Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-32 (2022) 119–123

Determining fresh tomato weight using depth images from an AR headset

Tim van Daalen*, Joseph Peller*, Jos Balendonck*

**Wageningen University & Research, Wageningen, The Netherlands*
(e-mail: {joseph.peller, tim.vandaalen, jos.balendonck}@wur.nl)

Abstract: There is a need for handsfree devices in greenhouses for training and harvesting assistance. Augmented Reality could offer this in the near future. We showed that it is possible to detect ripeness characteristics of tomatoes using the 3D scanning capabilities of the HoloLens. We verified this in an experimental setup with multiple tomato varieties. Our results show the possibilities and problems of this technique for future development.

Keywords: HoloLens, Depth imaging, Volume estimation, Handsfree measurement

1. INTRODUCTION

The global market of greenhouse horticulture is expected to double in the next five years (IMARC, 2022). Greenhouse horticulture is extremely important to the economies of several European countries including the Netherlands where an estimated 105 km^2 of the country is covered in greenhouses and that number is expected to grow (CBS, 2022). As the greenhouse industry expands, the adoption and retention of labour are two major bottlenecks of the industry. Work in greenhouses is hard, the climate is hot and requires expertise to correctly care and harvest the various plants grown indoors. Nowadays many new, un-skilled workers find a job in the greenhouse industry. Training these newcomers forms a considerable workload for greenhouse owners.

For greenhouse workers, assessing ripeness and determining the optimum time to harvest a product are major important skills to learn. In tomato crops, the correct time to pick a tomato is based on a combination of factors including colour, texture, and weight of the tomato. These factors are multimodal and have made automation of ripeness detection difficult in the past. There is a need for new tools and methodologies that can help train the next generation of greenhouse workers and improve their job quality and efficiency.

Augmented Reality (AR) is a hands-free solution that allows breeders, growers and personnel to interact intuitively with their crops and data. AR has been developed over the last 30 years and is a system that allows virtual objects to be placed in the real world. While initially limited due to the computational requirements of AR, the advent of smartphones, processing power and new more user friendly wearable peripherals has made AR applications more common. With the development of the Google glass in 2013, a new wave of AR headsets has entered the market with the added benefit of being handsfree.

In this paper we describe a tool we are developing to assist the grower. We used a HoloLens 2 headset (Microsoft) and worked with partners in tomato greenhouses in the Netherlands. We first developed the software and hardware and then validated it efficacy of our technique on multiple tomato varieties.

2. MATERIALS AND METHODS

2.1 Weight & Volume estimation

Weight is a key component to determining ripeness in tomatoes. However, weight is difficult to measure per tomato, taking additional labour to weigh each tomato from a crop. Furthermore, weighing can only be done after harvest. A better solution would be to estimate the weight of tomatoes using imaging, prior to harvesting, and leaving the option for not harvesting unripe fruits. Depending on growth conditions, tomatoes consist of between 92% and 95% water (Johnson, 1992). This water makes up most of the mass of the tomato, leading to a direct correlation between the weight and the volume. For this correlation, several papers have proposed different relations, e.g. Lee (2020) used machine learning based on the visible traits of tomatoes in images to estimate volume accurately but was limited by occlusion. Nyalala (2019) found that a power equation fit the relationship between volume and mass well, but only estimated the volume of the tomato using voxels from their depth camera. In both cases it was clear that there is a relationship between the volume and mass for the majority of tomato varieties.

Volume estimation from 2D images has also been explored in recent years. Du and Sun (2020) showed the volume of ham can be estimated from RGB images, and UluiŞik et al. (2018) showed that it is possible to estimate the volume of tomatoes using simple geometric modelling. Two dimensional horizontal and vertical cross sections of the fruit can be used to find the diameters of each projection. This will result in

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2022.11.125

four axis of which two are the same diameter from different projections. The volume of an ellipsoid can be computed as follows:

$$volume = \frac{\pi}{6} ABC \quad , \qquad (1)$$

where A, B, and C are the diameters of the constituent ellipses, as shown in figure 1.

