A Comparison of Three-Dimensional Data Acquisition Methods for Phenotyping Applications

Scholz, O. & Uhrmann, F. & Gerth, S. & Pieger, K. & Claußen, J.

Fraunhofer EZRT - Development Center for X-Ray Technology, Fürth, Germany

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

Abstract. *Currently Phenotyping is primarily performed using two-dimensional imaging techniques. While this yields interesting data about a plant, a lot of information is lost using regular cameras. Since a plant is three-dimensional, the use of dedicated 3D-imaging sensors provides a much more complete insight into the phenotype of the plant. Different methods for 3D-data acquisition are available, each with their inherent advantages and disadvantages. These have to be addressed depending on the particular application.*

In this paper we demonstrate a number of representative methods each with a distinct set of features which make them suitable for particular applications. Each method is presented along with example setups. Simple guidelines are shown to help the researcher select the best technology for a given application. Sample data is presented gained from various real life applications to help understand the data quality that can be expected using a particular method.

Keywords. Phenotyping, 3D-data, point cloud, sheet-of-light, triangulation, voxel, time-of-flight, CT, Helical-CT, Photogrammetry.

Motivation

Phenotyping can be defined as the determination of the observable physical or biochemical characteristics of a plant, as determined by both genetic makeup and environmental influences. Today most systems designed to automatically measure traits of a plant do so using two-dimensional optical approaches, for instance by taking one or more visible light or hyperspectral images of the plant to determine key parameters. One common task for example is the determination of the leaf area of a plant from a 2D color image employing image processing to separate the plant from the background and then counting the plant's pixel area, sometimes classifying pixels as leaf or non-leaf pixels before the final determination of the total pixel count. The resulting number can be a useful indicator when comparing it to the algorithm's results of similar plants, but it usually does not correspond with the actual leaf area because the two-dimensional projection of a leaf does not take into account the angle between leaf and camera for example. Even if the leaves are relatively flat and the camera is perpendicular in relation to the leaves there is an error and the resulting leaf area is not accurate.

