Non-destructive measurement of rice leaf dimension in 3D point cloud

P Rattanasuwan¹, I Kumazawa¹ and T Kasetkasem²

1 Department of Information and Communication Engineering, Tokyo Institute of Technology, Yokohama 226-0026, Japan
2 Department of Electrical Engineering, Kasetsart University, Bangkok 10900, Thailand
E-mail: rattanasuwan.p.aa@m.titech.ac.jp

Abstract. In order to try to improve the yield, knowing the expanding of a leaf by measuring its dimension can be one easy way to monitor the plant growth. While rice is considered to be one of the most important agricultural plant, their leaves are quite difficult to measure the size since they are mostly small and direct measurement could harm the plant. To do so, we deploy the non-destructive system to measure the dimension of the rice leaf from a point cloud. The proposed system takes only photos from a smartphone camera to produce the 3D model of a specific leaf using structure from motion and multi-view stereo. To measure the dimension of the rice leaf in metric units, the scale reference is placed into the scene before the photo taking process to involve it in the reconstruction result. From the reconstructed point cloud, the leaf veins can be extracted in order to find the accurate width of a leaf based on thresholding technique using combination of RGB and HSV spaces. By comparing with measured length of the leaf in point cloud with length of one block in scale reference, we can eventually obtain the distant between each leaf vein in the metric system unit.

1. Introduction
The fact that rice is one of the most consumed plant is not uncommon knowledge. For covering the demand of people in consuming, the improvement of yield is necessary. To do so, the monitoring of rice growth can be a factor to help understanding the best environment to grow rice and may can also help reducing the cost of growing. Despite the attempt to find the optimal environment for rice, but without the proper plant growth measurement, it is difficult to determine which is the best environment. There are many ways to perform the growth monitoring, the simplest one would be to see the expanding of a leaf. Generally, it can be done by directly measure a leaf from a tree with a simple measurement tool such as a ruler. However, with this approach, there is a possibility of a tree to get damaged since in some case the leaf may have to be torn or taken out. For example, Maloof et al. has proposed a method [1] that can accurately measure the leaves parameters but the leaves have to be taken out of the tree in order to be measured. As a result, the new approach of measuring a leaf without direct contacting and non-destructive has to be proposed.

The measurement from using of 2D or 3D images is not a new thing. For several past decades, the leaf measurement systems using images have been constantly published. Obliviously the using of 2D images is far simpler and more straight forward. However, in many cases, the results are heavily depending on relative position of the camera with respect to the plant which lead to unsatisfied result when applying to other environment, for instance, the method proposed by Easlon and Bloom [2] which takes 2D image from top view of a plant along with the calibration areas can accurately
measure under the only condition that the calibration area is kept in the same plane as the leaves in order to avoid perspective distortion. On the other hand, 3D system is far more accurate but usually come up with higher costs in term of money and computation. Thus, we would like to create a simple system that can measure a leaf without direct contact by using the 3D reconstruction technique. There are variety of 3D reconstruction methods, some use a ToF camera [3] or LiDAR sensor [4] while some use a laser scanning system [5]. But most widely used one is from a normal RGB camera since it can apply to many applications while keeping the cost low. The fundamental yet effective technique of 3D reconstruction is Structure from Motion (SfM) [6]. SfM reconstructs 3D model from two or more images using the concept of stereo vision. At the present, there are many programming libraries and software providing SfM function for free. Even though it can reconstruct 3D model effectively, SfM alone can only reconstruct a sparse model, which means it cannot reconstruct fine detail of an object surface. With the aid of Multi view stereo (MVS) [7], the sparse reconstruction result from SfM can be more detail which called a dense reconstruction. There are numbers of MVS based technique, one of the most effective is Pixel-wise Multi view stereo (PMVS) [8] which take pixel information from corresponding 2D images to help enhance the detail of the model surface. With both techniques, the 3D model of an object can finally be generated with a fine detail.

The objective of this research is to monitor the growth of a rice plant by measuring the expanding of a leaf with 3D measurement technique to see how much it grows in a period of time. Although there are many ways to acquire the 3D model of a rice leaf, in order to make the system compact, we chose the camera from a smart phone to use in the 3D reconstruction process. The SfM and MVS have been used to reconstruct the 3D model of a rice leaf. For evaluation of the leaf expanding, the distance between each leaf veins are estimated from the generated point cloud. To do so, the determining of which parts of point cloud are leaf veins or leaf surfaces is essential. Hence, the process of how to classify the leaf veins and the measurement of a distance between each leaf vein in a metric system is described in this paper.