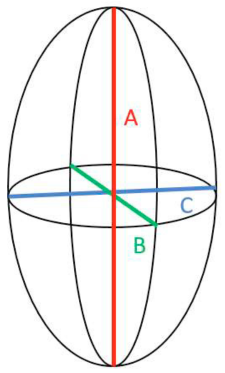

Fig 1. An example of an ellipsoid with diameter A, B, and C.

2.2 Image acquisition

AR headsets like the HoloLens 2 are typically used to create an augmented experience. They are designed to place virtual holograms in a real room, register user inputs and stream live video data. However, the HoloLens 2 has multiple sensors and cameras onboard that capture infrared and RGB images, depth maps, and magnetic fields that are normally locked from the developer. To use the headset as a measurement tool, more control over these sensors was needed.

In 2020, Microsoft gave direct access to the raw sensor streams through their Research Mode API (Ungureanu et al., 2020). This allowed us to write a program that used the short range, high resolution articulated hand-tracking (AHAT) depth camera and RGB camera of the Hololens 2 for the volume estimation. The depth camera returns both a depth and active brightness image. When the program starts, it activates the cameras. Voice commands and gestures can then be used to save images on the HoloLens. The depth and active brightness images are combined into a point cloud on an edge computer. The density of the point clouds was 50k voxels and covered a distance from 10 cm to 1 m from the HoloLens. At minimum distance from the camera, we had a feature resolution of 241 voxels/cm^2 and at maximum distance 2 voxels/cm^2.

2.3 Volume Verification Trial

To verify that our program was working accurately we designed a measurement trial. Tomatoes were imaged in 2020 at two demo greenhouses in the Netherlands (Syngenta, Tomato Vision; Letsgrow, Tomato World). Both trials were conducted with multiple varieties, from large beef tomatoes down to small cherries. In total 10 varieties were imaged, with at least 10 tomatoes per variety. Examples of each variety are shown in image 2.

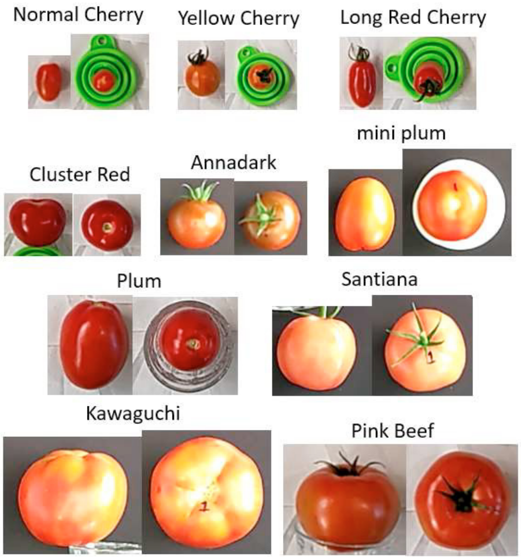

Fig 2. Examples of the images tomato varieties.

The tomatoes were first weighed using a standard scale. The tomatoes were then submerged in a graduated cylinder and held just below the surface using a pen. The displacement of the water was recorded as the true volume of each tomato.

Tomatoes were then imaged one after another using the HoloLens. The HoloLens was placed directly above the tomato on the table and each tomato was imaged in two positions, laying horizontally, and mounted vertically. The Hololens was held at roughly one arm length from the tomato between 0.5 m and 1 m.

After the trials, the images were downloaded from the HoloLens and saved with their respective weights for later analysis. Segmentation of the tomatoes in the point cloud were accomplished using the active brightness values of the point cloud. As seen in Fig. 3, the flat tabletop provided good contrast with the tomato and manual segmentations were performed.

A selection of 32 tomatoes were segmented by hand and their volumes computed using the method described in 2.1.

