Vision-based Potato Detection and Counting System for Yield Monitoring

Young-Joo Lee[1], Ki-Duck Kim[1], Hyeon-Seung Lee[1], Beom-Soo Shin[1]*

[1]Department of Biosystems Engineering, Kangwon National University, 1 Kangwondaehak-gil, Chuncheon, Gangwon-do, 24341, Republic of Korea

Abstract

Purpose: This study has been conducted to develop a potato yield monitoring system, consisting of a segmentation algorithm to detect potatoes scattered on a soil surface and a counting system to count the number of potatoes and convert the data from two-dimensional images to masses. **Methods:** First, a segmentation algorithm was developed using top-hat filtering and processing a series of images, and its performance was evaluated in a stationary condition. Second, a counting system was developed to count the number of potatoes in a moving condition and calculate the mass of each using a mass estimation equation, where the volume of a potato was obtained from its two-dimensional image, and the potato density and a correction factor were obtained experimentally. Experiments were conducted to segment potatoes on a soil surface for different potato sizes. The counting system was tested 10 times for 20 randomly selected potatoes in a simulated field condition. Furthermore, the estimated total mass of the potatoes was compared with their actual mass. **Results:** For a 640 × 480 image size, it took 0.04 s for the segmentation algorithm to process one frame. The root mean squared deviation (RMSD) and average percentage error for the measured mass of potatoes using this counting system were 12.65 g and 7.13%, respectively, when the camera was stationary. The system performance while moving was the best in L1 (0.313 m/s), where the RMSD and percentage error were 6.92 g and 7.79%, respectively. For 20 newly prepared potatoes and 10 replication measurements, the counting system exhibited a percentage error in the mass estimation ranging from 10.17−13.24%. **Conclusions:** At a travel speed of 0.313 m/s, the average percentage error and standard deviation of the mass measurement using the counting system were 12.03% and 1.04%, respectively.

Keywords: Counting system, Mass estimation equation, Potato yield monitoring, Segmentation algorithm

Introduction

Yield maps are an essential element in precision agriculture, because they act as indicators of the spatial and temporal cropping potential (Persson et al., 2004; Kumhála et al., 2009). Yield monitoring should first be performed to construct a yield map, and many studies have been conducted using load cells.

Vellidis et al. (2001) employed load cells to quantify the load of peanuts accumulated in a collection basket during harvest. However, the overall system was complicated and heavy, because it used a collection basket and a hydraulic cylinder. Ehlert (2000) used a bounce plate for the yield map, but additional conveyor belts were required to transport the potatoes. The method using load cells requires additional equipment, and so the size of the tractor should be large.

Computer-based vision technology provides a high level of flexibility and repeatability at a relatively low cost, with a high plant throughput and superior accuracy (Sylla, 2002; ElMasry et al., 2012). Therefore, the development of vision-based technologies and computers has broadened the possibility of employed machine vision outdoors. Hofstee and Molema (2002) conducted a study on machine vision-based yield mapping of potatoes, and they employed a similar system for estimating the volume of potatoes partly covered with dirt tare (Hofstee and Molema, 2003). They proposed the possibility of estimating potato masses from two-dimensional image

*Corresponding author: Beom-Soo Shin
Tel: +82-33-250-6493; Fax: +82-33-259-5561
E-mail: bshin@kangwon.ac.kr

information. However, they reported that further research on this method was required. Persson et al. (2004) used an optical sensor for tuber yield monitoring, and Gogineni et al. (2002) conducted a study to measure image-based sweet potato yield. However, the methods in both of these studies include conveyor belts.

The yield monitoring concept ultimately adopted in this study is a method of estimating the mass from two-dimensional image information, by employing a camera installed vertically for potatoes removed from the ground by a harvester. The advantage of this system is that it can be applied regardless of the size of the tractor, because the structure is considerably simpler than those in previous studies, and can greatly reduce the weight of the required equipment. Image processing technology is also important, because this method must be applied outdoors.

