Peduncle Detection of Sweet Pepper for Autonomous Crop Harvesting - Combined Colour and 3D Information

Inkyu Sa\textsuperscript{1*}, Chris Lehnert\textsuperscript{2}, Andrew English\textsuperscript{2}, Chris McCool\textsuperscript{2}, Feras Dayoub\textsuperscript{2}, Ben Upcroft\textsuperscript{2}, and Tristan Perez\textsuperscript{2}

\textbf{Abstract}—This paper presents a 3D visual detection method for the challenging task of detecting peduncles of sweet peppers (Capsicum annum) in the field. Cutting the peduncle cleanly is one of the most difficult stages of the harvesting process, where the peduncle is the part of the crop that attaches it to the main stem of the plant. Accurate peduncle detection in 3D space is therefore a vital step in reliable autonomous harvesting of sweet peppers, as this can lead to precise cutting while avoiding damage to the surrounding plant. This paper makes use of both colour and geometry information acquired from an RGB-D sensor and utilises a supervised-learning approach for the peduncle detection task. The performance of the proposed method is demonstrated and evaluated using qualitative and quantitative results (the Area-Under-the-Curve (AUC) of the detection precision-recall curve). We are able to achieve an AUC of 0.71 for peduncle detection on field-grown sweet peppers. We release a set of manually annotated 3D sweet pepper and peduncle images to assist the research community in performing further research on this topic.

I. INTRODUCTION

Automating agricultural processes such as planting, harvesting, weeding and inspection will play a key role in helping to improve farm productivity, increase crop quality and reduce input costs [1]. In particular, the harvesting of high value crops within horticulture still demands a large amount of hand labour costs due to the dexterity required to achieve the task. For these reasons, automating the task of harvesting is of great interest to the horticultural and robotics industries. Recent research in robotics has made impressive progress towards the goal of developing viable broad- acre [2], [3] and horticultural robots [4]–[8].

This paper focuses on peduncle detection in order to improve autonomous harvesting. It is highly desirable to be able to identify the peduncle prior to performing crop removal (e.g., cutting the peduncle or pulling of the crop using a custom end-effector) as this is one of the most challenging steps of the harvesting process. Accurate estimation of the peduncle reduces the potential of damaging the crop and other branches of the plant. Additionally, retaining the proper peduncle maximises the storage life and market value of each pepper. More importantly, accurate peduncle detection can lead to higher success rates for crop detachment, which in turn yields more harvested crops. Peduncle detection is challenging because of varying lighting conditions and the presence of occlusions by leaves or other crops as shown in Fig. 1. Additionally, while peduncles are usually green, crops such as sweet peppers vary in colour from green through to red (with other variations possible), making it difficult to detect peduncles based on colour information alone.

Prior work in peduncle detection has either identified the peduncle post-harvest or yielded accuracies that are too low to be of practical use. In [9], [10] the detection of the peduncle was performed using a Bayesian discriminant model of RGB colour information. However, such an approach is not practical in-situ for crops such as sweet pepper where the colour alone does not discriminate between peduncles, leaves and crops (see Fig. 1). In-situ detection of peduncle (multi-class detection) using multi-spectral imagery was performed for sweet peppers by Bac et al. [11]. Unfortunately their accuracy was too low to be of practical use. In order to address previous shortcomings, we use both colour and geometry shape features captured from multiple views by an RGB-D camera.

The aim and scope of this paper is peduncle detection from a 3D reconstructed model of a detected sweet pepper. In our previous work [12], we described a detection system based on deep neural networks which is capable of performing the detection step and in [6], we introduced a sweet pepper pose estimation method that combines point clouds from multiple viewpoints into a coherent point cloud using Kinect Fusion [13]. In this work, we thus assume that pre-registered

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Sweet pepper picking in operation showing a robotics arm equipped with an end-effector tool to harvest the pepper by cutting its peduncle. The photo highlights the presence of occlusions and varying lighting conditions of peduncles and sweet peppers grown in a field environment.}
\end{figure}
3D models of the scene containing the peduncle and sweet pepper are already obtained and the presented method is deployed afterwards.

