Publisher: CSIRO; Journal: FP:Functional Plant Biology
Article Type: research-article; Volume: ; Issue: ; Article ID: FP16167
DOI: 10.1071/FP16167; TOC Head:

Review

Approaches to three-dimensional reconstruction of plant shoot topology and geometry

Jonathon A. Gibbs[A,C], *Michael Pound*[A], *Andrew P. French*[A], *Darren M. Wells*[B], *Erik Murchie*[B] *and Tony Pridmore*[A]

[A]School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK.

[B]School of Biosciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, Leicestershire, LE12 5RD, UK.

[C]Corresponding author. Email: psxjg6@nottingham.ac.uk

There are currently 805 million people classified as chronically undernourished, and yet the World's population is still increasing. At the same time, global warming is causing more frequent and severe flooding and drought, thus destroying crops and reducing the amount of land available for agriculture. Recent studies show that without crop climate adaption, crop productivity will deteriorate. With access to 3D models of real plants it is possible to acquire detailed morphological and gross developmental data that can be used to study their ecophysiology, leading to an increase in crop yield and stability across hostile and changing environments. Here we review approaches to the reconstruction of 3D models of plant shoots from image data, consider current applications in plant and crop science, and identify remaining challenges. We conclude that although phenotyping is receiving an increasing amount of attention – particularly from computer vision researchers – and numerous vision approaches have been proposed, it still remains a highly interactive process. An automated system capable of producing 3D models of plants would significantly aid phenotyping practice, increasing accuracy and repeatability of measurements.

Additional keywords: image-based, plant modelling, reconstruction, three-dimensional.

J. A. Gibbs *et al*.

Reconstruction of plant shoot topology and geometry

The need for increased crop yields is becoming urgent as the amount of arable land available is reduced and environmental factors worsen, however, plant phenotyping has been identified as a key bottleneck in the process of improving crop yields. Here we review approaches to 3D shoot reconstruction to improve phenotyping using image-based methods. An automated system capable of producing 3D models of plants would significantly aid phenotyping practice, increase accuracy and repeatability of measurements and potentially aid the process of improved crop yields.

Introduction

Understanding the mechanisms underlying the growth of agriculturally important plant species is becoming increasingly critical to society, particularly as the quantity of food produced must double by 2050 if it is to meet the demands of the expanding global population, which is likely to exceed nine billion (Sticklen 2007; Faaij 2008; Paproki *et al.* 2012). The Food and Agriculture Organisation of the United Nations (FAO) already considers 805 million, or one in nine people 'chronically undernourished'. Moreover, population growth is not the sole contributor towards an increasing demand for food: the spread of prosperity throughout the world, predominantly in developing countries such as India and China, is increasing food intake per capita and driving demand for a richer, more varied diet (Kearney 2010; Bonhommeau *et al.* 2013). Consequently, increasing pressure is being placed on agriculture to improve crop yields (Sutton *et al.* 2011).

During the decades following the 'Green Revolution' (Evenson and Gollin 2003), annual improvements in crop yield were typically 2–5% (Gaud 1968). However, over the past two decades this has plateaued at around 1%, leading to concerns that some fundamental limit may have been reached (Khush 1996). The severity of the situation is such that rice demand recently exceeded supply for 2 years (2009–11), and world stocks of grains are now the lowest they have been for 45 years (Furbank *et al.* 2009; Furbank and Tester 2011).

Changes in climate and the shortage of arable land constitute further challenges for sustainable agriculture, as global warming has been shown to cause more frequent and severe flooding and drought, which destroy crops (Adeloye 2010). Recent work has shown that without crop climate adaption, crop productivity will actually deteriorate (Tester and Langridge 2010; Challinor *et al.* 2014). It is clear that a new approach to a sustainable increase in crop yield is necessary (Furbank and Tester 2011).

In the face of these challenges, an understanding of the relationship between genotype and environment on plant phenotype is invaluable to the agricultural community. An improved understanding of phenotypes would aid breeding and inform genetic modification, facilitating increased nutrient use and photosynthetic efficiency and thereby increasing crop yield and stability across hostile and changing environments (Quan *et al.* 2006). This would significantly alleviate a majority of problems defined by the FAO and help lift farmers out of poverty by generating additional income. In addition to pre-breeding applications, phenotyping currently constitutes a major bottleneck in basic research, particularly in the construction of quantitative models of plant development (Preuksakarn *et al.* 2010).

Phenotyping methods and technologies have attracted significant and rapidly increasing attention in recent years. Major phenotyping projects are now underway across Europe, Australia, Canada and the United States of America. Emphasis is being placed on fully-automatic, high-resolution, high-throughput, quantitative measurement of plant structure and function. Techniques have been proposed for the quantification of a wide range of properties of roots, shoots, leaves and seeds.

A majority of these methods are image-based (Fahlgren *et al*. 2015), relying on the automatic extraction of traits from, usually, colour images (Lobet *et al*. 2013). Simple analysis of colour can be important when examining plant response to biotic and abiotic stresses. When structural traits are needed, images are typically segmented to identify plant components, or key features identified, before measurements are made. These measurements are expressed on the (2D) image plane in pixel units. Conversion to real-world dimensions (e.g. mm) requires some pre-calibration of the image acquisition equipment, and a final pixel-to-mm conversion step. If angular measures are to be made, the camera must be arranged to ensure that angles measured in the image plane reflect the real-world angle of interest. It is common to find that the set of measurements obtainable from this type of system is determined by the relative placement of sample and camera.

The reconstruction of 3D models of the viewed plant provides an alternative approach. In this method, measurements are made across a representation of the 3D shape of the target object that is first reconstructed from sensor data rather than in the image plane. Assuming that a sufficiently accurate and detailed model can be created, a wide variety of traits can be computed. More importantly, if new traits are required at a later date they are likely to be computable from the same model. In the 2D, image-centred approach, some traits may not be recoverable from the available image(s). The features required may not be visible, or the calibration information needed to make real-world measurements might not have been recorded.

Access to 3D models that capture morphological and developmental data is also significant in the use of simulation approaches to study the ecophysiology of plants (Larcher 2003): for example, the modelling of photosynthesis. It is unclear whether plant species have an optimal arrangement for photosynthesis, and further studies using accurate plant representations need to be conducted to determine this (Pound *et al*. 2014). Detailed 3D representations of real plants allow numerous simulations, e.g. ray-tracing techniques to simulate illumination conditions, within a range of artificial canopies (Burgess *et al*. 2015).

It is clear that 3D models have the potential to provide the continued refinement of plant phenotyping methods required to quantify plant growth, development, tolerance and physiology. The cost associated with the 3D model-based approach is, however, that an appropriate reconstruction method is required.

In this review we appraise available approaches to the reconstruction of plant shoot topology and geometry from image data, reviewing their actual and potential contribution to

Publisher: CSIRO; Journal: FP:Functional Plant Biology
Article Type: research-article; Volume: ; Issue: ; Article ID: FP16167
DOI: 10.1071/FP16167; TOC Head:

the construction of accurate 3D models. The remainder of the paper is organised as follows: we begin by introducing the reader to 3D modelling in general, providing an overview of the various approaches before providing a more in-depth review of image-based modelling approaches; then we discuss how these have been applied to plants, and the challenges and opportunities facing plant modelling before adding our concluding remarks.

Background: three-dimensional modelling and plants

Three dimensional (3D) modelling has been applied to a wide range of scenarios from medical usage, creating a 3D representation of a brain using magnetic resonance imaging (MRI) (Lauterbur 1973), for example, to the creation of environments for films and animations. 3D models are ubiquitous, and becoming increasingly prevalent as modern, low-cost machines and sensors now have the capability to capture and render them.

Many 3D reconstruction methods focus on objects with relatively simple structures; those lacking occlusions and specularities but containing textured areas, or manmade objects with easily identifiable symmetry or shapes (Furukawa and Ponce 2010). Plants, however, are complex and challenging objects to model and, until the late 1960s, botanical drawings were the primary means of representing plant architecture. Today, with the use of high performance computers and the availability of portable cameras and sensors, many approaches exist, from those relying on depth data obtained by lasers to those drawn free-hand.

Approaches to model plant architecture typically fall into two categories, known as rule- and image-based approaches. Rule-based methods capture knowledge of plant structure and form in a set of user-defined rules, which can then be applied to generate example models consistent with that knowledge. There are many approaches to rule based modelling such as L-Systems (Lindenmayer 1968; Prusinkiewicz *et al*. 2000; Karwowski and Prusinkiewicz 2003; Prusinkiewicz 2003; Ole and Winfried 2008; Boudon *et al*. 2012), Relational Growth Grammars (Kurth 2007) and AMAP (de Reffye *et al*. 1988), which have been applied to a variety of problems (Lintermann and Deussen 1996; Deussen and Lintermann 1997; Shlyakhter *et al*. 2001; Boudon *et al*. 2003, 2012).

