Detection of Rows in Agricultural Crop Images Acquired by Remote Sensing from a UAV

Ramesh K N
Department of Electronics and Instrumentation, Bangalore Institute of Technology, Bangalore -560004, India
Email: rkestur@gmail.com

Chandrika N
Department of Computer Science and Engineering, SJBIT, Bangalore - 560060, India
Email: chandrika.n.chandu@gmail.com

Omkar S N
Department of Aerospace Engineering, Indian Institute of Science, Bangalore-560060, India
Email: omkar@aero.iisc.ernet.in

M B Meenavathi
Department of Electronics and Instrumentation, Bangalore Institute of Technology, Bangalore -560004, India
Email: meenavathi@gmail.com

Rekha V
Department of Computer Science and Engineering, SJBIT, Bangalore -560060, India
Email: rekhavgs@yahoo.in

Abstract—Detection of rows in crops planted as rows is fundamental to site specific management of agricultural farms. Unmanned Aerial Vehicles are increasingly being used for agriculture applications. Images acquired using Low altitude remote sensing is analysed. In this paper we propose the detection of rows in an open field tomato crop by analyzing images acquired using remote sensing from an Unmanned Aerial Vehicle. The Unmanned Aerial Vehicle used is a quadcopter fitted with an optical sensor. The optical sensor used is a vision spectrum camera. Spectral-spatial methods are applied in processing the images. K-Means clustering is used for spectral clustering. Clustering result is further improved by using spatial methods. Mathematical morphology and geometric shape operations of Shape Index and Density Index are used for spatial segmentation. Six images acquired at different altitudes are analysed to validate the robustness of the proposed method. Performance of row detection is analysed using confusion matrix. The results are comparable for the diverse image sets analyzed.

Index Terms—Row Crops, Unmanned Aerial Vehicle, KMeans, UAV Remote Sensing, Mathematical morphology.

I. INTRODUCTION

Today, UAVs are being increasingly used for Agriculture applications in remote sensing[1][18]. Remote sensing from a UAV known as Low Altitude remote Sensing (LARS) compliments traditional remote sensing using satellites. Several characteristics of UAV and remote sensing from UAV lend itself well to agricultural applications. UAVs images can be acquired even on cloudy days, they provide greater flexibility in scheduling imagery acquisition helping in temporal analysis. Other advantages of UAVs are their lower cost, and the lower probability of serious accidents compared with piloted aircraft. UAVs can fly at low altitudes, allowing for acquisition of high spatial resolution imagery which can be used to observe small individual plants and patches. Torres-Sanchez [2] used a conventional UAV with vision spectrum and multispectral cameras for vegetation detection in herbaceous crops of maize, wheat and sunflower. Berni [6] described remote sensing products using rotating-wing and fixed-wing UAVs equipped with commercial off-the-shell (COTS) thermal and narrow band spectral imaging sensors. Results obtained with a low cost UAV system for vegetation monitoring applications yielded comparable estimations, if not better, than those obtained with more traditional manned airborne sensors. Salim Malek [8] described a method based on computer vision for detection and localization of crop rows, especially of small-grain crops, was described. The method was intended for use in a system for automatic guidance of agricultural implements in selective treatment of rows and/or inter-row spaces.

Row detection and counting in crops is an important activity in site specific monitoring and management of farms. Row crop detection is fundamental to determining canopy, crop vigor and crop density. Rows detection
helps in detection of crop specific regions apriori to provide more reliable correlation of NDVI and other parameters [3]. It also finds application in autonomous path planning and navigation of ground robots. Manual detection of row crops is more precise however they are laborious and time consuming. Automated detection of rows by analysis of images acquired by remote sensing from a UAV provides an interesting option.

Several researches have carried out row detection using aerially acquired images. Comba [4] described a new algorithm for detection of grape vine rows based on dynamic segmentation, Hough space clustering and total least squares to automatically detect the grape vine rows. Puletti [5] used unsupervised classification for grape vine row detection. The red channel of RGB image was preprocessed using a high pass filter and was later analyzed using Ward’s modified technique.

Salim Malek [8] classified crop rows with in maize crop-field to distinguish small weed seedlings at early stages for in-season site specific herbicide treatment. An object-based image analysis (OBIA) procedure was applied. Senthilnath et. al. [7] used Spectral-spatial methods in processing of RGB images acquired by remote sensing from UAV. This technique [9] leverages spectral clustering of spectral vision band the results of spectral clustering is further improved by performing spatial segmentation.