The main contribution of this paper is the use of 3D geometric features, namely Point Feature Histograms (PFH) [14] in addition to colour information for robust and accurate peduncle detection. The proposed detection system operates with any sweet pepper type (red, green, and mixed), and achieves high accuracy. To encourage reproducibility, we share the manually annotated set of 3D capsicum and peduncle models along with their annotated ground truth used to generate the results in this paper.

The remainder of this paper is structured as follows. Section II introduces relevant related works. Section III describes sweet pepper peduncle detection and Section IV elaborates data collection environments and procedures. We present our experimental results in Section V. Conclusions are drawn in Section VII.

II. RELATED WORK/BACKGROUND

This section reviews existing methods for detecting peduncles and crops using 3D geometry and visual features. Such techniques are widely used for autonomous crop inspection and detection tasks. Cubero et al. demonstrated the detection of various fruit peduncles using radius and curvature signatures [15]. The Euclidean distance and the angle rate change between each of the points on the contour and the fruit centroid are calculated. The presence of peduncles yields rapid changes in these metrics and can be detected using a specified threshold. Blasco et al. [9] and Ruiz et al. [10] presented peduncle detection of oranges, peaches, and apples using a Bayesian discriminant model of RGB colour information. The size of a colour segmented area was then calculated and assigned to pre-defined classes. These methods are more likely suitable for the quality control and inspection of crop peduncles after the crop have been harvested rather than for harvesting automation, as they require an inspection chamber that provides ideal lighting conditions with a clean background, no occlusions, good viewpoints, and high-quality static imagery.

There has also been significant progress in crop detection and plant classification for in-field crop harvesting [16]. Recently, a similar work for autonomous sweet pepper harvesting in greenhouses was presented by [17]. The harvesting method within this project used only the centroid of each sweet pepper, with a harvesting tool that encloses the fruit and cuts the peduncle without its detection. Within this project, multi-spectral images (robust under varying lighting conditions) were tested to classify different parts of sweet pepper plants [11]. This work achieved 40% multi-classification for parts of the plant. Although these results were deemed insufficient for use on a real robot, they demonstrated significant progress towards a plant classification system based on hyper-spectral information.

Using an accurate laser scanner would be beneficial for plant classification. Paulus et al. [18] showed that the point feature histograms obtained from different plant structures, such as leaves and stems, are highly distinguishable. They, in turn, could easily classify plant organs with high accuracy using a supervised classification algorithm. Wahabzada et al. [19] demonstrated grapevine, wheat, and barley organ classification using data acquired from a 2D laser scanner mounted on a robotic arm. A 3D point cloud with an accuracy of 45 μm was produced from the scan using a k-means clustering algorithm. Our work differs from this approach in that we use a consumer-level affordable RGB-D camera with significant measurement noise and apply it in a field site.

Strawberry peduncle detection was reported by [20]. The region of interest (ROI) was pre-defined using prior knowledge and the boundary point between a fruit and peduncle was detected using colour information. The inclination of the peduncle - the angle between the vertical axis and the boundary point - was computed. It is possible to easily distinguish the boundary point for red strawberry but is a challenge to exploit this approach for green sweet pepper.

III. METHODOLOGY

Detecting the peduncles of sweet peppers is a challenging task. Unlike other crops, such as apples and mangoes, which have straight and vertical peduncles, the peduncles of sweet peppers are highly curved and sometimes even flattened against the top of the fruit as shown in Fig. 1 and Fig. 2.

In our work, we combine colour information that can distinguish red sweet peppers from their green peduncles with geometry features [21] that discriminate between peppers and peduncles that are both green. In this Section, we present a feature vector representation and the use of a supervised classifier for peduncle detection.

