# Plant Image Analysis

## Fundamentals and Applications

Edited by
S. Dutta Gupta and Yasuomi Ibaraki

*chapter nine*

# Anatomofunctional bimodality imaging for plant phenotyping: An insight through depth imaging coupled to thermal imaging

**Yann Chéné, Étienne Belin, François Chapeau-Blondeau, Valérie Caffier, Tristan Boureau, and David Rousseau**

## Contents

## 9.1   Introduction

The shoots of the plants constitute challenging scenes for computer vision. Leaves, distributed with multiple orientations and various sizes, are spatially arranged along the three-dimensional (3D) networked architecture of the plants. When observed with imaging systems, the shoots of the plants form complex 3D textures with various illuminations, depths,

or color details. These have been shown, for instance, to display fractal properties with scale invariance in 2D projections (Ruderman and Bialek, 1994) or in 3D (Boudon et al., 2013). As a result, the detection, or the segmentation, of items in such complex scenes raises specific questions that may require selection, adaptation, or even the developments of new imaging systems or of new image processing algorithms. These rather fundamental informational concerns also meet the need for plant scientists to perform noninvasive, rapid, automatic observations on large populations of plants to confront phenotypic data and genomic data now available at high throughput. Among applications requiring high-throughput phenotyping of plants with computer vision is the segmentation of individual leaves on the entire plant (Biskup et al., 2007; Omasa et al., 2007; Klose et al., 2009; Kraft et al., 2010; Fiorani et al., 2012; Chéné et al., 2012). Different techniques of depth cameras have been used in the researches with a variety of observation scales, including small plants (Klose et al., 2009; Kraft et al., 2010; Fiorani et al., 2012), more structured shoots of entire plants (Chéné et al., 2012), and canopy (Biskup et al., 2007; Omasa et al., 2007). It would therefore be interesting to push forward the proofs of feasibility given in the works (Biskup et al., 2007; Omasa et al., 2007; Klose et al., 2009; Kraft et al., 2010; Fiorani et al., 2012; Chéné et al., 2012) by comparing how the depth cameras compete for leaf segmentation on the same plants. We propose to contribute in this direction in this chapter. Leaf segmentation from depth information gives access to shape measurement as illustrated in several reports (Biskup et al., 2007; Omasa et al., 2007; Klose et al., 2009; Kraft et al., 2010; Fiorani et al., 2012; Chéné et al., 2012), but as underlined in Fiorani et al. (2012) and Chéné et al. (2012), it is also possible to couple the anatomic information extracted from depth imaging with functional information like water content, chlorophyll efficiency, nutrient content, or pathogen presence at the scale of the leaf. Anatomofunctional imaging providing registered bimodal information has been developed for biomedical imaging with the anatomic information of, for instance, computed tomography coupled with the functional imaging of magnetic resonance imaging or positron emitted tomography. Such anatomofunctional imaging modalities have been applied to plant imaging, but their large-scale application is rather limited by their cost and low throughput (Jahnke et al., 2009; Fiorani et al., 2012). In this report we demonstrate how the association of depth imaging with an existing functional imaging system could constitute a new direction of instrumentation development for a variety of low-cost anatomofunctional bimodal imaging in plant phenotyping. The chapter is organized as follows. Section 9.2 gives a classification and a comparison of depth cameras. We then review in Section 9.3 the existing functional imaging available in plant science and illustrate with a study case the monitoring of the development of a pathogen, apple scab, on an entire plant at the scale of the leaf with functional information extracted

from thermal imaging and anatomic information extracted from depth imaging. Section 9.4 details a registration approach between the anatomic and functional images with examples on apple scab. Section 9.5 concludes and gives a set of perspective opened by this work.

## 9.2 Contactless optical depth cameras

3D scanning methods can be classified in terms of physical principles such as contact or contactless, destructive or nondestructive, reflection or transmission, and optical or nonoptical, as organized in Figure 9.1. In this report, we focus on depth cameras based on contactless optical 3D scanning methods. These cameras use active or passive methods. In passive methods, the depth is estimated from scene properties, sensor position, and sensor geometry in ambient natural light (see Ballan et al., 2009, for a recent overview). In active methods, a controlled light source illuminates the scene and depth is estimated from information included in the scene reflectance (Blais, 2004, for a review with some manufacturers links). The choice of a given depth camera depends on various parameters (Sansoni et al., 2009, for a methodological approach), including depth accuracy, depth range, experimental conditions (indoor or outdoor, with or without natural light), acquisition time (object in movement or not), object properties (textured or not), and the allocated budget. In this chapter, we propose to review the physical principles at work for the depth estimation and the resulting limitations for each method highlighted in Figure 9.1, and discuss their practical interest for leaf segmentation on plants.

### 9.2.1 Passive depth cameras

#### 9.2.1.1 Multiview

##### 9.2.1.1.1 Stereovision methods. Stereovision, also named computational stereovision, reconstructs scenes in depth from two or more different points of view taken with standard cameras. Figure 9.2a provides the scheme of a stereovision system with two cameras. For one real point $p$ in the scene, each camera gives a projection of this point in its image plane (pixels $pl$ and $pr$ in Figure 9.2a). The projected points are conjugated. Extrinsic parameters of the cameras (relative positions of the projected points in camera images, global system geometry) and intrinsic parameters of the cameras (focal length, distortion) allow us to determine the real 3D position of the projected points. Intrinsic and extrinsic parameters can be determined thanks to a calibration step (Heikkila, 2000; Sturm and Maybank, 1999; Tsai, 1987; Zhang et al., 1999; Bouguet, 2004). The matching of the conjugated points in different images (two or more), as explained in Nalpantidis et al. (2007) and Brown et al. (2003), can be solved with
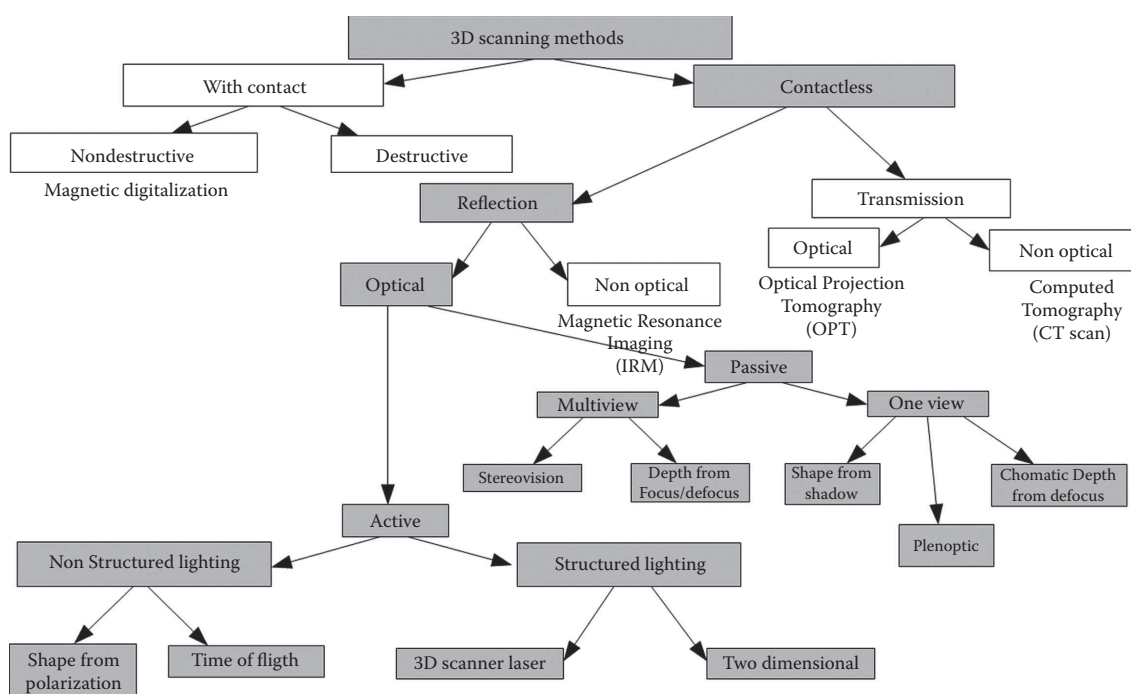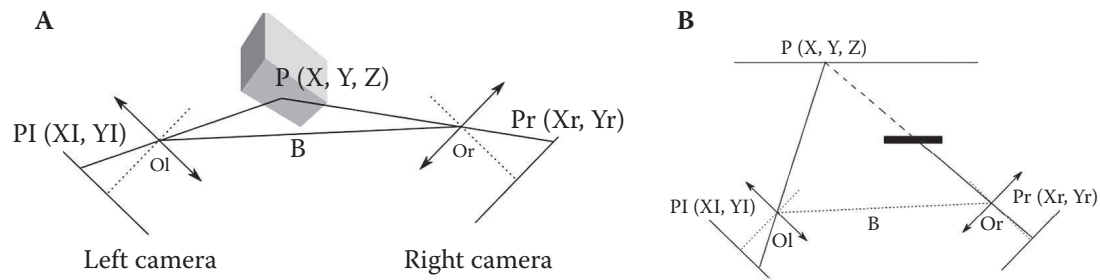
***Figure 9.1*** Nonexhaustive classification of depth imaging methods.

*Figure 9.2* The stereovision system and its possible occlusions. (a) Principle scheme of a stereovision system with two cameras. (b) Illustration of occlusions.

local methods or global methods. Local methods analyze the subwindow around the considered point (Bhat and Nayar, 1998; Birchfield and Tomasi, 1999; Kluth, et al., 1992). These methods are fast but not very accurate. Global methods (Birchfield and Tomasi, 1999; Tomasi and Manduchi, 1998; Boykov et al., 2001; Scharstein and Szeliski, 1998; Sun et al., 2003; Kutulakos and Steiz, 2000) are more accurate but slower than local methods. Certain scene points are visible with one camera but not with the other camera, as shown in Figure 9.2b. The matching algorithms alone cannot solve such occlusion problems. Occlusion points can be detected by different algorithms either before or after matching of nonoccluded points (Wildes, 1991; Fua, 1993; Silva and Santos Victor, 2000; Birchfield and Tomasi, 1999). The use of more than two cameras improves accuracy and reduces occlusion problems (Satoh and Ohta, 1996; Nakamura et al., 1996). Stereovision systems can be composed of low-cost cameras. The depth accuracy of a stereovision system is a function of the distance between the cameras (the baseline *b*) and of the camera resolutions, but can be very high when matching and occlusion problems are solved. Stereovision systems can be used in a field or greenhouse. The relative position of cameras must not be modified after the calibration steps. The frame rate of stereovision systems depends on the chosen algorithms, but it is currently rare to access real-time acquisition and display of the depth with stereovision.

*9.2.1.1.2  Depth from focus/defocus methods.*   Depth from focus methods and depth from defocus methods are based on a similar principle. As shown in Figure 9.3, an object point *P* of the scene is conjugated with the focus point $P_0$ across a lens. If the sensor plane is not superimposed with the focus plane, the image of *P* in the sensor plane is a disk of radius *r*. This translates in a blurry appearance in the image. The size of the disk depends on the focal length of the lens and on the distance *d* between the object plane and the lens. This blur can be modeled as a convolution by the point spread function (PSF). Depth from focus methods uses a measure of sharpness with translations of the object to determine its 3D shape. The object is set on a motorized translational stage (e.g., Minhas et al., 2009;
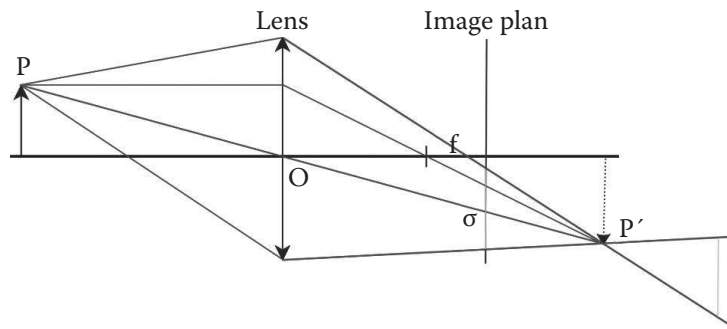
**Figure 9.3** Principle scheme of blur in camera.

