Development of a Machine Vision Yield Monitor for Shallot Onion Harvesters

Amanda A. Boatswain Jacques[1], Viacheslav I. Adamchuk[1], Guillaume Cloutier[2], James J. Clark[3], Connor Miller[1]

[1]Department of Bioresource Engineering, McGill University,
Ste-Anne-de-Bellevue, QC, H9X 3V9, Canada.
[2]Delfland, Inc., Napierville, QC, J0J 1L0, Canada.
[3]Department of Electrical and Computer Engineering, McGill University,
845 Sherbrooke Street West, Qc, H3A 0G4

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract.

Crop yield estimation and mapping are important tools that can help growers efficiently use their available resources and have access to detailed representations of their farm. Technical advancements in computer vision have improved the detection, quality assessment and yield estimation processes for crops, including apples, citrus, mangoes, maize, figs and many other fruits. However, similar methods capable of exporting a detailed yield map for vegetable crops have not yet been fully developed. A machine vision-based yield monitor was designed to perform identification and continuous counting of shallot onions in-situ during the harvesting process. The system is composed of a video and position logger, coupled with acomputer software, and can be used within the tractor itself. A modular camera bracket collected video data of the crops while positioned directly above the harvesting conveyor. Video data was collected in real-time with natural sunlight conditions and in a semi-controlled lighting environment using an artificial light source to enhance vegetable areas. Computational analysis was performed to track detected vegetables on the conveyor. The system is to be tested for a full continuous run during the summer 2018 harvesting season. Based on preliminary results, occasional occlusion of vegetables and inconsistent light conditions are the main limiting factors that may inhibit performance. Although further enhancements are envisioned for the prototype system developed, it has the potential to benefit many producers of small vegetable crops by providing them with useful harvest information in real time and can help to improve harvesting logistics.

Keywords.

Precision agriculture, yield estimation, machine vision, shape detection, shallot onions.

1. Introduction

Crop yield estimation and mapping are important tools that can help growers efficiently use their available resources and have access to detailed representations of their farm. Accurate yield estimation allows growers to efficiently manage their harvest logistics, crop storage and sales, and account for losses in a timelier manner (Nuske et al., 2014; Cheng et al., 2017; Bargoti and Underwood, 2017). Early and accurate predictions are also a key factor for market planning and trade (Bargoti and Underwood, 2017; Cheng et al., 2017). Currently, yield estimation is done by tedious manual sampling methods which are labour intensive, long and costly (Nuske et al., 2014; Dorj et al., 2017). Other methods rely heavily on imprecise historical or empirical data which is then extrapolated (Cheng et al., 2017). These calculations and measurements performed by humans are often prone to bias and sparsity, leading to false predictions (Bargoti and Underwood, 2017). Machine vision is a valuable tool that provides precise measurements by extracting useful information from digital images or videos using computational methods, and it does so in a non-destructive manner (AIA, 2017). It also remains easy to integrate in the various production processes of fruits and vegetables (Abdelhedi et al., 2012; Arakeri et al., 2016).

Fig. 1. Shallot onion harvesting machine and trailer (left) and onion field (right).

This study focuses primarily on the use of machine vision to perform detection, crop-yield estimation and yield mapping for the French grey shallot (*Allium oschaninii*). Shallot onions grow in clusters, where separate bulbs rest on the surface of the soil. They are harvested uniformly using a windrower and trailer (Fig. 1). Spatial variabilities in soil type, soil fertility and other cropping conditions result in great variability in onion size, and onion size is an important limiting factor when determining the percentage of the harvest destined to external suppliers. Quality assessment of shallots is done by human visual inspection and usually only after harvesting is fully completed. Knowing yield and size distribution early can help prevent large losses in money when matching quotas. Although studies have been performed for quality inspection of sweet onions (Shahin et al., 2002; Wang and Li, 2014; Wang and Li, 2015), similar work facilitating the yield estimation of shallots remains scarce. Therefore, an over-the-row machine vision system was created to accelerate yield estimation by running visual inspection on the go during the harvesting process of shallots.

The objective of this research was to develop an algorithm for the detection of shallot onions that will later be integrated in a yield mapping system for industrial harvesting conveyors. This paper is structured as follows: section 2 presents related work in the field of fruit and vegetable detection and mapping in outdoor scenes, section 3 outlines the experimental setup and structure of the algorithm, followed by section 4 presenting the results and deliberating on the relevance of the findings. Section 5 concludes the study.

2. Related Work

The design of diverse computer vision algorithms for fruit detection and counting has been an on-going process. Developing appropriate feature extractors that provide fruit or vegetable identification is a challenge for many researchers aiming to devise robust detection algorithms. The most predominant applications of computer vision in agriculture have been in fruit detection, where the goal is to detect individual fruits, segment them from scenes with branches, foliage, sky, and localize them in a space for yield estimation or as an initial step to the development of robotic harvesting systems (Karpach et al., 2012).