A. Feature representation

It is important to determine the discriminative features of an object we wish to detect. To achieve this, we utilise two features: colour and geometry shape.
Given two points \( p_n \) the query point and its \( k \)-neighbours respectively.

**1) Colour feature (HSV):** As shown in Fig. 2 (a), a highly distinguishable feature between a red sweet pepper and its peduncle is colour. Selecting a colour space for object detection is non-trivial [22] and we use the Hue, Saturation, and Value (HSV) colour space that expresses visual information in the dimensions of colour (H), lightness (S) and intensity (V). Since the HSV space has a component accounting for pixel brightness, its colour components are not significantly affected by varying shades of the same colour. This is useful for visual crop detection, given that 3D objects reflect light differently depending on the angle of incidence onto a surface as shown in Fig. 2 (a). Furthermore, using the RGB colour space may not be appropriate for peduncle detection due to the high correlation between its R, G, and B dimensions. This can result in problems under certain conditions where light reflects off differently from solid-colour objects and shadows obscured by the plants.

Although colour cues provide good distinguishable features for detecting red sweet pepper peduncles, they are only of limited use for green peppers as shown in Fig. 2 (b) since the fruit and its peduncle have similar colour responses. Furthermore, using only colour, it is difficult to distinguish between green peduncles and other plant parts, such as leaves or stems, which are also green.

Using geometry features, such as surface normal and curvature estimates, can be useful for identifying different classes with distinguishable 3D geometries. Fig 4 shows surface normals of red (left) and green (right) sweet peppers. Here, it can be qualitatively observed that the peduncle and the body have different curvature and surface normal directions. In other words, the curvature radius for a peduncle can be generally easily distinguished from that of a sweet pepper body.

**2) Geometry feature (Point feature histograms [14]):**

As shown in Fig. 3 (a), a highly distinguishable feature between a red sweet pepper and its peduncle is colour. Selecting a colour space for object detection is non-trivial [22] and we use the Hue, Saturation, and Value (HSV) colour space that expresses visual information in the dimensions of colour (H), lightness (S) and intensity (V). Since the HSV space has a component accounting for pixel brightness, its colour components are not significantly affected by varying shades of the same colour. This is useful for visual crop detection, given that 3D objects reflect light differently depending on the angle of incidence onto a surface as shown in Fig. 2 (a). Furthermore, using the RGB colour space may not be appropriate for peduncle detection due to the high correlation between its R, G, and B dimensions. This can result in problems under certain conditions where light reflects off differently from solid-colour objects and shadows obscured by the plants.

Although colour cues provide good distinguishable features for detecting red sweet pepper peduncles, they are only of limited use for green peppers as shown in Fig. 2 (b) since the fruit and its peduncle have similar colour responses. Furthermore, using only colour, it is difficult to distinguish between green peduncles and other plant parts, such as leaves or stems, which are also green.

Using geometry features, such as surface normal and curvature estimates, can be useful for identifying different classes with distinguishable 3D geometries. Fig 4 shows surface normals of red (left) and green (right) sweet peppers. Here, it can be qualitatively observed that the peduncle and the body have different curvature and surface normal directions. In other words, the curvature radius for a peduncle can be generally easily distinguished from that of a sweet pepper body.

2) Geometry feature (Point feature histograms [14]):

Although colour cues provide good distinguishable features for detecting red sweet pepper peduncles, they are only of limited use for green peppers as shown in Fig. 2 (b) since the fruit and its peduncle have similar colour responses. Furthermore, using only colour, it is difficult to distinguish between green peduncles and other plant parts, such as leaves or stems, which are also green.

Using geometry features, such as surface normal and curvature estimates, can be useful for identifying different classes with distinguishable 3D geometries. Fig 4 shows surface normals of red (left) and green (right) sweet peppers. Here, it can be qualitatively observed that the peduncle and the body have different curvature and surface normal directions. In other words, the curvature radius for a peduncle can be generally easily distinguished from that of a sweet pepper body.