2. Methodology
This section describes the theories involved in this research and the process lead to the result of distance between leaf veins. An overview of the proposed system is illustrated in figure 1. The proposed system starts with the collection of the 2D image dataset of a rice leaf. From a set of 2D images, SfM and MVS are used to generate the 3D point cloud, then use the generated point cloud to measure the distance between each leaf vein by choosing one leaf from the whole model, select points that stay in the same line of veins, determine whether each point is leaf vein or leaf surface with the thresholding of RGB and HSV, then measure the distance in a metric system respectively. The detail of each stage is described in this section.
2.1. 3D reconstruction from 2D images

In order to reconstruct the 3D point cloud from an object, a number of photos of an object is required. The most convenient way according to our past experiments is to take a video around the desired leaf in the most various angle possible and then extract the video frames to get the photos of the leaf. With this approach, the reconstruction result tends to be more accurate since the feature matching in SfM needs sufficient number of images and angles to accurately estimate the position of features in 3D space. Please note that when taking a video of a leaf, the distant between the leaf and camera should be enough for the camera focus to be exactly on the leaf and the amount of light must be enough to clearly see the leaf detail, i.e. leaf veins. In addition, the extracted photos should be at least 100 images in order to obtain fine reconstruction result. However, if only a leaf is the only object we capture a video of, the measurement of it alone would not be able to tell its size in the real-world unit, therefore, a paper with the chessboard pattern with 1 by 1 centimetre for each block, refer as a scale reference, is put beside the leaf before a video capturing to involve this scale reference in the reconstructed point cloud for further unit converting propose. Please note that the scale reference does not have to be in the chessboard pattern, it could be in any shape but it must contain a specific length we want to use as a reference for the real-world scale, e.g. 5 millimetres.

As described earlier, the 3D reconstruction to obtain a point cloud is performed by the SfM and MVS. While a number of open and free software are available with a variety of algorithms to be used, the authors chose COLMAP [8, 9] as it has a very clean and easy to use interface and also can customize in many ways. COLMAP can perform a sparse reconstruction by SfM and dense reconstruction by MVS. It also can create a mesh of an object with the triangulation of either Poisson or Delauney but only the dense reconstruction from MVS in ply file is enough for our study. Although COLMAP is selected for this work, any other approaches or software that can generate the fine enough point cloud, such as CMVS [10] or using of laser scanner or LiDAR, can also applicable instead. The using of COLMAP essentially gives a point cloud of a specific leaf the same as the one that has been taken photos of.

2.2. Leaf vein extraction and leaf width measurement

After the obtaining of a point cloud, the process of width measurement can be easily done by selecting coordinates of the leaf edge and the right edge of a leaf and then find distance between each two points. However, rice leaves usually have a slight curvature and directly find distant from left to right could lead to inaccurate result. Instead, we find a width between each leaf vein as they are aligned along the curvature of a leaf and such measuring could deal with this problem.

The next problem is how to distinguish between the leaf surface and leaf vein points in point cloud. Basically, it can be distinguished by the colour where the vein is in slightly brighter green colour than the surface. As a result, we extract the vein points by performing the thresholding using the RGB information. But as sometimes, thresholding on the RGB information alone cannot fulfil its purpose, the additionally use of the HSV information can help the system distinguish the vein from a leaf more accurate as HSV can tell the brightness and vividness of a specific colour better than RGB.