In this study, we developed a segmentation algorithm that separates the background (soil) from the objects (potatoes). Furthermore, we estimated the masses of potatoes using the volume of a potato calculated from two-dimensional image information and the density of a potato determined experimentally. A counting system was developed that estimates the number and masses of potatoes by applying the developed mass estimation equation to the counting algorithm. Performance tests were carried out to evaluate the mass measurement accuracy under moving conditions.

Materials and Methods

Camera

The camera (Prosilica GT1290, Allied Vision Technologies GmbH, Stadtroda, Germany) used in this study is equipped with an auto-iris function, allowing images to be received with a constant illuminance level up to 33.3 fps, and its image sensor is of the CCD progressive type. It was mounted on camera support at a height of 100 cm above the ground, and communicated with the computer via Gigabit Ethernet. The focal length of the camera lens (AG3Z3112KCS-MPIR, Computar, Cary, NC, USA) was fixed at 8 mm. The major specifications of the camera and lens are presented in Table 1 and Table 2, respectively.

Segmentation algorithm

A flow chart of the segmentation algorithm is presented in Figure 1. An acquired image is converted to grayscale. After top-hat filtering is first performed, it is converted into a binary image using Otsu's method. To remove noise, morphological opening is twice performed using the structuring element (SE) method on the binary image. Here, the first structuring element SE1 is square-shaped with a width of 10 pixels, and the second structuring element SE2 is disk-shaped with a radius of 10 pixels. Finally, the areas of the object remaining

Table 1. Specifications of camera	
Item	Specifications
Model/Manufacturer	Prosilica GT1290/Allied Vision
Resolution (pixel)	1280 (H) × 960 (V)
Sensor/type	Sony ICX445/CCD Progressive
Cell size (μm)	3.75 × 3.75
Lens mount	C-Mount
Max frame rate at full resolution (fps)	33.3

Table 2. Specifications of lens mounted on camera			
Item			Specifications
Model/Manufacturer			AG3Z3112KCS-MPIR/Computar
Focal Length (mm)			3.1–8
Iris	Operation Range		F1.2–F16C
	Control		Stepping Motor
Angle of View (°)	1/2.7 type (4:3)	D	123.1–48.3
		H	95.9–38.7
		V	71.0–29.1

in the image are calculated and recognized as a potato if the area is greater than 8000 pixel. If smaller than that, the area data is deleted. This minimum area value represents the smallest of the potatoes used in the experiment, and can be adjusted by the system operator. MATLAB Image Processing Toolbox (MATLAB, 2016) was employed in this study.

Counting system

Mass estimation

Using potato varieties (*Solanum tuberosum L.* subsp. *tuberosum*) widely cultivated in the Republic of Korea, a mass estimation model from two-dimensional surface images was developed. The potatoes purchased in a market were first divided into four groups (under 90, 90–120, 120–150, and over 150 g) based on mass. Twenty potatoes were randomly selected from each group, and a total of 80 potatoes were used to derive the mass estimation model.

The mass of a potato was estimated by multiplying the volume extracted by the volume calculation equation by the density, which was determined experimentally. First, a mathematical equation was verified to calculate the volumes of potatoes from two-dimensional images from a stationary camera. Assuming that the mass of a potato is proportional to the area projected onto an image, and that the shape of the potato is ellipsoidal (Pitts et al., 1987), the volume was calculated using equation (1) after rotating the ellipse about the major axis and setting the height and width of the rotating body to be the same. The major axis is determined as the axis corresponding to the longest distance in an ellipse with the same area as a two-dimensional projected potato image.

$$V = \frac{4}{3}\pi a b^2 \tag{1}$$

where V = volume of potato, m^3
$\quad\quad a$ = major axis length, m
$\quad\quad b$ = minor axis length, m

Second, the density of the potatoes was determined. The actual volume of each potato was measured by placing it in a cylinder filled with a certain amount of water, and the actual mass was measured using an electronic scales (MW II-N 3000, CAS Corp., Yangju-si, Gyeonggi-do, Rep. Korea). Tests were replicated three times, and the accuracy of the instrument was ± 0.1 g. Finally, the mass of a potato was estimated using equation (2). The mass coefficient is defined as the ratio of the estimated mass to the actual mass of the potato, and is used as a correction factor.

