DEVELOPMENT OF A MUSHROOM HARVESTING
ASSISTANCE SYSTEM USING COMPUTER VISION

A Thesis in
Agricultural and Biological Engineering
by
Cheng-Hao Lee

© 2020 Cheng-Hao Lee

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2020
The thesis of Cheng-Hao Lee was reviewed and approved by the following:

Daeun Choi  
Assistant Professor of Agricultural and Biological Engineering  
Thesis Advisor

John Pecchia  
Associate Research Professor of Plant Pathology and Environmental Microbiology

Long He  
Assistant Professor of Agricultural and Biological Engineering

Paul Heinemann  
Professor of Agricultural and Biological Engineering  
Head of the Department
Abstract

Conventional mushroom harvesting relies on manual labor, which is one of the major reasons for increased production costs. Different growing speed of individual mushrooms requires selective harvesting. To mechanize mushroom harvesting and assist human workers, a computer vision system was developed to detect individual mushrooms from mushroom clusters and evaluate the maturity of the mushroom. The goal of this study was to develop a machine vision system to assist automatic selective harvesting. The specific objectives to achieve the goal were to: (1) design an imaging platform for mushroom production beds, (2) develop algorithms to locate mushrooms, and (3) Define features to identify the maturity of mushrooms and quantify maturity stages. A Kinect v2 camera was used to capture color, near-infrared, depth, and point cloud images. Mushroom detection using a watershed algorithm and a Faster R-CNN object detector was developed to distinguish mushrooms from the substrate as well as other neighboring mushrooms. After the detection, the curvature, size, and shape were measured from collected images using various image processing techniques. The recall of mushroom detection was as high as 95.48%. The accuracy of maturity estimation was as high as 96.63%. This study can be further developed and re-scaled to be used in a commercial mushroom farm.
# Table of Contents

LIST OF FIGURES ........................................................................................................... V

LIST OF TABLES ............................................................................................................... VII

ACKNOWLEDGMENTS ....................................................................................................... VIII

1. INTRODUCTION ............................................................................................................. 1

2. LITERATURE REVIEW ................................................................................................. 4

   2.1 INTRODUCTION ....................................................................................................... 4

   2.2 COMPUTER VISION APPLICATIONS IN AGRICULTURE ..................................... 5

   2.3 GOALS ...................................................................................................................... 9

3. METHODOLOGY .......................................................................................................... 11

   3.1 HARDWARE DEVELOPMENT ................................................................................. 11

   3.2 SAMPLE PREPARATION AND IMAGE ACQUISITION ........................................... 13

      3.2.1 RGB-D image acquisition ............................................................................. 13

      3.2.2 Point clouds image acquisition and registration ........................................... 14

   3.3 MUSHROOM SEGMENTATION IN 2.5-D IMAGES AND POINT CLOUDS ............. 16

   3.4 MORPHOLOGICAL TRAITS OF MUSHROOM ....................................................... 18

   3.5 MATURITY IDENTIFICATION USING SUPPORT VECTOR MACHINE ............. 20

   3.6 MATURITY INDEX (MI) DEVELOPMENT USING THE VARIABLE SELECTION .... 21

4. RESULTS ...................................................................................................................... 22

   4.1 MUSHROOM DETECTION ....................................................................................... 22

   4.2 MORPHOLOGICAL FEATURES ............................................................................. 24

   4.3 MATURITY IDENTIFICATION ............................................................................... 25

   4.4 MATURITY INDEX ................................................................................................. 26

   4.5 RECOMMENDATIONS FOR FURTHER RESEARCH .............................................. 27

5. CONCLUSION .............................................................................................................. 29

REFERENCES .................................................................................................................. 30

APPENDIX ....................................................................................................................... 33
List of Figures

Figure 1. The developmental stages of the mushroom from immature to mature stage. .......................................................................................................................................................................................... 3

Figure 2. Demonstration of multiple layers of mushroom beds in the limited space. 
(pictures credit to Mushroom Farmers of Pennsylvania) ........................................... 5

Figure 3. The layout of the experiment system. (a) camera, (b) stepper motor, (c) temperature probe, (d) mushroom tubs, and (e) control system. ........................................... 11

Figure 4. Circuit diagram of the control system. .............................................................. 12

Figure 5. Example images used in this research. (A) RGB image, (B) NIR image, and (C) depth image. ...................................................................................................................... 14

Figure 6. Demonstration of point cloud image registration using a checkerboard. 
Point cloud images were taken at (a) 45 cm, (b) 30 cm, (c) 15 cm left-hand side away from the (d) middle point, and (e) 15 cm, (f) 30 cm, (g) 45 cm right-hand side away from the middle point. (h) Result of registration. ..................... 15

Figure 7. Flowchart of RGB-D image preprocessing and mushroom detection........... 16

Figure 8. Mushroom detection using the depth image. (A) depth image of mushroom tub (a grayscale indicates depth value of a pixel), (B) contrast-enhanced depth image, and (C) detected mushroom depth image (red circles). ....................... 17

Figure 9. Flowchart of point cloud image preprocessing and mushroom detection .......... 17