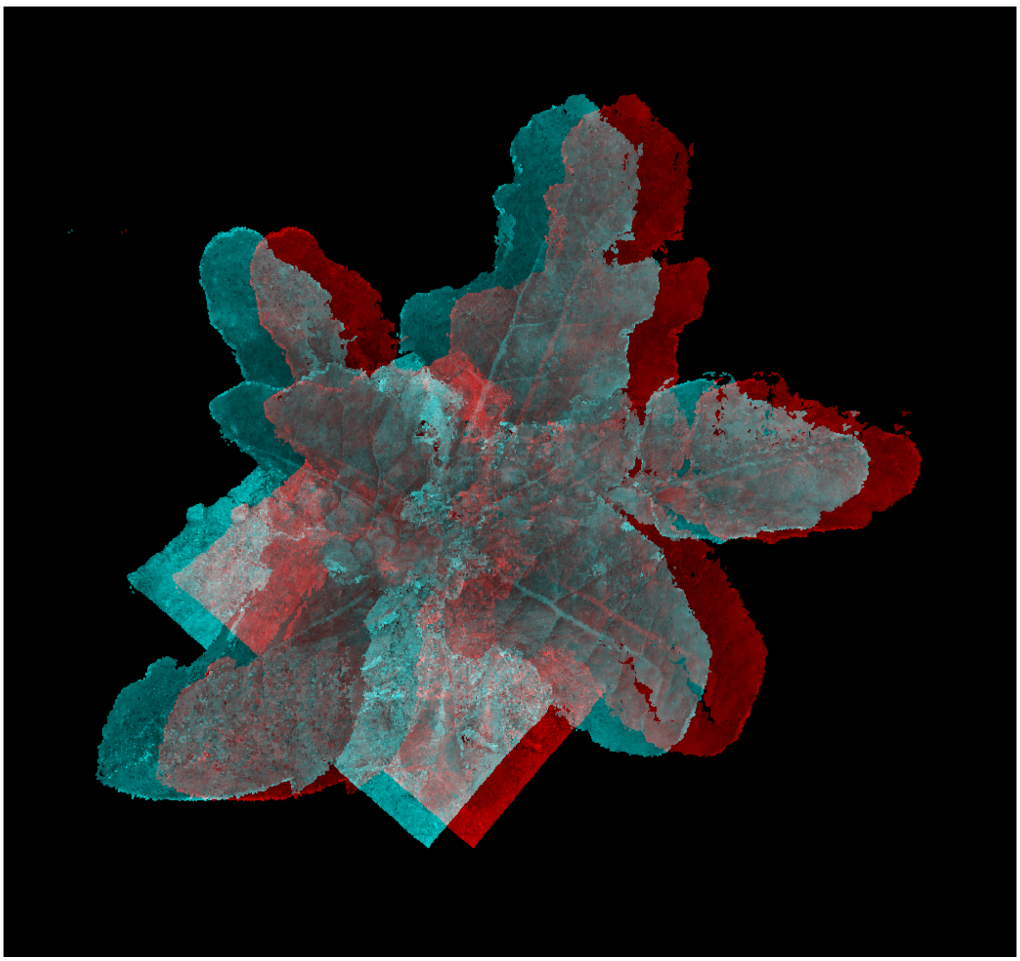

Fig 1: Anaglyph of sugar beet plant point cloud

A physical plant is of course three-dimensional (Fig. 1), so in order to gain a more accurate digital representation of a plant, three-dimensional image acquisition methods must be employed. Having accurate three-dimensional data is a prerequisite to create algorithms that can take into account the three-dimensional features of the plant allowing measurements of leaf angles, curvatures, and much more. The goal must be to acquire the plant with a level of detail that captures the relevant traits of the plant and transfers them to the digital domain for evaluation and possibly create new insights into the phenotype of the plant.

3D Data Acquisition Basics

Many different methods are available that generate 3D data. They can be divided into two main categories depending on the kind of data they generate:

- Point cloud generating methods
- Voxel generating methods

Point cloud generating methods produce a set of 3D-coordinates with a theoretically unlimited resolution. Another advantage of point clouds is that all data points are actually part of the measured scene. However, all points within a point cloud are inherently independent, so algorithms need to determine which points are actually neighbors and which ones are not.

Voxel data is based on the principle of dividing the measured volume into small cubes of (usually) equal width, height and depth called voxels (volume elements). The resulting voxel data file contains voxels for the entire measured volume, no matter if a voxel encodes an actual part of interest or empty space, although compression methods are available to increase entropy. Another disadvantage of voxel data is the quantization of space, i.e. the measured volume is quantized by the voxel grid regardless of the actual resolution of a scan.

A variation of the voxel data is a height image, which is actually a 2D image (employing a pixel grid) with each image pixel's value designating the distance of the pixel from a reference area. It is commonly used by surface scanning methods. It is quite memory efficient, but also quantizes the scanned space into pixels.

Each 3D data acquisition method has its unique advantages and disadvantages, and the selection of a particular method depends on the application and requires the consideration of a number of other parameters, the measurement situation and so on. In the following sections a number of 3D phenotyping methods are presented that were implemented or examined at the Fraunhofer Institute for Integrated Circuits in Germany. For each method, the acquisition principle is briefly explained and the setup and sample data are presented and discussed. Only systems generating high-resolution data for single plants preserving as much detail as possible are presented, not field scale systems providing only a low single plant resolution.

Optical Methods

Using optical methods for phenotyping is obvious, as digital cameras are readily available and inexpensive. In addition to conventional color or monochromatic digital cameras, infrared or hyperspectral cameras are available. Furthermore highly specialized cameras like e.g. polarization cameras providing polarization information for each pixel in addition to the pixel's luminance value exist today.

Stereoscopy

One of the most obvious methods is the use of two cameras to gain depth information using so-called photogrammetry. Its basic implementation uses two cameras capturing the scene at different angles. An algorithm then locates features from one image in the other image and based on the known camera parameters calculates the depth information for each feature. But correspondences cannot be calculated for every pixel, so depth information is sparse.

While this method generates 3D data of the plant, the coverage of the plant is limited: each feature must be captured by *both* cameras for the algorithm to be able to calculate the feature's depth information. In return, it is impossible to measure features visible by only one camera and obviously the rear of the plant is not accessible at all. But this method can be extended to more than two cameras to gain better coverage and more feature correspondences between the different cameras.

Fig.2: Schematic view (left), real world setup (right) of 3-camera sugar beet plant scanner

Fig. 2 shows a multi-view stereoscopic setup using three cameras for use in a greenhouse. A similar setup was used in the field for sugar beets. Using three cameras is a good compromise between coverage and cost, offering two sets of stereo image pairs for reconstruction.

An advantage of this method is the direct generation of high-resolution color 3D images. If enough light is available (e.g. in the field) the exposure times can be often shortened enough to limit wind induced motion blur to an acceptable level.