Nayar and Nakagawa, 1994). At each increment of the translation stage, the vision sensor acquires an image that is blurred for points out of the focused plane and sharp for points in the focus plane. In each point of the object, blur appearance is like the application of a low-pass filter on the focused image. The quantification of focus on each point of the object in an image can be done by quantifying the amount of high spatial frequencies (Xiong and Shafer, 1993; Nayar and Nakagawa, 1994; Subbarao and Choi, 1995; Martinez Baena et al., 1997; Choi and Yun, 2000; Helmli and Scherer, 2001; Ahmad and Choi, 2005, 2007; Malik and Choi, 2007, 2008; Minhas et al., 2009). All these focus measures work on highly textured objects. In each object point, the evolution of focus is computed as a function of the position of the translation stage. The position of the translational stage corresponding to the maxima of the focus measure gives the depth of the considered point. In depth from focus methods, the vision sensor can be a low-cost webcam. Due to the translation stage, depth from focus systems are cumbersome and hardly usable in the field. In depth from defocus methods the lens parameters (aperture, focal) are changed so that the scene does not have to be translated. The defocus measure consists in approximating the PSF in each point of the scene. The PSF can be approximated only for textured subwindows and for objects out of field depth. For each scene point, the PSF approximation is done locally in subwindows with a statistical framework (Rajagopalan and Chaudhuri, 1999; Schechner and Kiryati, 1999; Farid and Simoncelli, 1998) or with deterministic optimization (Xiong and Shafer, 1995; Gokstorp, 1994; Favaro and Soatto, 2000; Favaro et al., 2003; Trouvé et al., 2011). Depth from defocus methods must use vision sensors with well-known parameters (aperture, focal). The use of low-cost webcams for these methods is therefore prohibited. The depth from defocus systems are suitable for use in greenhouse on highly textured and nonmovable plants.

### 9.2.1.2   *One view*

*9.2.1.2.1   Chromatic depth from defocus.* The chromatic depth from the defocus method (Trouvé et al., 2012) uses depth from defocus

principles, except that the different images required to estimate depth are acquired at the same time without changing lens parameters. It uses the association of color vision sensor and a chromatic lens. A color vision sensor uses a Bayer filter to have the red, green, and blue (RGB) values in each pixel. Each channel (red, green, and blue) of a color image represents the image of a definite wavelength interval. The focal length of a chromatic lens varies with wavelength. The association of a color vision sensor with a chromatic lens leads to three different blurred images, one for each channel. Then, for each pixel of each channel, the PSF can be locally approximated with depth from defocus methods. So, in each pixel there is a unique triplet for the estimation of the radius $r$. Finally, this triplet, obtained with only one acquisition, is used to estimate the depth of the scene at each pixel. Chromatic depth from defocus methods use a vision sensor associated with a chromatic lens, which is not low cost. The chromatic depth from defocus systems can be used in a field or greenhouse with plants contrasted in color and in texture.

*9.2.1.2.2   Shape from shading methods.*   The shape of a continuous object influences the reflectance of light. Shapes from shading methods are based on the perfect knowledge of the position of the camera. The relative intensity in each camera pixel allows us to retrieve the value of the shape gradient norm. The shape gradient is equivalent to normal coordinates in each object point. The object shape is the integration of normal coordinates. To solve shape from shading problems, several approaches can be used. Minimization approaches solve the problem with an optimization process based on minimization of several constraints (Szeliski, 1991; Courteille et al., 2004; Crouzil et al., 2003). Constraints can be on brightness, smoothness, and gradient intensity. There are also propagation approaches (Bichsel and Pentland, 1992; Dupuis and Oliensis, 1992; Rouy and Tourin, 1992). These methods use reference points with a known shape or with unique solutions, like singular points (points with intensity maximum). Then, there are the local approaches where the type of shape expected is known (spherical, cylindrical) (Lee and Rosenfeld, 1985; Hayakawa et al., 1994). The normal direction to the shape at each pixel is determined with derivatives. Finally, there are the linear approaches where the nonlinear problem is solved as a linear problem (Pentland, 1990; Ping Sing and Shah, 1994). These methods linearize the luminance image. Zhang et al. (1999) and Durou et al. (2008) give an overview of shape from shading methods with comparisons. Shape from shading methods can be implemented with a low-cost webcam. However, its depth accuracy is the lowest among all presented depth cameras. These methods are disrupted for textured objects and by additional light. Objects must be continuous, and measures can be done in a field or greenhouse.

*9.2.1.2.3  Plenoptic.*   The plenoptic (Ng et al., 2005; Lumsdaine and Georgiev, 2009) camera uses, like conventional vision system, a lens and a vision sensor. An additional array of microlenses is fixed before the vision sensor plane at a distance *f* corresponding to the focal length of each microlens of the sensor plane. The $n \times n$ sensor pixels matrix beyond each microlens allows us to capture the subimage provided by the microlens. In one acquisition of a scene, a plenoptic camera captures a stack of images for each sensor position of a conventional camera. This stack of focused images can be obtained with the Fourier slice photography (Ng, 2005) or the discrete focal stack transform from the global image (Nava et al., 2008). The scene depth can be processed, from the stack of focused images, with methods of depth from focus or depth from defocus (Bishop and Favaro, 2009; Lüke et al., 2009). Plenoptic systems are very costly. The depth accuracy depends on the level of texture of the scene. Plenoptic cameras can be used in a field or greenhouse.

## 9.2.2   *Active depth cameras*

### 9.2.2.1   *Structured lighting methods*

*9.2.2.1.1  Three-dimensional laser scanner.*   Three-dimensional laser scanners (Amann et al., 2001; Pfeifer and Briese, 2007) use a laser as an optical source. The most widespread methods are based on triangulation, time of flight (TOF), phase shift, and frequency modulation. Similar to stereovision principle, the laser scanner triangulation methods (Veatch and Davis, 1990; Fu et al., 2012; Dorsch et al., 1994; Buzinski et al., 1992) rely on geometric properties. One camera of Figure 9.2a is substituted by the laser. The laser point is emitted toward the scene and the camera visualizes its reflection position. The relative position between the laser and the camera is known. Thus, the distance between the laser scanner reference and the scene can be computed thanks to the position of the reflection on the plane of the camera. With this method, the wider the baseline, the smaller the error on *z*. The error on *z* also depends on the knowledge accuracy of the relative position between the laser and the camera. Instead of the laser point, this method also works with laser stripes. Each range of the sampling stripe is defined by its position on the image sensor. The TOF method (Hiskett et al., 2008; Myllylä et al., 1999) uses the time between the emission and the reception of a laser pulse. The laser pulse is emitted at a known time. Then, it is scattered by the scene and captured by the photodiode. Knowing the speed of light, the distance between the laser scanner and the point scene can be computed. The minimum measurable depth is limited by the length of the emitted pulse. The depth error depends on the accuracy of the time calculator. The phase shift method (Nakamura et al., 2000; Poujouly and Journet, 2000) uses the phase shift between the

emitted and the reflected laser wave. The emitted wave is continuous and its amplitude is modulated at a given frequency. The photodiode senses the reflected wave. Then, the phase difference is computed. The distance between laser scanner and the scene point is processed from this phase shift. The phase ambiguity, with a value more than $2\pi$, limits the maximum measurable depth. The frequency modulation method (Piracha et al., 2010; Agishev et al., 2006) uses, like the phase shift method, a continuous laser wave with amplitude modulation, but its frequency is also modulated. The evolution of the frequency in time is used to determine depth. The reflected wave has the same frequency profile as the emitted wave, except that it is shifted in time. This shift in time corresponds to the time of flight of the wave, so that the depth can be estimated. In all these methods, to obtain a depth map, the laser source must scan the scene. This scanning is, in most cases, done with the association of a mirror fixed on a mechanical system that changes its orientation. The depth sensor space resolution depends on this system. Three-dimensional laser scanners give the highest depth accuracy among all the depth cameras presented here. However, they are also very high-cost 3D sensors. They can be used in a field or greenhouse. Objects must be fixed. Objects' color or texture did not disrupt depth measurement. Also, the time required for the mechanical scan limits the rate of image acquisition, which cannot be performed at video rates.

*9.2.2.1.2  2D structured lighting.*   The use of a 2D structured lighting (see Geng, 2011; Salvi et al., 2010, for overviews) pattern allows us to scan the entire scene at the same time. In each case, the source of light is associated with a vision sensor. This sensor visualizes the deformation of the 2D structured lighting pattern emitted in the scene. The specific form of this pattern allows us to link each measure point of the emitted pattern with its conjugate in the reflected pattern. Finally, each scene depth is calculated by triangulation. One camera of Figure 9.2a is substituted by the 2D structured lighting pattern. The spatial resolution depends on the resolution of the emitted pattern. When scenes are static, it is possible to use sequential projections of 2D structured patterns. A method relies on the binary code principle. The association of all patterns allows us to code each stripe with a unique binary code (Ishii et al., 2007; Valkenburg and McIvor, 1998). However, a large number of patterns is needed to code a wide surface with a good spatial resolution. Thus, to decrease the number of required patterns, it is possible to add gray-level stripes in the code (Krattenthaler et al., 1994). Another method with the sequential projections is the phase shift method (Huang and Zhang, 2006). All patterns have a space sinusoidal amplitude modulation with a given frequency, but they are phase shifted. The 3D shape of the scene changes the phase of each pattern. The association of at least

three patterns allows us to retrieve the phase deformation due to the arc tangent function. Finally, the 3D shape is retrieved by comparison with a reference. The use of arc tangent induces the need for an unwrapping step and forbids scanning of an the surface with wide depth skip. To improve the frame rate of the depth sensor and follow the moved object, continuous stripe indexing patterns or grid indexing patterns can be used. In continuous patterns, each stripe is coded by one gray level (Durdle et al., 1998) or color. The use of colors allows us to improve the stripe number and the stripes' segmentation. The rainbow 3D camera uses a continuous color pattern (Geng, 1996). Each stripe is defined by one color of the rainbow. Otherwise, stripes can be indexed by their composition. All stripes have the same color, but each stripe is segmented in several substripes, which have different lengths (Maruyama and Abe, 1993). For each stripe, the segmentation is unique. Moreover, stripes can be indexed in terms of their neighbors (Zhang et al., 2002; Monks and Carter, 1993). This indexation is based on the De Bruijn sequence. A De Bruijn sequence of rank $n$ with $k$ different intensity levels or colors allows us to code $k^n$ stripes. In a De Bruijn sequence, each stripe is at the center of a unique $n$-length word. For stripe detection, two consecutive stripes cannot have the same color, so the number of possible stripes is less than $k^n$. A last method is the grid indexing. A grid is filled with a pseudorandom law. So each subwindow of size $n \times m$ is unique. This uniqueness allows us to link emitted points and reflected points, and there are $2^{nm}$ different possibilities with black and white dots (Le Moigne and Waxman, 1988; Ulusoy et al., 2009). To improve coding strength, each dot can be substituted by a codeword pattern. Codeword patterns can be coded in color (Desjardins and Payeur, 2007) or shape (Griffin et al., 1992). However, it is difficult to ensure uniqueness of each subwindow. As mentioned above, there are several types of 2D structured lighting methods. Each method has its own advantages and disadvantages. In this report, we consider a low-cost 2D lighting structured system that uses a near-infrared random dot grid. Its depth accuracy of typically 1 cm is good enough for certain plants, as illustrated in Chéné et al. (2012) and later in this chapter. This system can be disrupted by external infrared light. So its use in a greenhouse is prohibited. However, measures in the field can be done with object isolation relative to sunlight. Objects' color and texture do not disturb depth measurement, and objects can move.