Rule-based methods are used to simulate plant growth, creating synthetic plant structures. These are exemplars of the class of plant simulated, but do not necessarily capture the detailed structure of any existing, real plant. They are, however, highly valuable as the basis of functional structural plant models (FSPMs). FSPMs are used to study the ecophysiology, how plants sense and respond to environmental change, of a plant by combining the 3D, structural representation with a model of some physiological function (Vos *et al*. 2010).

In contrast, image-based methods use real-world data to develop detailed 3D models of real plants, often relying on techniques developed by the computer vision community. These models can be used to support both simulations of plant function and the extraction of the trait measurements required for phenotyping. Although image-based modelling has made significant progress towards achieving photorealism, that is constructing a model as realistically as possible, over the past decade, creating accurate representations remains a research problem. This is, in part, due to the complexity of the plants and the environments they inhabit, and also the lack of a single definition of image-based modelling (McMillan and Bishop 1995): multiple approaches to the problem have been proposed, each with its own strengths and weaknesses. Fig. 1 provides an overview of current approaches, along with an indication of their current range of application in plant modelling.

Plant architecture, as defined by Godin (2000), is difficult to model due to the dynamic behaviour of plants, from short-term changes such as the reorganisation of foliage to long-term growth patterns, and intricate phyllotaxis (Ivanov *et al*. 1995; Tan *et al*. 2003; Reche-Martinez *et al*. 2004; Zeng *et al*. 2006; Kang and Quan 2009). A plant may consist of hundreds of leaves spanning arbitrary directions and angles – even a small plant could require a large number of polygons to define every facet digitally (Weber and Penn 1995).

Moreover, mature crop plants, which are of primary interest to the phenotyping and breeding communities, typically have a more complex 3D architecture than laboratory-based model plants such as *Arabidopsis thaliana*.

Despite these challenges, previous work (Tan *et al*. 2007) suggests that image-based approaches offer the best solution to 3D reconstruction. Image acquisition is usually straightforward, the tools involved have shown promising results and do not require their users to have high levels of expertise (Tan *et al*. 2007).

Image-based 3D modelling

Image-based approaches reduce, although do not eliminate, the complexity associated with rule-based approaches. They delineate real world plants by extracting geometry directly from images, with the elusive goal of achieving photorealism (Weber and Penn 1995). Capture techniques can be categorised as either active or passive, where active is significantly more expensive and requires specialist hardware to project some form of light into the scene. Light detection and ranging (LiDAR) and laser-based ‘digitisation’ are perhaps the best known active approaches.

Space carving, shape-from-silhouette (SFS), shape-from-shading (Cryer and Shah 1999), shape-from-contour, stereo vision and structure-from-motion (SFM) (discussed below) are passive approaches commonly conducted using standard hand-held cameras. The challenge for these methods is to produce 3D representations under normal, ideally natural, illumination conditions. Approaches such as shape-from-shading (Horn and Brooks 1989), shape-from-texture (Kender 1981) and shape-from-edges (Wahl 2001) are used but are uncommon in plant modelling due to the complexity of the object and their reliance on a single image, making them more susceptible to occlusion, a common occurrence in plants.

Image based approaches can be further categorised into those that begin with an existing, generic, plant model that is fitted to the image data, known as top-down, or those that apply a series of processes to the contents of images, to create an increasingly accurate and realistic plant model, known as bottom-up.

Top-down approaches use an existing model that is adjusted to fit the image data, so that the new plant representation is consistent with what is observed. The application of top-down approaches to inter-species is unclear, as differences between the expected and actual geometry of a plant or leaf increases. Bottom-up approaches, reviewed in this paper, are methods beginning with one or more images which reconstruct a plant model based only on the observed pixel data. We focus here on bottom-up approaches, as they provide the greatest opportunities for generic (species-independent) 3D reconstruction of plants. The top-down approach, although of interest, also suffers from a lack of models with which to guide analysis.

Active approaches

LiDAR, a remote sensing technology based on the extension of principles in radar technology, measures the distance between itself, the scanner, and the target object by illuminating the object with a laser and analysing the time it takes the reflected light to return (Northend 1967; Killinger 2014). LiDAR has two distinct fields of application; airborne LiDAR, in which the scanning device is commonly attached to a plane or helicopter, and terrestrial laser scanning (TLS), which is conducted on the ground and the scanner is either stationary or attached to a ground-based vehicle (Ullrich and Pfennigbauer 2011).

Laser scanning acquires information from an object by digitising selected co-ordinates and representing these as a 3D point cloud by recording the scanned distance to each. Just like cameras, they have a cone shaped field of view and capture multiple views in order to perform complete reconstruction. The main difference in resultant data between cameras and time-of-flight lasers is that the latter stores depth in each pixel whereas cameras store colour (Curless 1999).

'Structured light' techniques provide an alternative approach to depth measurement. Here the light source (usually laser, or near-infrared) is positioned a short distance from an imaging device (usually a camera fitted with appropriate filters). Light leaves the emitter and is reflected into the camera by the target object. Knowledge of the light source, and use of appropriate filters, makes the emitted light pattern easy to detect in the image. The relative positions and orientations of light emitter and imaging device are also known, allowing 3D data to be recovered from the position of key points of the emitted pattern by triangulation. A variety of light patterns have been used including spots, lines and 2D grids. Perhaps the most common example of a structured light device is the Microsoft Kinect, which emits a rectangular dot pattern in near-infrared. Microsoft's KinectFusion (Newcombe *et al*. 2011) software also allows depth data gathered from multiple views to be combined in a single model.

Structured light methods can be effective, and in recent years have become more easily obtainable and affordable, as components of RGB-D (red, green, blue, depth) devices such as the Kinect. RGB-D cameras combine depth sensing with common camera functionality, providing both 3D and colour measurements.

Unfortunately, however, structured light approaches suffer several drawbacks when applied to plants. They can be difficult to use in bright light, e.g. glasshouses, where background illumination makes the projected pattern hard to detect. Highly reflective leaf surfaces can also act as (partial) mirrors, reflecting a significant proportion of the emitted pattern away from the imaging device and again making it hard to detect. Narrow objects, e.g. rice leaves, can fall between the key points of the emitted pattern (e.g. Kinect's dots) and simply fail to reflect the pattern back.

With recent advances in technology such as readily available software to deal with the large computational requirements of these approaches and the development of 'multi-pulsed' LiDAR (Su *et al*. 2015), LiDAR is becoming more commonly used, and can easily be deployed in both airborne and ground-based forms. The airborne approach is particularly useful for reconstructing forest canopies and tree structure from dense forestry, enabling the reconstruction and acquisition of geometric properties from remote locations, which other image-based approaches may find difficult due to accessibility.

Passive approaches

Although LiDAR can be effective it requires expensive equipment that is out of reach of many. Passive approaches are therefore gaining an increasing amount of popularity, as they only require a standard 'off-the-shelf' digital camera to capture overlapped images, simultaneously or sequentially, and a basic computer to process them. As passive methods use only the radiation present in the scene, specialist lighting is often not required.

A variety of passive approaches exist which manipulate the 2D image information in various ways. One of these enables 3D objects to be reconstructed from 2D silhouettes by back-projecting them from their cameras' viewpoints and intersecting the resulting cones. SFS (shape-from-silhouette), introduced by Laurentini (1994), does exactly this. The aim is to construct a 3D model by projecting the 2D silhouette of the object from multiple images into a single 3D space in which intersecting projections produce the 3D model, known as the visual hull.

The visual hull determines the largest possible shape that is consistent with the available images. In many cases, where the number of input images is high, the resulting model will be a good approximation. However, as the scene becomes increasingly complex, for example, a scene with concavities and occlusions, the dissimilarity between the resulting model and the actual object will increase. A complex plant canopy consisting of multiple overlapping plants, for example, will produce poor results in which leaf thickness is overestimated and concavities are missed or underestimated.

SFS is simple to implement, requiring only a set of arbitrary views of an object from known camera positions, which can be obtained through camera calibration (Salvi *et al*. 2002). The biggest challenge lies in ensuring the foreground (object) and background can be separated to find the object's silhouette. In natural conditions this can be a challenging problem, however at present much phenotyping work is conducted in controlled environments where there exist several techniques for background and foreground separation, for example; the Canny algorithm (Canny 1986) or frame differencing (Piccardi 2004). A comprehensive review of SFS is provided by Dyer (2001).

Space carving was introduced by Kutulakos and Seitz (2000) as a solution to the difficulties associated with SFS. It starts with a bounding box big enough to encapsulate the entire object or scene, whose size is often pre-defined by the user. The bounding box is partitioned into a series of voxels, cubes in three-dimensional space represented by co-ordinates and size. The algorithm relies on measures of the photo-consistency of voxels, where a voxel is said to be photo-consistent if, and only if, the colour of the voxel appears to be (approximately) the same in all of the images in which it is visible. It is assumed that if some voxel is the same colour then it lies on the object's surface and is marked as seen. The set of voxels that are marked as 'seen' then make up the 3D model of the object.

The algorithm is again simple to implement, iterating through each voxel of the bounding box, projecting to each image and removing (carving) those voxels that are not photo-consistent. Each time a voxel is carved away it potentially uncovers a new voxel, which also requires evaluation for photo-consistency, and the process continues until all visible empty voxels are removed or some user defined stopping criteria is met.