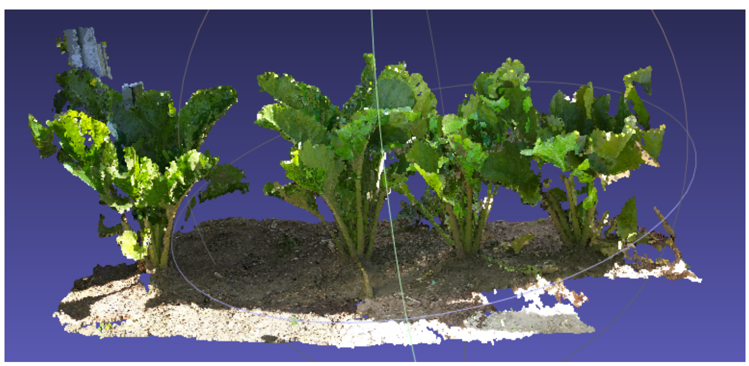

Fig 3: Color 3D data generated from 3-camera setup

Fig. 3 shows the color point clouds generated with the three-camera setup from Fig. 2 in the field. The resolution of this particular setup is approx. 1 mm, resulting in a good level of detail in the 3D data set. Moving the camera rig across the field allowed the quick acquisition of rows of plants in the field.

Sheet-of-Light

Another method to gain 3D information of a plant is the so-called "Sheet-of-Light" (SOL) method. It is based on a laser projecting a line on the object to be scanned (see Fig. 4).

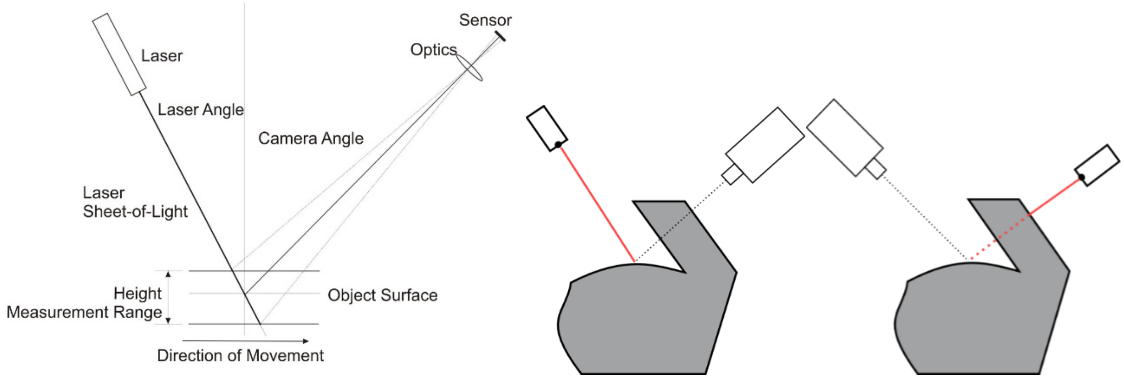

Fig. 4: (left) SOL principle of operation, (right) shadowing/occlusion

A camera captures the diffuse reflection of the line and the height profile of this line is calculated from the known measurement setup. The 3D surface of the object can then be reconstructed from the calculated profiles when moving object and sensor relative to one another generating profiles as fast and thus densely as possible.

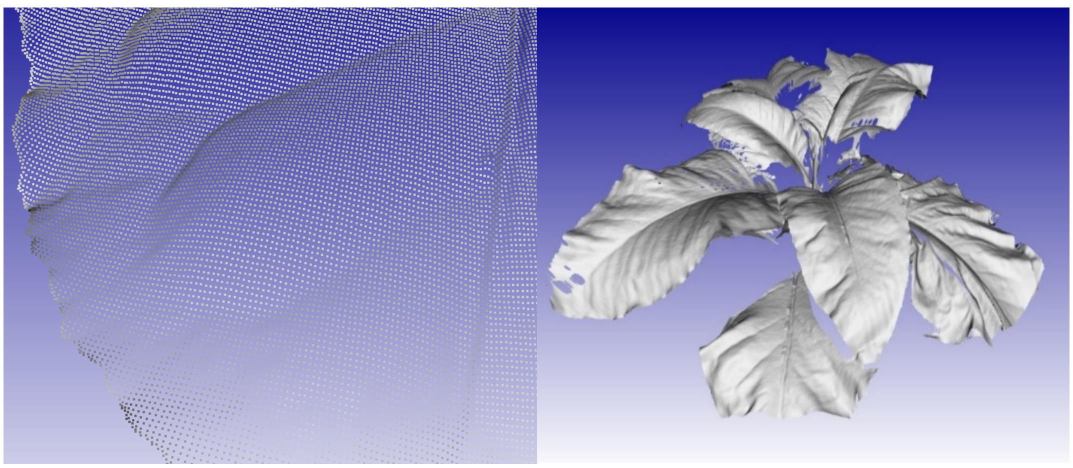

Fig. 5: SOL-scanned tobacco plant geometric data, no color (left: detail, right: overview)