**Support Vector Machines (SVMs) classifier tuning**

At this stage, we can build a feature vector for a single 3D point with 36 float elements extracted from HSV (3) and PFH (33). A concatenated feature vector \((n \times 36)\) where \(n\) is the number of points, serves as an input for a support vector machine (SVM) classifier. SVMs are one of the most successful and popular binary supervised learning algorithms (i.e. two-class classification) due to their simplicity, flexibility, effectiveness in dealing with high-dimensional spaces and ease of training. We provide insight into our experimental datasets in the following section, as well as choosing a kernel function and its parameters in section V-B. A more in-depth explanation of SVMs can be found in [24] chapter 4.

**1) Colour feature (HSV):** As shown in Fig. 2 (a), a highly distinguishable feature between a red sweet pepper and its peduncle is colour. Selecting a colour space for object detection is non-trivial [22] and we use the Hue, Saturation, and Value (HSV) colour space that expresses visual information in the dimensions of colour (H), lightness (S) and intensity (V). Since the HSV space has a component accounting for pixel brightness, its colour components are not significantly affected by varying shades of the same colour. This is useful for visual crop detection, given that 3D objects reflect light differently depending on the angle of incidence onto a surface as shown in Fig. 2 (a). Furthermore, using the RGB colour space may not be appropriate for peduncle detection due to the high correlation between its R, G, and B dimensions. This can result in problems under certain conditions where light reflects off differently from solid-colour objects and shadows obscured by the plants.

Although colour cues provide good distinguishable features for detecting red sweet pepper peduncles, they are only of limited use for green peppers as shown in Fig. 2 (b) since the fruit and its peduncle have similar colour responses. Furthermore, using only colour, it is difficult to distinguish between green peduncles and other plant parts, such as leaves or stems, which are also green.

Using geometry features, such as surface normal and curvature estimates, can be useful for identifying different classes with distinguishable 3D geometries. Fig 4 shows surface normals of red (left) and green (right) sweet peppers. Here, it can be qualitatively observed that the peduncle and the body have different curvature and surface normal directions. In other words, the curvature radius for a peduncle can be generally easily distinguished from that of a sweet pepper body.

**Support Vector Machines (SVMs) classifier tuning**

At this stage, we can build a feature vector for a single 3D point with 36 float elements extracted from HSV (3) and PFH (33). A concatenated feature vector \((n \times 36)\) where \(n\) is the number of points, serves as an input for a support vector machine (SVM) classifier. SVMs are one of the most successful and popular binary supervised learning algorithms (i.e. two-class classification) due to their simplicity, flexibility, effectiveness in dealing with high-dimensional spaces and ease of training. We provide insight into our experimental datasets in the following section, as well as choosing a kernel function and its parameters in section V-B. A more in-depth explanation of SVMs can be found in [24] chapter 4.

**Support Vector Machines (SVMs) classifier tuning**

At this stage, we can build a feature vector for a single 3D point with 36 float elements extracted from HSV (3) and PFH (33). A concatenated feature vector \((n \times 36)\) where \(n\) is the number of points, serves as an input for a support vector machine (SVM) classifier. SVMs are one of the most successful and popular binary supervised learning algorithms (i.e. two-class classification) due to their simplicity, flexibility, effectiveness in dealing with high-dimensional spaces and ease of training. We provide insight into our experimental datasets in the following section, as well as choosing a kernel function and its parameters in section V-B. A more in-depth explanation of SVMs can be found in [24] chapter 4.
IV. System setup and data collection

Image data training and system testing was performed using a prototype robotic sweet pepper harvester, "Harvey", we developed at QUT. As shown in Fig. 5, this device possesses a 6 DoF robotic arm from Universal Robots (UR5) mounted on a scissor lift platform. The robot arm has a customised harvesting tool mounted on its end effector that both grips each pepper with a suction cup and cuts the peduncle with an oscillating cutting blade. An Intel i7 PC records incoming RGB-D data (640×480) from an eye-in-hand depth camera.