Other less common voxel techniques used for 3D reconstruction include voxel colouring (Seitz and Dyer 1999) and generalised voxel colouring (Culbertson *et al*. 2000), which, like space carving, rely on the consistency of colours between images to determine whether some seen voxel lies on the surface of the object. However, unlike space carving, the camera positions are often constrained in order to determine colour consistency more easily, limiting the views that can be used, and so the complexity of the objects that can be modelled.

Stereo vision differs significantly from SFS and is based on key functionality of the human vision system – the ability to see the same scene but from slightly different viewpoints, achieved through the distinct lateral positioning of the eyes – known as binocular vision. Stereo vision aims to mimic this process, extracting 3D information by processing two 2D images captured simultaneously from slightly different horizontal angles, focusing on the same point in space.

Stereo vision has three main processing steps: stereo calibration, feature extraction and correspondence matching. These are discussed in turn below.

Stereo calibration finds the intrinsic parameters (focal length, principal point, radial and tangential distortion) of each camera and the extrinsic parameters (rotation matrix and translation vector) linking the two cameras. It allows 3D world co-ordinates to be mapped to 2D image co-ordinates.

Feature extraction identifies features of interest, independently, in each image. Features vary widely and range from simple image patches to extended straight lines, circles and regions corresponding to viewed objects. A common middle ground is to define features by their local image properties, most often their gradients. Edges and corners are widely used, these are points at which image values vary significantly (i.e. the gradient of image values is large) in one or more directions.

Correspondence matching links features found during feature extraction between views. If the image features associated with a particular object feature can be identified in multiple images, taken from different viewpoints, knowledge of the cameras' positions and orientations allow its 3D location to be determined. The disparity associated with each match – the difference in the image co-ordinates of the matched features – is obtained and can be used to create a disparity map which in turn can be used to acquire depth information.

Structure-from-motion (SFM) follows the same process. However, where stereo vision captures two images simultaneously, SFM captures images sequentially, estimating 3D points from an extended sequence of images. 3D data is then estimated either sequentially, by matching pairs of images, or globally, matching features between all images. A review of early vision dating back to the 1970s and 1980s can be found in work by Barnard and Fischler (1982) and Dhond and Aggarwal (1989), respectively, and Brown *et al*. (2003) provide a comprehensive review of the advances in modern stereo vision.

Binocular stereo and structure from motion rely on points on the target object projecting to different locations in each of a set of images. By finding image features arising from those points, and matching them between views, they can reverse the projection process to recover 3D. Photometric stereo (Woodham 1989) takes a different approach. Here, multiple images are taken from a fixed camera, but the lighting conditions are varied between each image acquisition. Object points therefore project to the same location in each image, but appear different due to changes in illumination. Knowledge of the lighting used, and of the image formation process, allows 3D information, usually surface orientation, to be computed from these variations on appearance.

Photometric stereo is less widely used in practise than binocular stereo and SFM, as it can be difficult to adequately control and quantify lighting conditions. Surface orientation must also be integrated to obtain depth estimates, which can pose further problems. Photometric stereo is, however, now attracting interest within the controlled environment phenotyping community.

Less common methods such as concept sketching, which is the process of digitally drawing 3D shapes or is the process of creating a 3D model from a 2D sketch, have also been applied to plant reconstruction (Masry and Lipson 2007), focusing more specifically on structure. The sketching technique is less relevant in modern times, as the available computing resources make methods based on real mages practicable.

Sketching does, however, have some advantages, such as the ability to use freehand drawing, allowing shapes to be accurately captured and contours to be easily identified (Anastacio *et al*. 2006). Sketching commonly uses an interface to enable direct manipulation of the plant simulation, allowing even novice users to create plant structures (Masry and Lipson 2007). Though, as with rule-based approaches, the model does not represent a real plant.

Representing 3D data

Though all the methods discussed here recover 3D information from images, different methods represent 3D data in different forms.

Voxel-based methods (SFS, voxel colouring, space carving) produce a volumetric description of the target object. This is a 3D array of cells – effectively a 3D image – in which each cell (voxel) contains one of two possible values. These values indicate whether or not that voxel is occupied by the object, effectively separating (3D) object material from (3D) space. Volumetric representations are compact, and their accuracy can be controlled by varying voxel size; larger voxels result in a more ‘blocky’ representation. The set of shape and other measures, i.e. traits, directly available from voxel descriptions is, however, limited.

Total object volume can be estimated by counting occupied voxels, and fitting a convex hull or similar structure around those voxels provides crude object dimensions. More detailed characteristics require further processes, however, and it is common to fit a surface over the object voxels using the marching cubes algorithm (Lorensen *et al*. 1987), or similar. Further measures and features can then be extracted from the surface description.

LiDAR, structured light, binocular stereo and structure from motion typically produce a point cloud representation: a set of unconnected x,y,z co-ordinates describing the locations of matched points. Again, coarse, summary traits can sometimes be obtained directly from this data structure, but it is common to first link nearby points to form a mesh, and fit some form of surface.

Photometric stereo is unusual, in that it typically produces local surface orientation estimates, from which depth must be recovered to produce a full surface representation. Whatever the route, surface-based representations are usually required in plant phenotyping and simulation work.

In a majority of cases, the final surface representation produced by 3D reconstruction methods is piecewise. Rather than fit a single, mathematically complex, surface over the whole object, a large set of simpler surfaces is used. These are linked together to produce a complete description. Small triangular planes are most commonly used, as these can be linked along their edges to describe a wide range of complex shapes.

Application to plants

It is crucial to construct precise 3D representations of plants to facilitate accurate assessments of physiology. With the use of accurate 3D plant models more subtle traits can be identified, leading to a greater amount of, and more useful, information with respect to plant architecture and growth. Models can be used to measure the geometric structural parameters of plants, which is of utmost importance in understanding the biological and physical processes of growth, a vital element in increasing crop yield (Wang *et al*. 2009). Height, dimensions, leaf area, angle and distribution are important parameters, all of which relate directly to the growth and photosynthetic properties of plants.

Plant architecture is known to be a determinant of the productivity of canopies. On a simple level this arises via the relationship between vertical leaf area index (LAI), leaf area distribution (LAD) and leaf angle. The penetration of light that results is mathematically described by the Mons–Saeki equation derived from Beer's law (Hirose 2005). Vertical distribution of leaf photosynthesis is dominated by the interaction between light gradients and the individual light response curve of each leaf. A vertical canopy thus permits a higher optimal LAI and a higher overall rate of canopy photosynthesis. Many existing productive crops have an 'erectophile' tendency. However, the dependence on a high LAI can lead to higher nutrient requirements and weed problems. Therefore, there is still a need to understand the relationship between photosynthesis dynamics and precise canopy architecture.

LAI and LAD estimates are two measurements that offer significant insight into the ability of a plant to capture radiation for photosynthesis. These measures can be obtained manually, though the process is often tedious and error prone, for example, an operator has to manually measure a leaf segment using callipers. As a result, observers may have varying opinions, and the approach tends to be intrusive and accuracy decreases compared with the automatic measurements. However, with the use of modern technology, approaches are becoming less interactive and are increasingly becoming more accurate and automated. One such image-based approach, which calculates the leaf area as the area of the surface of the 3D model by summing the area of triangles, is applied to corn plants by Wang (2009). Hosoi (2006) develop a method known as voxel-based canopy profiling to measure the LAI and LAD of small trees (namely *Camellia sasanqua* and *Deutzia crenata*) using both mobile ground-based and airborne LiDAR, obtaining results as accurate as 0.7 up to 17% for the minimum leaf thickness for the measurements of LAI and LAD. Automatic measurements were compared with those obtained by stratified clipping, where plant parts are manually measured in segments, one a plant segment has been manually measured it was removed to provide access to the next part, typically starting from the top of the plant and working downwards. Alternatively, a stereo vision approach can be used to obtain measurements and identify branch and leaf segments, for example, Paproki *et al.* (2011) applied this to cotton plants. Using a top-down approach, they recursively segment the plant into regions, at each iteration determining which segmentation algorithm to apply in order to extract a specific limb from the model. With this they accurately identified 20 out of 22 cotton plant segments.

The ability to automatically identify and extract single leaf data would significantly improve the process of calculating LAI and LAD. Biskup (2007) proposed an approach that uses stereo vision in a field setting to track the nocturnal and daytime movement of leaves and determine drought stress, with a particular focus on soybean plants. Some approaches use a skeleton representation of the plant to identify regions. The skeleton representation is a thin version of the shape emphasising its topological properties. In most cases the skeleton is a thin, connected, line aligned with the centre of the object. The process of creating a skeleton model is referred to as skeletonisation. Jin (2009) used a real-time stereo vision approach with a skeletisation algorithm to identify individual corn plants and highlight leaves from stems, they report that they were able to accurately detect 96.7% of corn plants and that they were within 1–5 cm accuracy when determining the plant centre.