This method generates precise and dense 3D information very quickly (see Fig. 5), but also has its own drawbacks. For one, it is possible for object parts to be inaccessible because they are either not illuminated (shadowing) or not visible to the camera (occlusion). Those regions will not be included in the resulting 3D dataset. However, since the type of object to be scanned is normally known in advance, these effects can be taken into account when designing the acquisition setup. Because this measurement technique only generates geometric surface data an additional color and/or hyperspectral camera can provide color information if desired (see Fig. 6). The setup must be calibrated to precisely map the color and 3D data sets into a common coordinate system. In addition, since SOL and color cameras have different angles of view, not all geometric data points have corresponding color data and vice versa.

Sometimes commercial video cameras are used in SOL systems, but this limits the number of profiles to be scanned to 25/30 per second, effectively limiting the resolution/throughput of such a system. For example, if a 1 mm resolution in forward direction is desired, the maximum speed may not exceed 25/30 mm/s. In contrast, commercially available SOL cameras provide 2000 profiles per second and more, providing high resolution and high throughput at the same time.

Because this method generates a sequence of 3D profiles while scanning an object in motion relative to it, it is ideally suited for scanning plants e.g. on a conveyor belt with high throughput.

Fig 6: SOL geometric data with color data mapped

Structured Light Measurement

Structured light method employs projection of a sequence of stripe patterns of varying frequencies/phases on the scene. One or more cameras capture the scene and an algorithm calculates the 3D information from this set of images (see Fig. 7).

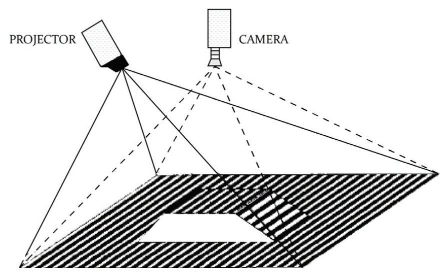

Fig 7: Structured Light projection principle (image credits: © 2016 Sam Van der Jeught et al.)

The main advantage of this method is that no geometric correspondences are needed, only one camera is required and the method generates color data without the need for calibration to map color to 3D data and the depth information generated is superior to stereoscopy data.

Fig. 8: Using structured light projection to measure tobacco plants

While this method does not require movement, the plant must remain still during measurement, because although no correspondences are required, the method is based on the prerequisite that each corresponding pixel in the sequence of images views the same part of the scene. The method is also sensitive to ambient light, because the ambient light will interfere with the projected patterns. Because of these two limitations, this method is not well suited for field use where wind movements and ambient lights quickly make measurements impossible. However, it is a very good method for lab use.

Like all optical methods, this method only covers the part of the scene captured by the camera, so in order to cover the plant completely it may be indicated to use multiple projectors/cameras to cover the plant from all sides. Fig. 8 shows a measuring setup using this method for scanning tobacco plants with a measuring volume of one cubic meter at a resolution of about 1 mm at a scanning time of 15 seconds per plant. Four sets of cameras/projectors were used in this setup to cover the tobacco plant from all sides.

Time of Flight Measurement

The final optical measurement method presented here is the so-called “Time-of-Flight” (TOF) method. The measurement principle is based on actively illuminating the scene and measuring the time, the light pulse takes to travel to each pixel. In effect, a TOF sensor not only generates luminosity information of the scene but also depth information directly. Since the processing takes place in the camera’s hardware, this method generates 3D data in real-time without additional processing. However, accurately evaluating time and phase is difficult, to achieve 1 mm depth resolution time must be measured with one femto second accuracy. The result is very noisy image data with low depth and spatial resolution as seen in Fig. 9.