$$M = e_m p V \tag{2}$$

where M = estimated mass of potatoes, kg
$\quad\quad e_m$ = average of the mass coefficient
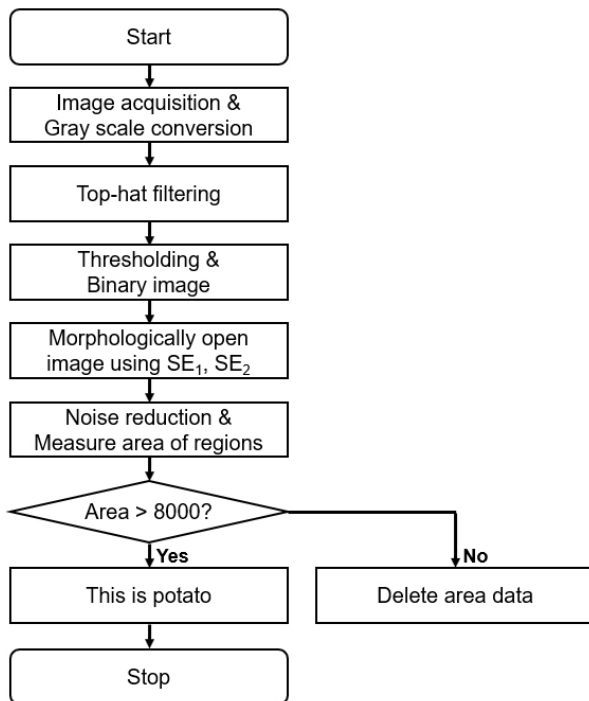$\quad\quad p$ = average of the measured potato density, kg/m^3

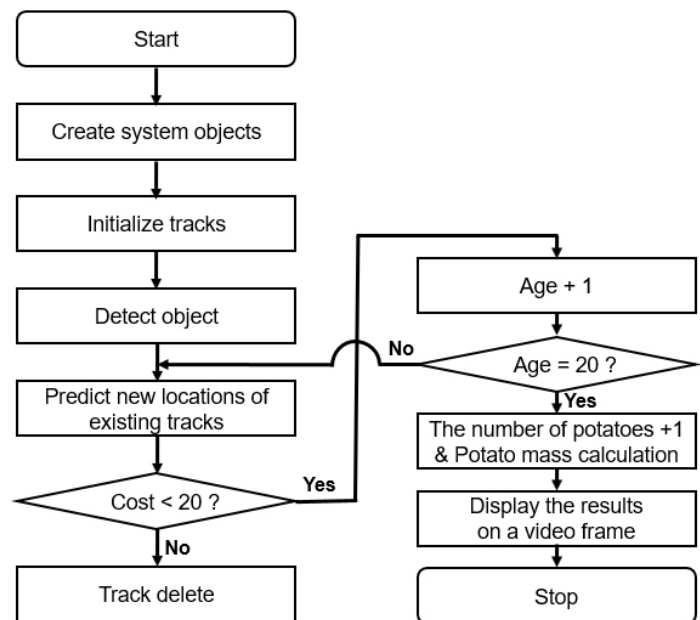**Figure 1.** Flow chart of segmentation algorithm.

Figure 2. Flow chart of counting algorithm.

Counting algorithm

A flow chart of the counting algorithm is presented in Figure 2. First, system objects are created to read the video frames, where foreground objects are detected. In the track initialization stage, an array of tracks is created, where each track is a structure representing a moving object in the video. The detection of moving objects employs a background subtraction algorithm based on Gaussian mixture models. When an object is detected, the characteristics of the object such, as the centroid coordinate, area, bounding box, major axis length, and minor axis length, are calculated.