Cai and Miklavcic (2012) used 2D skeletons to extract the 3D structure of cereal plants. They reported that they were able to deal with difficulties such as overlapping plant parts and broken segments resulting in smooth, connected 3D cereal structures. Stereo vision and SFM have been used to reconstruct plant models in many other similar scenarios, from the construction of trees to maize canopies (Ivanov *et al*. 1995; Andersen *et al*. 2005; Quan *et al*. 2006; Wang *et al*. 2009; Hartmann *et al*. 2011; Lou *et al*. 2014). Pound (2014) proposed a fully automated stereo vision approach to reconstruct plant shoots, namely wheat (*Triticum aestivum*) and rice (*Oryza sativa*). The reconstruction process works on segments of leaves and develops each individually using level sets, which optimises the model based on image information. The effects of occlusion are reduced by identification of the best image for each segment, requiring few assumptions to be made.

LiDAR has received a vast amount of attention in recent years because hardware has become more affordable and applicable to a range of plant species. For example, the geometric structure of white clover canopies has been assessed by Rakocevic (2000) using electromagnetic digitising apparatus. They used corner flags to aid calibration, thus improving the accuracy of the reconstruction, and applied a destructive approach. The canopy was pruned from the top downwards and scanned at each stage, with results showing that the semi-automated measurements varied between 5–20% in comparison to the manual measurements. The error in this work could, however, lie within either the manual or automatic measurements and without the use of an independent, confirmed ground truth it is not possible to tell.

Similarly, Paproki *et al*. (2012) presented a mesh-based, 3D LiDAR approach for reconstructing *Gossypium hirsutum*, which partitioned the plant into morphological regions. They stated that they were able to match leaves in 95% of the cases and that LAI accuracy was within 10% of manual measurements.

Aside from single leaf and small crop measurements, other larger plants have received a great deal of attention. Trees, for example, are particularly valuable due to their functional roles in the environment and have received considerable interest aimed at calculating the tree crown volume, 3D architecture and branching structure. LiDAR is the most common approach for the reconstruction and approximation of trees (Weber and Penn 1995; Sinoquet and Rivet 1997; Sakaguchi 1998; Shlyakhter *et al*. 2001; Boudon *et al*. 2003; Reche-Martinez *et al*. 2004; Phattaralerphong and Sinoquet 2005; Hosoi and Omasa 2006; Rutledge and Popescu 2006; Neubert *et al*. 2007; Omasa *et al*. 2007; Tan *et al*. 2007; Livny *et al*. 2010; Preuksakarn *et al*. 2010; Van Leeuwen *et al*. 2010; Tang *et al*. 2013), making it possible to estimate forest attributes, such as height, diameter and canopy closure, all of which are

Publisher: CSIRO; Journal: FP:Functional Plant Biology
Article Type: research-article; Volume: ; Issue: ; Article ID: FP16167
DOI: 10.1071/FP16167; TOC Head:

essential parts of forest management. Other modelling approaches are often limited in their capacity to retrieve individual tree and crown attributes due to occlusion or canopy gaps.

Skeletons can be used to represent the branching structure of trees, which can provide vital information, particularly when occluded by leaves. Tang (2013) used TLS to obtain skeletons from trees and Livny (2010) created a tree model from laser scans captured using a moving vehicle. They applied a series of global optimisations to the branching structure – a constraint ensuring branches are thicker closer to the root, for example, making it robust to noisy and incomplete data, before scans are employed to consolidate a point cloud representing one or more tree objects as skeletal structures. This optimisation aimed to reconstruct the major branches of the captured tree(s), resulting in a graph structure that they defined as the branch-structure-graph (BSG). The finer branching structures were then reconstructed from the BSGs, with the assumption that the finer parts of the tree structure are made up of the same branching structure as the core of the tree.

In the modelling of trees, canopy height models (CHMs), are used to represent horizontal and vertical properties of tree canopies. However, retrieving these characteristics is challenging and several difficulties have been identified, primarily the underestimating of height which can occur when the earth's surface is occluded by the tree canopy (Pitkänen *et al*. 2004; Zhao, Kaiguang 2007). Van Leeuwen (2010) proposed an airborne solution, the parametric height model (PHM), to overcome the problem of underestimating tree height in CHMs by describing the forest canopy as a series of cones fitted to the raw LiDAR point cloud (Illingworth and Kittler 1988).

Other approaches to tree modelling exist: Shlyakhter *et al*. (2001) used visual hulls to generate the skeleton of the tree augmented with an L-System approach, Neubert *et al*. (2007) used a space carving approach to estimate tree volume, and Reche-Martinez *et al*. (2004) combined volumetric opacity estimate with view-dependent texturing to reconstruct trees from images. LiDAR is seldom used in smaller plant representations due to high processing times but it is capable of producing adequate results, for example, Hosoi and Omasa (2009) estimated the vertical area of wheat canopies.

More recently, Apelt *et al*. (2015) introduced Phytotyping^{4D}, a light-field camera system which produced grey-scale images, depth information and a focus image, to measure plant features in 4D. They successfully monitored rosette and individual leaf growth in *Arabidopsis*.

Challenges and opportunities

With accurate 3D models various traits such as the tolerance, resistance, architecture, physiology and growth can all be easily obtained, and more complex traits such as LAI, LAD and photosynthesis measurements can be made. One recent method, proposed by Burgess *et al*. (2015), automatically obtains the light distribution in three different wheat (*Triticum aestivum*) lines without the need for manual measures. 3D models are captured using the stereo vision approach proposed by Pound *et al*. (2014). The methods reviewed here have also been shown to extract plant traits from 3D models that may otherwise have been tedious and error prone.

However, 3D reconstruction is a challenging problem and complications arise irrespective of the approach. Image-based models typically suffer from errors and omissions introduced by occlusion, in which aspects of the scene are obscured relative to the camera, or parallax, in which objects appear differently depending on their position relative to the camera (Kutulakos and Seitz 2000). Active approaches can struggle in natural illumination conditions and with reflective surfaces. These challenges, and others discussed here, make the complete reconstruction of scenes and objects, with any method, a complex task. Table 1 provides a summary of the advantages and disadvantages/challenges of these approaches.

Much of the previous work in this field has been focussed on single plant reconstruction, where some success has been achieved. More recently, however, there has been an increased interest in canopies, particularly those grown in the field, which is proving more difficult. In cases where plant structure has proved too complex, approaches have relied on semi-automatic reconstruction, i.e. (Rakocevic 2000), with a user guiding the reconstruction in areas of ambiguity.

Computer vision challenges

Despite advances in technology, resources and increased interest in plant-related problems from the computer vision community, approaches to the production of automated systems for 3D reconstruction are cumbersome. Few fully automated approaches – those capable of capturing data, performing the intermediate steps and producing an output as a 3D model – have been proposed. Many of the image-based approaches require user input, most commonly during segmentation (for example, separating the background from foreground or leaf from stem) or during image acquisition. However, the need for an automatic, robust and flexible image analysis tool for plant modelling clearly exists (Hartmann *et al*. 2011), as does a desire to extend these techniques to multiple plants and to install them in field environments.

For stereo vision, occlusion is perhaps the biggest challenge yet to be overcome. Images are often captured from only two viewpoints, which restricts the view of the rear of an object, resulting in a ‘2.5D’, rather than a complete 3D model. For this reason, stereo cameras are often used from above for canopy or rosette analysis where a detailed 3D structure is not necessary. Improved results may be obtained using multi-view stereo, or structure from motion (Dhond and Aggarwal 1989). Although techniques exist to make this process more computationally efficient, by e.g. exploiting epipolar geometry (Zhang 1998) or by using leaf orientation (Laga and Miklavcic 2013), it still remains challenging. The problem of occlusion is particularly common in plants where complex leaf structure may cause higher levels of occlusion than is often seen in other stereo vision tasks (Pound *et al.* 2014). A given leaf patch may not be visible in enough images, or its appearance may be so similar to that of its neighbours that it may not be possible to ensure the correct correspondence is made.

Silhouette-based approaches offer some advantages. They are often simple to implement and do not require a calibration target. Utilising multiple views, they form a complete model representing the plant being imaged. However, these approaches are also ill-suited to the high amounts of occlusion exhibited by some plants, and plant canopies (Mulayim *et al.* 2003), also failing to account for concave surfaces, which will be interpreted as solid.

As a result, a silhouette approach commonly has to be augmented with another approach that is capable of removing excess voxels (Mulayim *et al.* 2003). In extremely crowded scenes, the reconstruction will fail to adequately capture the scene, even with post processing, and an accurate reconstruction is impossible to obtain. Furthermore, silhouette approaches are a poor choice for reconstruction when surfaces are thin, as leaves often are. Silhouette-based plant reconstruction methods often result in blocky, overestimated data because the size of the voxels representing the object being larger than the object itself. Leaves are usually either poorly represented or, often, excluded.