Among the most popular and extensive applications are methods for counting apple fruits using canopy images (Tabb et al., 2006, Linker et al., 2011; Wang et al., 2012; Zhou et al., 2012; Gongal et al., 2015). Stajnko et al. (2004) developed a method for detecting apple fruit using thermal imaging. Images were collected at five time periods to model a temperature gradient between apple fruits and foliage. Correlation coefficients (R^2) between manually detected apples and the estimated number of apples ranged from 0.83 to 0.88. The images were also used to model average apple diameter over time with a maximum deviation from the mean of 6 mm. Wang et al. (2011) created a similar stereo-vision based system using a two-camera stereo rig. This system was stationed on an autonomous orchard vehicle designed to work at night with artificial lighting. It converted apples to the Hue-Saturation-Value (HSV) color space, and then used color segmentation and specular reflection to separate both red and green apples from foliage. The error obtained for crop yield estimation was -3.2 % for red apple trees, and 1.2 % in green apple trees with additional calibration due to significant fruit occlusion. Gongal et al. (2015) later developed an over-the-row machine vision system using both an RGB and stereo camera which captured dual images from both sides of the plant canopy and localized them in space. The experiment was performed in a controlled environment using a covered system with artificial lighting and a tunnel structure. Using image processing and clustering, apples were identified in the images based on shape and color with an accuracy of 78.9%. More state-of-the-art methods (Bargoti and Underwood, 2017) have adapted machine learning techniques such as Multi-Layered Perceptrons (MLPs) and Convolutional Neural Networks (CNNs) to perform pixel level fruit-segmentation under natural sunlight in orchards. The binary images were processed using both a Watershed Segmentation (WS) and Circle Hough Transform (CHT). The WS algorithm was able to detect and count apples with a R^2 value of 0.826 and output an apple yield map for

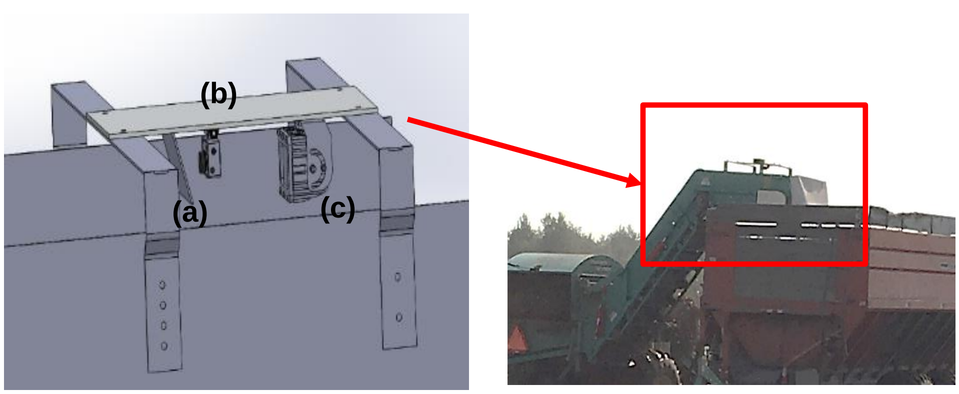


Fig. 2. Model of the machine vision camera bracket (left), (a) is a metal piece used to deflect incoming onions from the camera (b), and (c) is the external light source. Right image is the setup positioned on the farm harvester.

an orchard block using an integrated Global Positioning System (GPS) recording vehicle position with every image taken.

Other studies have focused on the identification of citrus fruit in similar conditions (Sengupta and Lee, 2014; Dorj et al., 2017). Sengupta (2014) and Dorj (2017) developed computer vision algorithms to count citrus fruits on trees using image processing and estimate early overall yield. Sengupta and Lee used shape and texture analysis to detect immature green citrus fruit in a canopy. Texture classification was performed using a Support Vector Machine (SVM), Canny edge detection and a graph-based connect component algorithm and Hough line detection. The algorithm accurately detected 80.4 % of citrus fruit. The study by Dorj et al. was based primarily on the color features of orange fruits. The algorithm consisted of converting the images to the HSV color space, thresholding, orange color detection, removal of noise using a median filter, watershed segmentation and counting. Overall, this algorithm obtained a high correlation (R^2 = 0.93) between the counting algorithm and human observation.

Other studies include work by Kondo et al. (2009), who developed a machine vision system for autonomous harvesting of tomato fruit clusters using stereo images of tomatoes in a greenhouse. The images were converted to the Hue-Saturation-Intensity (HSI) color space and then chromacity distribution plots of H-versus-I were used to cluster fruit region properties and develop a classifier. The research results showed a 73% success rate in locating the stems of clusters. In contrast, experiments performed on vegetable crops in the field are not as common, since most studies have been performed on industrial conveyor systems for sorting (Arakeri and Lakshmana, 2016; Wang and Li, 2015). However, Blok et al. (2016) developed a machine vision algorithm for identifying broccoli heads in the field for a fully autonomous harvester that could perform a selective harvesting process. Texture and color-based segmentation was used to isolate the heads from the background. The automatic segmentation was compared with results from two human experts by comparing the spatial overlap between both results and the individual broccoli head detection. The precision score of the segmentation was 99.5% and overall accuracy of the image segmentation was 92.4%.

Although computer vision systems have proven to have high detection rates and show promising results, the presence of many external factors has often negatively influenced detection. Farm image data is prone to large intra-class variation primarily due to variable illumination conditions, occlusion by other crops or foliage, clustering of crops, camera view-point, and seasonal maturity levels leading crops of varying size, shape or color (Hannan et al., 2009; Sengupta and Lee, 2014; Bargoti and Underwood, 2016; Gongal et al., 2016). Changes in object reflectance can cause object detection to be somewhat unreliable and may lead to incorrect or incomplete segmentation due to a non-uniform distribution of light intensity (Gongal et al., 2016). This problem can be addressed by creating a controlled, uniform lightning environment from which visual data is taken. Examples of controlled lighting environments include an over the row platform with integrated LED lights (Gongal et al., 2016) a wooden box with a painted black interior (Ohali, 2010) or simply performing the experiment at nightfall (Wang et al., 2012, Nuske et al., 2014). Other alternative solutions include using a perimeter-based detection method on top of basic color detection (Hannan et al., 2009; Payne et al., 2013) when variable lighting conditions are unavoidable. Other existing challenges are the multiple detection of the same object within sequential images, or occlusion by other objects or fruits which can lead to miscounting in yield calculation applications. Gongal et al. (2016) used a 2D and 3D imaging approach where apples identified in multiple images were mapped together in a common coordinate system that correctly identified and removed duplicates. The apples in the orchard were represented in a 3-dimensional space where apples registered with the same X, Y and Z coordinates were considered as one fruit. Wang et al. (2012) developed a similar software that calculated the distance between every two apples, and then merged the apples as one whenever this distance was below a given threshold. Hannan et al. (2009) used a centroid-based detection method to identify fruit clusters as a single fruit, and a perimeter-based detection method to locate the individual fruits which had a success rate of 93% and a false detection rate of 4%.