### 9.2.2.2   *Nonstructured lighting*

*9.2.2.2.1   Time of flight (TOF) depth camera.*   Time of flight (TOF) depth cameras (Kolb et al., 2009; Lange and Steiz, 2001) allow us to obtain an intensity map and a depth map at the same time. To our knowledge, two methods are currently used to determine the scene depth: the

continuous modulation wave method and the optical shutter method. Both use near-infrared LEDs as a light source and a light sensor. In the continuous modulation wave method (Foix et al., 2011; Gokturk et al., 2004; Lange et al., 2000), the emitted wave is continuous and modulated in time. When light is reflected by the scene, the wave is disrupted and phase shifted in terms of the scene shape. The reflected wave is sensed, in each sensor pixel, with a phase shift due to the depth. Then, to find the phase shift in each pixel, the sensor computes a correlation between the emitted wave and the reflected wave. The distance between the TOF depth camera and the scene in each pixel results from the phase shift. The maximal measurable distance depends on the modulation frequency. In the optical shutter method (Iddan and Yahav, 2001; Yahav et al., 2007), the emitted wave is a short light pulse. The time of flight is limited by a faster shutter. This shutter is in front of the light sensor and is opened during a given time interval. So, the duration of this time interval sets the measurable depth interval. During the time interval, the reflected light comes in each pixel of the sensor. Thus, in each sensor pixel, the scene depth depends on the amount of light received by the pixel and the total amount of light received by all pixels. These systems can be used in a field or greenhouse. Objects' color and texture do not disturb depth measurement and objects can move.

### 9.2.3   Comparison

In this chapter, we use active depth cameras to constitute anatomofunctional imaging of plants. For illustration, we have considered the two plants of Figure 9.4 that we observed with four different depth cameras based on structured lighting methods. In this chapter the anatomical information of interest is the segmentation of the upper leaves of plants, accessible from the top view of the plant. As visible in Figures 9.5 and 9.6, the depth maps are similar, demonstrating that various choices of technologies can produce similar results in terms of anatomical information, the choice of the technologies depending on the cost or the
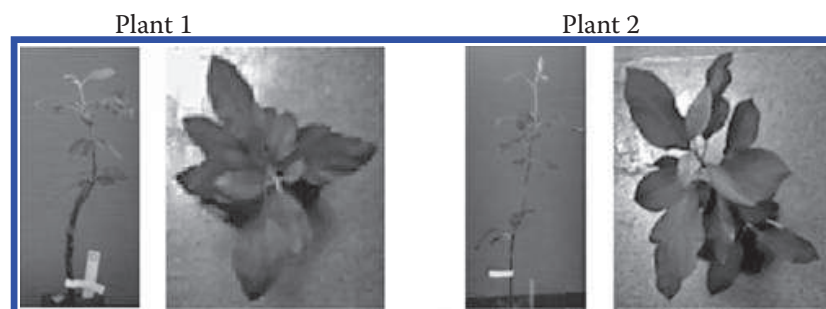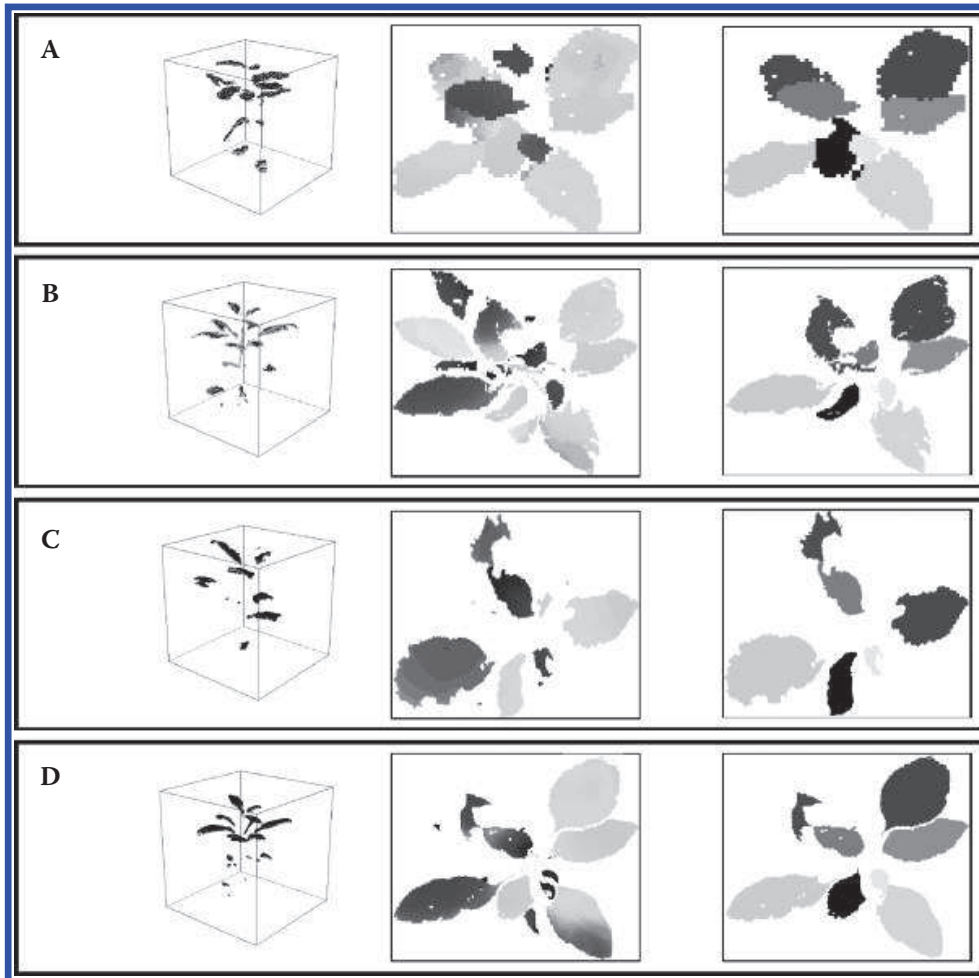


**Figure 9.4** Lateral and top views of the two plants used for depth camera comparisons.

*Figure 9.5* (**See color insert**.) Comparison of four depth cameras with the plant 1 of Figure 9.4. Each figure line is composed of the point cloud (left), the depth image (center), and the segmented image (right). In the image segmented with the algorithm of Chéné et al. (2012), only leaves presented in all depth images are represented with the same color for each depth camera. (a) Depth camera is a Microsoft Kinect that uses a pseudorandom dot grid as the structured lighting pattern. (b) Depth camera is a laser scanner that uses triangulation. (c) Depth camera is a hybrid depth camera that uses 2D structured lighting and stereovision (structured lighting pattern allows correspondences between two cameras). (d) Depth camera is a 2D structured lighting depth camera that uses a binary stripe sequential projection as the structured lighting pattern.

environmental constraints. A comparison of the four depth cameras used to produce Figures 9.5 and 9.6 is given in Table 9.1.

## 9.3   Functional imaging

We are now ready to consider the association of a functional imaging with the anatomical imaging reviewed in the previous section. In this report we illustrate anatomophysiological bimodality imaging with
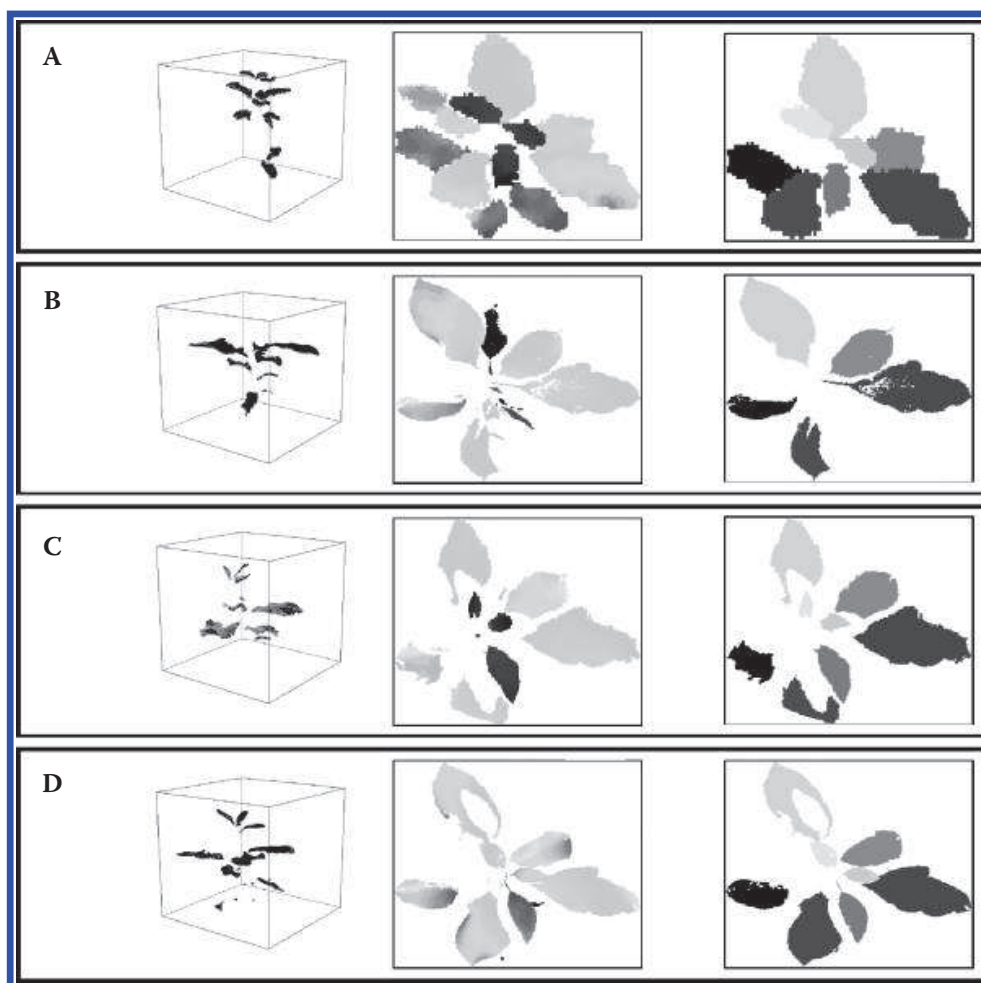
*Figure 9.6* (**See color insert**.) Same as in Figure 9.5 with plant 2 of Figure 9.4.

*Table 9.1* Characteristics of the Four Depth Cameras Used in Figures 9.5 and 9.6

|  | Kinect | Laser scanner | Hybrid method | 2D structured lighting method |
|---|---|---|---|---|
| Price (€) | 150 | 50,000 | 30,000 | 50,000 |
| Lighting type | Infrared | Visible | Visible | Visible |
| Depth resolution (cm) | 1 | 0.01 | 0.01 | 0.035 |
| Points number in cloud | 307,200 | 307,200 | 500,000 | 1,400,000 |
| Field of view (°) | 57 × 23 | Unknown | 23 × 21 | Unknown |
| Depth interval (cm) | 50–600 | 60–250 | 32.5–47.5 | 80 to unknown |
| Acquisition time (s) | 0.03 | 3 | 0.01 | 3.5 |
| Portability | + | − | + | − |
| Greenhouse used | No | Yes | Yes | Yes |

a problem of host-pathogen interaction detection on the entire plant. There are various existing imaging modalities providing functional information on plants. Several review articles have been proposed for general purposes (see, for instance, Fiorani et al., 2012, for recent state of the art) or for some specific purposes, like host-pathogen interaction (Chaerle and Straeten, 2001; Sankaran et al., 2010). Pathogens (virus, bacteria, fungus) can be indirectly detected by their impact on the plant physiology. The main functional imaging modalities reported for plants are thermography, chlorophyll fluorescence imaging, and hyperspectral imaging. These imaging modalities operate at distinct optical wavelengths and provide distinct information. Thermography, gazing around 10 µm, is linked to many structural and functional characteristics, such as leaf orientation, heat capacity, surface properties, infrared absorption, and transpiration rate. Chlorophyll fluorescence imaging (Papageorgiou, 2004) indirectly correlates to the amount of excited electrons in photosynthetic light reactions and is thus related to the photochemical and nonphotochemical charge separation in photosystems. Hyperspectral imaging gives access to various meaningful spectral regions, such as the transition from red to near infrared, the peak of green reflectance, and the water absorption bands around 970, 1600, and 2100 nm. For a given biological question, a comparison of the contrast given by each functional imaging is a preliminary task. We focus our attention in this report on the detection of apple scab. Apple scab, which is caused by the fungus *Venturia inaequalis*, requires more than 10 fungicide treatments per year to be controlled and can be considered the most serious disease for apple (Bowen et al., 2011). Quantifying the development of apple scab at the leaf scale is of major importance for studying the apple-scab interaction, as well as for analyzing the evolution of pathogenicity in *Venturia inaequalis* populations, and for breeding scab-resistant apple cultivars. Recent studies have undertaken the comparison of the detectability of apple scab with thermography, hyperspectral, and fluorescence chlorophyll (Delalieux et al., 2009; Oerke et al., 2011; Belin et al., 2013). Thermography has been demonstrated to provide a better detectability, i.e., higher contrast at earlier dates. Briefly, in the acquisition conditions described in Belin et al. (2013), a max-min difference of temperature greater than 0.21°C on a leaf gives a probability of good detection higher than 0.9 and a probability of false alarm less than 0.15. Such preliminary detection measurement can be done on a single leaf. To translate this detection diagnostic to an entire plant, it is then necessary to apply the detection test on images separating each of the individual leaves. In the following we therefore propose to associate a depth camera with thermal imaging to perform apple scab detection on an entire plant.