Fig. 6 shows the internal peduncle detection pipeline. The reconstruction of a dense sweet pepper point cloud from multiple views using Kinect fusion is implemented based on our previous work [6] (green). The point clouds are extremely large and noisy, consequently we use a statistical outlier remover and voxel grid down sampler (blue) supported from Point cloud library (PCL) [25]. Surface normals are calculated and fed into the PFH feature extractor. The PFH features are then concatenated with HSV colour information to create the 36 dimensional feature vector (red) used for classification. Finally, peduncle detection is performed using a SVM with the RBF kernel and the trained model from Section V-B.1 (black).

The Robot Operating System (ROS, Indigo) [26] is used for system integration on an Ubuntu 14.04 Linux system. The methods in the green, blue, and red boxes shown in Fig. 6 are implemented in C++ using PCL, while peduncle detection (black) is written in Python 2.7 [27].

A. Training/testing data collection

Field trials for data collection were conducted on a farm in North Queensland (Australia) over a 10-day period within a protected cropping system. Such a system involves growing plants under a structure (e.g., greenhouse) that aims to create an environment closer to the optimum for maximum plant growth and production as shown in Fig. 8. This facility provides relatively even lighting conditions by diffusing incoming sunlight through a semi-transparent plastic/mesh roof and wall.

Data is captured while the robot arm moves through a predefined triangle scanning trajectory around the centroid of a detected red pepper in order to acquire multiple viewpoints (see time of 28s-55s in the demonstration video). The scan takes 15s. During this process, not only is the red sweet pepper (in the centre) reconstructed in real-time but also the adjacent green crops. For each pepper, an RGB-D camera was moved in a scanning motion to view it from multiple directions and the Kinfu software library [25] was used to reconstruct a single point cloud.

We manually annotated a dataset with a total of 72 sweet peppers, example images are provided in Fig. 7. It can be seen that these images have a substantial amount of noise present, this is due in part to the low-cost RGB-D sensor as well as shades/reflections from the object surfaces. We randomly divide these 72 samples into two even groups for model training and testing. The output from our classifier is a binary prediction for each 3D point (i.e., 0 is for sweet pepper and 1 for peduncle).

Table I presents the first set split into a 50-50 ratio of training and test data whereas Table II demonstrates a 79-21 split. These datasets are used for system evaluation in Section V-A.
TABLE I
NUMBER OF ANNOTATED 3D MODELS FOR 50-50 RATIO FIELD TRIP DATASET.

<table>
<thead>
<tr>
<th></th>
<th>Train (peduncle+pepper)</th>
<th>Test (peduncle+pepper)</th>
<th>Num. 3D models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Trip 1</td>
<td>14 (50%)</td>
<td>14 (50%)</td>
<td>28</td>
</tr>
<tr>
<td>Field Trip 2</td>
<td>22 (50%)</td>
<td>22 (50%)</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>36 (50%)</td>
<td>36 (50%)</td>
<td>72</td>
</tr>
</tbody>
</table>

TABLE II
NUMBER OF ANNOTATED 3D MODELS FOR 79-21 RATIO FIELD TRIP DATASET.

<table>
<thead>
<tr>
<th></th>
<th>Train (peduncle+pepper)</th>
<th>Test (peduncle+pepper)</th>
<th>Num. 3D models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Trip 1</td>
<td>21 (75%)</td>
<td>7 (25%)</td>
<td>28</td>
</tr>
<tr>
<td>Field Trip 2</td>
<td>36 (82%)</td>
<td>8 (18%)</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>57 (79%)</td>
<td>15 (21%)</td>
<td>72</td>
</tr>
</tbody>
</table>

V. EXPERIMENTAL EVALUATION AND RESULTS

In this section, we present qualitative and quantitative peduncle detection results. To evaluate the performance of the 3D point segmentation, we use the Area-Under-the-Curve (AUC) measure of the precision-recall curve.