Figure 10. Mushroom detection using the point cloud image. (A) origin point cloud image, (B) results of detection overlaid on a point cloud image (top view). ............... 18

Figure 11. Examples of mushroom surface reconstruction. (a) Reconstruction using the completed point cloud (seven point cloud images), and (b) reconstruction using a single point cloud (the resolution is the same as the depth image). .......... 18

Figure 12. Four regions partitioned mushroom and normal vectors visualization. (A) top view, (B) side view, and (C) normal vectors visualization (red arrows). ............... 20

Figure 13. Flowchart of maturity identification and finding maturity index ................. 20

Figure 14. Color map of the identified maturity. The green, navy blue, bright blue, and red color represented the immature stage 1, 2, 3, and mature stage 4, respectively. .................................................................................................................. 21

Figure 15. Result of mushroom detection using the watershed algorithm. (a) Color, (b) near-infrared, and (c) depth image. .................................................................................... 23

Figure 16. Result of mushroom detection using Faster R-CNN in the point cloud top view projection image. The yellow box indicates detected mushrooms by the algorithm. ........................................................................................................ 24

Figure 17. Box plots of quantified features (defined in section 3.4) from depth and point
cloud images. ......................................................................................................................... 25

Figure 19. Results of the feature selection from the depth and point cloud image. ............... 27

Figure 20. The result of using different numbers of point cloud images and their accuracies. ...................................................................................................................................... 28
List of Tables

Table 1. The number of images collected to be used in this study. ........................................14
Table 2. Performance of detection algorithm using watershed segmentation. ...........................22
Acknowledgments

This work was supported by the USDA National Institute of Food and Agriculture and Multistate Research under Project #PEN04653 and Accession #1016510. I would first like to express my profound gratitude to Dr. Daeun Choi for the continuous support of my M.S study and research. Her willingness to give her time on my research and writing has been very much appreciated. In addition, I would like to express my appreciation to Dr. John Pecchia for providing mushrooms to my research. I could not imagine conducting this research without his help. Likewise, Dr. Paul Heinemann’s advice and insightful critiques are appreciated. I would also like to thank Dr. Long He for his valuable technical support and equipment. Thank you to the Penn State Mushroom Research Competitive Grant. Also, I would like to thank Penn State College of Agricultural Sciences as well as Giorgi Mushroom Company for their financial support. Finally, I wish to thank my parents and friends for their support and encouragement throughout my study.
1. Introduction

Mushrooms (*Agaricus bisporus*) have become more popular in the United States due to their nutritional value (Tsai et al., 2007). Mushrooms also play an important role in the economy, especially in Pennsylvania. Nearly two-thirds of mushrooms consumed in the United States are produced in Pennsylvania (Gorgo-Gourovitch, 2018; USDA, 2019). The mushroom industry in the United States has faced a critical labor shortage in recent years (USDA, 2019). Many tasks of mushroom production still rely on human labor. One hundred percent (891 million pounds) of mushrooms were manually harvested in 2017-2018 in the United States (USDA, 2019). However, not only is the harvesting labor-intensive, but the mushrooms are also selectively harvested due to the different growing speed of individual mushrooms. The selective harvesting of mushrooms requires a considerable amount of training of farm workers in order to distinguish which mushrooms have reached the appropriate maturity for harvesting.

Several mechanical harvesters have been developed and used by the mushroom industry in Europe. Typically, the mechanical harvester collects crops in a non-selective manner by cutting the stems in the growing bed regardless of the maturity of the mushrooms (Vandentop, 2019). Without selective harvesting, the quality and subsequent economic value of the crops significantly decreases and is used primarily for processed products. When harvesting overly mature mushrooms, mushrooms start to lose moisture which is not suitable for fresh market. On the other hand, harvesting mushrooms too early will prohibit mushrooms from reaching maximum weight causing economic loss (Schiau, 2013). Roy et. al. used the extent of cap opening to identify the maturity of mushrooms (Roy et. al., 1995). The veil developed from the tight, stretched and partially broken stage to the completely broken stage. However, this classification was performed...
manually. To achieve automatic selective harvesting, a machine vision system is needed to observe the cap opening. Another challenge of developing mechanical harvesters is that the harvester can damage mushrooms easily due to the fragility of mushroom’s basidiocarp or fruit body. To develop a mechanical harvester for selective harvesting, there are several technical challenges. First, the mechanical harvester should be able to locate every single mushroom and identify the maturity of the located fruit body. Second, the mechanical harvester should be able to pick mushrooms as accurate as human workers. Excessive bruising on mushrooms reduces marketability of value of the crop. Therefore, the overall goal of this study was focused on solving the first challenge: developing a mechanism for mushroom detection and maturity evaluation.