## 9.4   *Depth imaging coupled to thermal imaging*

The use of several imaging systems allows the combination of different scene information. To combine information, relative positions (extrinsic parameters) and internal properties of systems (intrinsic parameters) must be known. Intrinsic and extrinsic parameters of an imaging system can be found with a calibration method. The calibration of multiple color cameras has been widely studied (Heikkila, 2000; Zhang et al., 1999; Sturm and Maybank, 1999; Tsai, 1987; Bouguet, 2004). The same calibration procedure can be used to calibrate thermography (Lagüela et al., 2011; Luhmann et al., 2010), chlorophyll fluorescence imaging (Bellasio et al., 2012), and hyperspectral imaging. The calibration of a depth camera depends on the physical principle used to find the depth of the scene (Underwood et al., 2007; Sansoni et al., 2000; Krotkov, 1991; Antone and Friedman, 2007) for laser scanner calibration (Fuchs and Hirzinger, 2008; Kahlmann et al., 2006; Lichti, 2008; Lindner and Kolb, 2006), for TOF depth camera calibration (Chen et al., 2009; Li et al., 2008; Zhang et al., 2002; Zhou and Zhang, 2005), or for 2D structured lighting system calibration. Once the calibration of each imaging system is known, the association of both imaging systems can be done. The most widespread associations found in the literature are a depth camera with color camera (Zhang and Pless, 2004; Unnikrishnan and Hebert, 2005) for laser scanner depth cameras (Lindner et al., 2007; Kim et al., 2008), for TOF depth cameras (Herrera et al., 2012; Zhang and Zhang, 2011; Geiger et al., 2012; Fortenbury and Guerra Filho, 2012), and for a 2D structured lighting depth camera. The association of a depth camera and a physiological imaging has already been reported. In a biomedical context, Nagatani et al. (2012) and Skala et al. (2011) show the association of a thermal imaging respectively with a scanner laser and with a 2D structured lighting depth camera. In a plant science context, Bellasio et al. (2012) performs the association of a chlorophyll fluorescence imaging and a 2D structured lighting depth camera.

To perform the association of a 2D structured lighting depth camera with thermal imaging, we use the method presented in Herrera et al. (2012). The thermal imaging system and the depth camera can both be modeled by the pinhole model, also named the perspective projection model. Figure 9.7 gives the principle scheme of the pinhole model. Then, the model, described in Heikkila (2000), is used to fix intrinsic and extrinsic parameters of each imaging system. In the pinhole model, the transformation of a 3D real point $P_w$ in a 2D sensor pixel $p$ follows a homography with the projection center $O_s$. Extrinsic parameters allow us to pass from the world coordinates reference {$W$} to the system coordinates reference {$S$}. A point $P_w$ with 3D coordinates $[X_w, Y_w, Z_w]^T$ in the world coordinates
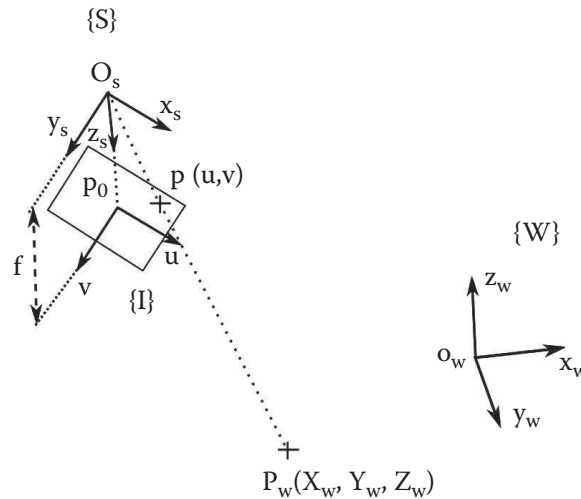
*Figure 9.7*  The principle scheme of the pinhole model.

reference $\{W\}$ becomes $P_s$ with the coordinates $[X_s, Y_s, Z_s]^T$ in the system coordinates reference $\{S\}$ following

$$P_s = RP_w + t \tag{9.1}$$

where $R$ is the $3 \times 3$ rotation matrix and $t$ is the $3 \times 1$ translation matrix. In the pinhole model, the extrinsic parameters are the rotation matrix $R$ and the translation matrix $t$.

Intrinsic parameters allow the transformation of $P_s$ in $p$ with 2D coordinates $[u, v]^T$ in the image coordinates reference $\{I\}$. First, the 3D point $P_s$ is transformed in 2D point $P_n$ by normalization, so coordinates of $P_n$ are

$$P_n = \left[ \begin{array}{c} X_n \\ Y_n \end{array} \right] = \left[ \begin{array}{c} \dfrac{X_s}{Z_s} \\ \dfrac{Y_s}{Z_s} \end{array} \right] \tag{9.2}$$

Next, owing to optical system distortions (radial and tangential), $P_n$ becomes $P_d$ such as

$$P_d = (1 + k_1 r^2 + k_2 r^4 + k_5 r^6)P_n + P_g \tag{9.3}$$

where $P_g = [2k_3 X_n Y_n + k_4 (r^2 + 2X_n^2), k_3 (r^2 + 2Y_n^2) + 2k_4 X_n Y_n]^T$, $r^2 = X_n^2 + Y_n^2$, and $k_i$, with $i = 1, \ldots, 5$, are the distortion coefficients. The pixel image of the 3D real point $P_w$ in the image coordinates reference $I$ is

$$p = \begin{bmatrix} f_{cx} & 0 \\ 0 & f_{cy} \end{bmatrix} P_d + P_0 \tag{9.4}$$

where $f_{cx}$ and $f_{cy}$ are the focal length in pixels in directions $x$ and $y$ and $P_0 = [u_0, v_0]^T$ is named the principal point (intersection between the optical axis $z_s$ and the image plane). In a pinhole model, intrinsic parameters are the focal length $f_c = [f_{cx}, f_{cy}]$, the principal point $P_0$, and the distortion vector $k = [k_1, \ldots, k_5]$. For depth cameras, depth calculation and depth distortion parameters can be added to intrinsic parameters. In the case of our 2D structured lighting depth camera, the depth is obtained from a disparity map. From Herrera et al. (2012), in each pixel, the disparity is distorted as follows:

$$d = d_k + D_\delta(u, v) \exp(\alpha_0 - \alpha_1 d) \tag{9.5}$$

with $d$ the undistorted disparity, $d_k$ the distorted disparity, $D_\delta(u, v)$ the spatial distortion pattern, and $\alpha = [\alpha_0, \alpha_1]$, which models the decay of the distortion effect. Then, the depth $z$ comes as

$$z = \frac{1}{C_1 d - C_0} \tag{9.6}$$

where $c = [c_0, c_1]$ is the depth calculation matrix. Thus, $\alpha$, $D_\delta(u, v)$, and $c$ are added to intrinsic parameters of the depth camera. Once intrinsic and extrinsic parameters are defined, the calibration of the anatomofunctional imaging system can be done. Figure 9.8 gives the principle scheme of coupling two imaging systems modeled by the pinhole model. The calibration step allows us to know the relation between imaging systems. Each calibration could be done individually, but the common calibration of both imaging systems gives better results. First, intrinsic and extrinsic parameters are initialized by use of independent calibrations. The physiological imaging individual calibration is done with Zhang et al. (1999). This calibration uses a specific pattern, a chessboard with known dimensions. The chessboard must be well contrasted in terms of the information delivered by the physiological imaging system. For the depth camera, intrinsic parameters are initialized with well-known manufacturer values. Next, the relative position between imaging systems is approximated by comparing the chessboard plane equation in both calibration images. Finally, intrinsic and extrinsic parameters are optimized by minimization of the measurement reprojection errors with all the parameters. A standard Levenberg-Marquardt algorithm was used to solve the nonlinear minimization problem.
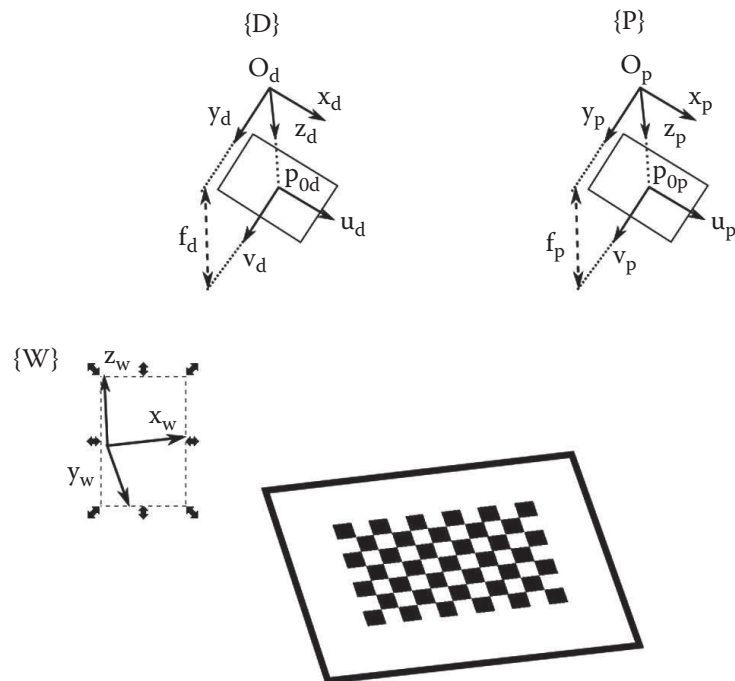
***Figure 9.8*** The principle scheme of coupling two imaging systems demonstrated by the pinhole model.

An example of an anatomofunctional imaging system is the coupling of our depth camera with a thermal imaging. Figure 9.9 (left) shows the rigid coupling. The whole system is calibrated with the method of Herrera et al. (2012). Figure 9.10 gives an example of the calibration of a scene. The calibration chessboard is in paper and sticks on a metallic plan. Use of a metallic plan allows us to have a well-contrasted calibration chessboard for thermal imaging, due to the strong difference in thermal reflectance between paper and metal. After calibration, the calibration file allows us
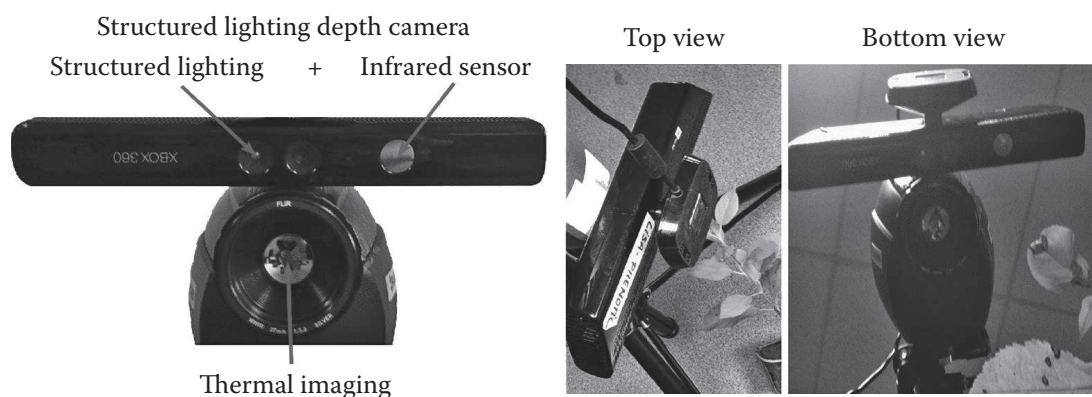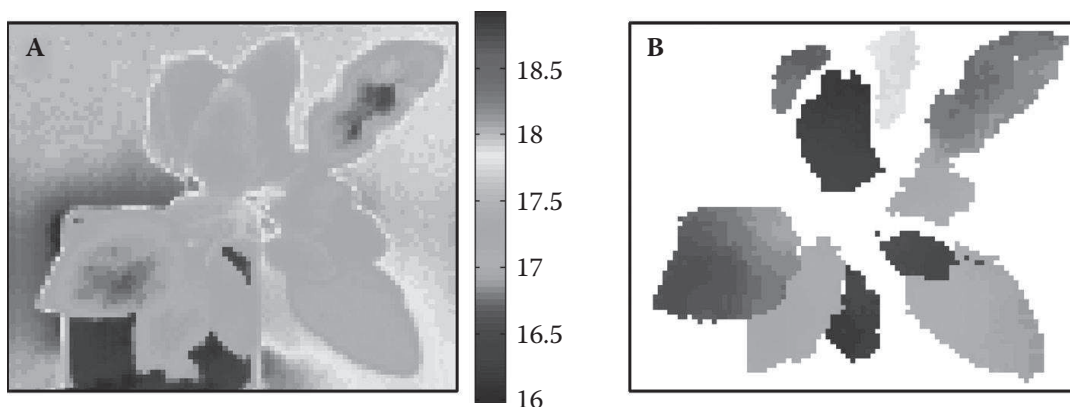