The threshold value of either RGB or HSV can simply be the average value. However, if the whole point cloud has been used in averaging, the thresholding would not be very efficient since the lighting of the whole leaf can be different in some part. Thus, the limiting the number of points to be only around the interested part of the leaf. To do so, the proposed system took the coordinates of 2 points of leaf edges, one from left and one from right, in the part we want to measure and make the linear equation \( L \) out of them. The next step is to assign the number of total points \( n \) which we want to limit to and find \( n \) point that are closest to the \( L \) that has been calculated earlier. The thresholding will be performed only on these \( n \) points for both RGB and HSV and, obliviously, the results from these two are not necessary to be the same. If RGB and HSV thresholding are both state the same single point is the leaf vein, we label the point as the leaf vein. On the other hand, If the two distinguish points are selected by each thresholding, we will not label any of them as the leaf vein. But the problem is sometimes these two distinguish point from RGB thresholding \( p_1 \) and from HSV thresholding \( p_2 \) are very close to each other, it would not be fair to label them as a leaf surface. This can be solved by measuring the average distance from \( p_1 \) to the 10 closest point and compare it to the distance between
If the distance between $p_1$ to $p_2$ is shorter, we can assure that these two points are close to each other and then can be safely labelled as the leaf vein. After the thresholding, the vein point can be obtained and now the distance between each leaf vein can be measured. But as there could be several points in a single leaf vein, we decided to use the centroid of each individual vein in the measurement. At this stage, the measured width will not be in the metric unit and thus cannot be understandable the true dimension. To tackle this problem, the system measures the width of one block of the scale reference, which each one has a length of one centimetre, and compare it with the obtained veins distance to convert it into the centimetre unit. By this approach, the system can finally generate the true dimension of a leaf in centimetre that can be understandable and validated with the real plant.

3. Experimental result

In this paper, the experiment for validating our approach was conducted on a fake rice tree model made of paper, shown in figure 2(a). The model has been created with metal wires for stems and green papers for leaves. The wires were wrapped with the green paper tape and the paper leaves were wrote with green ink for making the leaf veins.

![Figure 2. (a) the paper rice tree model, (b) the marked area](image)

There are several reasons why the authors conducted an experiment on paper model. First, the verifying the result would be a lot easier than doing with the real rice plant since it is not growable, which mean we do not have to concern of the growing of the leaf in the meantime the experiment is being conducted and validating afterward could be inaccurate. Second, the setting up of an experiment is much easier than the real rice as it will not break apart easily or get affected by external environment.

In this experiment, the authors marked a faint line across the specific leaf where the measurement is going to be made on with a pencil, shown is figure 2(b). The leaf was recorded along with the scale reference in a form of chessboard pattern flag with an iPhone 7 camera in 4k 30fps for 15 seconds. From the video, we extract it into 159 pictures, each size 3840 x 2160 pixels, in a form of video frame and use them in the 3D reconstruction with COLMAP. The reconstructed point cloud contains 308,861 points as shown in figure 3.
After obtaining the point cloud, two points were selected whereas located in the left and right leaf edge on the marked area and created the linear equation from these points. Then, the 1,500 nearest points to the line were selected and used in thresholding. The thresholding results are displayed as the red points in figure 4. The measured distances are then calculated and compared with the results from Easy Leaf Area (ELA) [2] which has a similar idea but executed in 2D top view image.

The pictures we used for measuring with ELA contains three pictures: the direct top view picture, a picture taken with a tilted camera angle, and a picture where the leaf is not on the same level as the scale reference. These pictures are shown in Figure 5.
The leaf model in this experiment contains 9 veins as shown in figure 6. The measuring result between each vein of each picture comparing with the measuring in point cloud is shown in table 1.

![Figure 6. The number of each vein](image)

Table 1. The measuring result between each vein for each experiment in millimetres

| Distance between veins (mm) | 1st to 2nd | 2nd to 3rd | 3rd to 4th | 4th to 5th | 5th to 6th | 6th to 7th | 7th to 8th | 8th to 9th |
|-----------------------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ground truth               | 2.05       | 1.03       | 1.25       | 0.86       | 1.83       | 0.86       | 1.45       | 0.72       |
| 2D: Top view               | 1.96       | 0.89       | 1.19       | 0.72       | 1.66       | 0.78       | 1.27       | 0.67       |
| 2D: Tilt angle             | 1.59       | 0.68       | 0.93       | 0.66       | 1.30       | 0.68       | 1.05       | 0.54       |
| 2D: Different level        | 3.10       | 1.37       | 1.82       | 1.20       | 2.61       | 1.11       | 2.04       | 0.97       |
| 3D: Point cloud            | 2.17       | 1.04       | 1.19       | 0.90       | 1.82       | 0.78       | 1.45       | 0.81       |

4. Discussion
The accuracy of Easy Leaf Area is highly depending on the camera angle and the level between the scale reference and the associated leaf as evidenced in the experiments where the camera angle is tilted and when the scale reference is place lower than the leaf. On the other hand, 3D space allows the independency of camera angle and makes the accuracy of our approach the highest among all experiments. The root mean square error (RMSE) for each experiment is shown in table 2.