Computer vision is the technology to mimic the human visual system by acquiring, processing, and analyzing digital images (Gomes and Leta., 2012). Not only can the computer vision observe much more subtle details than humans, but the efficiency of tasks can be increased (Broin nan and Sun, 2004). Computer vision has been used in various applications for specialty crops such as mushroom, strawberry, and tomato (Concha-Meyer et. al, 2018). Machine learning is used in computer vision for object detection and classification. Recently, deep learning, one of the machine learning techniques, became popular because of its outstanding ability for solving complicated problems with large amounts of data efficiently (LeCun et. al, 2015). Images taken in an agricultural field usually involve complex backgrounds and objects with large variability in visual characteristics. For example, in a mushroom bed there could be various stages of mushrooms such as immature, mature, defective, and crowded ones on the bed. In many studies, computer vision techniques has demonstrated the ability to predict yield, harvest the fruit, detect diseases and grade agricultural products.
However, not many studies have addressed automated solutions for the mushroom industry.

Due to the labor shortage, the US mushroom industry would like to transition to automated production. In this study, the harvesting assistance system was developed to facilitate mechanization and automation. To be more specific, the objectives of this system were: (1) collect RGB, NIR, depth, and point clouds images of mushroom on a mushroom production bed using a practical image acquisition platform, (2) develop a machine vision algorithm for detecting mushrooms in a production tub and (3) estimate the maturity stage of mushrooms according to the shape, cap opening, and size (Figure 1).

![Figure 1. The developmental stages of mushrooms from immature to mature stages.](image)
2. Literature Review

2.1 Introduction

Pennsylvania is the largest producer of Agaricus mushrooms in the United States. In 2018, a total of 572 million pounds of mushrooms, which counts for almost two-thirds of production in the U.S., were produced in Pennsylvania (USDA, 2019). According to a statistic report from National Agricultural Statistics Service (NASS) in season 2017-18, the number of mushroom growers in the U.S. decreased 18 percent over the past decade (USDA, 2019), which indicates the problems of labor shortage.

In the Netherlands, a commercial company (Van den Top, Barneveld, the Netherlands) developed a single layer mushroom picking machine, called Zig-Zag. Although the company claimed the speed of harvesting using the machine is four times faster than human workers, the single-layer design of the system requires considerable space and money. In Pennsylvania, growers use multiple layers of mushroom beds, trays, or bulk in the limited space (Figure 2). Also, the mechanized mushroom harvesters in European countries use a non-selective harvesting strategy. In the fresh market, non-selective harvesting is ineffective since the growing speed of every single mushroom is different. For example, while some mushrooms matured, it is possible that other mushrooms just appear. A non-selective harvesting system does not consider the variabilities in maturity stages and swipes the whole layer to harvest the entire crops. Consequently, immature mushrooms are also harvested at the same time. On the other hand, the overly matured mushrooms start to lose moisture, which is not preferable in the fresh markets. In manual harvesting, human workers check the beds, trays, or bulk every day to pick the mature ones, which is labor-intensive and requires considerable training.
Figure 2. Demonstration of multiple layers of mushroom beds in the limited space. (picture credit to Mushroom Farmers of Pennsylvania)

2.2 Computer vision applications in agriculture

Computer vision plays an essential role in agricultural industries (Brosnan and Sun, 2004). With recent development in computer vision, not only can the cameras capture much more subtle details than human eyes, but the sensors can also observe features beyond the visible range spectrum. For example, moisture, chemical composition, and sugar content can be detected by the hyperspectral camera and used for monitoring fruits in orchards, plant disease, and sugar content in fruits (ElMasry, 2007; Wu et. al., 2012). Embedded sensors with internet connectivity such as IoT became an feasible option to be applied in the field. Embedded sensors can be mounted on various hardware platforms such as an unmanned aerial vehicle (UAV). The signal from these sensors can be transmitted by the internet and be analyzed by computers. With different types of cameras and analysis, computer vision has unlimited potential to help farmers to achieve automated solutions.
In general, mushrooms can be harvested 15–17 days after casing. Peat moss is applied to spawn compost. Then, mushrooms are harvested for 3 to 5 days per flush depending on the maturity. Several previous studies (Concha-Meyer et. al., 2018; Wang et. al., 2018; Tillett and Batchelor, 1991; Williams, 1999) attempted to defined and quantify the maturity of mushrooms based on visual characteristics. Concha-Meyer et al. (2018) created a volume estimation system for mushrooms using machine vision. The optical imaging system measured the volume of the objects and the results of the estimation were compared to weight and density. Wang (2018) applied watershed methods, and morphological algorithms to acquire the pileus diameter of fresh white button mushrooms. The results showed that the average maximum grading speed, accuracy, and damage rate of the grading system were 102.41 mushrooms/minute, 97.42% and 0.05%, respectively. Tillett et al. (1991) used colors of mushrooms to develop a vision processing algorithm for identifying the size and location of mushrooms in the growing bed. In the algorithm, two threshold values were set for different purposes. $T_{\text{Low}}$ was used for segmenting the mushrooms and the background. $T_{\text{High}}$ was used for picking out the brighter parts corresponding to the nearly horizontal parts of the mushroom surface. Hue, saturation, and intensity (HSI) color space is used to quantify the color of the mushroom cap. Heinemann et al. (1994) graded mushrooms using the size, stem length, veil, shape, color, and stem trim of mushrooms. Collected mushrooms were partitioned into two groups according to the quality. Then, several thresholds were determined by the linear regression method. The average accuracy is 80%.