Extensive work has been done to perform detection of fruit in orchard environments such as for apples (Linker at al., 2012; Zhou et al., 2012; Wang et al., 2012; Gongal et al., 2016; Mizushima et al., 2013), oranges (Hannan et al., 2009; Dorj et al., 2017), mangoes (Payne et al., 2013), and berries (Nuske et al., 2014; Pothen and Nuske, 2016). Sorting processes have also been developed for fruits on conveyor systems (Sofu et al., 2016; Ohali, 2010); however, none have attempted to develop a system directly linked to industrial harvesters that can generate a yield map. The initiative to develop better automated crop-estimation systems for vegetables, such as a machine vision-based yield monitor for vegetable crops, is one that has yet to reach its full potential. Incorporating machine vision into everyday agricultural practices allows farmers to significantly reduce their labor requirement and processing time while providing better consistency and uniformity (Sun, 2008). It also allows them to analyze their fields on a higher level of precision.

3. Materials and Methods

3.1 System Design

The system was mounted on a shallot onion harvesting conveyor on the agricultural farm Delfland, Inc. located in Napierville, Québec, Canada. Data acquisition was done in mid and late August (2017) during the end of the harvesting season. A customized bracket (Fig. 2) provides a vertical camera orientation, capturing an image where the camera is facing downwards and directly on the conveyor. The bracket is positioned on the harvester right before the onions are deposited in the trailer to reduce the amount of onions falling backwards and being detected more than once by the algorithm. An on-board positioning system provides the geographic coordinates of the harvester which are used by the software to determine the coordinates of every detected onion in the field. Fig. 3 illustrates the system and its primary components. Shock absorbing pads made of Sorbothane™, a synthetic viscoelastic urethane polymer, are placed beneath all of the top pieces of the bracket to reduce vibration effects of the conveyor and help stabilize the camera.

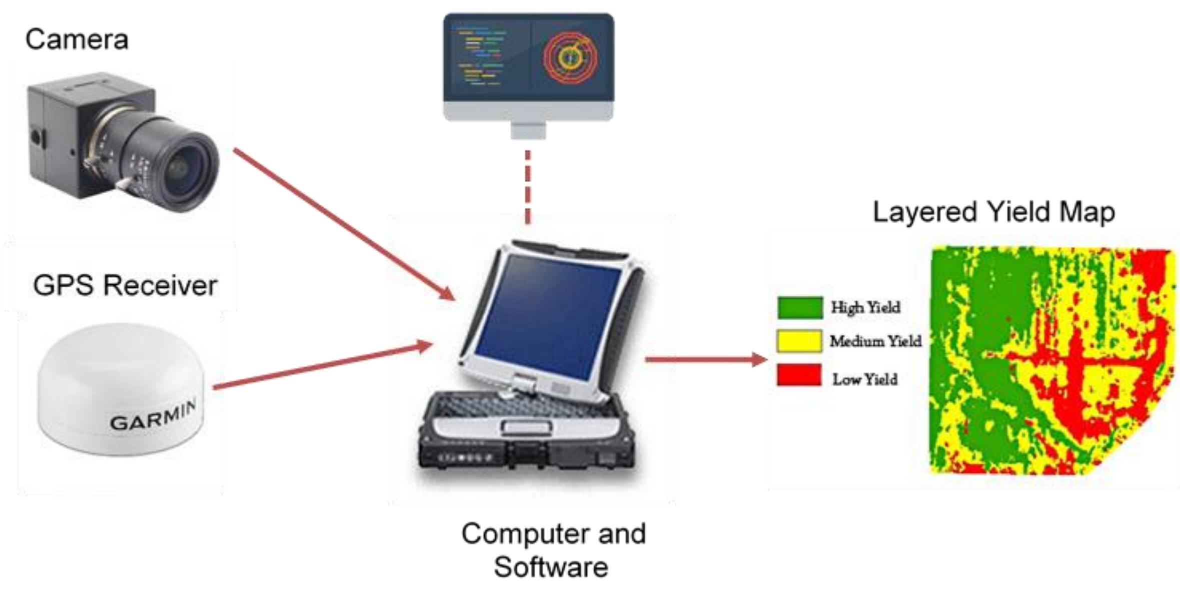


Fig. 3. Diagram of the machine vision yield mapping system. The camera records continuous video of the crops being collected by the harvester. Each frame is coupled with a location tag given by the GPS receiver. The data is analyzed using a computer vision software and exports a yield map showing the size distribution of the crops per unit area.

An accompanying software detects the onions. The software is written in the Python (version 3.5.2) coding language (Python Software Foundation, Welmington, Delaware, USA) and the OpenCV (version 3.2.0) libraries (Itseez, Inc., San Francisco, California, USA). The software runs on a 64-bit PC with an Intel® Core™ i7-7500 CPU processor with a 2.70 GHz clock speed and 8GB of RAM.

3.2 Algorithm

The proposed algorithm is developed using a pipeline similar to that stated in Gongal et al. (2015). The yield-monitoring system performs four main steps: 1) image acquisition, 2) image processing, 3) differentiation between the shallot onions and background using segmentation and 4) noise filtering.