As the next step, the Kalman filter is employed to predict the centroid of each track in the current frame, and its bounding box is updated accordingly for display purposes. Then, the cost is calculated. The cost is defined as the negative log-likelihood of a detection corresponding to a track, and this employs the Euclidean distance between the predicted centroid of the track and the centroid of the detection.

If the cost is under 20, then the previous and current objects are recognized as being the same, and the age of the object is increased by one. However, if it is larger than 20, then the track is deleted, as it is judged to be too damaged to be recognized as the same object. The process returns to predicting the location of the track until the age of the object reaches 20. At this point, the object is recognized as a perfect potato, its mass is calculated, and the number of potatoes is counted. Finally, the result is displayed on a video frame and the algorithm ends.

Evaluation of overall system performance

Experiments were conducted to confirm whether the developed algorithm could be applied under moving conditions. After preparing a potato harvesting field indoors by placing potatoes randomly on a strip of soil surface, images were captured using a camera vertically installed on an agricultural vehicle (track-type speed sprayer) at three different travel speeds, 0.313 m/s, 0.618 m/s, and 0.959 m/s, according to gear shifts L1, L2, and H2, respectively (Park, 2008). The tests were replicated three times at each travel speed level.

For validation purposes, a new set of 20 potatoes were prepared and the performance of the counting system was evaluated at a fixed travel speed corresponding to gear shift L1. The test run was replicated 10 times.

Results and Discussion

Segmentation algorithm

Figure 3 presents the images corresponding to each step of the segmentation algorithm. The original image contains 11 potatoes, and has been marked with green circles to clearly show the locations of the potatoes. Through top-hat filtering, we found that potatoes appear darker than the soil. Because of these differences, it was possible to distinguish soil and potatoes when constructing a binary image using Otsu's method. It was confirmed that the noise in an image is reduced by applying morphological opening using the
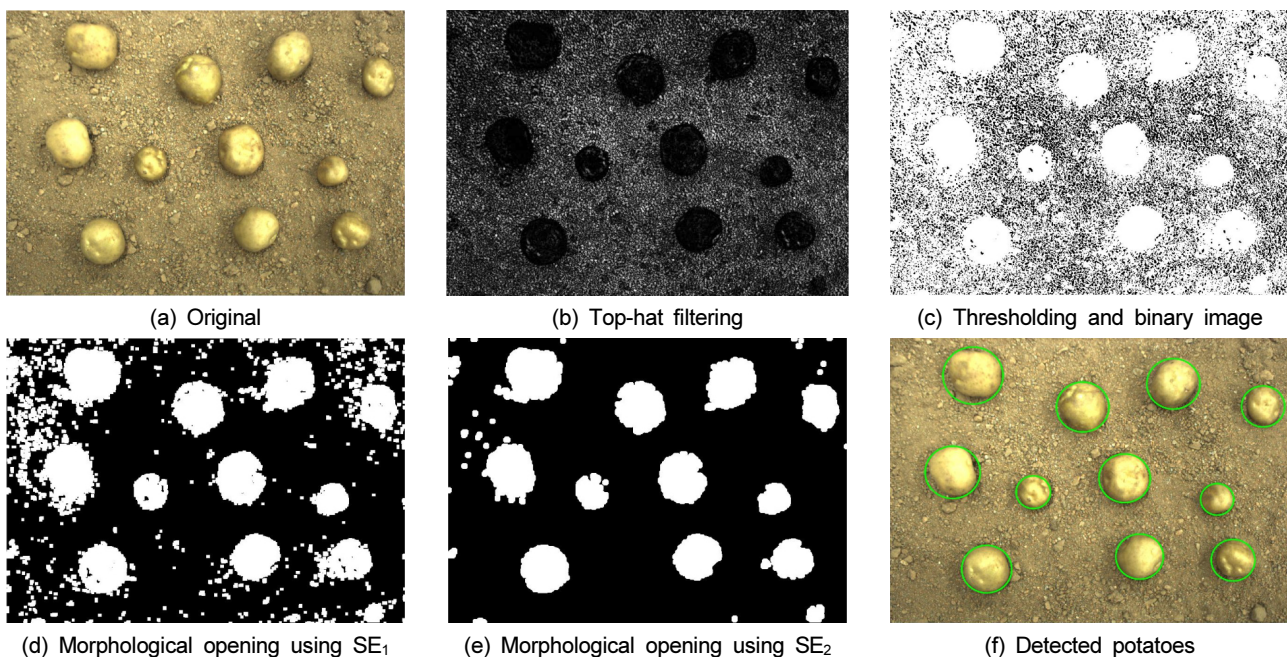

