Studies on high-throughput phenotyping by use of time series crop images taken under natural environments

（自然環境下で撮影した作物時系列画像を用いた高速フェノタイピングに関する研究）

東京大学大学院　農学生命科学研究科
農学国際専攻
平成 23 年博士課程入学　郭　威
指導教員　東京大学教授
二宮　正士

Content

List of figures

Chapter 1 Introduction.

1.1 Research background.

The world's population has grown tremendously over the past 50 years, with over 200,000 people still being born every day in the 21st century. According to the latest figures from the United Nations, the current global population of 7.2 billion is predicted to increase by 1 billion over the next 12 years, and to reach 9.6 billion by 2050 (UN Report, 13 June 2013, http://www.unpopulation.org). Experts have predicted that crop production must at least double before that time in order to support almost 10 billion people (Furbank, 2009; Tilman et al., 2011). However, achieving this goal will pose a formidable challenge to plant breeders, because the average rate of crop yield increases at only 1.3% per year for the four key global crops — maize, rice, wheat and soybean while productivity must increase 2.4% to meet that demand of the world population (Ray et al., 2013). Furthermore, the impacts of climate change on global temperatures and rainfall patterns are likely to reduce crop productivity (Lobell and Gourdji, 2012). The primary task is therefore, with very limited land and water resources, to advance cultivation technology and breeding development in order to minimize climate change impacts on crop production, by increasing both the quantity and quality of crop production enough to meet the increasing nutritional demands of the growing human population.

To improve crop varieties, a fundamental advance would be to dramatically accelerate the rate of both genotypic and phenotypic analysis. Based on the development of high speed sequencing technology over the past decade, genotyping technologies are developing rapidly, and the genomic sequences of numerous major crops have been determined, such as rice which is the one of the world's most important food plants (Sequencing ProjectInternational Rice Genome, 2005). However, the analysis of plant phenotypes remains at a stage of conventional qualitative analysis, which creates a major bottleneck for future advances in plant breeding.

A plant's phenotype is formed during plant growth and development under the dynamic interaction between its genotype and surrounding environment. Plant phenotyping is a comprehensive assessment of complex plant traits such as growth, development, architecture, physiology, tolerance, resistance, ecology, yield, and the basic measurement of individual quantitative parameters that form the basis for the

more complex traits (http://www.lemnatec.com). The main reason that phenotyping is a time-limiting step is that it involves qualitative, destructive and labor-intensive measurement.

Image analysis technology is considered as a powerful tool for improving the quantitative analysis of crop phenotypes. The technology involves a mathematical process to extract, characterize, and interpret useful information of the targets or Regions of Interest (ROI) from digital or pixel elements of an image. By analyzing the information extracted from images, multiple phenotypic characteristics can be evaluated quantitatively. Several studies using this approach have been published in recent decades. For example, after Ninomiya and Shigemori (1991) defined 18 shape parameters that were extracted from digital soybean images, several methods were proposed to quantitatively evaluate soybean shape (Ambuel et al., 1997; Jack et al., 1997; Ninomiya and NguyenCong, 1998; Ninomiya, 2009; Ninomiya et al., 1998; Oide and Ninomiya, 1998) — the results led to similar conclusions as those achieved with visual inspection by experienced soybean breeders. Several studies have analyzed the contour shape of crop organs using elliptic Fourier descriptors (EFDs, Kuhl and Giardina, 1982), which enable normalized rotation and size invariant shape analysis. Using EFDs, researchers have successfully classified soybean varieties based on quantitative measurements of leaf shape (Oide and Ninomiya, 2000), proposed a new QTL analysis of soybean leaflet shape (Yamanaka *et. al.*, 2001), examined the genotype ×environment interactions among the variations of citrus by their leaf shape (Iwata et al., 2002), and quantitatively evaluated the petal shape of *Primura*, the effect of distance on phenotypic variation in the petal shape, and petal shape transition from wild populations to modern cultivars over 300 years of selective breeding (Yoshioka et al. (2007a, 2006, 2005)). Similarly, the technique has been used to develop a texture-based method to quantitatively evaluate the fineness of wrinkles on grains of malting barley (Ninomiya et al., 1992), and to evaluate a major quality trait of rice (chalkiness) with high accuracy, using gray-scale images (Yoshioka et al., 2007b).