3.2.1 Data Acquisition

Video data is recorded at a resolution of 1920 x 1080 pixels and frame speed of 60 fps, but images are later downsized to a third of their width (640 x 360) to reduce processing time. A Nikon ® KeyMission 170 action camera is used due to its high reliability (weatherproof, waterproof, dustproof and shockproof). The camera has a 1/2.3” Red-Green-Blue (RGB) complementary metal-oxide semiconductor (CMOS) sensor, an aperture of f/2.8 mm and a 15 mm focal length. The camera’s field of view is 33.2 cm by 62.87 cm, capturing the distance between two conveyor paddles. The speed of the onions on the conveyor is set to 0.711 m/s.

3.2.2 Image Preprocessing

The captured images contain a large amount of distortion, where lines which are straight in the real-world deviate from their rectilinear projection in the image (Fig. 4). This distortion must be corrected to extract appropriate quantitative measurements corresponding to real-world dimensions (Balletti et al., 2014). These are corrected using distortion calibration methods in computer vision libraries such as OpenCV. This effect is due to the wide-angle lens of the action camera which is designed to have a large field of view despite its small focal length. The most predominant form of distortion observed was radially symmetric distortion or barrel distortion.

3.2.3 Segmentation

Segmentation is a process where regions of interest are extracted from an image by separating the foreground objects (in this case, shallot onions) from the background (conveyor). Accurate segmentation is crucial for it is a starting point for the succeeding steps such as size classification and counting (Bargoti and Underwood, 2016; Mizushima and Lu, 2013). Challenges including highly variable illumination and shadowing effects can significantly affect the segmentation process and make it ineffective. Performing data collection in a controlled lighting environment (i.e. nightfall) can help achieve better segmentation results (Wang et al., 2012; Nuske et al., 2014;

Fig. 4. The distorted image (left) from the Nikon® camera and its undistorted (right) version. The undistorted image was created by calculating the intrinsic properties of the camera and using remapping functions in OpenCV.

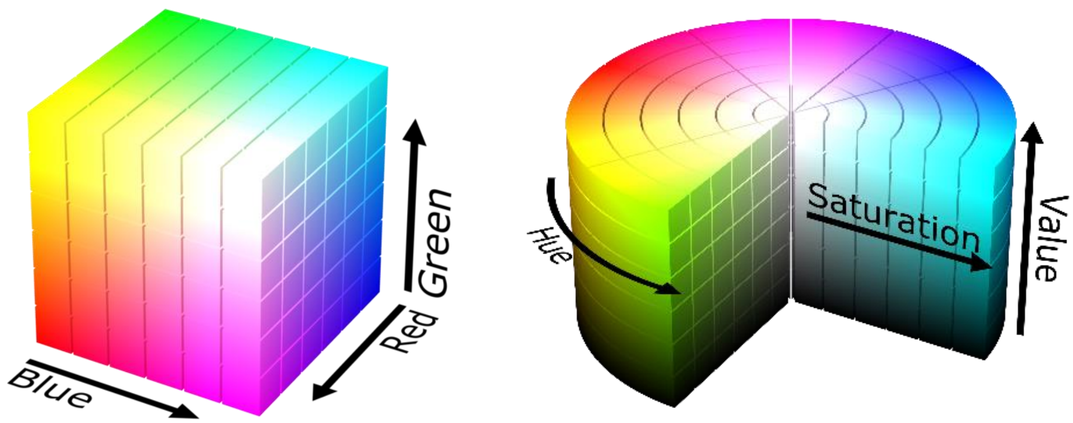


Fig. 5. RGB (left) and HSV (right) color models.

Pothen and Nuske, 2016). However, in practice, onion harvesting usually occurs in natural daylight and incorporating cameras onto tractors will be easier for growers if large experiments are performed during normal operation times (Bargoti and Underwood, 2016).

Digital cameras usually capture images in the RGB format, where each channel corresponds to the intensity of the three primary colors of light (red, green and blue). All colors are then created by the additive reproduction process of various amounts of red, green and blue, with brightness values ranging from 0 to 255. For example, red, green and blue are defined by the vectors (255, 0, 0), (0, 255, 0), and (0, 0, 255), respectively. White can be represented by combining all three components at their highest intensity (255, 255, 255), and black is the absence of all colors in each channel (0, 0, 0). Fig. 5 (left) shows a model of the cartesian RGB color space. The RGB model is not the most intuitive model for discerning color from a perceptual point of view for it is difficult to extract characteristics such as lightness and intensity (Wang et al., 2011; Gongal et al., 2015). Therefore, images are converted to the Hue-Saturation-Value (HSV) color space illustrated in Fig. 5 (right) using the following conversion formulae (Nishad and Manicka, 2013; Kobalicek and Bliznak, 2011).

- The *hue* of a color is the pure color we are examining. All tones and shades of a given color correspond to the same unique hue. Hues are defined using an angle ranging between 0 and 360 along the horizontal cross-section of the cylinder.
- The *saturation* of a color describes how much white is present within the color. A fully saturated color is strong in pigment. For example, tints of red have saturations ranging between 0 and less than 1, while white has a saturation of 0.
- The *value* of a color describes its lightness, or how much black is present within the color. A value of 0 would be black, where lightness increases gradually as value goes towards 1.