A. Peduncle detection results

Fig. 9 shows the quantitative results of peduncle detection. We note that Fig. 9(a) uses 50% of the available data for training and testing (Table I) whereas Fig. 9(b) has an uneven ratio of 79% for training and 21% for testing (Table II). It can be seen that Field Trip 1 yields a consistent AUC result of 0.73 but that Field Trip 2 varies significantly from 0.54 to 0.67. This is mainly due to the higher number of more challenging green and mixed sweet peppers in the latter. Fig. 9(a) contains 4 green sweet peppers in the training set and 5 green and 2 mixed in the testing set. Fig. 9(b) encapsulates 5 green and 2 mixed for training and 4 greens for testing. Table III summarises different types of sweet peppers (colour coded: mixed=blue) mixed in the training and testing data for various experiments in the two field trips. It is challenging to estimate the optimal ratio of training-testing datasets, especially when the number of samples is relatively low [28]. As a rule-of-thumb in the community, we present results with ratios of 50-50 and 80-20.

We also perform an analysis over three different types of sweet peppers red, green, and mixed colour (red and green). Using the model trained with the 50-50 ratio dataset, sweet peppers in the test dataset are clustered by their colour (red: 29, green: 5, and mixed: 2). There are two mixed sweet peppers which have different amounts of red and green colouration (red=32.5% and 75%, green=67.5% and 25% respectively). We then perform the detection shown in Fig. 10. As expected, while red sweet peppers can be easily distinguished, detection results are poorer for green ones (AUC: 0.52). This supports our previous argument on the performance drop for Field Trip 2. Mixed sweet peppers show the intermediate result of only 2 testing samples.

1) Qualitative evaluation of detection: Fig.11 displays the qualitative results of our sweet pepper peduncle detection method. We assess whether these outcomes can be utilised in the field. The columns indicate different types of sweet peppers and rows are colour and detection views. For red and mixed sweet peppers, the proposed algorithm reports impressive results. Although sufficient performance is achieved to detect the peduncles of green sweet peppers, there are noticeable false positives, $F_X$, that mainly stem from a small margin of sweet peppers and peduncles in colour and PFH features. This paper only presents one view point of 3D model, but the accompanying video shows entire views for further inspection.

B. Processing time and parameter tuning

Our testing methodology highlighted issues in processing time. For example, it takes about 40 seconds to predict the 6,094 input points, which is unlikely to be practical for autonomous sweet pepper picking. Since predicting input features is fully independent, we thus apply parallel processing, which can submit jobs across multiple processes [29] as shown in Fig. 12. Feature vectors are evenly divided between the number of processes and quadrupole speed-ups can be achieved with quad cores.

1) Model (Kernel function) selection and parameter tuning: The aim of this study is to identify the contributions of features (i.e., colours and 3D point clouds). To achieve this, we set a dataset only with red sweet peppers and vary kernels and combinations of features. Note that the dataset used in this model selection study differ to those in Table II.

We use the Machine Learning Python package [27] that supports a variety of kernel function options such as linear, polynomial, radial basis function (RBF), and even custom kernel functions. It is challenging to pick the optimal kernel and associated parameter sets for all possible combinations since training is very time-consuming (e.g., model training takes about 10 hours for linear SVMs with 50K sample points on a modern Intel i7 computer.). We thus evaluate 16 different cases for 3 common kernel functions (SVC, RBF, and linear) and 2 dominant parameters such as Gaussian kernel and regularisation terms $^3$ (named gamma and $C$ respectively). Precision-recall curves are calculated for all cases and AUC is used as their comparison metric, with higher AUC numbers implying better detection performance. The RBF kernels with $\gamma=0.01$ and $C=100$ report the best AUC score of 0.971, where $\gamma$ is the inverse of the radius of a single training point accounting for hyperplane estimation and $C$ defines the impact of mis-classification on the cost.