One of the applications of machine learning is to solve the problems of classification. For example, plant species can be identified by classifying the leaf images (Lee et. al.,
2019). Object detection can be implemented by classifying an object from other background images; and face recognition can be achieved by identifying a specific face from the dataset. Likewise, supervised learning can solve various problems based on how the dataset and labeling are created (Jin and Ghahramani, 2003). For example, evaluating the maturity of apples in images can be achieved by creating a dataset containing thousands of labeled apple images with different maturity stages. The color and other visual characteristics of apples in the images can be used as features, then, the maturity recognition model can be developed by analyzing the relationship between color and maturity.

Deep learning is a branch of machine learning techniques (LeCun et. al., 2015). Conventional machine learning methods need to quantify features (Bishop, 2006), then the objects are classified based on the quantified features, for example, the color of the apple. However, deep learning methods can classify objects without defining and quantifying traits due to built-in feature extraction capabilities. The model can find the traits by itself while training. The basic deep learning method is Convolutional Neural Network (CNN), which imitates the structure of the neuron connection (Szegedy et. al., 2013). In computer vision using CNNs, the input (stimulation) is an image. The input images are then convolved by thousands of kernel functions (activation of the neuron). The number of outputs is determined by the number of classes. When training CNN models, the dataset would be split into several batches (subset of training data) so that the parameters of kernel functions can be adjusted according to the results of every batch. This is called back propagation, which provides deep learning methods high flexibility to adapt to the diverse applications. When CNN models finish a batch training, another batch is back-to-back trained by the same CNN models. This is called
iterations, which makes deep learning methods have robust generalization to solve complicated problems.

Previous studies have addressed the specialty crop localization in the field using deep CNNs. Bargoti and Underwood (2017) applied Faster R-CNN to detect specialty crops in orchards. Apple, mango, and almond images were captured using the high-resolution RGB camera with small exposure time (~ 70 μs). The ground truth annotation for fruits was labeled using a rectangular bounding box. Due to the difficulty of labeling a large number of small objects, they only randomly sampled smaller sub-image patches from the pool of larger images. ZF and VGG16 networks were then applied to implement transfer learning (Ren et. al., 2015). After 5000 iterations for apples and almonds and 40,000 iterations for mangos, the model converged in detection performance. The F1 scores of apple, mango, and almond detection are 0.904, 0.908, and 0.775, respectively. The F1 score is the harmonic mean of recall and precision, which is an indicator to consider type I (false positive) and type II (false negative) errors at the same time. The results demonstrated the potential for applying deep learning to the mushroom industry.

The point cloud is a set of data points which stores 3-D coordinates and other information and is used to construct a 3-D image. Unlike RGB-D images storing data in pixels, point cloud data can represent the 3-D coordinate system. Previous studies have described that deep learning can implement to the data structure such as point cloud to segment and recognize objects (Qi et al., 2017). In the past, the point cloud usually needed to be downsampled and use grids to represent points. Then, the grids could be used as input. This downsampling process caused two problems. First, the point cloud is usually sparse spatial data. After downsampling, most values in the grid
were zero and it was not helpful for learning features. Secondly, using grid values as inputs is a sort of human-defined feature, which limited the deep learning model capacity. Hence, Qi et al. (2017) proposed the deep learning model structure (PointNet), which can use a raw point cloud as input. The model performed better compared to conventional 3D CNN methods. For example, the accuracy of identifying a mug from a chair could achieve 93% and 89.6%, respectively. Some researchers proposed object detection in the point cloud using the multiview method (Chen et al., 2017). The point cloud was projected to the front view and bird view (top view). The detected regions of interest (ROI) were obtained via ROI pooling in the proposed fusion network. These proposed models can be implemented in agricultural applications, especially in the mushroom model construction to increase data point resolution and accuracy of detection algorithms.

2.3 Goals

In manual mushroom harvesting, farmers need a significant training period prior to picking. Mushroom identification systems using machine vision in the previous studies did not use comprehensive and robust indicators to identify the maturity of mushrooms. In most studies, the size of mushrooms was only used as the visual characteristic. The shape of the cap should be considered at the same time because the small mushroom could also be mature. The proposed system and methods in this research intend to fill the gap in knowledge of definitions of maturity and quantifying visual characteristics. The goal of this research was to develop a machine vision system to assist automatic selective harvesting. Three specific objectives were to:

1. Design an imaging system platform for mushroom production beds,
2. Develop algorithms to locate mushrooms, and

3. Define features to identify the maturity of mushrooms and quantify maturity stages.
3. Methodology

3.1 Hardware development

The hardware consisted of the imaging system, rail system, and control system. Figure 3 shows an entire setup of hardware in the developed system.

![Image: The layout of the experiment system. (a) camera, (b) stepper motor, (c) temperature probe, (d) mushroom tubs, and (e) control system.]