From authors survey of literature it is observed that there is no evidence of use of spatial methods for detection of row crops in RGB image acquired by remote sensing from a UAV. In this paper, we propose an unsupervised clustering method followed by spatial method detecting and counting rows. K-means algorithm is used to spectrally cluster the RGB image data set. Clustering is further improved by use of spatial segments which includes the mathematical morphological operations of Skeletonization, region growing and segmentation based on geometric properties of shape index and density index. Further Houghline transform is used to detect the lines. The lines are aggregated to arrive at the number of rows in the crop using a novel aggregation technique.

The paper is organized as follows. Section 2 describes the methodology. The results are discussed in section 3. The conclusion is summarized in section 4.

A. UAV Description and Data Acquisition

In this study, UAV is built using components off-the shelf. In this work, the UAV is used a quadcopter which comes under the multi-rotor aircraft category [7]. The camera used in UAV is a Raspberry Pi which is having 2592x1944, 5 mega pixels sensor resolutions with very high HD video resolutions. Images are taken in an open field tomato crop acquired by UAV with different altitudes. It is having very low cost and low endurance. To have a stable platform during image acquisition VTOL is used in a quadcopter. The payload is of 0.1kg and weight during take-off is 2.1kg.

The spatial resolution of the frames extracted in an image are only based on the altitudes took from an UAV. Due to change in an altitude as the quadcopter flies there may be vary in the altitudes, depth and in spatial resolutions. The video is analyzed only in an offline mode to have a better aerial view of the tomato plant crop. Only when tomato crops shed their leaves, fruits will be visible in aerial view. From the video taken using UAV, the images are cropped when the row crops are clear to analyze.

II. METHODOLOGY

In this paper, a Spectral-spatial method is proposed used for detection and counting of row crops in the RGB image dataset acquired using remote sensing from a UAV. Unsupervised clustering is used for spectral classification. The results of clustering are improved by the use of spatial segmentation. Houghline transform is used to detect lines. Several line more than the number of rows is detected. The lines are aggregated using a novel technique to arrive at the number of crows. The proposed method is depicted in (Fig 1).

![Fig 1. Spectral-Spatial Method](image)

A. Spectral Methods:

K-means Clustering Algorithm:

Kmeans algorithm is used for spectral clustering [11], [12], [13]; The Kmeans clustering is used because it is a very robust and also simple clustering algorithm. The image is an RGB image of 676 X 530 pixels. The Kmeans algorithm assigns every pixel to a cluster. The cluster centers are updated. The cluster assignment and cluster centers are updated until the exit criteria is met. The Kmeans clustering calculated as shown below in the equation (1).
\[ I = \sum_{i=1}^{n} \| y_i^j - c_i \|^2 \]  

(1)

Where \( \| y_i^j - c_i \|^2 \) is the distance measure between the cluster centre and the data points, \( k \) denotes the number of clusters, \( i \) is the number of pixels in the image.

### Spatial Methods:

Since the RGB image is a vision spectrum with low spectral resolution of 3 bands i.e., Red, Green and blue bands, the extraction of the row is incomplete. In such case there is a high misclassification of the row crop pixels which leads to the incorrect and will not produce the clustering efficiently. To address misclassification from K-means clustering, spatial segmentation is used [7].

The key spatial methods used in the row crop detection are as follows:

(i) **Mathematical Morphology:** The main goal of the morphology [10] in image processing is to remove the imperfections such as noise, texture which is in different forms and the shapes. Some of the spatial features are applied to the clustered image to eliminate the misclassified regions by using the morphological operations discussed as below.

**Erosion:** The white pixels in the binary image which are present in the foreground are eroded. The structuring element is a very important parameter in the morphology [7]. The structuring element is a very small range of pixels each with size of either zero or one. The eroding of the binary image is done by the structuring element called ‘disk’. The pixels with size zero and one specify the exact shape of the binary image. The erosion of the image \( b \) is defined by \( b \odot d \), where \( d \) is the structuring element. A new binary image is produced \( h = b \odot d \in (Y) \) of a set. \( Y \) is defined by the structuring element \( d \). It is calculated as shown in the equation (2).

\[ Y = \{ y \mid \forall d \in d, d + (-b) \in Y \} \]  

(2)

**Closing:** The closing of a binary image is dilation followed by erosion. Closing is applied here to remove the thin interiors in the image in which each row is divided into individual contiguous figure. In the binary image the counters which are in white pixel is always increased by one unit. A ‘line’ structuring element is used. Here \( P \) is the original image and \( Q \) is the closed image. The holes in foreground of an image will be filled only if it is smaller than \( Q \). We are defining \( A \) for closed operation. The closing operation is defined as shown in equation (3).