To solve the destructive and labor-intensive limitations of phenotyping, a new approach known as “Phenomics” has become widely used; it requires speedy extraction of high-dimensional phenotypic data on an organism-wide scale. The term phenomics was originated and made popular by Dr. Steven A. Garan (Garan, 2003). In plant species, it has been defined as the study of plant growth, performance and composition, through integration with biology of a suite of new technologies including photonics, computers and robotics to accelerate progress in understanding gene function and environmental responses (Fiorani and Schurr, 2013; Furbank and Tester, 2011; Houle et

al., 2010). Thus, over the past decade, high-throughput plant phenotyping platforms have been established worldwide (Duan et al., 2011; Granier et al., 2006; Hartmann et al., 2011; Iyer-Pascuzzi et al., 2010; Montes et al., 2007; Reuzeau et al., 2010). Examples include the "TraitMill™" from "CropDesign" (http://www.cropdesign.com), the "Scanalyzer discovery platform" from "PhenoFab Europe" (http://www.phenofab.com) and "LemnaTec", the "Jülich Plant Phenotyping Centre" (http://www.fz-juelich.de) and the "Australian Plant Phenomics Facility" (http://www.plantphenomics.org.au/) — all fully automated facilities in greenhouses or growth chambers, with robotics, precise environmental control, and image processing techniques, to assess the growth and performance of individual plants. The basis of the rapid phenotyping includes the following key technologies:

1. High-resolution images, stored digitally to document the growth process and reactions to time-dependent stress factors.
2. Automation technology to allow the transport of plants through multisensory detection systems under high-throughput conditions, enabling the screening of thousands of plants per day.
3. Image processing algorithms to enable the extraction of hundreds of phenotypic parameters from a set of images.
4. Molecular biology, genomics, proteomics, and metabolomics approaches.

Based on these technologies, automated and image-based plant phenotyping is being conducted successfully.

A major part of agricultural research is conducted in outdoor fields, so that the genetic analysis and breeding of most crops are usually carried out under natural conditions. Because the plant phenotype is formed under the dynamic interaction between genotype and environment, phenotypes generated from indoor experiments do not always correlate well with typical field behavior of plants. Moreover, different from the situation of indoor plants grown individually in pots, plants in the field do not grow isolated but instead are free to interact with neighboring plants, for example via their root systems. Therefore, phenotypic characteristics such as canopy configuration measured in outdoor plant communities are different those of individual plants (Prashar et al., 2013). Thus, there is a strong need to establish high-throughput phenotyping methods that can be used to screen crop populations under natural environmental conditions in the field. Studies of such an approach have become very popular in the last few years (Andrade-Sanchez et al., 2013; Araus and Cairns, 2013; Cabrera-Bosquet et al., 2012; Cairns et al., 2012; Cobb et al., 2013; Gleadow et al., 2013; Pereyra-Irujo et al., 2012; Sparks et al., 2012; White et al., 2012). However, because of the complex

background and various light conditions that exist outdoors, exiting image processing techniques cannot be applied to extract phenotypic characteristics from images taken under real field conditions. Therefore, the development of methods to analyze images captured in the field is a big challenge for field-based high-throughput phenotyping.

Actually, images acquired with some optical sensors, such as near-infrared (NIR) cameras, can avoid the effects of illumination change and simplify the image analysis process. However, it is generally difficult for people to interpret visible characteristics of crops through images acquired from such sensors for invisible bands. Moreover, being precision instruments, those camera sensors are very costly and it is difficult to be using them continuously in field conditions. So, invisible band-based image sensors have not been able to be considered in this study.

Recently, several low-cost field monitoring systems such as "field sever" (Fukatsu and Hirafuji, 2005; Fukatsu et al., 2012, 2011; Mizoguchi et al., 2010) have become available. In addition to meteorological data, those monitoring systems automatically provide time-series RGB images, so their suitability for use in phenotyping in the field has been proposed. In this study, we adopted such field monitoring systems in order to establish a high-throughput phenotyping system that would be practical to use in the field. We anticipate that time series images can provide totally new phenotypic characteristics of crops for analysis. Namely, in addition to ordinal "static characteristics" that are currently evaluated, "dynamic characteristics" along a time line will add a new component to plant phenotyping in crop fields.

In this study, we aimed to develop a field phenotyping system that will be usable in natural environments, particularly based on time-series images collected in field monitoring systems with RGB cameras. First, we needed to solve problem of the weakness of RGB images caused by complex natural backgrounds and light conditions, which make image analysis difficult. We established an illumination invariant crop segmentation method for RGB images captured in the field. Then, we evaluated the applicability of the crop segmentation method for a typical time-series characteristic, i.e., canopy coverage ratios for wheat and paddy rice. We also tried to improve the crop segmentation method so that it can be used when weeds and algae confound the crop canopy. Then, utilizing time-series images, we attempted to detect flowering timing and the relative amount of flowers present on plants, which are major characteristics in crop evaluation. Finally, we critiqued the methods that we developed and discuss the possibility to utilizing those methods to construct a field based high-throughput phenotyping system.

