," *European Journal of Agronomy*, vol. 70, pp. 112-123, 2015.
- [11] L. Comba, P. Gay, J. Primicerio, and D. R. Aimonino, "Vineyard detection from unmanned aerial systems images," *computers and Electronics in Agriculture*, vol. 114, pp. 78-87, 2015.
- [12] K. Ramesh, N. Chandrika, S. Omkar, M. Meenavathi, and V. Rekha, "Detection of rows in agricultural crop images acquired by remote sensing from a UAV," *International Journal of Image, Graphics and Signal Processing*, vol. 8, no. 11, p. 25, 2016.
- [13] M. Pérez-Ortiz, J. Peña, P. A. Gutiérrez, J. Torres-Sánchez, C. Hervás-Martínez, and F. López-Granados, "A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method," *Applied Soft Computing*, vol. 37, pp. 533-544, 2015.
- [14] Y. Pang et al., "Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery," *Computers and Electronics in Agriculture*, vol. 178, p. 105766, 2020.
- [15] N. Samet, S. Hicsonmez, and E. Akbas, "Houghnet: Integrating near and long-range evidence for bottom-up object detection," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16, 2020: Springer*, pp. 406-423.