***Figure 9.9*** (**See color insert**.) Left: The anatomofunctional imaging system in this chapter is composed of a depth camera using the 2D structured lighting (binary dot pseudorandom grid) method and a thermal imaging. Right: An acquisition on a plant is done by one shot taken in plant top view.
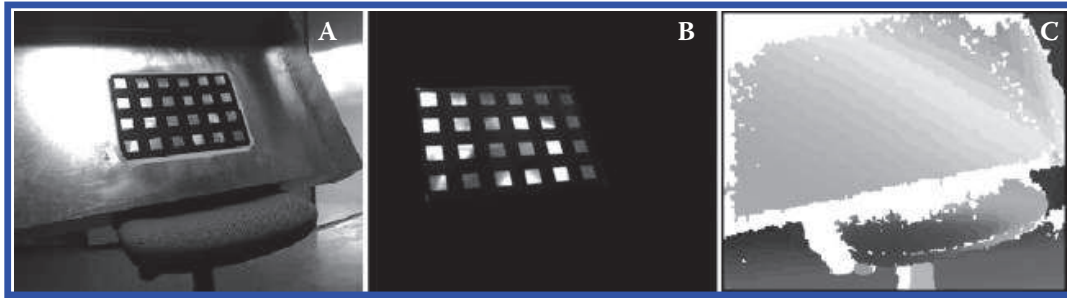
*Figure 9.10* (**See color insert**.) A calibration scene. (a) The RGB image of the calibration chessboard on the metallic plane. (b) The thermal image of the calibration chessboard on the metallic plan. (c) The disparity map of the calibration chessboard plan.

to convert disparity in depth expressed in meters and to register the depth information on the thermal image. We illustrate the anatomofunctional imaging system with the example of plant 2 of Figure 9.4. As shown in Figure 9.9 (right), an acquisition on a plant is done by one shot taken from the plant top view. Figure 9.11 shows the thermal image and the registered depth map of the plant top view. The segmented image of leaves, visible in Figure 9.12, is obtained by application of the depth segmentation algorithm of Chéné et al. (2012). Application of the edges of the leaves of the plant on the thermal image is given to visualize calibration results. In this image, each leaf is associated with a number. This number allows the creation of a mask for each plant leaf. Figure 9.13 shows the application of each leaf mask on the thermal imaging. Figure 9.14 gives individual thermal images of each segmented leaf of a plant. Each image is the result of application of the leaf mask on the thermal image. The calibration step gives a good registration of depth image on the thermal image. There



*Figure 9.11* (**See color insert**.) An anatomofunctional imaging composed of depth camera and a functional imaging allows us to have two registered images: the depth map and the functional image. In this chapter, the anatomofunctional imaging system is composed of (a) a thermal imaging providing a thermal image and (b) a depth camera providing a depth map.

*Figure 9.12* (**See color insert**.) Plant leaves are segmented by application of the depth segmentation algorithm of Chéné et al. (2012). Like the depth map, the segmented image is registered with the thermal image. In this image, each leaf is associated with a number. These numbers allow us to create a mask for each individual leaf.
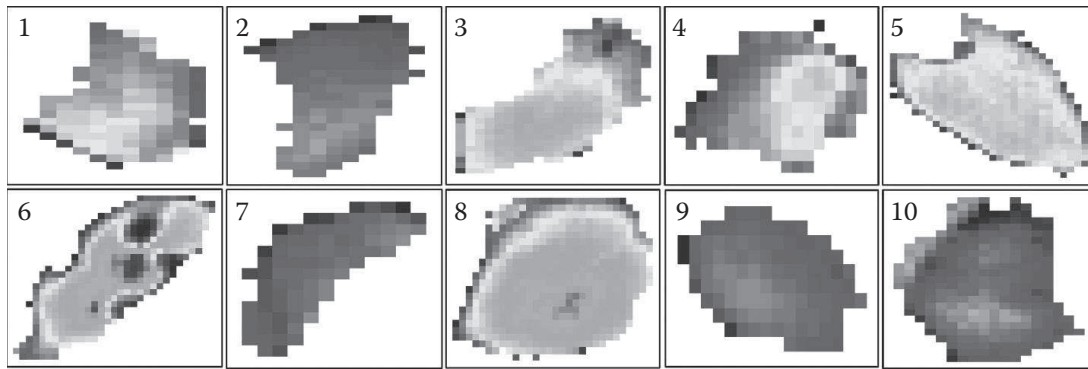


*Figure 9.13* (**See color insert**.) The edge of each leaf mask is applied on the thermal image. Colors of these edges are equal to the color of the leaf-segmented image of Figure 9.12. In this case only leaf 6 is hosting the pathogen.

are few errors on leaves' contour, visible in Figure 9.14, which are due to Kinect and calibration precision.

## 9.5   Conclusion

In this chapter, we have presented the potential of depth cameras for anatomic imaging modality of plants in association with functional imaging to deliver anatomofunctional imaging for plant science applications. We have detailed the various technologies of depth cameras currently available and discussed their advantages and limitations for plant imaging. Overall, a simple but useful message is that there is no single solution for plant phenotyping with depth cameras. The suitability of each technology depends on the biological question raised. We have illustrated this fact with a host-pathogen interaction where a thermal camera is associated with a depth camera to detect the presence or absence of a pathogen from a single top view image on each visible leaf of the upper part of the plant. We have provided practical information concerning the image registration procedure necessary to couple both the cameras. We also

*Figure 9.14*  (**See color insert**.) Thermal images of each individual leaf segmented from the plant. Numbers correspond to numbers in Figure 9.12. The image resolution is the functional image resolution.

demonstrated that for such a biological question, low-cost depth cameras can be well adapted if the image acquisition is under controlled light conditions. Anatomofunctional imaging similar to the one illustrated here could also serve in extracting information from the branching structure. However, it would require higher depth resolution and possibly multiple views to allow a full reconstruction of the branching structure of bushy plants. It is likely that the grail of an ultimate phenotyping system for such more demanding applications may not exist, and that there will be an opportunity for multiple technological approaches depending on the measurement constraints and the biological question raised.

## Acknowledgments

## References

Agishev, R., Gross, B., Moshary, F., Gilerson, A., and Ahmed, S., Range-resolved pulsed and CWFM lidars: potential capabilities comparison, *Appl. Phys. B Lasers Optics*, 85, 149–162, 2006.

Ahmad, M., and Choi, T., A heuristic approach for finding best focused shape, *IEEE Trans. Circ. Syst. Video Technol.*, 15, 566–574, 2005.

Ahmad, M., and Choi, T., Application of three dimensional shape from image focus in LCD/TFT displays manufacturing, *IEEE Trans. Consum. Electron.*, 53, 1–4, 2007.

Amann, M., Bosch, T., Lescure, M., Myllyla, R., and Rioux, M., Laser ranging: a critical review of usual techniques for distance measurement, *Opt. Eng.*, 40, 10–19, 2001.

Antone, M., and Friedman, Y., Fully automated laser range calibration, in *Proceedings of British Machine Vision Conference*, Warwick, UK, 2007, pp. 1–11.

Ballan, L., Brusco, N., and Cortelazzo, G., 3D content creation by passive optical methods, in *3D online multimedia and games: processing, visualization and transmission*, ed. I. Cheng, G.M. Cortelazzo, A. Basu, and S.K. Tripathi, 231–270, World Scientific Publishing, 2009.

Belin, E., Rousseau, D., Bourreau, T., and Caffier, V., Thermography versus chlorophyll fluorescence imaging for detection and quantification of apple scab, *Comput. Electron. Agric.*, 90, 159–163, 2013.

Bellasio, C., Olejníčková, J., Tesař, R., Šebela, D., and Nedbal, L., Computer reconstruction of plant growth and chlorophyll fluorescence emission in three spatial dimensions, *Sensors*, 12, 1052–1071, 2012.

Bhat, D., and Nayar, S., Ordinal measures for image correspondence, *IEEE Trans. Pattern Anal. Mach. Intell.*, 20, 415–423, 1998.

Bichsel, M., and Pentland, A., A simple algorithm for shape from shading, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 459–465.

Birchfield, S., and Tomasi, C., Depth discontinuities by pixel-to-pixel stereo, *Int. J. Comput. Vision*, 35, 269–293, 1999a.

Birchfield, S., and Tomasi, C., Multiway cut for stereo and motion with slanted surfaces, in *Proceedings of the 7th IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999b, pp. 489–495.

Bishop, T., and Favaro, P., Plenoptic depth estimation from multiple aliased views, in *IEEE 12th International Conference on Computer Vision*, 2009, pp. 1622–1629.

Biskup, B., Scharr, H., Schurr, U., and Rascher, U., A stereo imaging system for measuring structural parameters of plant canopies, *Plant Cell Environ.*, 30, 1299–1308, 2007.

Blais, F., Review of 20 years of range sensor development, *J. Electron. Imaging*, 13, 231–240, 2004.

Boudon, F., Godin, C., Puech, O., Pradal, C., and Sinoquet, H., Estimating the fractal dimension of plants using the two-surface method: an analysis based on 3D-digitized tree foliage, *Fractals*, 14, 149–163, 2006.

Bouguet, J., Camera calibration toolbox for MatLab, http://www.vision.caltech.edu/bouguetj/calib_doc/, 2004 (accessed August 27, 2013).

Bowen, J., Mesarich, C., Bus, V., Beresford, R., Plummer, K., and Templeton, M., *Venturia inaequalis*: the causal agent of apple scab, *Mol. Plant Pathol.*, 12, 105–122, 2011.

Boykov, Y., Veksler, O., and Zabih, R., Fast approximate energy minimization via graph cuts, *IEEE Trans. Pattern Anal. Mach. Intell.*, 23, 1222–1239, 2001.

Brown, M., Burschka, D., and Hager, G., Advances in computational stereo, *IEEE Trans. Pattern Anal. Mach. Intell.*, 25, 993–1008, 2003.

Buzinski, M., Levine, A., and Stevenson, W., Laser triangulation range sensors: a study of performance limitations, *J. Laser Appl.*, 4, 29–36, 1992.

Chaerle, L., and Straeten, D., *Venturia inaequalis*: the causal agent of apple scab, *Biochim. Biophys. Acta*, 1519, 153–166, 2001.

Chen, X., Xi, J., Jin, Y., and Sun, J., Accurate calibration for a camera-projector measurement system based on structured light projection, *Opt. Lasers Eng.*, 47, 310–319, 2009.

Chéné, Y., Belin, E., Rousseau, D., and Chapeau-Blondeau, F., Multiscale analysis of depth images from natural scenes: scaling in the depth of the woods, *Chaos Solit. Fractals*, 54, 134–149, 2013.

Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caffier, V., Morel, P., Belin, E., and Chapeau-Blondeau, F., On the use of depth camera for 3D phenotyping of entire plants, *Comput. Electron. Agric.*, 82, 122–127, 2012.

Choi, T., and Yun, J., Three-dimensional shape recovery from the focused-image surface, *Opt. Eng.*, 39, 1321–1326, 2000.