The images were taken by the imaging system. To capture images from three tubs on the shelf, the camera (Figure 3a) was carried by the motor (Figure 3b) on the rail system. The speed and direction of the motor were controlled by the control system (Figure 3e). A Windows laptop (Processor, Intel i7-8700; RAM, 16G; Windows 10 version; GPU, GTX 1080 Ti) was used to drive the camera and the rail system. The rail system was operated using C++ language and the imaging system was controlled using MATLAB.
An RGB-D camera (Kinect for Xbox, Redmond, WA) was used to capture near-infrared, depth and color images. The camera was transported to multiple image acquisition points by the motor in the rail system so that the field of view was increased. Not only can the top view of mushroom be captured but the side view also can be acquired. The motor (NEMA 17 Stepper Motor, Openbuilds, Monroeville, NJ) provided 5.5 kg-cm torque to move a 1.4 kg Kinect camera. The motor required 24 volts and 1.68 amp as power input so the rail system was equipped with a 24V power supply (Meanwell power supply, Monroeville, NJ). The power supply not only provided the power but also changed alternating current (AC) to direct current (DC), which can be also used for motor control. An H bridge motor driver was equipped so that the signal (5V) from Arduino can be transformed to the duty cycle (24V) to drive the motor. There were two digital output pins in Arduino controlling the speed and the direction of the motor. Finally, the Arduino board was driven by MATLAB (Natick, MA) and all components were integrated together as imaging, rail, and control system (Figure 4).

![Circuit diagram of the control system.](image)
3.2 Sample preparation and image acquisition

The white button mushrooms (Agariaus Bisporus) were spawned (“seeded”) in a tub at Penn State Mushroom Research Center (MRC, University Park, PA) and were cased 16 days after spawning (application of pH-buffered peat moss). Crops were placed in environmentally controlled growing rooms and pictures were taken on March 26th, April 19th, May 12th, June 24th, September 6th, and October 3rd, 2019. Time-lapse images were captured every ten minutes until harvesting of the third break. Four thousand and seven hundred mushrooms of different growing stages were collected by time-lapse images. The size of a mushroom tub is 24.76 cm height, 48.26 cm width, and 62.23 cm length. Each shelf is 60.96 cm deep by 152.4 cm wide (3 shelves per rack, Figure 3). The mushrooms in the collected images were partitioned into four maturity groups according to how close to the harvesting time, starting from four days prior to harvesting. The duration between each maturity group was 24 hours (Figure 1). In this research, both RGB-D and point cloud images were captured and analyzed.

3.2.1 RGB-D image acquisition

The Kinect camera was mounted 0.6 m above three mushroom tubs. The location of the camera was fixed and aligned with the middle tub when capturing RGB-D images. The camera automatically took the time-lapse images by using MATLAB code. A set of near-infrared (NIR), depth, and color images were acquired (Figure 5). A total of 14,739 time-lapse image sets (NIR, depth, and color) were acquired, and 4991 point cloud images were captured (Table 1). The resolution of NIR, depth, and color images were 512x424, 512x424, and 1980x1080 pixels, respectively. The raw image sets were used for mushroom detection and performing maturity identification for each detected mushroom.
### Table 1. The number of images collected to be used in this study.

<table>
<thead>
<tr>
<th>Date</th>
<th>RGB-D images (color, NIR, and depth)</th>
<th>Point cloud images</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 26th</td>
<td>2731</td>
<td>N/A</td>
</tr>
<tr>
<td>April 19th</td>
<td>2875</td>
<td>N/A</td>
</tr>
<tr>
<td>May 12th</td>
<td>2358</td>
<td>N/A</td>
</tr>
<tr>
<td>June 24th</td>
<td>2120</td>
<td>N/A</td>
</tr>
<tr>
<td>September 6th</td>
<td>2527</td>
<td>2863</td>
</tr>
<tr>
<td>October 3rd</td>
<td>2128</td>
<td>2128</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14739</strong></td>
<td><strong>4991</strong></td>
</tr>
</tbody>
</table>

![Example images used in this research. (A) RGB image, (B) NIR image, and (C) depth image.](image)

**Figure 5.** Example images used in this research. (A) RGB image, (B) NIR image, and (C) depth image.

### 3.2.2 Point clouds image acquisition and registration

The Kinect camera provided the color and depth images. Time-of-Flight (ToF) method was performed when capturing depth images. The camera decides the depth, in pixels, by measuring the time emitted light takes to travel from the camera to the object and back. After acquiring the depth and color images from the Kinect, the point cloud image was generated by matching two images pixel by pixel. However, the resolution and the location of the depth sensor are different from the color sensor. To properly register the images, the extrinsic and intrinsic properties for both sensors were used for calibrating the point cloud image. The intrinsic property included the focal length and the principal point. The extrinsic parameters included the rotation and translation matrix, which
performed the coordinate system transformation from world coordinates to camera coordinates.

Seven point cloud images from different views of angle were registered together. A checkerboard was used for finding the camera extrinsic property (Figure 6). The rotation and translation matrices were calculated by matching the feature points on the checkerboard in different images. Three different algorithms were compared when registering images, which were the iterative closet point (ICP), coherent point drift (CPD), and normal distributions transform (NDT) algorithm. The ICP algorithm provided the best result with a 0.0682 RMSE (root mean squared error) value.