Active methods such as LiDAR have the advantage of avoiding the correspondence problem often seen in stereo imaging, and can deal well with complex object boundaries. A primary concern with laser-based approaches is that their scanning time is directly related to the resolution required. For example, LiDAR struggles with single leaf analysis, where the required resolution dramatically increases the scanning time. This has been highlighted in much of the work where high resolution scans are required. For example, Watanabe *et al.* (2005) modelled small rice plants using a continuous plant and fractal generator (CPFG) approach with a 3D sonic digitiser to capture the initial point cloud. The digitisation process can take up to an hour to complete for each rice plant. As a result, capturing high resolution scans can only be achieved in a controlled environment where wind is avoided and other environmental conditions can be monitored and controlled (Biskup *et al.* 2007). Rakocevic (2000) claimed that the digitisation process for their approach to reconstruct white clover canopies required between 3 and 7 h, which also involves a destructive approach to obtain a complete reconstruction. This eliminates the possibility of repeating the experiment using the same plant. The initial cost of hardware is also often prohibitive.

Non-laser approaches can also suffer from high processing requirements if too much information is acquired. When using image-based reconstruction, determining the optimal number of samples (images) is often problematic. Collecting excess samples is known as 'oversampling', and will inevitably lead to a more intensive data acquisition model, higher capacity requirements and greater redundancy (Shum and Kang 2000). In many cases oversampling will lead to significantly higher computational requirements, without notable benefits in output quality. Indeed, in some cases oversampling can lead to degradation in reconstruction quality.

In contrast, incomplete and inaccurate reconstruction is a classic consequence of 'undersampling', where an inadequate number of images fail to deal with the issues of occlusion in the scene, and some regions of the model remain unobserved. The issues of under or oversampling can be partly addressed by a robust image acquisition strategy using an automated capture system. This can be quickly adapted to a variety of plant species or experimental requirements, and the number optimal number of images derived.

The determination of an appropriate image acquisition strategy is challenging, particularly given the dynamic structure of plants. Existing approaches typically rely on the use of manually captured images or static camera positions that do not change, regardless of plant species. With the use of active vision more flexible image acquisition approaches can be adopted, dynamically changing to reflect the size of the plant. Gibbs *et al*. (2015), for example, developed an active vision system that is capable of capturing images of plants using a robot arm and a turntable overcoming the problems of static camera positioning. This approach improves the data in comparison to fixed camera positions and produces a more detailed point cloud, thus enabling a more accurate reconstruction.

Some plants may have to be moved if the camera position is static, for thin plants this can cause difficulties in reconstruction as the leaf setup may vary between images. Though the problem can be alleviated; for example, Kumar *et al*. (2012) reconstructed a plant using two cameras and twin mirrors enabling the back of the plant to be seen from a front view and as a result the plant does not need to be moved from its original setup. Alternatively, Kumar *et al*. (2014) proposed a method in which the plant remains static and the camera rotates at a fixed height around it.

Some image-based approaches require a calibration target – an object in the scene that is used as a reference point to determine correspondence between two images – that is ideally visible in each image. This can limit the types of plants modelled as they may occlude the calibration target. Approaches that require a calibration target add further challenges to field based phenotyping, where they are harder to include.

Moreover, phenotyping methods in general often make over-simplifying assumptions, such that the object is of a specific shape or size, that the background is a certain colour, that the object is green, or that each leaf is the same shape. With these specific conditions the approaches lack robustness and struggle to deal with varying plant species. The approach by Pound *et al*. (2014) provides a more robust approach with respect to plant species and is able to reconstruct a variety of plants due to the ability to work on smaller areas (patches), manipulate image data and lacks plant specific constraints which often reduce the robustness of reconstructions.

Phenotyping is receiving an increasing amount of attention and is now recognised on a global scale. Computer vision experts are becoming more involved, offering insights to biologists. Conferences such as Computer Vision Problems in Plant Phenotyping (CVPPP) and the International Workshop on Image Analysis Methods for Plant Science (IAMPS) are becoming increasingly popular and provide a way to collaboratively improve approaches. Training courses for biologists are also becoming more easily and frequently available.

Validation challenges

3D reconstruction has been a topic of interest in the wider computer vision community for many years. As new reconstruction methods have been proposed it has been increasingly important that some objective evaluation and comparison criteria be adopted. Several approaches present themselves. First, standard test objects, of which at least some dimensions have been precisely measured, can be used. Evaluation then becomes measurement of the difference (error) between those measurement and corresponding values reported by the proposal reconstruction method. This approach can be used to assess plant reconstruction methods, but the complex and flexible nature of plant shoots can make it difficult to provide appropriate ground-truth measurements.

An alternative approach is to create artificial images from existing 3D plant models (e.g. Pound *et al*. 2014). Here, computer graphics techniques are used to produce images which can be re-analysed by competing techniques. Evaluation is performed by comparing the newly reconstructed and original 3D models. Once again, the complex and variable properties of plant shoots (this time their appearance) can make this method challenging.

Regardless of the approach taken, there is a pressing need for sizeable plant reconstruction datasets, including both images and ground truth, to be created and made available to the development community. Recently, Minervini *et al*. (2015) released a first of its kind dataset to investigate approaches in state of the art leaf segmentation. Scharr *et al*. (2016) provided a collation of previous segmentation approaches and applied these to the CVPPP dataset, discussing the methods and findings of the application.

From laboratory to field

At present phenotyping experiments are commonly conducted in controlled environments where natural conditions such as light and wind can be monitored and manipulated. Much of the work focuses on single plant reconstruction, though small canopies are now being used in controlled environments too.

When constructing a dense plant, or a canopy, approaches to 3D modelling often require intrusive, (moving the plant foliage in order to obtain further information), and destructive, (the removal of plant parts), approaches to plant reconstruction in order to acquire plant geometry. This allows image capture of aspects of a plant or canopy that may not otherwise be seen, but makes repeat experiments, or capture of time series data, impossible. Destructive approaches often require manual pruning of plants, adding additional time to the acquisition process and increasing the potential for irreversible error, i.e. pruning too low, resulting in an incomplete acquisition process. Despite these drawbacks, destructive methods continue to be one of the few reliable methods for extending reconstruction approaches to dense canopies, where occlusion is at its highest level. Indeed, most existing image-based approaches will fail quickly as the number of plants is increased – a problem for which a reliable solution is yet to be found. In principle, a surface based reconstruction approach could be extended to denser canopies, but any results have yet to be presented. Field based phenomics still proves challenging in this regard due to the ever changing environment and the need to reconstruct crowded scenes containing multiple plants and many leaves. White *et al*. (2012) explain the difficulties associated with field based phenomics, concluding that it provides too much of a challenge for existing technology and that advances need to be made.

Directly related to field based phenomics are the difficulties associated with tree reconstruction. Tree height, dynamic surroundings and the inability to conduct investigations in controlled environments make modelling trees difficult. Key difficulties lie within physical accessibility, availability of objective and efficient measurement techniques and the associate costs (Lovell *et al*. 2003). Furthermore, Jin and Tang (2009) found that during experiments in natural conditions the acquisition of images under direct sunlight turned out to be severely saturated when compared with those taken under cloudy lighting conditions.

Using LiDAR in field environments is challenging as daylight can make it difficult to capture data where the sun interferes with the reflection back to the scanner. If the illumination of a single object changes during data acquisition further difficulties arise, such as the colour of the object changing. Most LiDAR hardware is also affected by nearby metal structures and magnetic sources, making experiments in urban environments challenging.

With respect to stereo vision, the matching problem is further complicated by issues of illumination changes and poorly textured surfaces. Illumination is a key area that prevents correct matching between a left and right view of the scene, in many cases adding noise, or preventing parts of the 3D model being recovered (Paproki *et al*. 2011). Furthermore, approaches such as space carving and voxel colouring that rely on colour consistency between images become impractical reconstruction choices. Even in a controlled environment it is often overlooked that when using a turntable with fixed lighting and a rotating object, the light hitting the surface will change at each rotation and as a result produces different shades in each image.

Although field based phenomics is still challenging, experiments in controlled environments show promising results and the use of robotics and active vision to automatically capture images of plants used to perform reconstruction are further enhancing the process improving both the quality and control.

Concluding remarks

A variety of methods have been proposed that seek to recover quantitative data on plant traits from image sensor data captured in laboratories, glasshouses and field environments. Some important plant traits, such as plant height, can be extracted directly from carefully acquired images. Others, for example, capturing the detailed shape of wheat spikes or leaves, require intermediate representations to be acquired first. Although phenotyping techniques based on 3D representations are beginning to appear (Vadez *et al*. 2015; Cabrera-Bosquet *et al*. 2016), the construction of 3D models of real plants remains a challenge. The ability to recover physically correct representations of the 3D shape and structure of plants and plant components from image data would underpin the measurement of rich sets of plant traits, and thus accurate phenotypic information.

Different approaches to the 3D reconstruction of plants have been examined and it is clear that reconstruction remains a challenging computer vision problem in which advances in technology and optimal data acquisition techniques are required. Reductions in the cost of equipment with regards to laser scanners and computers offering extensive computational power, along with reduced costs in outdoor sensing equipment, is one area that is actively improving, though the size of 3D models and the required detail is also increasing.