TABLE III
TYPES OF SWEET PEPPERS FOR EACH FIELD TRIP. (*THE TEXT COLOUR INDICATES THE SWEET PEPPER COLOUR, MIXED=BLUE*).

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50-50</td>
<td>79-21</td>
</tr>
<tr>
<td>Field Trip 1</td>
<td>14, 0, 0</td>
<td>21, 0, 0</td>
</tr>
<tr>
<td>Field Trip 2</td>
<td>18, 4, 0</td>
<td>29, 5, 2</td>
</tr>
</tbody>
</table>

Fig. 8. A panoramic view of a sweet peppers within a protected cropping facility (left) and the crop layout within it (right). It is evident that the brightness of the left and right regions is significantly different.

Fig. 9. Quantitative detection results of precision-recall curves and their AUC when using 50% training and 50% testing dataset (a) and 79%-21% ratio dataset (b).

function. Again, we empirically choose these parameters by performing parameter sweeping (16 evaluations with varying parameters for each kernel (i.e., SVC, RBF, and linear)).

Given this optimised kernel function and parameters, we investigate the impact of each feature on peduncle detection in Fig. 13. We examine the impact of using just colour information (HSV, denoted as green and black in Fig. 13) and colour combined with geometry features (HSV+PFH, in red and blue). Moreover, we study the effects of different kernels with the same features. The plots demonstrate that RBF kernels are superior to their linear counterparts, and colour is the dominant feature for red sweet pepper detection. PFH contributes to improving the detection performance (see Fig. 13 at recall level around 0.1). Further experimental results on different sweet peppers types, such as green and mixed colours, are presented in Section V-A.

VI. DISCUSSION

This paper presents initial, and encouraging, results of peduncle point-level segmentation, which aims at enhancing the performance of a sweet pepper robotic harvester. The integration of the segmentation with the actual harvester and the assessment of its performance is outside the scope of the paper and part of on-going work.
A noticeable attribute of the proposed method is that the trained model yields false positives for other stems and leaves that exhibit similar colour and shape to the peduncle as shown in Fig. 14. To overcome this issue, future work could examine the potential of using context features that account for the geometrical relationship between a sweet pepper and a peduncle (i.e., a sweet pepper usually grows the direction of gravity and locates the top of the fruit). The orientation of each segmented peduncle point and its bounding volume can then be computed. With this information, the supervised learning algorithm may be able to determine that peduncles are usually above the centroid within a range of distances, reducing false positive rates. The accompanying video demonstration displays qualitative results with different viewing points.

**VII. CONCLUSIONS AND FUTURE WORK**

We present a detection system for sweet pepper peduncles in a farm field site. Colour and shape obtained from a RGB-D sensor, are utilised as dominant features for the detection and we discuss the impact of selecting features and their tuning approaches.

The novel use of PFHs provides further discriminative power and demonstrates feasibility in using depth information for agricultural applications. We present promising qualitative results and quantitatively achieve an AUC of 0.71 for point-level segmentation.

Future work will consider ways of utilising GPU parallel processing for detection speed improvement. Other classifiers such as Random Forest or Gaussian Processes will also be compared regarding their prediction speed and detection rate. With more training data, detection performance could be improved by incorporating different weather conditions in future field trips during harvesting season. Also, to avoid intensive labour work associated with manual annotation, we plan to conduct further data collection field trips in near future. We are interested in applying this system to other sweet pepper varieties or different peduncle types.
ACKNOWLEDGEMENT

The authors would like to thank Raymond Russell and Jason Kulk of QUT Strategic Investment in Farm Robotics Program for their key contributions to the design and fabrication of the harvesting platform. We also would like to acknowledge Elio Jovicich and Heidi Wiggenhauser from Queensland Department of Agriculture and Fisheries for their advice and support during the field trips. We thank Marija Popović for comments and proofreading that greatly improved the quality of manuscript.

REFERENCES