Figure 6. Demonstration of point cloud image registration using a checkerboard. Point cloud images were taken at (a) 45 cm, (b) 30 cm, (c) 15 cm left-hand side away from the (d) center, and (e) 15 cm, (f) 30 cm, (g) 45 cm right-hand side away from the center, and (h) result of registration.
3.3 Mushroom segmentation in 2.5-D images and point clouds

A flooding-based watershed algorithm was applied to partition an acquired image into an individual region-of-interest (ROI) mushroom image. Also, a series of image pre-processing algorithms using depth filter and connected component analysis was applied to remove background and noise pixels in images. Figures 7 and 8 illustrate the flowchart and demonstration of the image preprocessing and the mushroom detection algorithm. First, the depth filter was applied to raw depth images to identify the location of the tub and remove pixels located outside of the tub. The upper and lower threshold value was 700 mm and 500 mm, respectively. Then, the range of depths was enhanced so that the intensity of the pixels of the mushrooms and compost could be differentiated. The flooding-based watershed algorithm was then applied to separate individual mushrooms from clusters. The size and shape of detected objects were measured using connected components analysis. Detected objects with less than 30 pixels were regarded as noise (i.e., sparkles, composts) and removed. Also, the detected objects with an aspect ratio larger than 3 or smaller than 0.3 were regarded as cluster mushrooms. The morphological erode operation were then applied to cluster mushrooms and followed by an additional flooding-based watershed algorithm. This iteration was run until the aspect ratio of detected objects fell between 0.3 and 3. This range was determined by trial and error.

Figure 7.  Flowchart of RGB-D image preprocessing and mushroom detection
Figure 8. Mushroom detection using the depth image. (A) depth image of mushroom tub (a grayscale indicates depth value of a pixel), (B) contrast-enhanced depth image, and (C) detected mushroom depth image (red circles).

Mushrooms in the point cloud image were identified using the Faster R-CNN deep learning method. Figure 9 and 10 illustrate the flowchart and demonstration of mushroom detection in point cloud image. First, the top view of the single point cloud image was projected to the flat plane, which generated a 2D image. The 2D image was used as the input of the Faster R-CNN model. The outputs of the Faster R-CNN model were several bounding boxes and every bounding box contained a detected mushroom. The model was pre-trained by the VGG-16 model and the weights of neurons were transferred. The VGG-16 was developed for the ImageNet Challenge 2014 as the classifier for ImageNet’s 1000 classes. As a result, the network learned the rich feature representation. When retraining the model, the stochastic gradient descent with momentum (SGDM) algorithm was used as the optimizer method. The learning rate and max epoch were set as 0.0001 and 50, respectively. The results of detection in seven point clouds were registered as a completed point cloud.

Figure 9. Flowchart of point cloud image preprocessing and mushroom detection
Figure 10. Mushroom detection using the point cloud image. (A) origin point cloud image, (B) results of detection overlaid on a point cloud image (top view).

The registered point-cloud image was used for reconstructing the cap surface. The cubic spline interpolation (CSI) method was applied to the registered point cloud image. The CSI used a series of cubic functions to fit the points to reconstruct the cap surface (Figure 11). The shape was then analyzed using the fitted cubic function.

Figure 11. Examples of mushroom surface reconstruction. (a) Reconstruction using the completed point cloud (seven point cloud images), and (b) reconstruction using a single point cloud (the resolution is the same as the depth image.).

3.4 Morphological traits of mushroom

Eight morphological traits of detected mushrooms were quantified using the mushroom point cloud images (Figure 11a): the mushroom average radius (mm), cap height (mm), cap area (mm²), perimeter (mm), circularity, eccentricity, solidity, and extent. The mushroom average radius \( r_{\text{avg}} \) was calculated by using the area of the ROI (mushroom) and was determined as follows:
\[ r_{\text{avg}} = 1.75 \cdot \sqrt{\frac{A}{\pi}} \]  

(1)

where \( A \) is the area of the region of interest of the mushroom and the constant (1.75) is the coefficient for unit conversion from pixel to the millimeter. Cap height was defined as the maximum depth value in the region of interest. The perimeter was computed by calculating the distance between each adjoining pair of pixels around the border of the region. The roundness of the cap was described using the circularity (Eq. 2).

\[ \text{Circularity} = \frac{4 \cdot A \cdot \pi}{\text{Perimeter}^2} \]  

(2)

The shape of the cap also quantified using the eccentricity, solidity, and extent, which were computed using Eq. 3, 4, and 5, respectively.

\[ \text{Eccentricity} = \frac{\text{Distance between the foci}}{\text{Major axis length}} \]  

(3)

\[ \text{Solidity} = \frac{A}{\text{Convex hull area}} \]  

(4)

\[ \text{Extent} = \frac{A}{\text{Bounding box area}} \]  

(5)