\[ A = (P \odot R) \oplus R \]  

(3)

(ii) **Skeletonization:** The noise in the image is removed. The threshold value in the image depends on the resolution of the entire image. The Skeletonization [14], [15] gives the exact representation of the shape but the original size of the object remains the same and also the topology of the image is same. The eroded image is then dilated and subtracted from the eroded image to get the exact skeletonized image. \( Y \) is defined by the structuring element. \( S_j \) indicates Skeletonization of an image \( Y \); \( k \) is successive erosions applied for original image by using structuring element. \( Q \) is the closed image. The Skeletonization of an image \( Y \) is carried out by equation (4).

\[ S(Y) = [Y \ominus kQ] - [(Y \ominus kQ) \ominus Q] \]  

(4)

(iii) **Region Growing:** Some of the connected objects in the skeletonized image are disconnected. To connect the disconnected objects region growing methodology [14], [15] is applied. The image is dilated with the structuring element ‘line’. The neighboring pixel is always coordinate from the centre of the pixel in the image. The region growing of an image \( A \) is carried out as shown in equation (5). Where \( T_i \) is the connected region, \( i = 1, 2, 3 \ldots n \) and \( T \) is the region.

\[ T_i = T \]  

(5)

(iv) **SI & DI Index:** Shape and density index are 2 geometric properties used in image segmentation. The connected objects in the image due to more white pixels which are unwanted for the process are removed by applying the SI & DI [14], [15]. The threshold is used with some specific value for the process. By applying the SI & DI, the complete noise and the unwanted texture is removed. The SI and DI are calculated as shown in the following equations (6) and equation (7).

\[ DI(j, l) = \sqrt{\frac{\text{Area}}{1 + \sqrt{\text{var sum}}}} \]  

(6)

\[ SI(j, l) = \frac{n \times \text{Perimeter}}{4 \times \sqrt{\text{Area}}} \]  

(7)

(v) **Houghline Transform:** To extract the lines in the image Houghline Transform is applied for the images. Line detection is a major part because each row is identified by an individual line. Canny edge detection is used with the specified threshold value. Here the detected lines are specified in color red. Based on the threshold value the lines are incremented and counted in a right way. The Houghline Transform is applied using the equation (8), where \( P \) is the length from the origin, \( \theta \) is the orientation, for any point on the line \( (a, b) \), \( \theta \) is constant.

\[ P = A \cos \theta + B \cos \theta \]  

(8)

(vi) **Line Aggregation:** At the final step, the proposed method is designed to count the line count of each individual detected line. In this, the line count is done by taking the middle row with the pixels 240 and traversing each other. The line count is increased only when number of black pixels is greater than 50. The threshold value of
black pixel is considered to be >50 which is in between two red lines and increases for correct counting of lines.

III. RESULTS AND DISCUSSIONS

In this section, the results obtained from the UAV images are discussed. The proposed method is applied to six images acquired at different altitudes. To check the consistency, we are taking several images and applying the spectral-spatial classification. The following figures show the different results with the different datasets. The images taken for the implementation are the heterogeneous images at different image altitudes which in turn resulted in image pixel resolution. The original image is an RGB image of 676 X 530 pixels. The Kmeans algorithm assigns every pixel to a cluster. The cluster centers are updated. The KMeans algorithm clustered image is shown in the figure 2(b). Due to the misclassification in the cluster the segmentation is not accurate so we are using morphological operations like erosion, dilation, opening and closing. The eroding of the binary image is done by the structuring element called ‘disk’ with value 6. For a close operation a ‘line’ structuring element is used with the radius 6. The morphological operations are shown in the figure 2(c) and 2(d). The resulted images shows contiguous figure and all rows are separated individually. Due to discontinuities present in the image we eliminate the discontinuities by applying skeletonization. The structuring element ‘Morph Cross’ is used with the radius 3. The skeletonized image is shown in the figure 2(f). Since there are many disconnected objects in the image by applying region growing all disconnected objects are connected. A preprocessing step called dilation is used in region growing. The image is dilated with the structuring element ‘line’ with radius 2. Successive dilation is performed with the structuring element ‘line’. The dilated image is show in the figure 2(g). The region grown image is shown in the figure 2(h). The noise which is present in the image is removed by applying Shape Index and Density Index to the image. The threshold value for SI must be greater than 0.88 and for Density Index the threshold value must be less than 0.33. The noise is removed in the image and is shown in the figure 2(i). As a next step Houghline transform is applied to extract each line in the image. The threshold value is set to 50. Lines are fairly delineated in the image and are shown in the figure 2(j). Finally line aggregation is done and number of rows present is counted. The lines are aggregated to the actual number of rows and are shown in the figure 2(j). The number of rows are counted and shown in 2(k).