(a) Original   (b) Top-hat filtering   (c) Thresholding and binary image

(d) Morphological opening using SE$_1$ (e) Morphological opening using SE$_2$ (f) Detected potatoes

Figure 3. Each of the steps of the segmentation algorithm.

structuring elements SE_1 and SE_2. As a result, 11 potatoes were completely separated from the soil. When the image size is 640 × 480, the processing speed for one frame is measured to be about 0.04 s.

Counting system

Mass estimation

The densities of potatoes in four different groups are presented in Table 3. These turned out to be relatively uniform. The range of the 95% confidence interval varied slightly between 1.3% and 2.3%. The average total density of the potatoes was 1213.5 kg/m³, the standard deviation was 55.1 kg/m³, and the 95% confidence interval was ± 12.1 kg/m³ (1.0%).

The mass coefficients of the four different groups are presented in Table 4. It was confirmed that this decreased as the mass of a potato increased. The masses of potatoes were overestimated, because the shapes of potatoes were assumed to be ellipsoidal. Therefore, the mass coefficient is less than 1. For the mass estimation equation, the average value for each group was employed.

The deviations and percentage errors between the actual masses of potatoes and the masses obtained using the estimation equation are presented in Table 5. As the mass of the potato increased, the root mean squared deviation (RMSD) tended to increase, and the total RMSD was 12.65 g. On the other hand, the percentage error did not increase even when the mass of a potato increased. It was the largest in the group over 150 g, and the total average percentage error was 7.13%.

Table 3. Densities of potatoes in four different groups

Classification by mass (g)	N	Density (kg/m³)					
		Max	Min	Average	Std. dev.	95% conf. int.	
below 90	20	1382.5	1112.1	1231.3	64.0	± 28.0	2.3%
90 – 120	20	1418.3	1139.1	1228.3	63.8	± 28.0	2.3%
120 – 150	20	1289.0	1144.2	1204.9	40.0	± 17.5	1.5%
over 150	20	1294.3	1140.6	1189.7	34.5	± 15.1	1.3%
Total	80	1418.3	1112.1	1213.5	55.1	± 12.1	1.0%

Table 4. Mass coefficient of four different groups (e_m)

Classification by mass (g)	N	e_m
below 90	20	0.6432
90–120	20	0.6429
120–150	20	0.5884
over 150	20	0.5760
Total	80	0.6126

Table 5. Deviations and percentage errors between the actual masses of potatoes and the masses obtained using the mass estimation equation

Classification by mass (g)	N	Deviation (g)		Percent error (%)		
		Average	RMSD	Max	Min	Average.
below 90	20	3.92	5.18	20.07	0.20	6.04
90–120	20	7.48	9.06	16.81	0.23	7.16
120–150	20	8.80	11.83	17.03	0.24	6.26
over 150	20	16.56	19.79	19.52	0.53	9.05
Total	80	9.19	12.65	20.07	0.20	7.13

Table 6. Deviation and percentage error according to travel speed

Travel speed (m/s)	Deviation (g)		Percent error (%)		
	Average	RMSD	Max	Min	Average.
L1 (0.313)	5.78	6.92	21.38	0.13	7.79
L2 (0.618)	5.48	6.35	22.07	0.07	7.89
H2 (0.959)	6.36	8.31	21.56	0.07	8.99

Counting algorithm

Table 6 presents the deviations and percentage errors of the counting system according to the travel speed of the machine, where the selected 20 potatoes were not replaced. The RMSD and average percentage error at 0.313 m/s and 0.618 m/s were not significantly different, although slight differences are observed. However, the system performance was lowest at a travel speed of 0.959 m/s.