Although image-based approaches can produce realistic looking plant models, they still remain highly interactive. A fully-automated system is clearly a necessity. However, an active vision approach, that is an approach capable of manipulating the camera viewpoint in order to investigate the environment, is required along with the ability to determine objects of importance without user interaction or assumptions being made beforehand. Advanced computing and algorithms and a reduction in hardware costs are necessary before this becomes a reality and until then semi-automated approaches must be used.

Field-based phenomics are especially challenging due to environmental challenges and data acquisition processes. Capturing data on a large crop is intrusive and requires modification to the land setup, providing space to access the plants along single rows. Furthermore, the process of acquiring data is resource intensive with multiple vehicles required in order to capture rows more than once per day. With the lack of arable land it isn't feasible to approach FBP like this and improving current crop yields is necessary beforehand.

It is encouraging to see phenotyping receiving increasing attention, particularly from computer vision researchers, and as a result several conferences, workshops and training courses are now available. Utilising 3D data will aid phenotyping practice and we expect to see an increase in the development and uptake of 3D approaches in the near future.

References

<jrn>Adeloye A (2010) Global warming impact: flood events, wet-dry conditions and changing scene in world food security. *Journal of Agricultural Research and Development* **9**(1), doi:10.4314/jard.v9i1.56128</jrn>

<conf>Anastacio F, Sousa MC, Samavati F, Jorge JA (2006) Modeling plant structures using concept sketches. In 'Proceedings of the 3rd international symposium on non-photorealistic animation and rendering'. pp. 105. (ACM Press: New York)</conf>

<jrn>Andersen HJ, Reng L, Kirk K (2005) Geometric plant properties by relaxed stereo vision using simulated annealing. *Computers and Electronics in Agriculture* **49**(2), 219–232. doi:10.1016/j.compag.2005.02.015</jrn>

<jrn>Apelt F, Breuer D, Nikoloski Z, Stitt M, Kragler F (2015) Phytotyping^{4D}: a light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. *The Plant Journal* **82**(4), 693–706. doi:10.1111/tpj.12833</jrn>

<jrn>Barnard ST, Fischler MA (1982) Computational stereo. *ACM Computing Surveys* **14**(4), 553–572. doi:10.1145/356893.356896</jrn>

<jrn>Biskup B, Scharr H, Schurr U, Rascher U (2007) A stereo imaging system for measuring structural parameters of plant canopies. *Plant, Cell & Environment* **30**(10), 1299–1308. doi:10.1111/j.1365-3040.2007.01702.x</jrn>

<jrn>Bonhommeau S, Dubroca L, Le Pape O, Barde J, Kaplan DM, Chassot E, Nieblas AE (2013) Eating up the world's food web and the human trophic level. *Proceedings of the National Academy of Sciences of the United States of America* **110**(51), 20617–20620. doi:10.1073/pnas.1305827110</jrn>

<jrn>Boudon F, Prusinkiewicz P, Federl P, Godin C, Karwowski R (2003) Interactive design of bonsai

tree models. *Computer Graphics Forum* **22**(3), 591–599. doi:10.1111/1467-8659.t01-2-00707</jrn>

<jrn>Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C (2012) L-Py: an L-system simulation framework for modeling plant architecture development based on a dynamic language. *Frontiers in Plant Science* **3**, 76. doi:10.3389/fpls.2012.00076</jrn>

<jrn>Brown MZ, Burschka D, Hager GD, Member S (2003) Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(8), 993–1008. doi:10.1109/TPAMI.2003.1217603</jrn>

<jrn>Burgess AJ, Retkute R, Pound MP, Foulkes J, Preston SP, Jensen OE, Pridmore TP, Murchie EH (2015) High-resolution three-dimensional structural data quantify the impact of photoinhibition on long-term carbon gain in wheat canopies in the field. *Plant Physiology* **169**(2), 1192–1204. doi:10.1104/pp.15.00722</jrn>

<jrn>Cabrera-Bosquet L, Fournier C, Brichet N, Welcker C, Suard B, Tardieu F (2016) High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. *New Phytologist* doi:10.1111/nph.14027</jrn>

<jrn>Cai J, Miklavcic S (2012) Automated extraction of three-dimensional cereal plant structures from two-dimensional orthographic images. *IET Image Processing* **6**(6), 687–696. doi:10.1049/iet-ipr.2011.0281</jrn>

<jrn>Canny J (1986) A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**, 679–698. doi:10.1109/TPAMI.1986.4767851</jrn>

<jrn>Challinor AJ, Watson J, Lobell DB, Howden SM, Smith DR, Chhetri N (2014) A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*. doi:10.1038/nclimate2153</jrn>

<jrn>Cryer JE, Shah M (1999) Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(8), 690–706. doi:10.1109/34.784284</jrn>

<edb>Culbertson WB, Malzbender T, Slabaugh G (2000) Generalized voxel coloring. In 'Vision algorithms: theory and practice. Vol. 1883'. (Ed. B Triggs, A Zisserman, R Szeliski) pp. 100–115. (Springer: Berlin)</edb>

<jrn>Curless B (1999) From range scans to 3D models. *Computer Graphics* **33**(4), 38–41. doi:10.1145/345370.345399</jrn>

<jrn>de Reffye P, Edelin C, Françon J, Jaeger M, Puech C (1988) Plant models faithful to botanical structure and development. *Computer Graphics* **22**(4), 151–158. doi:10.1145/378456.378505</jrn>

<jrn>Deussen, Oliver and Bernd Lintermann (1997) A modelling method and user interface for creating plants. *Graphics Interface* **97**, 189–198.</jrn>

<jrn>Dhond UR, Aggarwal JK (1989) Structure from stereo – a review. *IEEE Transactions on Systems, Man, and Cybernetics* **19**(6), 1489–1510. doi:10.1109/21.44067</jrn>

<edb>Dyer C (2001) Volumetric scene reconstruction from multiple views. In 'Foundations of image understanding'. (Ed. LS Davis) pp. 469–489. (Springer: Boston, MA, USA)</edb>

<jrn>Evenson RE, Gollin D (2003) Assessing the impact of the Green Revolution, 1960 to 2000. *Science* **300**(5620), 758–762. doi:10.1126/science.1078710</jrn>

<jrn>Faaij A (2008) 'Bioenergy and global food security.' (WBGU: Utrecht, Berlin)</jrn>

<jrn>Fahlgren N, Gehan MA, Baxter I (2015) Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology* **24**, 93–99. doi:10.1016/j.pbi.2015.02.006</jrn>

<jrn>Furbank RT, Tester M (2011) Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science* **16**(12), 635–644. doi:10.1016/j.tplants.2011.09.005</jrn>

<jrn>Furbank RT, von Caemmerer S, Sheehy J, Edwards G (2009) C_4 rice: a challenge for plant phenomics. *Functional Plant Biology* **36**(11), 845–856. doi:10.1071/FP09185</jrn>

<jrn>Furukawa Y, Ponce J (2010) Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(8), 1362–1376. doi:10.1109/TPAMI.2009.161</jrn>

<other>Gaud WS (1968) 'The Green Revolution: accomplishments and apprehensions.' Discurso perante a Society for International Development. Available at http://www.agbioworld.org/biotech-info/topics/borlaug/borlaug-green.html [Verified 30 July 2016].</other>

<conf>Gibbs JA, *et al*. (2015) Three-dimensional reconstruction of plant shoots from multiple images using an active vision system. In 'Proceedings of the IROS workshop on agri-food robotics, Hamburg'. (Eds G Kootstra, Y Edan, E van Henten, M Bergerman) Available at https://agrifoodroboticsworkshop.com/accepted-papers/ [Verified 30 July 2016].</conf>

<jrn>Godin C (2000) Representing and encoding plant architecture: a review. *Annals of Forest Science* **57**(5), 413–438. doi:10.1051/forest:2000132</jrn>

<jrn>Hartmann A, Czauderna T, Hoffmann R, Stein N, Schreiber F (2011) HTPheno: an image analysis pipeline for high-throughput plant phenotyping. *BMC Bioinformatics* **12**(1), 148. doi:10.1186/1471-2105-12-148</jrn>

<jrn>Hirose T (2005) Development of the Monsi-Saeki theory on canopy structure and function. *Annals of Botany* **95**(3), 483–494. doi:10.1093/aob/mci047</jrn>

<bok>Horn BKP, Brooks, MJ (1989) 'Shape from shading.' (MIT Press: Cambridge, MA, USA)</bok>

<jrn>Hosoi F, Omasa K (2006) Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning LiDAR. *IEEE Transactions on Geoscience and Remote Sensing* **44**(12), 3610–3618. doi:10.1109/TGRS.2006.881743</jrn>

<jrn>Hosoi F, Omasa K (2009) Estimating vertical plant area density profile and growth parameters of

a wheat canopy at different growth stages using three-dimensional portable LiDAR imaging. *ISPRS Journal of Photogrammetry and Remote Sensing* **64**(2), 151–158. doi:10.1016/j.isprsjprs.2008.09.003</jrn>

<jrn>Illingworth J, Kittler J (1988) A survey of the Hough transform. *Computer Vision Graphics and Image Processing* **44**(1), 87–116. doi:10.1016/S0734-189X(88)80033-1</jrn>