To convert a color in the HSV space, we must first determine the maximum (M) and minimum (m) intensities of each pixel, and the difference between them also known as the *chroma* (Δ).

$$\begin{gathered} M = \max(R, G, B) \\ m = min(R, G, B) \\ \Delta = M - m \end{gathered} \tag{1}$$

H is represented by a step-wise function where the chromatic intensity is determined by a two-color difference component. The function relies on the value of M, which gives the angular position of the color on the cylinder.

$$H' = \begin{cases} undefined, \ if \ \Delta = 0 \\ \frac{G-B}{\Delta} mod 6, \ if \ M = R \\ \frac{B-R}{\Delta} + 2, \ if \ M = G \\ \frac{R-G}{\Delta} + 4, \ if \ M = B \end{cases} \tag{2}$$

$$H = 60° \times H'$$

V is defined as the largest component of a color, *M.*

$$V = M \tag{3}$$

Finally, to determine *S*, we divide Δ by *M*.

$$S = \begin{cases} 0, \ \ if \ V = 0 \\ \frac{\Delta}{M}, \ \ otherwise \end{cases} \tag{4}$$

Once the images are converted to HSV, color thresholding is done using two methods. The first is Otsu's thresholding selection method (1979) which has been largely used in computer vision applications in agriculture (Abdelhedi et al., 2012; Mizushima and Lu, 2013; Mollazade et al., 2012; Gongal et al., 2016). Otsu's method automatically determines a threshold using the histogram of a grayscale image. This threshold minimizes the weighted intra-class variance and is defined as a weighted sum of variance of the two classes:

$$\sigma_w^2(t) = \ \omega_0(t)\sigma_0^2(t) + \ w_1(t)\sigma_1^2(t) \tag{5}$$

where ω_0 and ω_1 are the probabilities of the two classes separated by a threshold *t,* and σ_0^2 and σ_1^2 are the variances of these two classes. The algorithm was performed using the hue channel for it assumes the image contains two classes of pixels following the bi-modal histogram. In the second method, thresholding is done by applying a band pass filter to the hue channel, and a high pass filter on the saturation channel. A binary image is then presented where the partial solutions of both channels intersect. The threshold value was determined by analyzing the hue channel histogram and selecting the hue region corresponding to where most onions were located. Both methods are then joined with a segmentation based on texture properties using the magnitude of the red color intensity (Stajnko et al., 2009), and Canny edge detection finds the contour lines of the onions. Finally, shape properties are extracted from the binary image to identify regions corresponding to onions and eliminate those that are not.

3.2.4 Noise Filtering

Morphological operations, including erosion and dilation, were used to close holes within the vegetables and remove noise (Sonka et al., 2015). The kernel was shaped elliptically to preserve the circular shapes of the onions but was not chosen too large to avoid the merging of various regions together. Other preprocessing steps included Gaussian filtering using a 9x9 kernel with a sigma of 0 and standard deviation of 1, and median filtering with a 9x9 kernel.

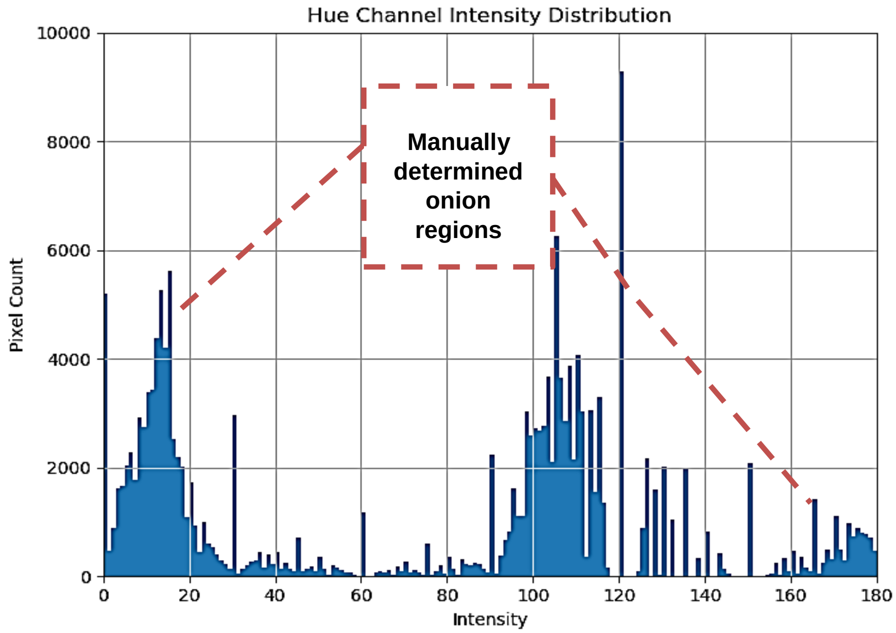
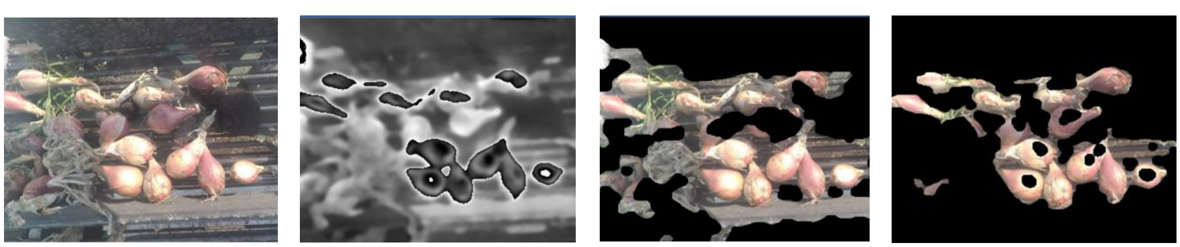


Fig. 6. Hue intensity distribution of a sample image.

4. Results and Discussion

Fig. 6. shows a hue channel histogram of a sample image. Since the range 0 to 360 cannot be represented using only 8-bit integers, in OpenCV hue values range from 0 to 180. The high bin count situated at the value 121 corresponds to a shade of blue violet, representing the surface behind the conveyor belt. The two modes in the histogram are located between 0-30 (left region) and 90-120 (right region). Otsu's method determines an average threshold of 60, situated roughly in the middle of the two peaks, where the left region represents the shallot onions and the right region represents the light-green portions of the paddle and conveyor (see online version for color images). The manually determined threshold combined the left hue peak and the hue values from 165-180 and are shown in Fig. 6.