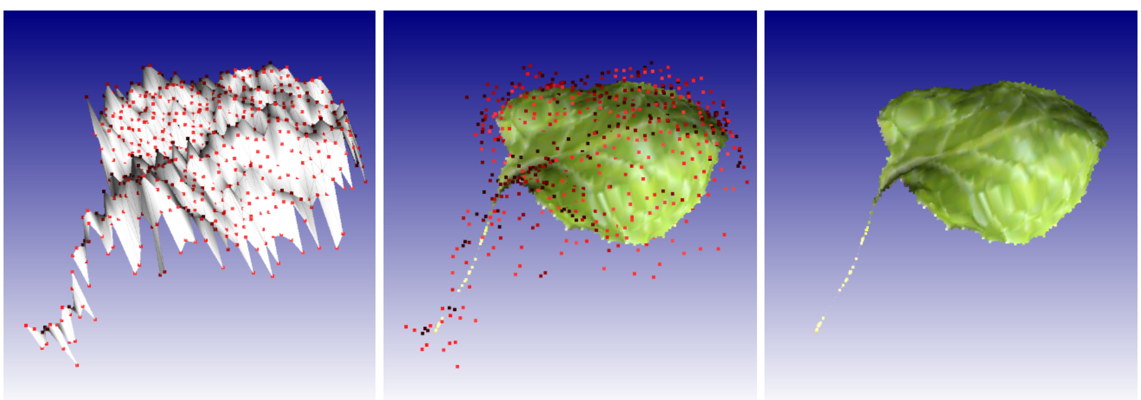

Fig. 9: TOF leaf measurement (left: noisy data, center/right: model leaf fit to data points)

Without further processing, the data is not usable, but using a model approach and using the noisy data points as input to the model optimization algorithm it was possible to recreate data similar to the leaf used in the data acquisition. However, this implies the application of prior knowledge, which is often not available and assumes properties of the object that may or may not be present and may lead to generation of false data.

It was also observed that the TOF cameras are sensitive to surface properties of the scenery. For instance, a test object with a black and white flat surface was captured with very different depth values for white and black surface points respectively where in reality all surface points should have been at the same depth.

Finally, no color TOF sensors are available yet, so if color data is required, the TOF sensor must be combined with a color camera and calibration between the cameras must be implemented.

X-Ray Methods

If parts of the plant are obscured by the plant itself, if inner structures need to be visualized or if the area of interest is otherwise visually not accessible (like the root of a potted plant) X-ray methods are well suited to provide three-dimensional data. While one of the most expensive methods, the voxel data generated by such a system provides information about the entire scanned volume, including absorption data for each voxel.

Computed Tomography

The main method of generating 3D voxel data of a plant is naturally Computed Tomography (CT). While CT is an established method for medical diagnostics, it is increasingly used in phenotyping.

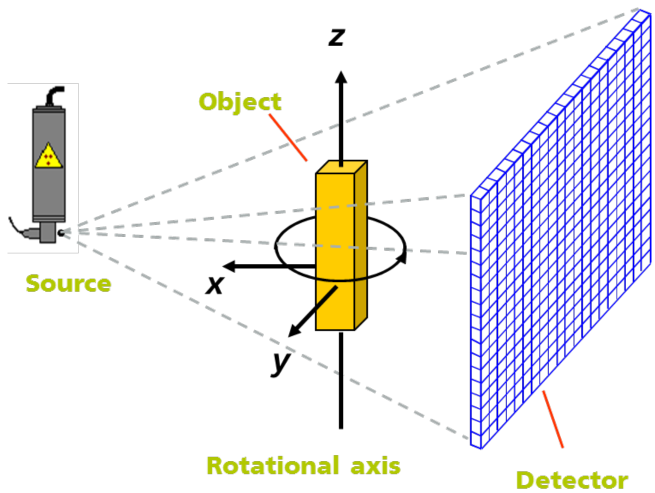

Fig. 10: CT principle of operation

The operating principle (see Fig. 10) is based on taking a series of 2D-X-ray images of the object while rotating the object. From these 2D images, a 3D volume can be reconstructed algorithmically. While other X-ray based methods exist e.g. for 3D acquisition of flat objects conventional CT is best suited for phenotyping purposes. It reveals all internal structures and does not suffer from occlusion effects, although certain materials can produce artifacts.

Fig. 11: Measurement system for automated root analysis

Fig. 11 shows a measuring system for automated root analysis, capturing 800-2400 individual 2D images of a plant resulting in a reconstructed volume of typically 2000x2000x2000 voxels. The time required for one measurement varies from minutes to hours, so this method is neither low-cost nor high-throughput, but it reveals information not available by any other means. Fig. 12 shows potato tubers of potted plants in soil with the soil algorithmically removed.