<jrn>Ivanov N, Boissard P, Chapron M, Andrieu B (1995) Computer stereo plotting for 3-D reconstruction of a maize canopy. *Agricultural and Forest Meteorology* **75**(1–3), 85–102. doi:10.1016/0168-1923(94)02204-W</jrn>

<jrn>Jin J, Tang L (2009) Corn plant sensing using real-time stereo vision. *Journal of Field Robotics* **26**(6–7), 591–608. doi:10.1002/rob.20293</jrn>

<bok>Kang SB, Quan, L (2009) 'Image-based modeling of plants and trees.' (Morgan & Claypool Publishers: London)</bok>

<jrn>Karwowski R, Prusinkiewicz P (2003) Design and implementation of the L+C modeling language. *Electronic Notes in Theoretical Computer Science* **86**(2), 134–152. doi:10.1016/S1571-0661(04)80680-7</jrn>

<jrn>Kearney J (2010) Food consumption trends and drivers. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**(1554), 2793–2807. doi:10.1098/rstb.2010.0149</jrn>

<bok>Kender JR (1981) Shape from texture. Computer science technology report CMU-CS-81-102. Carnegie-Mellon University, Pittsburgh, PA, USA.</bok>

<edb>Khush GS (1996) Prospects of and approaches to increasing the genetic yield potential of rice. In 'Rice research in Asia: progress and priorities'. pp. 59–69. (CAB International: Wallingford, UK)</edb>

<bok>Killinger DK (2014) 'LiDAR (light detection and ranging) In 'Laser spectroscopy for sensing: fundamentals, techniques and applications'. pp. 292–312. (Elsevier Science: Amsterdam, The Netherlands)</bok>

<jrn>Kniemeyer O, Kurth K (2008) The modelling platform GroIMP and the programming language XL. *Applications of Graph Transformations with Industrial Relevance* **5088**, 570–572. doi:10.1007/978-3-540-89020-1_39</jrn>

<conf>Kumar P, Cai J, Miklavcic S (2012) High-throughput 3D modelling of plants for phenotypic analysis. In 'Proceedings of the 27th conference on image and vision computing New Zealand'. pp. 301–306. (ACM Press: New York)</conf>

<conf>Kumar P, Connor J, Mikiavcic S (2014) High-throughput 3D reconstruction of plant shoots for phenotyping. In '13th International conference on control automation robotics and vision (ICARCV)'. pp. 211–216.</conf>

<jrn>Kurth W (2007) Specification of morphological models with L-systems and relational growth

grammars. *Journal of Interdisciplinary Image*</jrn>

<jrn>Kutulakos KN, Seitz SM (2000) A theory of shape by space carving. *International Journal of Computer Vision* **38**(3), 199–218. doi:10.1023/A:1008191222954</jrn>

<jrn>Laga H, Miklavcic SJ (2013) Curve-based stereo matching for 3D modeling of plants. In ‘20th International congress on modelling and simulation, Adelaide, Australia, 1–6 December 2013’. pp. 524–520.</jrn>

<bok>Larcher W (2003) ‘Physiological plant ecology: ecophysiology and stress physiology of functional groups.’ (Springer-Verlag: Berlin)</bok>

<jrn>Laurentini A (1994) The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(2), 150–162. doi:10.1109/34.273735</jrn>

<jrn>Lauterbur PC (1973) Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature* **242**(5394), 190–191. doi:10.1038/242190a0</jrn>

<jrn>Lindenmayer A (1968) Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology* **18**(3), 280–299. doi:10.1016/0022-5193(68)90079-9</jrn>

<bok>Lintermann B, Deussen O (1996) ‘Interactive modelling of branching structures.’ (Plant International BV: Wageningen, The Netherlands)</bok>

<jrn>Livny Y, Yan F, Olson M, Chen B, Zhang H, El-Sana J (2010) Automatic reconstruction of tree skeletal structures from point clouds. *ACM Transactions on Graphics*, **29**(5):151:1–151:8.</jrn>

<jrn>Lobet G, Draye X, Périlleux C (2013) An online database for plant image analysis software tools. *Plant Methods* **9**(1), 38. doi:10.1186/1746-4811-9-38</jrn>

<conf>Lorensen WE, Cline HE, Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. In ‘Proceedings of the 14th annual conference on computer graphics and interactive techniques’. pp. 163–169. (ACM Press: New York)</conf>

<edb>Lou L, Liu Y, Han J, Doonan JH (2014) Accurate multi-view stereo 3D reconstruction for cost-effective plant phenotyping. In ‘Image analysis and recognition’. (Eds A Campilho, M Kamel) pp. 349–356. (Springer: Berlin)</edb>

<jrn>Lovell JL, Jupp DLB, Culvenor DS, Coops NC (2003) Using airborne and ground-based ranging LiDAR to measure canopy structure in Australian forests. *Canadian Journal of Remote Sensing* **29**(5), 607–622. doi:10.5589/m03-026</jrn>

<bok>Masry M, Lipson H (2007) A sketch-based interface for iterative design and analysis of 3D objects. In ‘SIGGRAPH ’07 ACM SIGGRAPH 2007 courses’. pp. 31. (ACM Press: New York)</bok>

<conf>McMillan L, Bishop G (1995) Plenoptic modeling. In ‘Proceedings of the 22nd annual conference on computer graphics and interactive techniques’. pp. 39–46. (ACM Press: New York)</conf>

<jrn>Minervini M, Fischbach A, Scharr H, Tsaftaris SA (2015) Finely-grained annotated datasets for image-based plant phenotyping. *Pattern Recognition Letters* doi:10.1016/j.patrec.2015.10.013</jrn>

<jrn>Mulayim AY, Yilmaz U, Atalay V (2003) Silhouette-based 3-D model reconstruction from multiple images. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics* **33**(4), 582–591. doi:10.1109/TSMCB.2003.814303</jrn>

<bok>Neubert B, Franken T, Deussen O (2007) Approximate image-based tree-modeling using particle flows. In ‘ACM SIGGRAPH 2007 papers on – SIGGRAPH ’07. Vol. 26’. p. 88. (ACM Press: New York)</bok>

<conf>Newcombe RA, *et al*. (2011) KinectFusion: real-time dense surface mapping and tracking. In ‘10th IEEE International symposium on mixed and augmented reality’. pp. 127–136. IEEE.</conf>

<jrn>Northend CA (1967) LiDAR, a laser radar for meteorological studies. *Naturwissenschaften* **54**(4), 77–80. doi:10.1007/BF00608760</jrn>

<jrn>Omasa K, Hosoi F, Konishi A (2007) 3D LiDAR imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany* **58**(4), 881–898. doi:10.1093/jxb/erl142</jrn>

<conf>Paproki A, *et al*. (2011) Automated 3D segmentation and analysis of cotton plants. In ‘International conference on digital image computing: techniques and applications’. pp. 555–560 IEEE.</conf>

<jrn>Paproki A, Sirault X, Berry S, Furbank R, Fripp J (2012) A novel mesh processing based technique for 3D plant analysis. *BMC Plant Biology* **12**(1), 63. doi:10.1186/1471-2229-12-63</jrn>

<jrn>Phattaralerphong J, Sinoquet H (2005) A method for 3D reconstruction of tree crown volume from photographs : assessment with 3D-digitized plants. *Tree Physiology* **25**(10), 1229–1242. doi:10.1093/treephys/25.10.1229</jrn>

<conf>Piccardi M (2004) Background subtraction techniques: a review. In ‘International conference on systems, man and cybernetics’. pp. 3099–3104.</conf>

<jrn>Pitkänen J, Maltamo M, Hyyppä J, Yu X (2004) Adaptive methods for individual tree detection on airborne laser based canopy height model. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(8), 187–191.</jrn>

<jrn>Pound MP, French AP, Murchie EH, Pridmore TP (2014) Automated recovery of three-dimensional models of plant shoots from multiple color images. *Plant Physiology* **166**(4), 1688–1698. doi:10.1104/pp.114.248971</jrn>

<conf>Preuksakarn C, *et al*. (2010) Reconstructing plant architecture from 3D laser scanner data. In ‘Proceedings of the 6th international workshop on functional–structural plant models’. pp. 12–17 </conf>

<bok>Prusinkiewicz P (2003) ‘Introduction to modeling with L-systems.’ (University of Calgary: Calgary, Canada)</bok>

<jrn>Prusinkiewicz P, Hanan J, Měch R (2000) An L-system-based plant modeling language. *Applications of Graph Transformations with Industrial Relevance* **1779**, 395–410. doi:10.1007/3-540-45104-8_31</jrn>

<jrn>Quan L, Tan P, Zeng G, Yuan L, Wang J, Kang SB (2006) Image-based plant modeling. *ACM Transactions on Graphics* **25**(3), 599. doi:10.1145/1141911.1141929</jrn>

<jrn>Rakocevic M (2000) Assessing the geometric structure of a white clover (*Trifolium repens* L.) canopy using3-D digitising. *Annals of Botany* **86**(3), 519–526. doi:10.1006/anbo.2000.1209</jrn>