Fig. 2. UAV Image 1: (a) Original image (b) Kmeans (c) Erosion (d) Dilation (e) Close (f) imfill (g) Skeletonization (h) Succesive Dilation (i) Region Growing (j) SI & DI Index (j) Houghline Transform (k) Actual Count = 5 ; Line count=5.

To check the robustness of the proposed method several UAV images at different image altitudes which in turn resulted in image pixel resolution are considered and the spectral-spatial classification is applied. The results for different datasets are as shown below.

UAV Image 2:

Fig. 3. UAV Image 2: (a) Original image (b) Kmeans (c) Dilation (d) Close (e) Skeletonization (f) Succesive Dilation (g) Region Growing (h) Skeletonization (i) Shape and Density Index (j) Houghline Transform (k) Actual Count = 13; Line count=12.
The table [2] below shows the performance validation [19] using confusion matrix [7] [9]. Accuracy is checked for all six images and results are presented in the following table. Images 1, 3 and 4 give 100% accuracy. Image 2 is approximately accurate. Image 5 and 6 are acquired at a higher altitude so the results are not accurate. In the Table 1, for all images TP, FP, TN, FN [16], [17] gives the actual number of rows present in each image. In Table 2, TPR, FPR, FNR, TNR and Accuracy is been calculated.

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Image 2</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Image 3</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Image 4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Image 5</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Image 6</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
</tbody>
</table>
Table 2. Calculation of Performance Validation

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>TPR</th>
<th>TNR</th>
<th>FPR</th>
<th>FNR</th>
<th>F1 Score</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Image 2</td>
<td>0.84</td>
<td>0</td>
<td>0.07</td>
<td>0.96</td>
<td>0.96</td>
<td>92.3</td>
</tr>
<tr>
<td>Image 3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Image 4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Image 5</td>
<td>0.62</td>
<td>0</td>
<td>0</td>
<td>0.37</td>
<td>0.77</td>
<td>62.5</td>
</tr>
<tr>
<td>Image 6</td>
<td>0.64</td>
<td>0</td>
<td>0</td>
<td>0.36</td>
<td>0.78</td>
<td>64</td>
</tr>
</tbody>
</table>

Where,

TP: True Positive
FP: False Positive
TN: True Negative
FN: False Negative

TPR: True Positive Rate = TP/ (TP+FN)
FNR: False Negative Rate = FN/ (TP+FN)
TNR: True Negative Rate = TN/ (TN+FP)
FPR: False Positive Rate = FP/ (TN+FP)
F1 Score: 2TP/ (2TP+FP+FN)

For a good classifier TPR, F1 score and Accuracy is close to 1 or 100% and TNR an FNR is are zero. Fig 8 is a box plot for visualization of the results. It can be observed that the performance parameters are lower for image 5 and image 6. This is because images 5 and 6 are acquired at a higher altitude as compared to the other images. Image 1 is taken at the lowest altitude as compared to the other images hence the performance results are close to the expected results of a good classifier.

The performance of row crop detection for all images is statistically visualized using box plots [20] as shown in Fig 8. It can be observed that TPR, F1SCORE and ACCURACY are closer to 100% indicating good detection of rows. Further, the range and inter quartile range is low for F1SCORE as compared to TPR and ACCURACY. FNR is towards 0 indicating good performance on not detecting false negatives.

IV. CONCLUSION

In this study, we successfully carried out detection of rows in open field tomato crop using our proposed spectral-spatial method. Six images acquired using LARS carried out from quadcopter UAV are analyzed using the proposed method. In the proposed Spectral-spatial method, K-means algorithm is used for spectral clustering and spatial methods of mathematical morphology and geometrical operations are used. Six images acquired at various altitudes are analyzed. Results are encouraging; however, this work can be extended to detect rows in vast stretch of crops by analyzing mosaiced images of larger crop extents.

ACKNOWLEDGEMENT

The authors thank Dr. Gautham Anand from IISc for the help in providing UAV Remote Sensed Images. We also thank Mr. Nandigam Hari Krishna from NIT, Srinagar for his help in line aggregation sub step.