Because the average percentage error was lowest for L1 in the previous experiment, the travel speed was fixed at 0.313 m/s. Figure 4 presents the results of 10 repetitions for a random extraction test. As a result, the average and standard deviation were 12.03% and 1.04%, respectively, and exhibited relatively constant results. Figure 5 presents a scene in which the results of the counting system are displayed on a video frame. Because the segmentation was assumed to be perfectly successful, the images were acquired with white pixels.

Conclusions

In this study, the possibility of employing a segmentation algorithm to detect potatoes from raw images of a potato harvesting field was identified, and a counting system was developed to count the number of potatoes and calculate the mass of each potato.

Through the images, we verified the process of detecting potatoes while reducing the noise at each step of the segmentation algorithm. When the size of the image is adjusted to 640 × 480, the time required for the segmentation algorithm is about 0.04 s, and it guarantees image processing up to 25 fps.

The average and standard deviation of the potato density used in the mass estimation equation were 1213.5 kg/m^3 and 55.1 kg/m^3, respectively. However, to reduce the deviation, a new method of measuring the densities of potatoes is required.
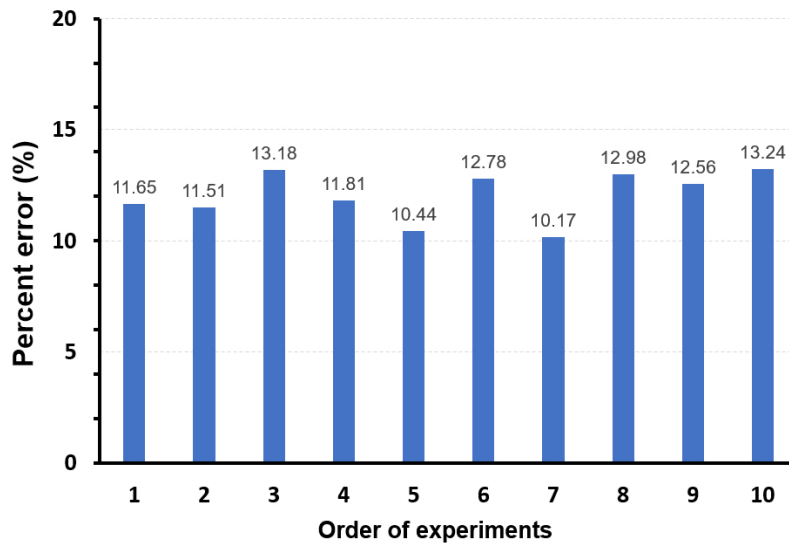

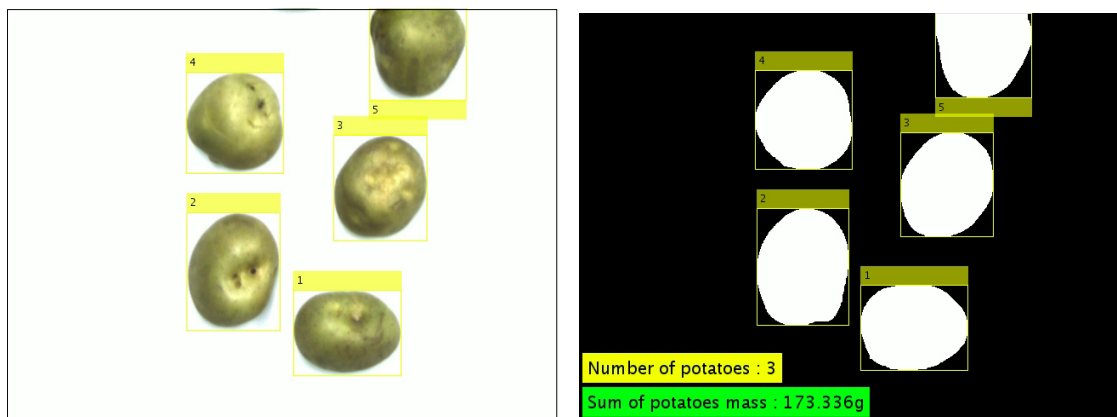**Figure 4.** Measurement percentage error over 10 replications with 20 potatoes.