Naturally, CT does not provide color information either, but again if color information is desired, a color camera can be added and with a suitable calibration, color and 3D data can be combined in a common coordinate system.

Because of the size and nature of an X-Ray CT system, it is not suitable for use in the field, but well suited for lab use for e.g. potted plants.

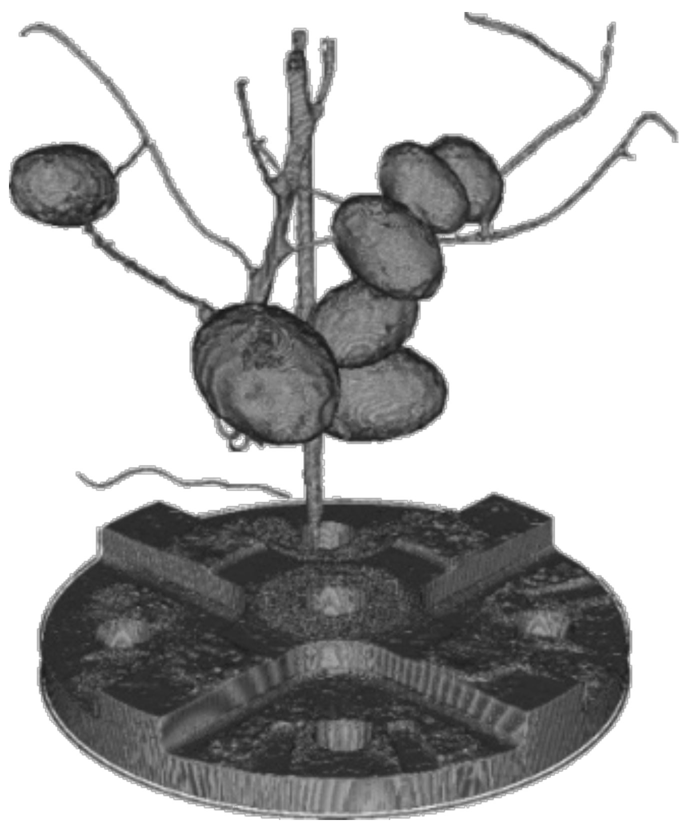

Fig. 12: 3D voxel data of potato tubers

Other Methods

Many other 3D acquisition methods are available, but they are either not suitable for generating plant level with sufficient detail for phenotyping or have another sort of serious drawback:

- MRT (Magnetic Resonance Tomography) – Expensive and high effort
- LIDAR (Light Detection and Ranging) – Similar to TOF, only for a single point, with all disadvantages of TOF
- (Ground penetrating) Radar – Detects e.g. water in soil, but level of detail not suitable for geometric plant analysis

- Ultrasonic methods – Transducer must be in physical contact with object being tested, only localized data
- Impedance Tomography – Very low resolution

Basically none of these methods provide the required level of detail or the cost is prohibitive.

Simple Guidelines

Although many parameters are to consider, a few rules of thumb can narrow the search for the best acquisition method for a given application. First, it must be determined if hidden structures are required or if otherwise inaccessible structures like roots need to be acquired. If that is the case, X-ray methods – especially CT – should be considered if applicable.

If optical methods are sufficient, the first criterion is the choice of indoor or field use. For field use, if wind is an issue (multi-)stereoscopic methods offer acceptable color data even under adverse conditions. If wind is not an issue, SOL offers superior 3D quality and can be augmented with color data if desired.

For indoor use all methods are applicable, so the capabilities and limitations of each method must be taken into account when deciding for or against a method.

Summary and Conclusion

Acquiring a plant three-dimensionally instead of two-dimensionally is naturally capable of creating a much more detailed digital representation of a plant thus providing a basis for a wide range of analyses not possible with two-dimensional data. A number of methods are available delivering 3D data of varying quality, each with distinct advantages and disadvantages.

The acquisition method must be carefully selected and optimized depending on the situation and possibly other parameters in order to gain the best possible data quality for subsequent algorithms and analyses. Especially the type of application, i.e. indoor or field use as well as factors such as wind, ambient light but also required resolution/point density or data type (point cloud/voxel volume) must be considered when choosing an acquisition method.

Acknowledgements

Some of the work of this paper is based on the “For3D” research project in cooperation with Strube D&S, partially funded by the “Bayerische Forschungsstiftung” (BFS). We would also like to thank Dr. Peter Schmitt for his support in creating this contribution to the 14^{th} ICPA.