Research on Node Detection Algorithm of Tomato Seedlings Based on Digital Image

Xue-guan ZHAO1,2, Xiu WANG1,2*, Jin-wei ZHANG1,2, Kai JIANG1,2 and Shuai SU1,2

1Beijing Research Center for Intelligent Equipment Technology in Agricultural
2National Engineering Research Center Information Technology in Agricultural
*Corresponding author

Keywords: Image processing, Seedling detection, Classification algorithm, Machine vision.

Abstract. According to a large number of studies, the internode length is sensitive to environmental stress. In the quality testing of tomato seedlings, optimization algorithm is proposed to extract tomato seedling stem node, using line scan algorithm to extract the main stem area, improve the accuracy of the main stem region selection, and finally through the bag of words model extract main stem node, this method has good robustness. Through the test, it is found that the accuracy of main stem node detection is lowest at fourth nodes, which is 81%. These results demonstrate that our method has the ability to evaluate the vigor of tomato seedlings quickly and accurately.

Introduction

Intelligent agriculture is the integration of agriculture with artificial intelligence technology and modern information technology, and it is also an urgent technology in the current global agricultural development. The crop growth information acquisition technology is an important technical support of wisdom agriculture, through the establishment of mathematical model of growth of multi information fusion based on scientific evaluation of crop, realize the monitoring and evaluation of crop growth, it is possible to make scientific and intelligent management and decision making by using computer for crop production.

At home and abroad, crop growth information collection includes the acquisition of apparent information (such as leaf area, plant height and biomass measurement, leaf morphological identification) and internal information (such as physical and chemical nutrition information monitoring, leaf and canopy temperature, leaf water potential, chlorophyll content, etc. with the help of external means of access the information)1-5.

Such as Zhai Ruifang of Huazhong Agricultural University using the shape parameter method to identify different rape at seedling stage, through several typical characteristics of rape seedling cotyledon geometry calculation, can automatically identify the rape seedling stage, three leaf and four leaf stage, the research results can provide decision basis for field management6. Jia Biao of Shihezi University using digital camera for real-time tracking and monitoring in cotton, cotton growth monitoring and screening of N nutrition diagnosis features of color sensitive parameters, established the model between different color feature parameters and cotton agronomic parameters between the implementation of cotton growth and nitrogen nutrition status information for rapid monitoring and diagnostic accuracy7.

The Life Science Research Institute of University of Tokyo has carried out image analysis based tomato nodal detection and pitch measurement, and achieved the purpose of node detection by

380
visual bag model, which is a machine learning method. The relative error of internode length estimation is less than 15.4% \[1\]. Based on the existing algorithm, optimization algorithm is proposed to extract the tomato seedling stem node, and the use of linear scanning algorithm in the trunk region selection, and further improve the accuracy of the main stem area selection, finally by extracting the bag of words model of main stem nodes based on the calculation of stem pitch, and the method has good robustness. It solves the errors caused by the traditional manual and manual measurement methods, and reduces the subjective factors, which is of great significance for the development of precision agriculture and intelligent agriculture.

**Extraction Algorithm of Main Stem Region Based on Linear Scanning**

In order to get the main stem node of tomato seedling, we first use ossification and thinning algorithm to get skeleton diagram of seedling image. The intersection point of skeleton line contains main stem branch point, that is main stem node. Through certain way, we can find main stem nodes in many intersection points, first we eliminate the obvious non main stem nodes.

The main stem line is considered to be the position where the main stem is located in the image, the next step is to remove the nodes which are far away from the main stem line \[15\]. These nodes are not on the main stem and can not be used as the candidate node of the branch point. In the binary image of tomato seedlings, the main stem of the seedling can be attached to a straight line from the top to the root. According to this, this paper uses the main stem area extraction detection algorithm based on linear scanning, the principle and process of detection are as Figure 1:

![Figure 1. Schematic diagram of linear scanning.](image)

(1) Line scanning area selection: because the main stem of each image is affected by accidental factors, it cannot guarantee its vertical, often appear as shown in the picture, linear scanning directly to cause inefficiency. Therefore, it is proposed to first judge the lowest point of the main stem., which is the most low-end middle position of O in diameter, then the most left end position A and the right end position B were extracted by scanning method, with OC as the bisector of angle AOB, through the statistics on seedling inclination angle, this paper selects the OC as the center line, the left and right symmetry is scanned by 25 degrees angle.

(2) Crop line detection algorithm: take O point as the endpoint when scanning, the angle increment of scan line is 2 degrees. It produces different slope straight lines, statistics the number of target points falling on the straight line, and finally takes the line contain most target point as the main stem line.