Fig. 7 shows an original image of the onions on the conveyor (Fig. 7a) and the results obtained from the various segmentation methods tried. A red color intensity texture image (Fig. 7b) is used to further extract regions that are high in red chroma after initial global thresholding. In most cases,



(a) **(b)** **(c)** **(d)**

Fig. 7. Segmentation results. (a) Original conveyor image and results (b, c, d) of the image processing and segmentation algorithm. (b) is the red color intensity image, (c) is the image segmented using Otsu's method (d) is the segmentation performed using the manually selected hue and saturation thresholds.

Fig. 8. Onion detection results. Onions identified by the algorithm are located on the image using ellipses. Colors represent size ranges. In this case, blue corresponds to small and green to medium sized onions.

Otsu's method (Fig. 7c) leads to over segmentation, capturing not only the onion regions but also much of the conveyor system and stems. This may be due to the varying number of onions in each image which affects the histogram distributions, and in some cases, making it unimodal when onion counts were unusually low. Combining the HSV color threshold with the red texture image reduces the number of false positives for the algorithm first checks whether the object is within the proper color range, and then analyses the image further and searches for the distinct red intensity corresponding to onions. This method is often used in apple detection, since apple fruit have a very distinct red color when compared with green foliage in trees (Zhou et al., 2012; Stajnko et al., 2009).

The approach used to validate the system consisted of gathering 34 random screenshots comparing the algorithm's number of detected shallot onions and the true shallot onion count performed by manual observation. Examples of screenshots and results from this process are shown in Fig. 8. The detection rate proved to be relatively low: the mean number of manually counted onions per paddle section was equal to 16.05, ranging from 4 to 37 onions in observed examples with a standard deviation of 6.16, and the mean from the automatically detected onions was 7.17, with ranges between 1 and 15 correctly observed onions and standard deviation of 3.10. Performance of the algorithm is illustrated in Fig. 9. Although the algorithm underestimated the true crop load (regression slope of 0.445), it is important to know that there was a high correlation between the manual count and algorithm count, with an R^2 value of 0.762. By taking

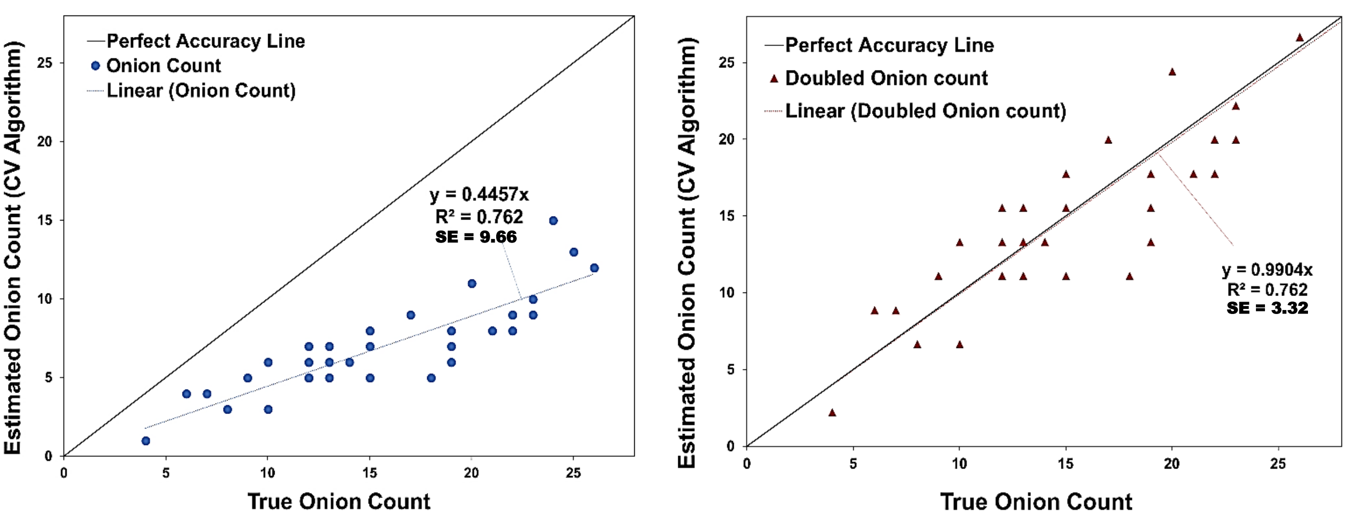


Fig. 9. Onion detection accuracy of the current machine vision algorithm (left) and accuracy obtained by doubling the output (right).