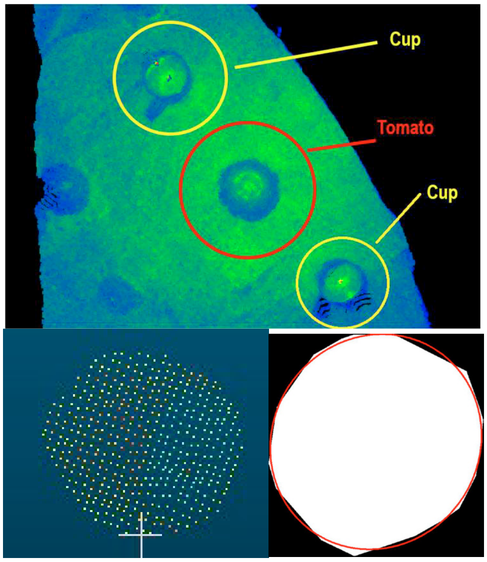

Fig 3. Top, active brightness point cloud with a tomato in the centre and two cups on the sides. Bottom left, segmented tomato. Bottom right, rendered top view of the contour and ellipse fit in red.

3. RESULTS

3.1 Density measurements

The weight and volume of all 116 tomatoes derived through submersion is plotted in figure 4. The trend appeared linear, so a linear regression was performed to show the relation between weight and volume. The derived weight is shown in (2) and had an R^2 value of 0.988.

$$Weight\ (g) = 0.987 * \text{Volume} + 1.116 \qquad (2)$$

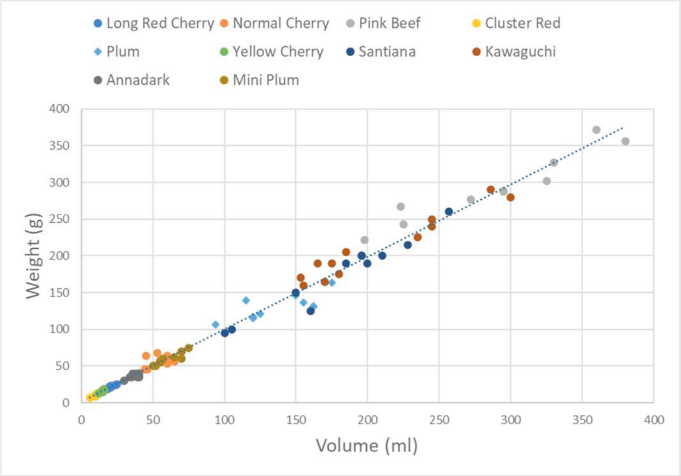

Fig 4. Measured tomato weight versus submerged volume per variety measured.

3.2 Volume Comparison

A subset of 32 tomatoes were segmented in the active brightness images. The tomatoes where manually segmented in these images and an exterior contour was created.

As visible in figure 3 this segmentation is not perfect, due to the low density of points within the point cloud. To correct for this, an ellipse fit from the opencv library (Bradski, 2000) was performed. The function returns both the major and minor axis of the ellipse, which are two of the diameters of the ellipsoid shown in figure 1. The pixels are transformed into millimetres using the depth information and the ellipsoid volume computation described in section 2.1.

The ellipsoidal volume was compared to the submerged volume as can be seen in Figure 5. Again, a linear relation seemed to exist between the ellipsoidal (V_{elli}) and submerged volume (V_{sub}), and a linear regression was performed. The relationship is shown below and has an R^2 value of 0.964.