Courteille, F., Crouzil, A., Durou, J., and Gurdjos, P., Towards shape from shading under realistic photographic conditions, in *Proceedings of the 17th International Conference on Pattern Recognition*, 277–280, 2004.

Crouzil, A., Descombes, X., and Durou, J., A multiresolution approach for shape from shading coupling deterministic and stochastic optimization, *IEEE Trans. Pattern Anal. Mach. Intell.*, 25, 1416–1421, 2003.

Delalieux, S., Auwerkerken, A., Verstraeten, W., Somers, B., Valcke, R., Lhermitte, J., and Coppin, P., Hyperspectral reflectance and fluorescence imaging to detect scab induced stress in apple leaves, *Remote Sensing*, 1, 858–874, 2009.

Desjardins, D., and Payeur, P., Dense stereo range sensing with marching pseudo-random patterns, in *4th Canadian Conference on Computer and Robot Vision*, 2007, pp. 216–226.

Dorsch, R., Häusler, G., and Herrmann, J., Laser triangulation: fundamental uncertainty in distance measurement, *Appl. Opt.*, 33, 1306–1314, 1994.

Dupuis, P., and Oliensis, J., Direct method for reconstructing shape from shading, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 453–458.

Durdle, N., Thayyoor, J., and Raso, V., An improved structured light technique for surface reconstruction of the human trunk, in *IEEE Canadian Conference on Electronics and Computer Engineering*, 1998, pp. 874–877.

Durou, J., Falcone, M., and Sagona, M., Numerical methods for shape-from-shading: a new survey with benchmarks, *Comput. Vision Image Underst.*, 109, 22–43, 2008.

Farid, H., and Simoncelli, E., Range estimation by optical differentiation. *J. Opt. Soc. Am. A*, 15, 1777–1786, 1998.

Favaro, P., Mennucci, A., and Soatto, S., Observing shape from defocused images, *Int. J. Comput. Vision*, 52, 25–43, 2003.

Favaro, P., and Soatto, S., Shape and radiance estimation from the information divergence of blurred images, in *European Conference on Computer Vision*, 2000, pp. 153–158.

Fiorani, F., Rascher, U., and Jahnke, S., Imaging plants dynamics in heterogenic environments, *Curr. Opin. Biotechnol.*, 23, 227–235, 2012.

Foix, S., Alenya, G., and Torras, C., Lock-in time-of-flight (ToF) cameras: a survey, *IEEE Sens. J.*, 11, 1917–1926, 2011.

Fortenbury, B., and Guerra Filho, G., Robust 2D/3D calibration using RANSAC registration, *Adv. Visual Comput.*, 7431, 179–188, 2012.

Fu, G., Menciassi, A., and Dario, P., Development of a low-cost active 3D triangulation laser scanner for indoor navigation of miniature mobile robots, *Robotics Auton. Syst.*, 60, 1317–1326, 2012.

Fua, P., A parallel stereo algorithm that produces dense depth maps and preserves image features, *Machine Vision Appl.*, 6, 35–49, 1993.

Fuchs, S., and Hirzinger, G., Extrinsic and depth calibration of ToF-cameras, in *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, AK, 2008, pp. 1–6.

Geiger, A., Moosmann, F., Car, O., and Schuster, B., Automatic camera and range sensor calibration using a single shot, in *IEEE International Conference on Robotics and Automation*, St. Paul, MN, 2012, pp. 3936–3943.

Geng, J., Structured-light 3D surface imaging: a tutorial, *Adv. Opt. Photon.*, 3, 128–160, 2011.

Geng, Z., Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems, *Opt. Eng.*, 35, 376–383, 1996.

Gokstorp, M., Computing depth from out-of-focus blur using a local frequency representation, in *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Conference A: Computer Vision and Image Processing*, 1994, pp. 153–158.

Gokturk, S., Yalcin, H., and Bamji, C., A time-of-flight depth sensor—system description, issues and solutions, in *Conference on Computer Vision and Pattern Recognition*, 2004, pp. 35–35.

Griffin, P., Narasimhan, L., and Yee, S., Generation of uniquely encoded light patterns for range data acquisition, *Pattern Recogn.*, 25, 609–616, 1992.

Hayakawa, H., Nishida, S., Wada, Y., and Kawato, M., A computational model for shape estimation by integration of shading and edge information, *Neural Networks*, 7, 1193–1209, 1994.

Heikkila, J., Geometric camera calibration using circular control points, *IEEE Trans. Pattern Anal. Mach. Intell.*, 22, 1066–1077, 2000.

Helmli, F., and Scherer, S., Adaptive shape from focus with an error estimation in light microscopy, in *Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis*, 2001, pp. 188–193.

Herrera, C., Kannala, J., and Heikkilä, J., Joint depth and color camera calibration with distortion correction, *IEEE Trans. Pattern Anal. Mach. Intell.*, 34, 2058–2064, 2012.

Hiskett, P., Parry, C., McCarthy, A., and Buller, G., A photon-counting time-of-flight ranging technique developed for the avoidance of range ambiguity at gigahertz clock rates, *Opt. Expr.*, 16, 13685–13698, 2008.

Huang, P., and Zhang, S., Fast three-step phase-shifting algorithm, *Appl. Opt.*, 45, 5086–5091, 2006.

Iddan, G., and Yahav, G., Three-dimensional imaging in the studio and elsewhere, in *Proceedings of the SPIE Three-Dimensional Image Capture and Applications IV*, 2001, pp. 48–55.

Ishii, I., Yamamoto, K., and Tsuji, T., High-speed 3D image acquisition using coded structured light projection, in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 925–930.

Jahnke, S., Menzel, M., Van Dusschoten, D., Roeb, G., Buhler, J., Minwuyelet, S., Blumler, P., Temperton, V., Hombach, T., and Streun, M., Combined MRI–PET dissects dynamic changes in plant structures and functions, *Plant J.*, 59, 634–644, 2009.

Kahlmann, T., Remondino, F., and Ingensand, H. Calibration for increased accuracy of the range imaging camera SwissRanger™, in *International Society for Photogrammetry and Remote Sensing*, Dresden, Germany, 2006, pp. 136–141.

Kim, Y., Chan, D., Theobalt, C., and Thrun, S., Design and calibration of a multi-view TOF sensor fusion system, in *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, AK, 2008, pp. 1–7.

Klose, R., Penlington, J., and Ruckelshausen, A., Usability study of 3D time-of-flight cameras for automatic plant phenotyping, *Bornimer Agrartechnische Berichte*, 69, 93–105, 2009.

Kluth, V., Kunkel, G., and Rauhala, U., Global least squares matching, in *Geoscience and Remote Sensing Symposium*, Houston, TX, 1992, pp. 1615–1618.

Kolb, A., Barth, E., Koch, R., and Larsen, R., Time-of-flight sensors in computer graphics, *Comput. Graph. Forum*, 29, 141–159, 2010.

Kraft, M., Saloma De Freitag, N., and Munack, A., Test of a 3D time of flight camera for shape measurements of plants, in *CIGR Workshop on Image Analysis in Agriculture*, 2010, pp. 108–116.

Krattenthaler, W., Mayer, K.J., and Duwe, H.P., 3D-surface measurement with coded light approach, in *Proceedings of the 17th Meeting of the Austrian Association for Pattern Recognition on Image Analysis and Synthesis*, Munich, Germany, 1993, pp. 103–114.

Krotkov, E., Laser rangefinder calibration for a walking robot, *in IEEE International Conference on Robotics and Automation*, Sacramento, CA, 1991, pp. 2568–2573.

Kutulakos, K., and Seitz, S., A theory of shape by space carving, *Int. J. Comput. Vision*, 38, 199–218, 2000.

Lagüela, S., González Jorge, H., Armesto, J., and Herráez, J., High performance grid for the metric calibration of thermographic cameras, *Meas. Sci. Technol.*, 23, 15402–15410, 2011.

Lange, R., and Seitz, P., Solid-state time-of-flight range camera, *IEEE J. Quantum Electron.*, 37, 390–397, 2001.

Lange, R., Seitz, P., Biber, A., and Lauxtermann, S., Demodulation pixels in CCD and CMOS technologies for time-of-flight ranging, in *Proceedings of SPIE*, 2001, pp. 177–188.

Le Moigne, J., and Waxman, A., Structured light patterns for robot mobility, *IEEE J. Robot. Autom.*, 4, 541–548, 1988.

Lee, C., and Rosenfeld, A., Improved methods of estimating shape from shading using the light source coordinate system, *Art. Intell.*, 26, 125–143, 1985.

Li, Z., Shi, Y., Wang, C., and Wang, Y., Accurate calibration method for a structured light system, *Opt. Eng.*, 47, 053604, 1–9, 2008.

Lichti, D., Self-calibration of a 3D range camera, *Int. Arch. Photogram Rem. Sens. Spatial Inform. Sci.*, 37, 927–932, 2008.

Lindner, M., and Kolb, A., Lateral and depth calibration of PMD-distance sensors, in *Advances in Visual Computing, 2nd International Symposium, ISVC*, Lake Tahoe, CA, 2006, pp. 524–533.

Lindner, M., Kolb, A., and Hartmann, K., Data-fusion of PMD-based distance information and high-resolution RGB-images, in *International Symposium on Signals, Circuits and Systems*, Lasi, Romania, 2007, pp. 1–4.

Luhmann, T., Ohm, J., Piechel, J., and Roelfs, T., Geometric calibration of thermographic cameras, *Int. Arch. Photogram Rem. Sens. Spatial Inform. Sci.*, 38, 411–416, 2010.

Lüke, J., Pérez Nava, F., Marichal Hernández, J., Rodríguez Ramos, J., and Rosa, F., Near real-time estimation of super-resolved depth and all-in-focus images from a plenoptic camera using graphics processing units, *Int. J. Digit. Multimed. Broadcast.*, 2010, 12–15, 2009.

Lumsdaine, A., and Georgiev, T., The focused plenoptic camera, in *IEEE International Conference on Computational Photography*, 2009, pp. 1–8.

Malik, A., and Choi, T., Application of passive techniques for three dimensional cameras, *IEEE Trans. Cons. Electron.*, 53, 258–264, 2007.

Malik, A., and Choi, T., A novel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise, *Pattern Recogn.*, 41, 2200–2225, 2008.

Martinez Baena, J., Fdez Valdivia, J., and Garcia, J., A multi-channel autofocusing scheme for gray-level shape scale detection, *Pattern Recogn.*, 30, 1769–1786, 1997.

Maruyama, M., and Abe, S., Range sensing by projecting multiple slits with random cuts, *IEEE Trans. Pattern Analysis Mach. Intell.*, 15, 647–651, 1993.

Minhas, R., Mohammed, A., Wu, Q., and Sid Ahmed, M., 3D shape from focus and depth map computation using steerable filters, in *Proceedings of the 6th International Conference on Image Analysis and Recognition*, 2009, pp. 573–583.

Monks, T., and Carter, J., Improved stripe matching for color encoded structured light, in *5th International Conference on Computer Analysis of Images and Patterns*, 476, 1993.

Myllylä, R., Marszalec, J., Kostamovaara, J., M̈antyniemi, A., and Ulbrich, G., Imaging distance measurements using TOF lidar, *J. Opt.*, 29, 188, 1999.

Nagatani, K., Otake, K., and Yoshida, K., Three-dimensional thermography mapping for mobile rescue robots, presented at *8th International Conference on Field and Service Robotics*, Matsushimas, Japan, 2012.

Nakamura, K., Hara, T., Yoshida, M., Miyahara, T., and Ito, H., Optical frequency domain ranging by a frequency-shifted feedback laser, *IEEE J. Quantum Electron.*, 36, 305–316, 2000.

Nakamura, Y., Matsuura, T., Satoh, K., and Ohta, Y., Occlusion detectable stereo-occlusion patterns in camera matrix, in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 1996, pp. 371–378.

Nalpantidis, L., Sirakoulis, G., and Gasteratos, A., Review of stereo matching algorithms for 3D vision, in *Proceedings of the 16th International Symposium on Measurement and Control in Robotics*, Warsaw, Poland, 2007, pp. 116–124.