Figure 5. A scene that counts the number of potatoes and calculates the masses of potatoes using an algorithm.

As the mass of potatoes increased, the mass coefficient tended to decrease. The average mass coefficient used in the mass estimation equation was 0.6126. The RMSD and the average percentage error for the counting system in measuring the masses of potatoes were 12.65 g and 7.13%, respectively, when the camera was stationary.

The system performance was strongest in the L1 condition (0.313 m/s), where the RMSD and percentage error were 6.92 g and 7.79%, respectively. The results of 10 test runs showed that the average percentage error and standard deviation in estimating the masses of 20 potatoes were 12.03% and 1.04%, respectively, and exhibits relatively uniform results.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgement

References

Ehlert, D. 2000. Measuring mass flow by bounce plate for yield mapping of potatoes. Precision Agriculture 2(2): 119-130.
https://doi.org/10.1023/A:1011469429338

ElMasry, G., S. Cubero, E. Moltó and J. Blasco. 2012. In-line sorting of irregular potatoes by using automated computer-based machine vision system. Journal of Food Engineering 112(1-2): 60-68.
https://doi.org/10.1016/j.jfoodeng.2012.03.027

Gogineni, S., J. G. White, J. A. Thomasson, P. G. Thompson, J. R. Wooten and M. Shankle. 2002. Image-based sweetpotato yield and grade monitor. In: 2002 ASAE Annual Meeting, ASAE Paper No. 021169. St. Joseph, MI: ASAE.
http://doi.org/10.13031/2013.10586

Hofstee, J. W. and G. J. Molema. 2002. Machine vision based yield mapping of potatoes. In: *2002 ASAE Annual Meeting*, ASAE Paper No. 021200. St. Joseph, MI: ASAE.
http://doi.org/10.13031/2013.9699

Hofstee, J. W. and G. J. Molema. 2003. Volume estimation of potatoes partly covered with dirt tare. In: *2003 ASAE Annual Meeting,* ASAE Paper No. 031001. St. Joseph, MI: ASAE.
http://doi.org/10.13031/2013.15380

Kumhála, F., V. Prošek and J. Blahovec. 2009. Capacitive throughput sensor for sugar beets and potatoes. Biosystems Engineering 102(1): 36-43.
https://doi.org/10.1016/j.biosystemseng.2008.10.002

MATLAB. 2016. Image Processing Toolbox User's Guide. Ver. R2016a. Natick, MA, USA: The MathWorks, Inc.

Park D. S. 2008. Development of steering controller for autonomous-guided orchard sprayer. Unpublished MS thesis. Chuncheon, Gangwon-do, Rep. Korea: Department of Biological Systems Engineering, Kangwon National University.

Persson, D. A., L. Eklundh and P.-A. Algerbo. 2004. Evaluation of an optical sensor for tuber yield monitoring. Transaction of the ASAE 47(5): 1851–1856.
http://doi.org/10.13031/2013.17602

Pitts, M. J., G. M. Hyde and R. P. Cavalieri. 1987. Modeling potato tuber mass with tuber dimensions. Transactions of the ASAE 30(4): 1154-1159.
http://doi.org/10.13031/2013.30536

Sylla, C. 2002. Experimental investigation of human and machine-vision arrangements in inspection tasks. Control Engineering Practice 10(3): 347-361.
https://doi.org/10.1016/S0967-0661(01)00151-4

Vellidis, G., C. D. Perry, J. S. Durrence, D. L. Thomas, R. W. Hill, C. K. Kvien, T. K. Hamrita and G. Rains. 2001. The peanut yield monitoring system. Transactions of the ASAE 44(4): 775-785.
http://doi.org/10.13031/2013.6239