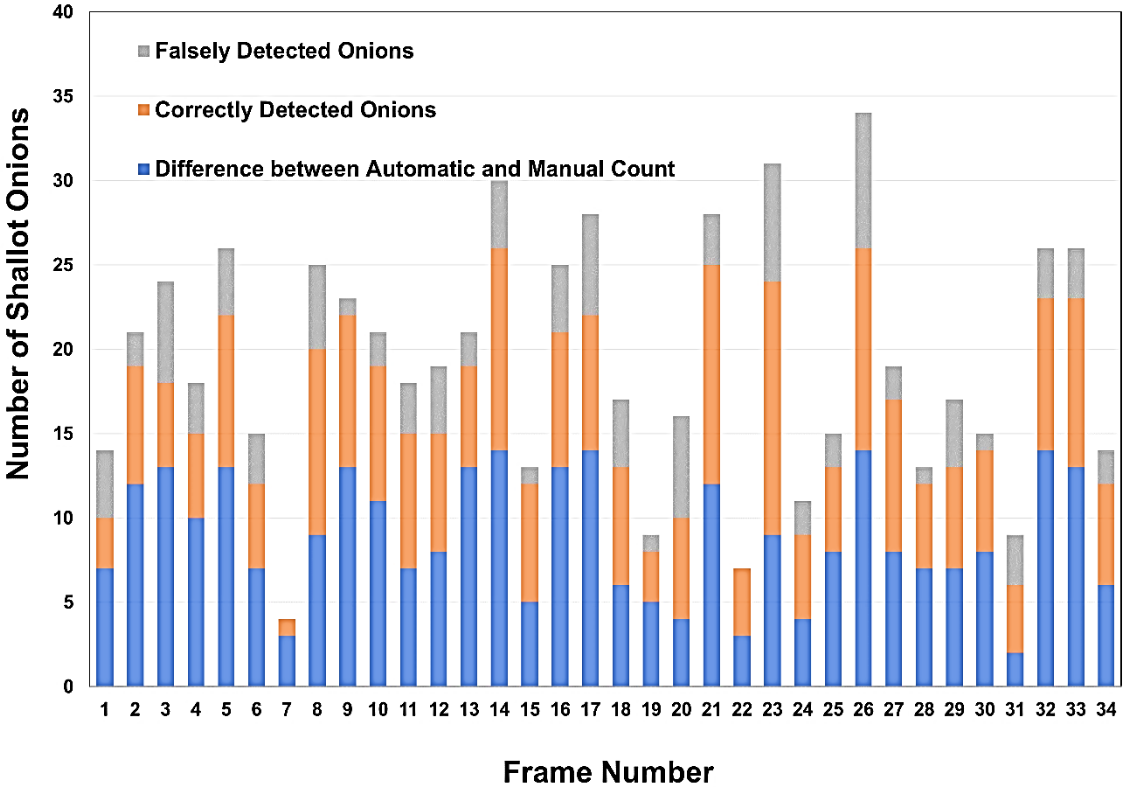


Fig. 10. Shallot onions detected by the computer vision algorithm vs. manual count.

the automatically determined onion count and dividing it by 0.45, this increased the efficiency of predictions dramatically, raising the detection regression slope to 0.99.

Under low onion count, the algorithm shows better results, missing at most 3 onions per frame, but falsely detected onion count remained between 0 and 7 (Fig. 10). High detection rates occur when the onions are not clustered together or superimposed, causing them to be segmented as a joint object and thereafter, making them difficult to isolate. The algorithm also misses onions located in shadowy regions which affects color thresholding, and onion stems that protrude and occlude other vegetables also contribute to inaccuracies in the detection rate. Most cases of falsely detected crops were onions that were detected twice, which was caused by improper boundary definition by edge detection. More precise analysis and stricter limitations on boundary properties such as shape and curvature could enhance object delimitation, as well as more rigorous image preprocessing to normalize the lighting influence. Despite these drawbacks, false detection was on average relatively low (3.11) and average deviation between the detection algorithm and manual count was 8.88.

Further enhancements of the algorithm must be made to better separate individual onion regions and increase overall accuracy. This is rendered difficult due to the high amount of overlap between individual onions and natural lightning or shade which blur the regions between adjacent onions and makes them hard to define. A way to enhance this could be to perform semantic segmentation of the onions by using a CNN structure like that of Bargoti and Underwood (2017). With enough examples of onions in clusters, in shade or occluded by stems, the algorithm might be able to learn features that could accommodate for all variabilities in onion appearance.

4. Conclusions

A machine vision system for quantifying and sorting shallot onion crops during the harvesting process was developed. Based on preliminary evaluation, the total onion count per frame was about 45% lower than the actual onions that should have been detected. With compensation for this difference, standard error for onion count prediction was 3.32. Through interpolation, this is acceptable to generate overall yield production maps. However, correct accuracy is not sufficient to document spatially variable onion size distribution, which is important for improved farm management logistics. The next step in this research is to apply artificial intelligence to a library of preprocessed images exposing the unloading conveyor during various production and ambient conditions. Preliminary analysis of this methodology suggests the potential for a significant improvement in onion detection accuracy but will require substantial implementation efforts.

Acknowledgements

This research was supported in part by Mitacs Accelerate internship program under support and in partnership with Delfland, Inc. (Naperville, Quebec, Canada).

References

Abdelhedi, S., Taouil, K., Hadjkacem, B. (2012). Design of Automatic Vision-based Inspection System for Monitoring in an Olive Oil Bottling Line. *International Journal of Computer Applications,* 51(21).

Arakeri, M. P., Lakshmana. (2016). Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture Industry. Procedia Computer Science, 79, 426-433.

Automotive Industries Association of Canada (AIA). (2017). *Vision Education*. Retrieved from Global Association for Vision Information: http://www.visiononline.org/vision-resources-details.cfm/vision-resources/Computer-Vision-vs-Machine-Vision/content_i. Accessed October 2017.

Balletti, C., Guerra, F., Tsioukas, V., Vernier, P. (2014). Calibration of Action Cameras for Photogrammetric Purposes. *Sensors,* 14, 17471-17490.

Bargoti, S. and Underwood, J. P. (2017) Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards. *Journal of Field Robotics*, doi:10.1002/rob.21699.

Blok, P. M., Barth, R., Berg, W. (2016). Machine vision for a selective broccoli harvesting robot. *IFAC- PapersOnLine*, 49(16), 66-71.

Cheng, H., Damerow, L., Sun, Y., Blanke, M. (2017). Early Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks. *Journal of Imaging,* doi:10.3390/jimaging3010006.