$$V_{sub} = 1.509 * V_{elli} - 10.607 \qquad (3)$$

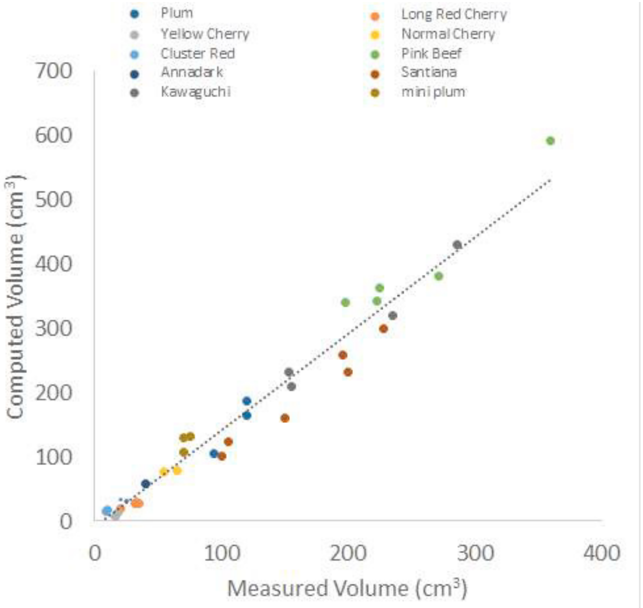

Fig 5. The Computed elliptical volume versus Measured Volume.

4. DISCUSSION

The results of the density measurements indicate that the relationship between weight and volume is linear for these varieties. This is largely in line with literature, our R-Squared value is nearly the same as Nyalala et al. (2019) found with their power relation. We chose to use the linear function as it fits the data well and is less complex than the power function. It may be that this discrepancy with the literature was caused by the different tomato varieties investigated in this study. Nyalala et al. focused on cherries while we analysed multiple varieties. These varieties were some of the normal varieties grown in the Netherlands but do not include exceptions such as hollow tomatoes which we expect will not follow the same trend.

Because we compared two quantitative measurements in our volume comparison tests, we wanted to be sure there is no bias in the volume measured by the HoloLens. To check this a Bland Altman (BA) plot (Bland and Altman, 1986; Giavarina, 2015) was used to evaluate the two methods (figure 6). BA plots are used to identify systematic biases that may not be as apparent while using a least squares regression. In our case, the two measurements show a systematic underestimation of larger tomatoes compared to smaller ones. However, more importantly is the trend of the data in the BA plot. This shows that at small volume there is good agreement between the measures, but as volume increases the bias grows. In reviewing the data this appears to be caused by the more "squished" non elliptical shapes of larger tomatoes, as can be seen in figure 2. The single outlier in particular is due to the shape of the side ellipse encompassing an additional 20% area outside of the tomato.

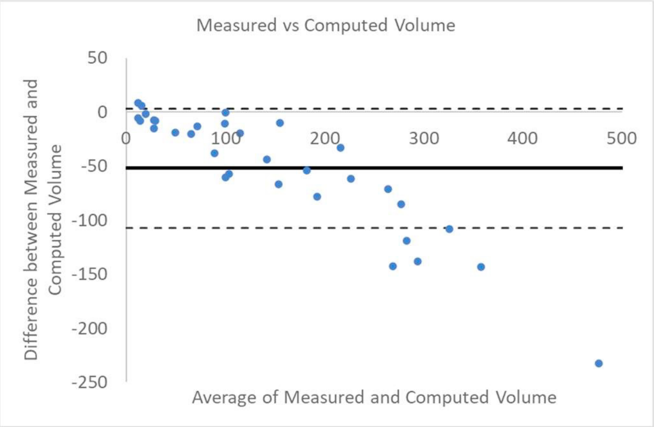

Fig 6. Bland–Altman plot of the Measured and Computed Volume.

This is more than a little concerning since in this trial most tested varieties were roughly ellipsoidal. No tomato varieties such as Coeur de Boeufs or heavily ribbed varieties were included in the trials. Even for these varieties the results of our volume comparison show there is a difference between the volume recorded by the HoloLens and the volume derived through submersion. More tomatoes should be analysed to investigate this further and determine if variety specific models are necessary. A possible approach to this problem would be to use the shape of the tomato in each orientation to construct a 3D projection, from which the volume can be determined. This would eliminate the errors from the ellipsoidal fit for most symmetric tomatoes.