Table 2. Root mean square error for each experiment

| Experiment          | 2D: Top view | 2D: Tilt angle | 2D: Different level | 3D: Point cloud |
|---------------------|--------------|----------------|---------------------|-----------------|
| RMSE                | 0.0115       | 0.0326         | 0.0518              | 0.0051          |

For the 2D measured distance between leaf veins, direct top view image is the most accurate among 2D image. The curvature of the leaf makes the result slightly incorrect. For the tilt angle image, the tilt angle causes the reference pattern to be more toward the camera while the leaf is further and affect the conversion of unit. In different level image, because of the different level between the leaf and the pattern makes the leaf look bigger, which leads to the larger measurement result than a ground truth.

For 3D point cloud measurement, there are slight errors at the distance between veins which may caused by the thresholding result. As can be seen in figure 4, all the red points are not perfectly aligned in the form of linear line while the measuring distance between veins at the linear line is the way the ground truth is created. Despite the error, the results indicate that this approach is still more accurate than using only 2D images.

Although the cell phone camera, specifically iPhone 7 with 12-megapixel resolution camera, is used in this study, there is a possibility that the result could get more accurate when using the professional camera instead of a cell phone camera since the image resolution is higher which can lead to finer surfaces of the reconstructed point cloud. However, as the number of people with smart phone is much
higher than the people who have professional camera, the authors decided to use the cell phone camera as it makes the system more compact and can be easily used by any person.

5. Conclusion
The system uses 3D reconstruction techniques to measure the distance between leaf veins by taking a set of 2D images of a rice leaf with SfM and MVS to generate 3D point cloud. From the point cloud, we extract venation with RGB and HSV thresholding, and lastly, converting the measured unit into millimetres with the help of a scale reference. Our system successfully generates the 3D point cloud and estimate the width between each vein with higher accuracy than the result from using 2D images. As the method for the leaf veins classification in point clouds based on a thresholding result introduced in this paper still has a room for improvement in term of the accuracy, the authors plan to propose a better way to selecting the leaf veins without heavily depending on the colour information in future.

6. References
[1] Maloof J N, Nozue K, Mumbach M R and Palmer C M 2013 LeafJ: an ImageJ plugin for semi-automated leaf shape measurement. JoVE (Journal of Visualized Experiments) (71), e50028.
[2] Easlon H M and Bloom A J 2014. Easy Leaf Area: Automated digital image analysis for rapid and accurate measurement of leaf area. Applications in plant sciences 2(7), apps.1400033. https://doi.org/10.3732/appps.1400033
[3] Vázquez-Arellano M, Reiser D, Paraforos D S, Garrido-Izard M, Burce M E C and Griepentrog H W 2018 3-D reconstruction of maize plants using a time-of-flight camera. Computers and Electronics in Agriculture 145 pp 235-247.
[4] Weiss U and Biber P 2011 Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. Robotics and autonomous systems, 59(5), pp 265-273.
[5] Seidel D, Beyer F, Hertel D, Fleck S and Leuschner C 2011 3D-laser scanning: A non-destructive method for studying above-ground biomass and growth of juvenile trees. Agricultural and Forest Meteorology, 151(10), pp 1305-1311.
[6] Ullman S 1979 The interpretation of structure from motion. Proceedings of the Royal Society of London. Series B. Biological Sciences, 203(1153), pp 405-426.
[7] Bailer C, Finckh M and Lensch H P A 2012 Scale Robust Multi View Stereo. Computer Vision – ECCV 2012. Lecture Notes in Computer Science, vol 7574
[8] Schönberger J L, Zheng E, Frahm J M and Pollefeys M 2016 Pixelwise view selection for unstructured multi-view stereo. In: European Conference on Computer Vision. Springer, Cham, pp 501-518.
[9] Schönberger J L and Frahm J M 2016 Structure-from-motion revisited. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 4104-4113.
[10] Furukawa Y, Curless B, Seitz S M and Szeliski R 2010Towards internet-scale multi-view stereo. In: 2010 IEEE computer society conference on computer vision and pattern recognition. pp 1434-1441.

Acknowledgments
We would like to thank you Hiroki Yano for helping us beautifully create the rice leaf model and assisting in numerous experiments.