REFERENCES

in the department of Electronics and Instrumentation at Bangalore Institute of Technology, Bangalore. His research interests include Image processing; machine learning and remote sensing from UAVs.

Chandrika N received her B.E degree in Computer Science and Engineering from SKSEM, Bangalore, Karnataka in 2014. Currently she is a MTech final year student pursuing Computer Engineering from SJBIT, Bangalore, Karnataka. Her research interests are Image Processing, Remote Sensing, Design and Analysis of Algorithms.

Dr. S.N. Omkar received B.E degree in mechanical engineering from the University Visvesvaraya College of Engineering in 1985, M.Sc. (Eng.) degree in aerospace engineering from Indian Institute of Science in 1992 and Ph.D. degree in Aeronautics Engineering from the Indian Institute of Science, Bangalore, in 1999. Currently he is working as a Chief Research Scientist at the Department of Aerospace Engineering at the Indian Institute of Science, Bangalore. His research interests include Biomechanics, helicopter dynamics, fuzzy logic, Image Processing, Neural Networks and parallel computing.

Dr. Meenavathi M.B. received her B.E. degree in Electronics and Communication Engineering from the Mysore University in 1989. ME. degree in Digital Techniques and Instrumentation from university of Indore, Madhya Pradesh 1994 and Ph.D. degree in Electronics and Communication Engineering from Dr MGR university, Chennai in 2010. Currently she is working as a Professor and Head of department of Electronics and Instrumentation Engineering at Bangalore Institute of Technology, Bangalore. Her research interests include Image processing, signal processing, Neural Networks and fuzzy logic.

Rekha V received B.E. degree in computer science and engineering from the VTU in HKBK college of Engineering in 2002, MTech in Computer Networking from VTU in BMSCE in 2009 and pursuing Ph.D. degree in computer networking from VTU in BMSCE Research Centre. Currently working as an Associate Professor at the Department of Computer Science and Engineering at the SJBIT, Bangalore. Research interests include Image processing, Networking, Cloud Computing, Wireless and Mobile Computing.

Authors’ Profiles

Ramesh K.N received B.E degree in Electronics engineering from BMSE college of Engineering in 1995, the ME degree in Electronics and Communication in 2015 from University Visvesvaraya college of Engineering. Ramesh K.N has 17 years of work experience at Infosys Technologies Ltd, Bangalore. He was a Group Project Manager at Infosys. He has worked in development of features for Core switching and Digital Centrex products in the area of communication protocol ‘development, call processing, translations and routing software. He is currently a PhD student in the department of Electronics and Instrumentation at Bangalore Institute of Technology, Bangalore. His research interests include Image processing; machine learning and remote sensing from UAVs.

Chandrika N received her B.E degree in Computer Science and Engineering from SKSEM, Bangalore, Karnataka in 2014. Currently she is a MTech final year student pursuing Computer Engineering from SJBIT, Bangalore, Karnataka. Her research interests are Image Processing, Remote Sensing, Design and Analysis of Algorithms.

Dr. S.N. Omkar received B.E degree in mechanical engineering from the University Visvesvaraya College of Engineering in 1985, M.Sc. (Eng.) degree in aerospace engineering from Indian Institute of Science in 1992 and Ph.D. degree in Aeronautics Engineering from the Indian Institute of Science, Bangalore, in 1999. Currently he is working as a Chief Research Scientist at the Department of Aerospace Engineering at the Indian Institute of Science, Bangalore. His research interests include Biomechanics, helicopter dynamics, fuzzy logic, Image Processing, Neural Networks and parallel computing.

Dr. Meenavathi M.B. received her B.E. degree in Electronics and Communication Engineering from the Mysore University in 1989. ME. degree in Digital Techniques and Instrumentation from university of Indore, Madhya Pradesh 1994 and Ph.D. degree in Electronics and Communication Engineering from Dr MGR university, Chennai in 2010. Currently she is working as a Professor and Head of department of Electronics and Instrumentation Engineering at Bangalore Institute of Technology, Bangalore. Her research interests include Image processing, signal processing, Neural Networks and fuzzy logic.

Rekha V received B.E. degree in computer science and engineering from the VTU in HKBK college of Engineering in 2002, MTech in Computer Networking from VTU in BMSCE in 2009 and pursuing Ph.D. degree in computer networking from VTU in BMSCE Research Centre. Currently working as an Associate Professor at the Department of Computer Science and Engineering at the SJBIT, Bangalore. Research interests include Image processing, Networking, Cloud Computing, Wireless and Mobile Computing.