Curvature was used based on the concept of the openness of the mushroom cap. The openness of the cap indicated the maturity since caps start to open and veil stretch to release spores when mushroom matures. The curvature can be quantified using normal vectors of the mushroom cap surface in the depth images. Normal vectors consist of \( i, j, \) and \( k \) vector components (unit vectors in \( x, y, \) and \( z \)-axis directions, respectively). Only \( k \) component values (magnitude of the normal vector in the \( z \)-direction) were used in this method. \( i \) and \( j \) vector components were ignored because they indicate the angularity. The angularity is associated with how mushrooms grow (clusters or sparse), not based on the maturity. From the acquired data, the author found that a sphere-shaped object had the value of normal vectors in the \( z \)-direction at the edge pixels of cap approaching 0. On the other hand, the value of \( z \) unit vectors approaching 1 indicated
that the surface of the cap was flatter. To calculate z unit vectors, the centroid of the cap was defined by the center of mass and the cap of mushroom was partitioned into four regions according to the distance from the centroid (Figure 12). The mean of z unit vectors in every region was calculated as curve traits.

Figure 12. Four regions partitioned mushroom and normal vectors visualization. (A) top view, (B) side view, and (C) normal vectors visualization (red arrows).

3.5 Maturity identification using support vector machine

Figure 13 demonstrated the flowchart of model development. Soft-margin SVM classifiers with radial basis function kernels were developed to identify the maturity stage from the previously defined traits in section 3.4. The margin and kernel parameters of the classifiers were determined using grid search and 10-fold cross-validation. The performance of the classifiers was also evaluated using 10-fold cross-validation. The results were also visualized as a color map (Figure 14)

Figure 13. Flowchart of maturity identification and finding maturity index
Figure 14. Color map of the identified maturity. The green, navy blue, bright blue, and red color represented the immature stage 1, 2, 3, and mature stage 4, respectively.

3.6 Maturity Index (MI) development using the variable selection

Neighborhood component analysis (NCA) (Yang et. al., 2012) was used to select important features. NCA selects features by learning feature weights for minimization of the objective loss function. The objective function was used for classifying groups and was determined by the previously developed SVM model. The learning rate was set using stochastic gradient descent (SGD). The result of feature weights was verified using leave-one-out cross-validation.
4. Results

4.1 Mushroom detection

The results of the mushroom detections using the watershed algorithm (section 3.3) are shown in Figure 15 and Table 2. The color image provided the best, which was 95.28 percent for the F1 score (harmonic mean of FP and FN). The main reason was due to the relatively higher resolution of the color images compared to the near-infrared and depth images. The edge of objects in images plays a vital role in object detection. With the same size of mushrooms, the higher resolution image used more pixels to depict the mushroom, which made the edge of mushrooms clearer. On the other hand, the near-infrared image had the lowest precision. The reason was the “salt and pepper” noise problem. The near-infrared image had lower resolution and more noise, which made noise removal more difficult. Consequently, more false-positive detection was produced. On the other hand, the depth image had the lowest recall. The reason is the limited height variance, which made the edge of compost and mushroom blurred, and some mushrooms were not detected. Also, some crowded mushrooms were detected as a single mushroom with a larger diameter. As a result, the value of recall was the lowest for the depth image. For all types of images, more missed mushrooms were found in the early growing stages due to not enough variance in sizes and heights.

<table>
<thead>
<tr>
<th></th>
<th>Color</th>
<th>NIR</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>93.81%</td>
<td>73.68%</td>
<td>87.95%</td>
</tr>
<tr>
<td>Recall</td>
<td>96.81%</td>
<td>93.33%</td>
<td>82.95%</td>
</tr>
<tr>
<td>F1 score</td>
<td>95.28%</td>
<td>82.35%</td>
<td>85.37%</td>
</tr>
</tbody>
</table>
Figure 15. Result of mushroom detection using the watershed algorithm. (a) Color, (b) near-infrared, and (c) depth image.

To improve the performance of mushroom detection, 3D point cloud images were used. The result of mushroom detection in the point cloud using Faster R-CNN is shown in Figure 16. The precision was 98.8%. The false-positive detections occurred when the mushrooms were crowded in the corners of the tub. The recall was from 88.13% to 90.22% depending on the location of the camera. If the mushroom tub was at the edge of the image, the resolution would be lower than in the middle of the image. However, because the seven cloud images were a set for the three tubs, the missed detection in a point cloud image could be made up by other point cloud images. Once the missed mushroom was detected in any of the other point cloud images, the mushroom would be regarded as detected in the final registered point cloud. Hence, the overall recall for a set of seven point cloud images was 95.48%.
4.2 Morphological features

The quantified features defined in section 3.4 were visualized using box plots (Figure 17). Features were quantified from the depth and point cloud images. Larger variances and more outliers were observed in the features from the depth image. The reconstructed surface had noise from the uncertainty of the depth measurement and the interpolation. The noise caused the poor performance of the curvature, shape, and size measurements. Quantified features from the point cloud images are visualized in Figure 18. The obvious changes of values were observed, especially in the curvature, radius, and area feature. With the higher resolution, the reconstructed surface was more continuous than the surface from the depth image. The avg1, radius, and area had relatively smaller variances and the means were distinguishable. These features had the potential to be maturity indicators. The maturity index was determined in the feature selection section.
4.3 Maturity identification