Nava, F., Marichal Hernández, J., and Rodríguez Ramos, J., The discrete focal stack transform, in *Proceedings of the 16th European Signal Proceedings Conference*, 25–106, 2008.

Nayar, S., and Nakagawa, Y., Shape from focus, *IEEE Trans. Pattern Anal. Mach. Intell.*, 16, 824–831, 1994.

Ng, R., Fourier slice photography, *ACM Trans. Graph.*, 24, 735–744, 2005.

Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., and Hanrahan, P., Light field photography with a hand-held plenoptic camera, *Comput. Sci. Tech. Rep. CSTR*, 2, 1–11, 2005.

Oerke, E., Fröhling, P., and Steiner U., Thermographic assessment of scab disease on apple leaves, *Prec. Agric.*, 12, 699–715, 2011.

Omasa, K., Hosoi, F., and Konishi, A., 3D lidar imaging for detecting and understanding plant responses and canopy structure, *J. Exp. Bot.*, 158, 881–898, 2007.

Papageorgiou, G.G., *Chlorophyll a fluorescence: a probe of photosynthesis*, Kluwer Academic, Dordrecht, The Netherlands, 2004.

Pentland, A., Linear shape from shading, *Int. J. Comput. Vision*, 4, 153–162, 1990.

Pfeifer, N., and Briese, C., Laser scanning—principles and applications, presented at *3rd International Exhibition and Scientific Congress on Geodesy, Mapping, Geology, Geophisics*, Cadaster GEO-SIBERIA, 2007.

Ping Sing, T., and Shah, M., Shape from shading using linear approximation, *Image Vision Comput.*, 12, 487–498, 1994.

Piracha, M., Nguyen, D., Mandridis, D., Yilmaz, T., Ozdur, I., Ozharar, S., and Delfyett, P., Range resolved lidar for long distance ranging with sub-millimeter resolution, *Opt. Expr.*, 18, 7184–7189, 2010.

Poujouly, S., and Journet, B., Laser range finding by phase-shift measurement: moving towards smart systems, in *Proceedings of the SPIE Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology*, 2000, pp. 152–160.

Rajagopalan, A., and Chaudhuri, S., Simultaneous depth recovery and image restoration from defocused images, in *IEEE Conference on Computer Vision and Pattern Recognition*, 1999, pp. 1348–1353.

Rouy, E., and Tourin, A., A viscosity solutions approach to shape-from-shading, *SIAM J. Numer. Anal.*, 29, 867–884, 1992.

Ruderman, D.L., and Bialek, W., Statistics of natural images: scaling in the woods, *Phys. Rev. Lett.*, 73, 814–817, 1994.

Salvi, J., Fernandez, S., Pribanic, T., and Llado, X., A state of the art in structured light patterns for surface profilometry, *Pattern Recogn.*, 43, 2666–2680, 2010.

Sankaran, S., Mishra, A., Ehsani, R., and Davis, C., A review of advanced techniques for detecting plant diseases, *Comput. Electron. Agric.*, 59, 634–644, 2010.

Sansoni, G., Carocci, M., and Rodella, R., Calibration and performance evaluation of a 3D imaging sensor based on the projection of structured light, *IEEE Trans. Instrum. Meas.*, 49, 628–636, 2000.

Sansoni, G., Trebeschi, M., and Docchio, F., State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation, *Sensors*, 9, 568–601, 2009.

Satoh, K., and Ohta, Y., Occlusion detectable stereo-systematic comparison of detection algorithms, in *Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, Austria, 1996, pp. 280–286.

Scharstein, D., and Szeliski, R., Stereo matching with nonlinear diffusion, *Int. J. Comput. Vision*, 28, 155–174, 1998.

Schechner, Y., and Kiryati, N., The optimal axial interval in estimating depth from defocus, in *Proceedings of the 7th IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, pp. 843–848.

Silva, C., and Santos Victor, J., Intrinsic images for dense stereo matching with occlusions, *Comput. Vision*, 1842, 100–114, 2000.

Skala, K., Lipić, T., Sović, I., Gjenero, L., and Grubišić, I., 4D thermal imaging system for medical applications, *Period. Biol.*, 113, 407–416, 2011.

Sturm, P., and Maybank, S., On plane-based camera calibration: a general algorithm, singularities, applications, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, CO, 1999, pp. 432–437.

Subbarao, M., and Choi, T., Accurate recovery of three-dimensional shape from image focus, *IEEE Trans. Pattern Anal. Mach. Intell.*, 17, 266–274, 1995.

Sun, J., Zheng, N., and Shum, H., Stereo matching using belief propagation, *IEEE Trans. Pattern Anal. Mach. Intell.*, 25, 787–800, 2003.

Szeliski, R., Fast shape from shading, *Comput. Vision Graphics Image Proc.*, 53, 129–153, 1991.

Tomasi, C., and Manduchi, R., Stereo matching as a nearest-neighbor problem, *IEEE Trans. Pattern Anal. Mach. Intell.*, 20, 333–340, 1998.

Trouvé, P., Champagnat, F., Le Besnerais, G., Druart, G., and Idier, J., Chromatic depth from defocus: a theoretical and experimental performance study, in *Computational Optical Sensing and Imaging*, CM3B.3, 2012.

Trouvé, P., Champagnat, F., Le Besnerais, G., and Idier, J., Single image local blur identification, in *18th IEEE International Conference on Image Processing*, 2011, pp. 613–616.

Tsai, R., A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE J. Robotics Autom.*, 3, 323–344, 1987.

Ulusoy, A., Calakli, F., and Taubin, G., One-shot scanning using de Bruijn spaced grids, in *IEEE 12th International Conference on Computer Vision Workshops*, 2009, pp. 1786–1792.

Underwood, J., Hill, A., and Scheding, S., Calibration of range sensor pose on mobile platforms, in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, CA, 2007, pp. 3866–3871.

Unnikrishnan, R., and Hebert, M., *Fast extrinsic calibration of a laser rangefinder to a camera*, Robotics Institute Technical Report, 2005, http://repository.cmu.edu/robotics/339/ (accessed August 27, 2013).

Valkenburg, R., and McIvor, A., Accurate 3D measurement using a structured light system, *Image Vision Comput.*, 16, 99–110, 1998.

Veatch, P., and Davis, L., Efficient algorithms for obstacle detection using range data, *Comput. Vision Graphics Image Proc.*, 50, 50–74, 1990.

Wildes, R., Direct recovery of three-dimensional scene geometry from binocular stereo disparity, *IEEE Trans. Pattern Anal. Mach. Intell.*, 13, 761–774, 1991.

Xiong, Y., and Shafer, S., Depth from focusing and defocusing, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1993, pp. 68–73.

Xiong, Y., and Shafer, S., Moment filters for high precision computation of focus and stereo, in *IEEE/RSJ International Conference on Intelligent Robots and Systems, Human Robot Interaction and Cooperative Robots*, 1995, pp. 108–113.

Yahav, G., Iddan, G., and Mandelboum, D., 3D imaging camera for gaming application, in *International Conference on Consumer Electronics, Digest of Technical Papers*, 2007, pp. 1–2.

Zhang, C., and Zhang, Z., Calibration between depth and color sensors for commodity depth cameras, in *IEEE International Conference on Multimedia and Expo*, 2011, pp. 1–6.

Zhang, L., Curless, B., and Seitz, S., Rapid shape acquisition using color structured light and multi-pass dynamic programming, in *Proceedings of the 1st International Symposium on 3D Data Processing, Visualization, and Transmission*, 2002, pp. 24–36.

Zhang, Q., and Pless, R., Extrinsic calibration of a camera and laser range finder (improves camera calibration), in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2301–2306, 2004.

Zhang, R., Tsai, P., Cryer, J., and Shah, M., Shape-from-shading: a survey, *IEEE Trans. Pattern. Anal. Mach. Intell.*, 21, 690–706, 1999.