**Main Stem Extracted Based on the Word Bag Algorithm**

By selecting candidate nodes in the main stem area, most of the interference nodes were removed, but some nodes were still not branching points in the main stem area. although there are existence differences between the different branching point region of seedlings, but we can still find some
common places at these branch points region. We don't see too much difference, we can extract the characteristics of the main stem branch points and the characteristics of non-main stem branches between different seedlings, then take the characteristics as the target recognition of visual vocabulary, namely through the bag of words model (Bag-of-words) classification.

SIFT algorithm is the most widely used algorithm in image local invariant feature extraction. So we can manually select the feature area of the main stem node (35x35 pixels) and non-stem node regions, extracting invariant feature points. The K-Means algorithm is used to combine visual vocabulary with similar meaning, and a word list containing K words is constructed.

Finally, count the number of times each word appears in the image, so that the image is represented as a K dimensional numerical vector, and the word bag model histogram is generated, and the classification algorithm based on SVM algorithm is implemented. The algorithm flow chart is as Figure 4:

![Algorithm Flow Chart](image)

**Figure 4. Flow chart of the generation of the word bag model.**

**Test Result**

The program was developed based on Visual Studio 2012 on the Windows 7 system (Inter Pentium 5, CPU 2.80 GHz, 4G Memory), implemented by VC++ language. In this study, all seedlings are taken from the same distance to determine a uniform threshold. Compare of the location algorithm recognize and actual node position of by artificial, determine whether the main stem node.

The skeleton line intersection point of the captured image is shown in Fig 5A, due to the limitation of the size and the field of view, producing areas that are easily interfered as shown, but it will not affect the final extraction result of this research. The image of the main stem area after the straight line scanning is shown in Figure 5B, when the interference point is mainly concentrated in the top area of the seedlings. The main stem nodes extracted by the word bag model algorithm are shown in Figure 5C.

![Image of skeleton lines](image)
Figure 5. Extraction results of main stem nodes.

Recall accuracy = \( \frac{\text{Number of nodes identified by algorithm}}{\text{Actual number of main stem nodes}} \)

Table 1. Identification results of node and pitch of main stem.

<table>
<thead>
<tr>
<th>Node</th>
<th>Recognition accuracy</th>
<th>Time consuming</th>
<th>Pitch accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skeleton pruning</td>
<td>Skeleton pruning</td>
<td>(skeleton pruning + linear scan)</td>
</tr>
<tr>
<td></td>
<td>Linear scan</td>
<td>Distance alg.</td>
<td>Linear scan</td>
</tr>
<tr>
<td>1</td>
<td>95.8</td>
<td>93.7</td>
<td>18.12</td>
</tr>
<tr>
<td>2</td>
<td>94.7</td>
<td>90.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>92.8</td>
<td>87.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>78.3</td>
<td>79.8</td>
<td></td>
</tr>
</tbody>
</table>

Discussion and Conclusion

In this study, an image processing method is proposed to detect the main stem nodes of plants, which combining ordinary industrial camera with machine learning. Because the proposed method is carried out by the computer, it is faster and more accurate than the traditional visual inspection.

From the table, in the two main stem region extraction methods, the first node in the main stem node have the highest recognition accuracy, decrease in sequence. The time consuming of linear scanning method decreased by 5.87 ms. However, the accuracy of the main stem node detection was not significantly improved, and accuracy of linear scanning method is 4.2 percent higher than that of distance algorithm. This study further verifies the accuracy of pitch detection algorithm, count the distance of measured values of the main stem nodes and the actual main stem nodes, through the results found that the accuracy rate of first and second points is the highest, can reach 92.8%, were consistent with the product of two node recognition accuracy.

The pitch detection of this method includes three processes: ossification and thinning, and the elimination of the main stem region points in line scanning, and the acquisition of the main stem branch points based on word bag model, the main stem is not necessarily a vertical state when tomato seedlings were photographed in the field of view, when the stem is tilted, we regard the path of the main stem as a straight line, using linear scanning method in a certain range, this method is successful. Because the mass grown seedlings used in this study grow in greenhouse, the stems are straight. Compared to the literature, determine the main stem node by the distance between the node and the main stem, improve the adaptability of the algorithm, from the table, each node point
detection accuracy is higher through line scanning method, the recognition accuracy of nodes 1 and 2 is up to 95%.

The leaves of tomato plants increase with the increase of growth. Because of the leaf occlusion node, the method of node detection developed here is only suitable for seedling stage. As shown in the picture, As shown in the map, the location of the fourth main stem points is prone to background hiding or foreground occlusion in the top area of the seedlings, it is easy to cause misjudgement, so the detection accuracy of the fourth nodes is also low, which is 86.3%. So in the process of plant detection try to keep the fourth node in the spot where it is easy to see.

Acknowledgement

This research was financially supported by the Youth Research Fund of Beijing Academy of agricultural and Forestry Sciences (QNJJ201726).

Reference