The next step for this system is to implement a true segmentation network and test it in the greenhouse on tomatoes on the vine. Our original intention was to use the colour images provided by the HoloLens to segment the tomatoes from the canopy. Unfortunately, a software bug in the HoloLens prevented automatic segmentation, which meant our tests could only be conducted with a fixed background. Once the bug is fixed automatic segmentation of tomatoes will be accomplished by using the colour images with a neural network and then combining those annotations with the depth images to perform in greenhouse segmentation.

To this end, the issue of orientation needs to be addressed. All tomatoes imaged were situated on a table post-harvest. The ellipsoidal method requires the user to photograph at least two sides of the tomato. While this is easy to perform post-harvest, it is challenging on the vine. The orientation of the tomato has to be detected first. A two images approach was used, which is more labour intensive than a single image approach.

There are several possible solutions, such as continuous streaming to create a 3D object, or adding orientation as a class to the neural network. Another approach could be to determine per variety a correlation between the first and second diameters and the third. This was observed in our trial, and if it held true could mean an approximate volume could be determined by a single image in the greenhouse. But a larger dataset needs to be investigated to confirm this relationship first.

5. CONCLUSION

In this trial we created a system to perform handsfree tomato fruit weight measurements. By measuring the volume of the tomato using depth cameras built into the Microsoft HoloLens 2 the weight could be estimated. A determination of tomato density was performed. The computed volume from the HoloLens was found to have good agreement with the actual volume in the form of a linear relationship. This baseline shows the potential for hands free machine vision applications in the future and for further development with augmented reality systems.

ACKNOWLEDGMENTS

The authors would like to extend thanks to Jos Ruizendaal, Anna Petropoulou, Arjan Vroegop, and Menno Sytsma from Wageningen University and Research for their assistance in this project and development of the HoloLens application. This work "Augmented horticulture – Understanding plants via augmented reality" was funded by the Dutch Topsector Tuinbouw en Uitgangsmaterialen (reference: TU18145).

REFERENCES

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The lancet, 327(8476), 307-310.

Bradski, G. (2000). The openCV library. Dr. Dobb's Journal: Software Tools for the Professional Programmer, 25(11), 120-123.

CBS, (2022). Greenhouse horticulture by size class, 2021, Assessed (16-2-2022) https://www.cbs.nl/nl-nl/maatwerk/2022/06/glastuinbouw-naar-omvangsklasse-2021.

Du, C. J., & Sun, D. W. (2006). Estimating the surface area and volume of ellipsoidal ham using computer vision. *Journal of Food Engineering*, *73*(3), 260-268.

Giavarina D, (2015). Understanding Bland Altman analysis. *Biochemia medica, 25(2),* 141–151. https://doi.org/10.11613/BM.2015.015.

IMARC, (2022). Greenhouse Horticulture Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2022-2027. Assessed (16-2-2022) https://www.imarcgroup.com/greenhouse-horticulture-market.

Johnson, R. W., Dixon, M. A., & Lee, D. R. (1992). Water relations of the tomato during fruit growth. Plant, Cell & Environment, 15(8), 947-953.

Lee, J., Nazki, H., Baek, J., Hong, Y., & Lee, M. (2020). Artificial intelligence approach for tomato detection and mass estimation in precision agriculture. Sustainability, 12(21), 9138.

Nyalala, I., Okinda, C., Nyalala, L., Makange, N., Chao, Q., Chao, L., Yousaf K. & Chen, K. (2019). Tomato volume and mass estimation using computer vision and machine learning algorithms: Cherry tomato model. *Journal of Food Engineering*, *263*, 288-298.

Uluişik, S., Yildiz, F., & Özdemİr, A. T. (2018). Image processing based machine vision system for tomato volume estimation. In *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, pp. 1-4, doi: 10.1109/EBBT.2018.8391460.

Ungureanu, D., Bogo, F., Galliani, S., Sama, P., Duan, X., Meekhof, C., Stühmer J., Chashman T.J., Tekin B., Schönberger J.L., Olszta P. & Pollefeys, M. (2020). HoloLens 2 research mode as a tool for computer vision research. *arXiv preprint arXiv:2008.11239*.