<bok>Reche-Martinez A, Martin I, Drettakis G (2004) Volumetric reconstruction and interactive rendering of trees from photographs. In ‘ACM transactions on graphics (ToG). Vol. 23’. pp. 720–727.</bok>

<conf>Rutledge AM, Popescu SC 2006. Using LiDAR in determining forest canopy parameters. In ‘ASPRS 2006 annual conference’.</conf>

<bok>Sakaguchi T (1998) Botanical tree structure modeling based on real image set. In ‘ACM SIGGRAPH 98 conference abstracts and applications on – SIGGRAPH ’98’. p. 272. (ACM Press: New York)</bok>

<jrn>Salvi J, Armangué X, Batlle J (2002) A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition* **35**(7), 1617–1635. doi:10.1016/S0031-3203(01)00126-1</jrn>

<jrn>Scharr H, Minervini M, French AP, Klukas C, Kramer DM, Liu X, Luengo I, Pape J-M, Polder G, Vukadinovic D, Yin X, Tsaftaris SA (2016) Leaf segmentation in plant phenotyping: a collation study. *Machine Vision and Applications* **27**(4), 585–606. doi:10.1007/s00138-015-0737-3</jrn>

<jrn>Seitz SM, Dyer CR (1999) Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision* **35**(2), 151–173. doi:10.1023/A:1008176507526</jrn>

<jrn>Shlyakhter I, Rozenoer M, Dorsey J, Teller S (2001) Reconstructing 3D tree models from instrumented photographs. *IEEE Computer Graphics and Applications* **21**(3), 53–61. doi:10.1109/38.920627</jrn>

<edb>Shum H-Y, Kang SB (2000) A review of image-based rendering techniques. In ‘Visual communications and image processing 2000’. (Eds KN Ngan, T Sikora, M-T Sun) pp. 2–13. (International Society for Optics and Photonics)</edb>

<jrn>Sinoquet H, Rivet P (1997) Measurement and visualization of the architecture of an adult tree based on a three-dimensional digitising device. *Trees* **11**(5), 265. doi:10.1007/s004680050084</jrn>

<jrn>Sticklen MB (2007) Feedstock crop genetic engineering for alcohol fuels. *Crop Science* **47**(6), 2238. doi:10.2135/cropsci2007.04.0212</jrn>

<jrn>Su J, Wang Y, Liang D (2015) Long range detection of line-array multi-pulsed coding LiDAR by combining the accumulation coherence and subpixel-energy detection method. *Optics Express* **23**(12), 15174–15185. doi:10.1364/OE.23.015174</jrn>

<bok>Sutton MA, *et al*. (2011) 'The European nitrogen assessment: sources, effects and policy perspectives.' (Cambridge University Press: Cambridge, UK)</bok>

<jrn>Tan P, Yuan L, Wang J (2003) 'Image-based plant modeling overview of plant modeling system.' pp. 599–604. (ACM Press: New York)</jrn>

<jrn>Tan P, Zeng G, Wang J, Kang SB, Quan L (2007) Image-based tree modeling. *ACM Transactions on Graphics* **26**(3), 87. doi:10.1145/1276377.1276486</jrn>

<jrn>Tang S, Dong P, Buckles BP (2013) Three-dimensional surface reconstruction of tree canopy from LiDAR point clouds using a region-based level set method. *International Journal of Remote Sensing* **34**(4), 1373–1385. doi:10.1080/01431161.2012.720046</jrn>

<jrn>Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. *Science* **327**(5967), 818–822. doi:10.1126/science.1183700</jrn>

<jrn>Ullrich A, Pfennigbauer M (2011) Echo digitization and waveform analysis in airborne and terrestrial laser scanning. *Photogrammetric Week* **11**, 217–228.</jrn>

<jrn>Vadez V, Kholová J, Hummel G, Zhokhavets U, Gupta SK, Hash CT (2015) LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. *Journal of Experimental Botany* **66**(18), 5581–5593. doi:10.1093/jxb/erv251</jrn>

<jrn>Van Leeuwen M, Coops NC, Wulder MA (2010) Canopy surface reconstruction from a LiDAR point cloud using Hough transform. *Remote Sensing Letters* **1**(3), 125–132. doi:10.1080/01431161003649339</jrn>

<jrn>Vos J, Evers JB, Buck-Sorlin GH, Andrieu B, Chelle M, de Visser PH (2010) Functional-structural plant modelling: a new versatile tool in crop science. *Journal of Experimental Botany* **61**(8), 2101–2115. doi:10.1093/jxb/erp345</jrn>

<jrn>Wahl, S Winkelbach, FM (2001) Shape from 2D edge gradients. *Pattern Recognition* **•••**, 377–384.</jrn>

<jrn>Wang H, Zhang W, Zhou G, Yan G, Clinton N (2009) Image-based 3D corn reconstruction for retrieval of geometrical structural parameters. *International Journal of Remote Sensing* **30**(20), 5505–5513. doi:10.1080/01431160903130952</jrn>

<jrn>Watanabe T, Hanan JS, Room PM, Hasegawa T, Nakagawa H, Takahashi W (2005) Rice morphogenesis and plant architecture: measurement, specification and the reconstruction of structural development by 3D architectural modelling. *Annals of Botany* **95**(7), 1131–1143.

doi:10.1093/aob/mci136</jrn>

<conf>Weber J, Penn J (1995) Creation and rendering of realistic trees. In 'Proceedings of the 22nd annual conference on computer graphics and interactive techniques. Vol. 22'. pp. 119–128. (ACM Press: New York)</conf>

<jrn>White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA, French AN, Heun JT, Hunsaker DJ, Jenks MA, Kimball BA, Roth RL, Strand RJ, Thorp KR, Wall GW, Wang G (2012) Field-based phenomics for plant genetics research. *Field Crops Research* **133**, 101–112. doi:10.1016/j.fcr.2012.04.003</jrn>

<jrn>Woodham RJ (1989) Photometric method for determining surface orientation from multiple images. *Optical Engineering* **19**(1), 139–144.</jrn>

<conf>Zeng J, Zhang Y, Zhan S (2006) 3D tree models reconstruction from a single image. In 'Sixth international conference on intelligent systems design and applications. Vol. 2'. pp. 445–450. IEEE.</conf>

<jrn>Zhang Z (1998) Determining the epipolar geometry and its uncertainty : a review. *International Journal of Computer Vision* **27**(2), 161–195. doi:10.1023/A:1007941100561</jrn>

<conf>Zhao K, Popescu S (2007) Hierarchical watershed segmentation of canopy height model for multi-scale forest inventory. In 'Proceedings of the ISPRS working group'. pp. 436–442.</conf>

Manuscript received 4 May 2016, accepted 23 July 2016

Fig. 1. Three-dimensional modelling classification and uses for plant reconstruction shaded according to the key. (Best viewed in colour).

Fig. 2. 3D plant reconstruction using structure-from-motion (SFM); (*a*) one of the original images of the plant; (*b*) the point cloud generated by SFM; and (c) the final reconstructed model of the plant.

Table 1. Summary of advantages and disadvantages of methods for 3D plant reconstruction

Advantages	Disadvantages/challenges	Notes
	Shape-from-silhouette	
Easy to implement and use	Unable to deal with concavities	Applicable for simple non-occluded plants with no concavities. Best conducted in a controlled environment
Supports arbitrary view points	Quality depends on depth of data structure	
No calibration target required	Can fail to reconstruct crowded scenes	
–	Difficulties with thin surfaces	
	Space carving	
Easy to implement and use	Relies on photo consistent measures	Can deal with more complex plants than SFS but relies on photo consistent measures. Most suited for controlled environments and textured surfaces
Guarantees the entire object will be captured	Quality depends on depth of data structure	

Arbitrary viewpoints	Requires a bounding boxing is specified in advance	
No calibration target required	Can fail to reconstruct crowded scenes	
	Stereo vision	
Arbitrary viewpoints	Struggles with occlusions	Ability to reconstruct more complex plants but not well suited for high levels of occlusion. Most suited for controlled environments.
Ability to deal with concavities	Does not guarantee the entire object will be faithfully represented	
Can work on complex objects	Over/under sampling	
Affordable - requires only a standard handheld camera	Potentially high computational requirements	
–	Correspondence and parallax	
	Structure-from-motion	
Arbitrary viewpoints	Requires a calibration target	Suitable for complex plants and can deal with occlusions given an efficient image section strategy. Potential for field, but currently best suited for controlled environments
Ability to reconstruction complex objects	Over/under sampling	
Requires only a standard handheld camera	Potentially high computational requirements	
Deals with concavities	Does not guarantee the entire object will be faithfully represented	
–	Correspondence and parallax	
	LiDAR	
Can be deployed as both airborne and ground-based	Struggles with highly reflective surfaces	Suitable for moderately complex objects and is conducted in both controlled and field environments. More suitable for trees outdoors and would struggle with crops
Can handle concavities	Difficult to conduct under natural conditions (sunlight)	
Ability to reconstruct complex objects	Initial setup is still expensive	
No correspondence problem	Large computational requirements	
–	–	

1002