Zhou, F., and Zhang, G., Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations, *Image Vision Comput.*, 23, 59–67, 2005.
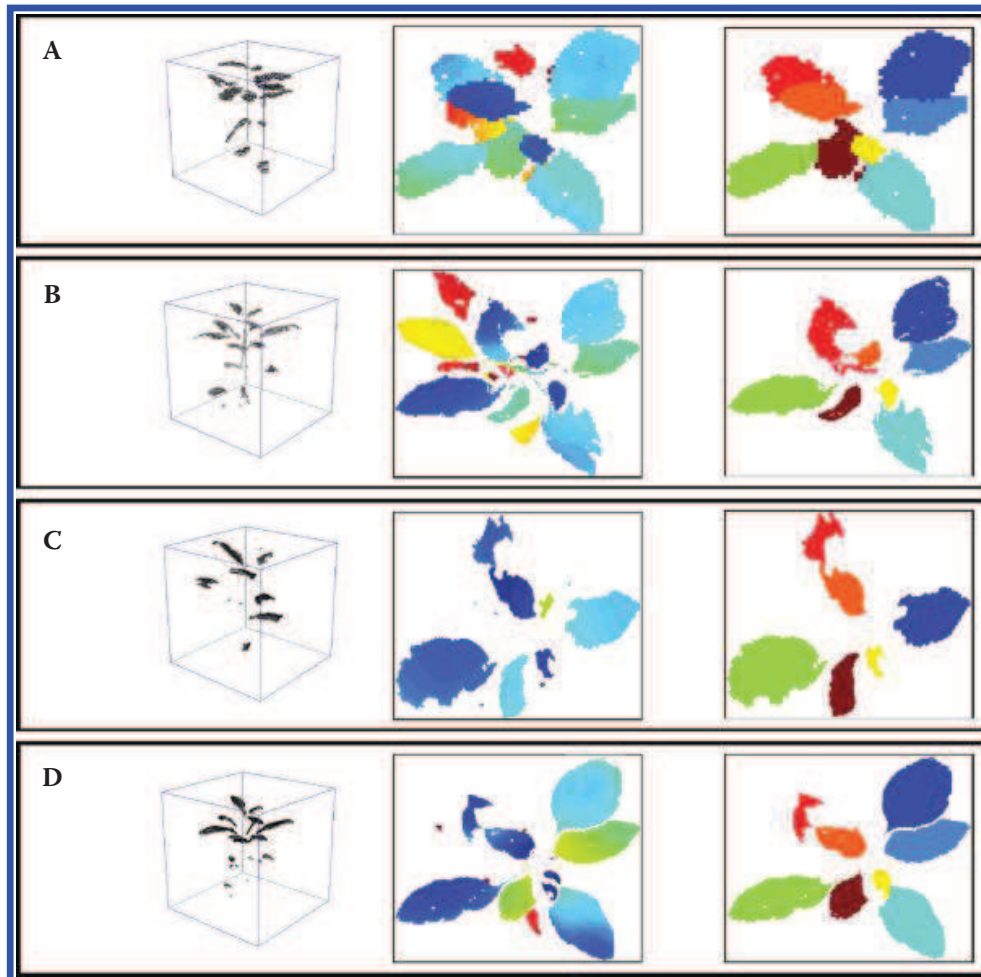
*Figure 9.5* Comparison of four depth cameras with the plant 1 of Figure 9.4. Each figure line is composed of the point cloud (left), the depth image (center), and the segmented image (right). In the image segmented with the algorithm of Chéné et al. (2012), only leaves presented in all depth images are represented with the same color for each depth camera. (a) Depth camera is a Microsoft Kinect that uses a pseudorandom dot grid as the structured lighting pattern. (b) Depth camera is a laser scanner that uses triangulation. (c) Depth camera is a hybrid depth camera that uses 2D structured lighting and stereovision (structured lighting pattern allows correspondences between two cameras). (d) Depth camera is a 2D structured lighting depth camera that uses a binary stripe sequential projection as the structured lighting pattern.
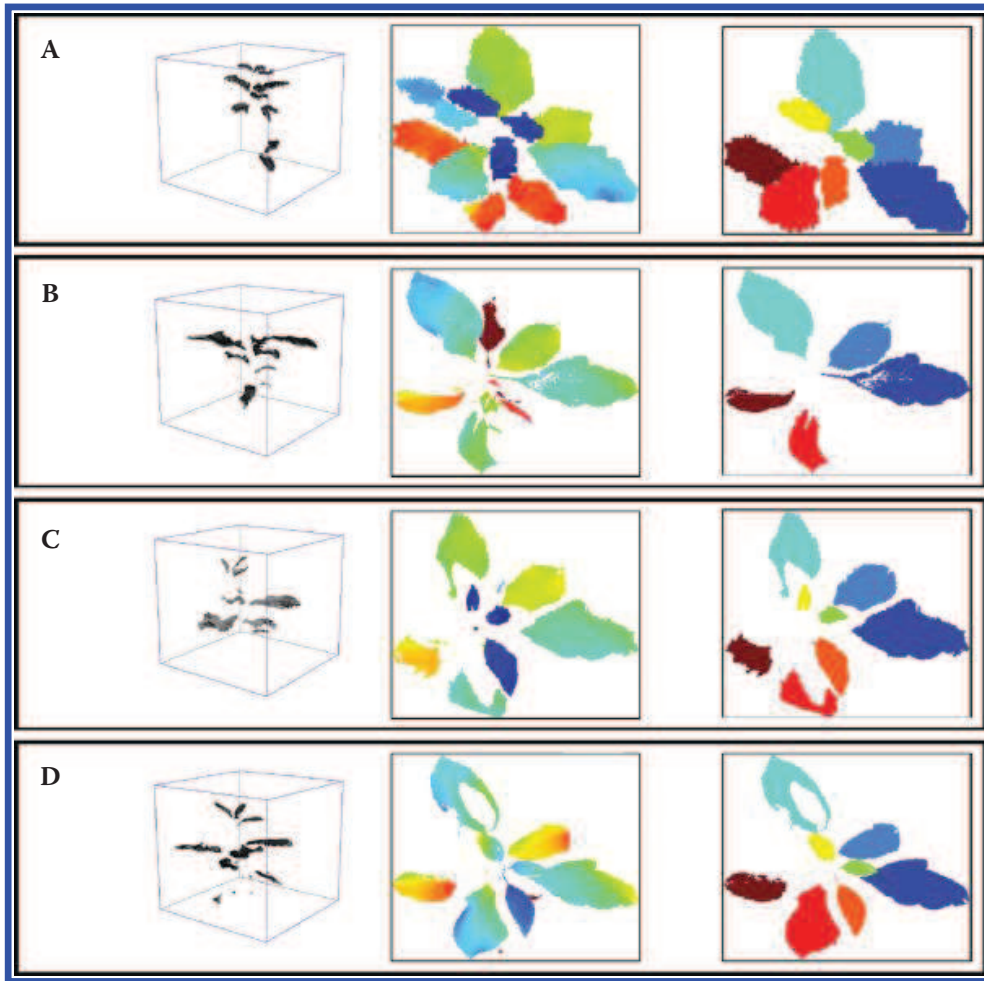
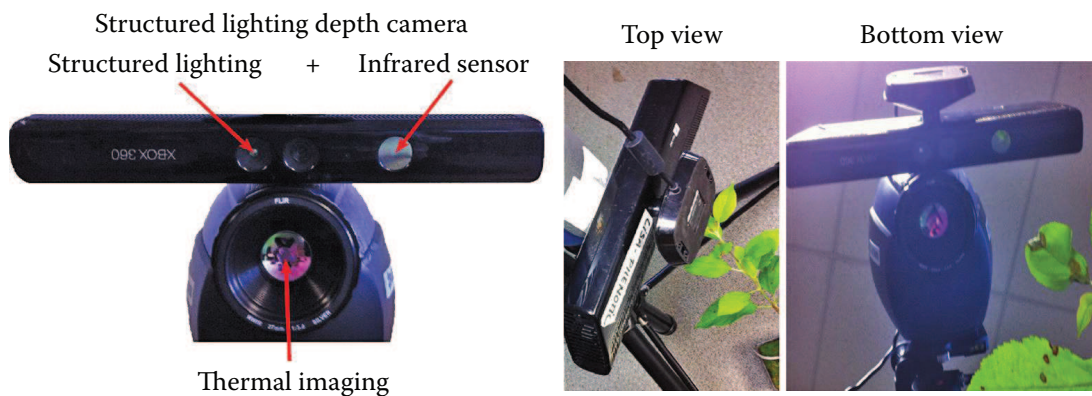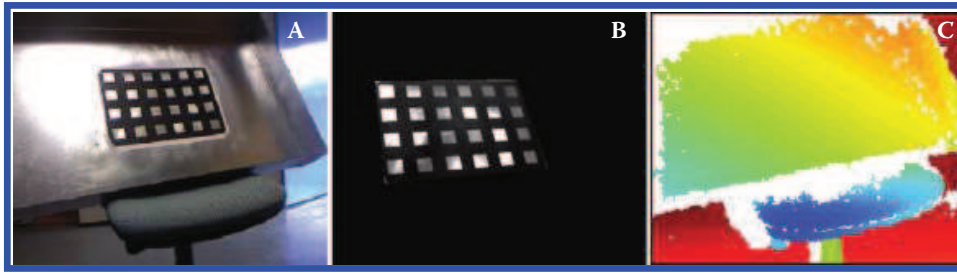*Figure 9.6* Same as in Figure 9.5 with plant 2 of Figure 9.4.



*Figure 9.9* Left: The anatomofunctional imaging system in this chapter is composed of a depth camera using the 2D structured lighting (binary dot pseudo-random grid) method and a thermal imaging. Right: An acquisition on a plant is done by one shot taken in plant top view.

*Figure 9.10* A calibration scene. (a) The RGB image of the calibration chessboard on the metallic plane. (b) The thermal image of the calibration chessboard on the metallic plan. (c) The disparity map of the calibration chessboard plan.
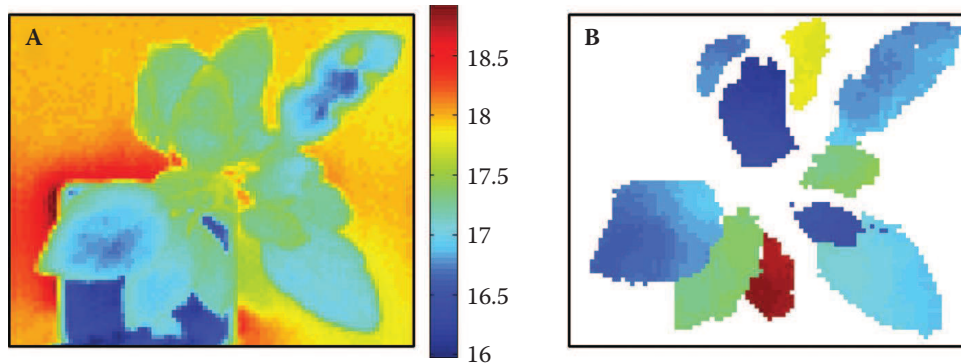


*Figure 9.11* An anatomofunctional imaging composed of depth camera and a functional imaging allows us to have two registered images: the depth map and the functional image. In this chapter, the anatomofunctional imaging system is composed of (a) a thermal imaging providing a thermal image and (b) a depth camera providing a depth map.



*Figure 9.12* Plant leaves are segmented by application of the depth segmentation algorithm of Chéné et al. (2012). Like the depth map, the segmented image is registered with the thermal image. In this image, each leaf is associated with a number. These numbers allow us to create a mask for each individual leaf.



*Figure 9.13* The edge of each leaf mask is applied on the thermal image. Colors of these edges are equal to the color of the leaf-segmented image of Figure 9.12. In this case only leaf 6 is hosting the pathogen.
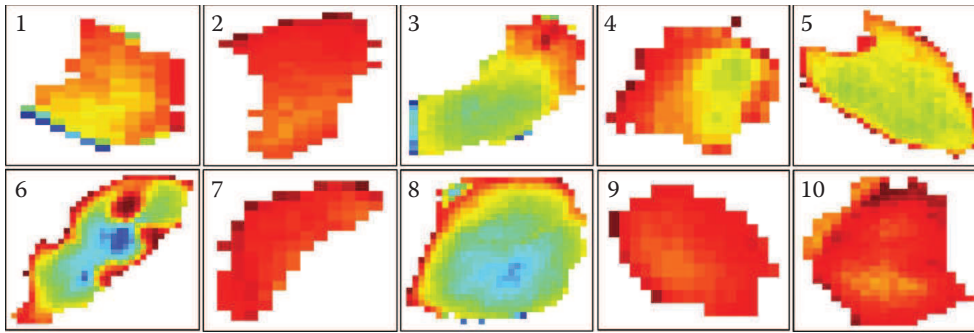
**Figure 9.14** Thermal images of each individual leaf segmented from the plant. Numbers correspond to numbers in Figure 9.12. The image resolution is the functional image resolution.
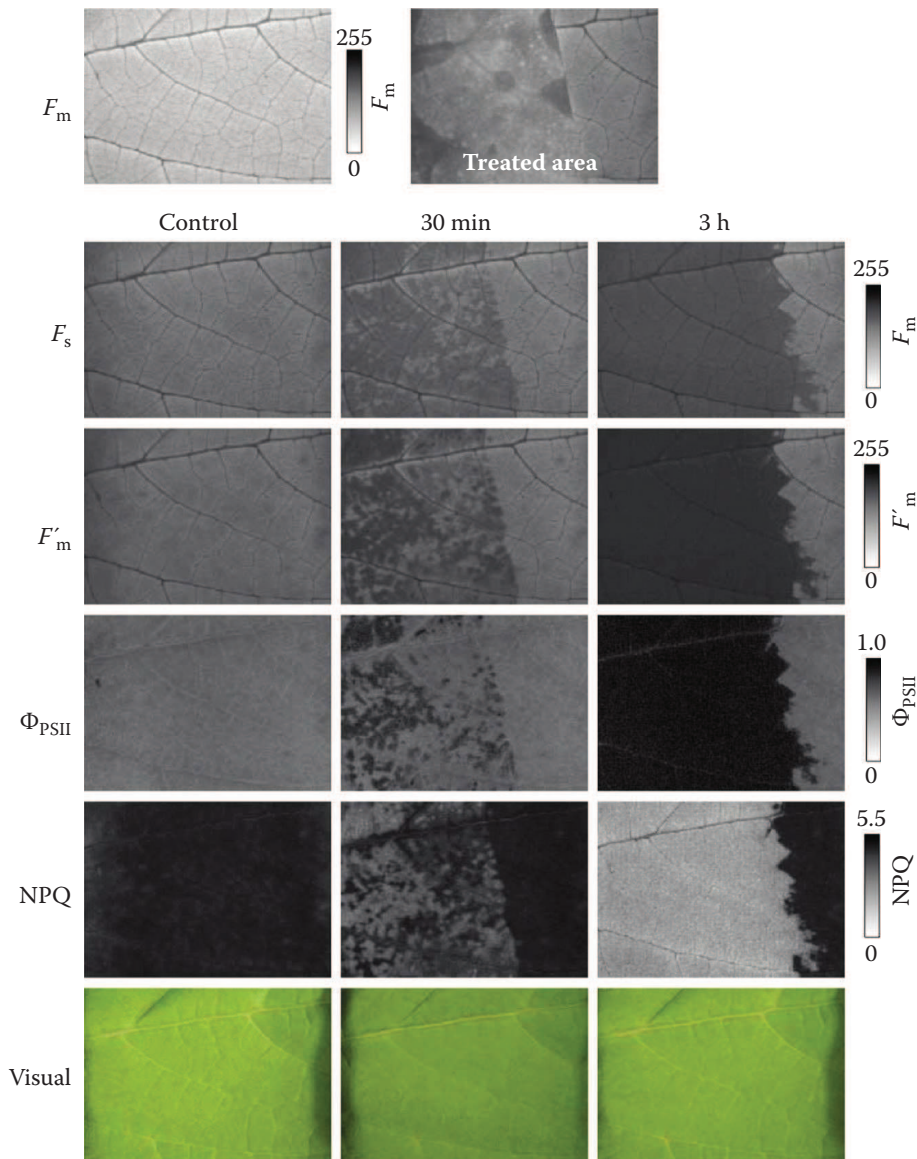


**Figure 10.7** Effects of Basta treatment on a *Phaseolus vulgaris* L. leaf. The fluorescence image titled "Treated area" represents the Basta treatment. Chlorophyll fluorescence intensity images ($F_s$ and $F_m$) were captured before (control), 30 min, and 3 h after treatment, and $\delta_{PS\alpha\alpha}$ and NPQ images were calculated each time.