Figure 18 shows the confusion matrix of the maturity classification of mushrooms. The x and y-axis are the actual and predicted labels, respectively. The actual labels were determined by the definition in Figure 1. The grayscale in the matrix indicates accuracy. The left matrix was the result of using a depth image for classification. The overall accuracy was 82.6%. The right matrix was the result of the point cloud images which was created using seven depth images for classification. The overall accuracy was 96.63%. This indicated the better performance of the point cloud dataset with a higher resolution. The higher resolution provided a better measurement of shape and size. The developed model from point cloud images was used to identify the maturity (Figure 1) and verified in the field by an expert. The results of test accuracy were 97.7%. Verification was completed on the 7th and 8th Nov 2019. Out of the 348 mature mushrooms, 340 mature mushrooms were successfully identified.
4.4 Maturity index

Traits for maturity identification were selected using the NCA method. The NCA algorithm was applied to both the traits collected from the depth and point cloud images. For the traits of depth images, the selected trait set contained the difference between normal vectors, radius, height, extension, area, and perimeter. Among the selected features, shape and size were found to be the most important in maturity identification. Only one curvature feature was included in the selected features set because the resolution of images was low and low resolution of depth data produced unreliable gradient vector values. Size and the 2D shape feature were easier to be measured accurately. The curvature using gradient vectors required a more continuous surface with higher resolution depth data. On the other hand, for the traits of point cloud images, the selected trait set contained normal vectors in the cap edge, radius, height, extent, and area. Among the selected features, curvature and size played the most important roles in maturity identification. The curvature at the edge of the cap and size changed significantly during a growing period. Except for increasing the size, the edge of the cap gets angular when it becomes mature. With high resolution point cloud, curvature and size features were accurately measured. In addition, the features from the point...
cloud image provided higher accuracy. As a result, this provided evidence that the curvature at the cap edge and the mushroom size were used as the maturity indicators.

![Feature selection from depth and point cloud images](image)

**Figure 18. Results of the feature selection from the depth and point cloud image.**

4.5 Recommendations for further research

The results of this study demonstrate an improved performance using high resolution point cloud images. Furthermore, detecting and identifying the maturity of the mushrooms were demonstrated with promising results for future applications in selective harvesting. However, several challenges need to be addressed to develop a commercial scale system for the mushroom industry.

First, the author suggests using a different camera in commercial mushroom farms due to space limitation of production facility. Although the Kinect camera provides various types of images, it requires a minimum of 50 centimeters for acquiring reliable depth data. Usually, the mushroom facilities do not have such space for installing the imaging system due to multiple beds layer for efficiency. The working range of the camera of 15 centimeters will be preferable, which can be feasible using a stereo vision camera or a line scanner. Second, the position of the mushrooms in the corner was usually not straight, which affects the performance of the algorithms. The maturity-misclassified
mushrooms were most often this case. A further investigation of effects of mushroom position to morphological traits in point cloud images will be needed to improve the accuracy. In addition, a flexible maturity definition will need to be investigated to further strengthen this study. Different mushroom facilities require different maturity criteria for various demands of customers. Some may need the mushrooms to be picked when mushrooms are still small to meet the demand of the fresh market. Finally, a further investigation of how many point clouds are required to achieve the highest accuracy is needed. The author used at most 21 point cloud images to reconstruct the mushroom surface and achieve 92.26% accuracy. However, this included only 162 mushrooms. The author suggests collecting more images for acquiring the reliable results.

![Figure 190. The result of using different numbers of point cloud images and their accuracies.](image-url)
5. Conclusion

In this study, a machine vision system to detect mushrooms and evaluate maturity stages was developed. A total of 4719 mushrooms in different growing stages were monitored by time-lapse images. The point cloud images were used for reconstructing the high-resolution cap surface. The locations of mushrooms were detected using a watershed and Faster R-CNN algorithm at 85.37% and 95.48% accuracy, respectively. The maturity was identified using quantified curvature, size, and shape features. The SVM classifier was then applied to identify maturity from the depth and the point cloud images. The training accuracies were 82.6% and 96.63%, respectively. When testing the model in the field, the accuracy can reach 97.7%. After developing the model, the feature selection was applied to find the most important features for identifying maturity. The curvature at the cap edge and the size of mushrooms were proved to be the maturity indicators. This study can be extended to field applications and to be used in a commercial scale mushroom farm.
References


Lee, C. H., Hsu, H. C., Yang, C. K., Tsai, M. J., & Kuo, Y. F. (2019). Identifying Fagaceae Species in


[https://usda.library.cornell.edu/concern/publications/r781wg03d?locale=en](https://usda.library.cornell.edu/concern/publications/r781wg03d?locale=en)


Wu, D., Shi, H., Wang, S., He, Y., Bao, Y., & Liu, K. (2012). Rapid prediction of moisture content of
dehydrated prawns using online hyperspectral imaging system. *Analytica Chimica Acta*, 726, 57-66.

Appendix

The code and the raw data can be found online at https://github.com/harry83017622/mush_detect