Dorj., U.-O., Lee, M., Yun, S.-S. (2017). An Yield Estimation in Cirtrus Orchards via Fruit Detection and Counting using Image Processing. *Computers and Electronics in Agriculture*, 140, 103-112.

Golzarian, M.-R., Lee, M.-K., Desbiolles, J. M. A. (2012). Evaluation of Color Indices for Improved Segmentation of Plant Images. *Transactions of the American Society of Agricultural and Biological Engineers*, 55(1), 261-273.

Gongal, A., Silwal, A., Amatya, S., Karkee M., Zhang, Q., Lewis, K. (2016). Apple crop-load estimation with over-the-row machine vision system. *Computers and Electronics in Agriculture,* 120, 26–35.

Gongal, A., Amatya, S., Karkee M., Zhang, Q., Lewis, K. (2015). Sensors and systems for fruit detection and localization: A review. Computers and Electronics in Agriculture, 116, 8-19.

Hannan, M. W., Burks, T. F., Bulanon, D. M. (2009). A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection. *Agricultural Engineering International XI: CIGR Journal, 9*, Manuscript 1281.

Kapach, K., Barnea, E., Mairon, R., Edan, Y. and Ben-Shahar, O. (2012) Computer vision for fruit harvesting robots – state of the art and challenges ahead. *Int. J. Computational Vision and Robotic*s, 3, 4–34.

Kobalicek, Petr, Bliznak, Michal. (2011). Optimized RGB to HSV color conversion using SSE technology. *Annals of DAAAM for 2011 & Proceedings of the 22nd International DAAAM Symposium*, 22(1), pp. 1591–1592.

Kondo, N., Yamamoto, K., Shimizu, H., Yata, K., Kurita, M., Shiigi, T., Monta, M., Nishizu, T. (2009). A machine vision system for tomato cluster harvesting robot. *Engineering in Agriculture, Environment and Food*, 2(2), 60-65.

Linker, R., Cohen, O., Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agricutlure*, 81, 45-57.

Mizushima, A. and Renfu, L. (2011). Cost Benefit analysis of in-field presorting for the apple industry. *American Society of Agricultural and Biological Engineers,* 21(1), 33-40.

Mizushima, A., & Renfu, L. (2013). An image segmentation method for apple sorting and grading using support vector machine and Otsu's method. *Computers and Electronics in Agriculture,* 94, 29–37.

Mollazade, K., Omida, M., Arefi, A. (2012). Comparing data mining classifiers for grading raisins based on visual features. *Computers and Electronics in Agriculture*, 84, 124–131.

Nishad, P.M., Manicka Chezian, R. (2013). Various colour spaces and colour space conversion algorithms. J. Glob. Res. *Computer Science*, 4 (1), 44–48.

Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S., Singh, S. (2014). Automated Visual Yield Estimation in Vineyards. *Journal of Field Robotics*, 31(5), 837-860.

Ohali, Y. A. (2010). Computer vision based date fruit grading system: Design and implementation. *Computer and Information Science, 23*, 29–36.

Otsu, N.. (1979). A threshold selection method from gray-level histograms. *Institute of Electrical and Electronics Engineers Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66.

Payne, A. B., Walsh, K. B., Subedi, P. P., & Jarvis, D. (2013). Estimation of mango crop yield using image analysis – Segmentation method. *Computers and Electronics in Agriculture, 91*, 57–64.

Pothen, Z., Nuske, S. (2016). Automated Assessment and Mapping of Grape Quality through Image-based Color Analysis. IFAC- PapersOnLine. Volume 49, Issue 16, Pages 72-78.

Sengupta, S., Lee, W. S. (2014). Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions. *Biosystems Engineering,* 117, 51-61.

Shahin, M.A., Tollner, E.W., Gitaitis, R.D., Sumner, D.R., Maw, B.W. Classification of Sweet Onions Based on Internal Defects Using Image Processing and Neural Network Techniques. *Transactions of the American Society of Agricultural Engineers*, 45(5), 1613-1618.

Stajnko, D., Lakota, M., Hocevar, M. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture*, 42, 31-42.

Stajnko, D., Rakun, J., Blanke, M. (2009). Modelling Apple Fruit Yield Using Image Analysis for Fruit Color, Shape and Texture. *European Journal of Horticultural Science*, 74(6), 260-267.

Sonka, M., Hlavac, V., Boyle, R. (2015). *Image Processing, Analysis and Machine Vision, 4th ed.* USA: Cengage Learning.

Sun, D. (2008). *Computer Vision Technology for Food Quality Evaluation.* Elsevier Inc.: New York.

Wang , Q., Nuske, S., Bergerman, M., Singh, S. (2012). Automated Crop Yield Estimation for Apple Orchards. *Proceedings of the International Symposium on Experimental Robotics*, 1-14.

Wang, W., Li, C. (2014). Size estimation of sweet onions using consumer-grade RGB-depth sensor. *Journal of Food Engineering, 142*, 153–162.

Wang, W., Li, C. (2015). A multimodal machine vision system for quality inspection of onions. *Journal of Food Engineering, 166*, 291–301.

Zhang, B., Huang, W., Lia, J., Zhao, C., Fan, S. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International, 62*, 326-343.

Zhou, R., Damerow, L., Sun, Y., Blanke, M. (2012). Using color features of cv. “Gala” apple fruits in an orchard in image processing to predict yield. *Precision Agriculture*, 13, 568-580.

Nomenclature

CHT	Circular Hough Transform
CNN	Convolutional Neural Network
CV	Computer Vision
GPS	Global Positioning System
HSI	Hue-Saturation-Intensity
HSV	Hue-Saturation-Value
MLP	Multi-Layer-Perceptron
RGB	Red-Green-Blue
WS	Watershed Transform