1.2 Objectives of this study.

1. To develop a new image segmentation method that can extract crop vegetation from large time series of images taken under natural light conditions.
2. To evaluate the applicability of that proposed crop image segmentation method to phenotyping wheat fields and rice paddies.
3. To further develop the crop image segmentation method proposed in Objective 1 to solve the problems caused by weeds and algae that are present in crop fields and thus confound vegetation images.
4. To develop a new method to automatically detect the timing of flowering and the relative amount of flowers present in rice paddies, using time-series images.
5. To discuss the criteria for a low-cost, field-based, high-throughput phenotyping system based solely on time series of RGB images.

1.3 Outline of this study

Chapter 1 is an overview of background information and the research objectives of this study.

Chapter 2 develops a robust method to isolate the vegetation segments of images from other background elements. This involves two major challenges for images taken in outdoor fields. First is how to overcome diverse natural light conditions that can strongly affect the profile of crop images taken outdoors, causing specular reflections and shadows over the vegetation. Second is to determine a generalized model that can be used for a large number of time series images whose properties change under varying natural light conditions, so that model modification is not necessary for each image of the time series. In order to overcome these two problems a Decision Tree Segmentation Model (DTSM) is proposed, which is broadly applicable to time-series images taken under natural light conditions without the need for individual treatment of each image.

Chapter 3 evaluates the applicability of DTSM for phenotyping. First, the ability of DTSM to estimate individual wheat plant canopy coverage ratios is examined. Then, using a set of time series images taken of rice paddies in the field, the success of estimating plant community canopy coverage ratios during the whole growth stage is assessed.

Chapter 4 reports on trials used to solve the problems with image interpretation posed by weeds and algae, which complicate the background of crop images and thus make it very difficult to isolate crop vegetation, particularly under organic and natural paddy

farming conditions.

Chapter 5 proposes a method to automatically detect flowering time and relative amount of flowers from time-series images of rice paddies. Currently, such detection relies fully on human judgments that are rather qualitative. After surveying both the biological and physical properties of rice flowers, we developed a feature-based model to automatically detect flowering panicles, using only RGB images taken in outdoor fields. This is the first successful trial of automatic flower detection, to the best of our knowledge.

Finally, Chapter 6 summarizes the overall conclusions of this research and discusses the high-throughput field phenotyping system that was developed, along with priorities for future research.

Chapter 2 Illumination invariant Decision tree based vegetation segmentation model (DTSM).

2.1 Introduction.

Crop phenotyping plays an important role to assess crop characteristics and its dynamic condition. Most of the crop phenotyping processes are still manually and visually conducted, taking huge amount of time and money. To accelerate the processes, several scientists have made efforts to utilize image data to extract useful phenotypic information by image analysis, expecting that crop images are one of the resources which can generate rich phenotypic information in a non-destructive manner (Furbank and Tester, 2011; Ishizuka et al., 2005; Louhaichi et al., 2010; Ninomiya, 2009). Particularly, image data can easily provide us time series information along with crop growth, using an automated photographing device. We can expect that such a series of information in total can generate new knowledge which would not be obtainable only with the observations on a single timing or a few timings. For example, low cost field monitoring systems such as Field Sever (Fukatsu and Hirafuji, 2005; Sudharsan et al., 2012) facilitate the easy acquisition of such sequential images in crop fields.

In utilizing crop images, a substantially important step is to effectively segment only vegetation part of an image from its background, because images of crops usually include non-vegetation parts as backgrounds. In this process, one of the challenges is to reduce erroneous detections under various natural light conditions which sometimes strongly affect the profile of the crop images taken outdoor, causing specular reflections and shadow over the vegetation, for example (Figure 2-1). This often makes the segmentation of vegetation parts from the background in an image difficult (Jafari et al., 2006; Jeon et al., 2011; Lati et al., 2011).

The vegetation segmentation in images has been an important research target and several methods have been proposed. Some studies (Liu et al., 2011; Panneton and Brouillard, 2009; Philipp and Rath, 2002) tried to discover the optimal color spaces for the segmentation from existing ones such as RGB, HSV, CIE L*a*b* and Ohta's color space (Ohta et al., 1980), but the recommended features differed depending on target plants and there is no common set of the features yet. Some other studies proposed new color features to segment vegetation using green color emphasizing formulae based on

the chromatic characteristic of the vegetation. The excess green ExG (Woebbecke et al., 1995), ExG-ExR (Meyer and Neto, 2008), Modified ExG (abbreviated to MExG in this paper) (Burgos-Artizzu et al., 2011), NDI (Pérez et al., 2000) are examples of such newly defined features and some of them are widely applied in vegetation segmentation presently (Burgos-Artizzu et al., 2010; Guijarro et al., 2011; Riomoros et al., 2010; Tang et al., 2009; Tellaeche et al., 2008).

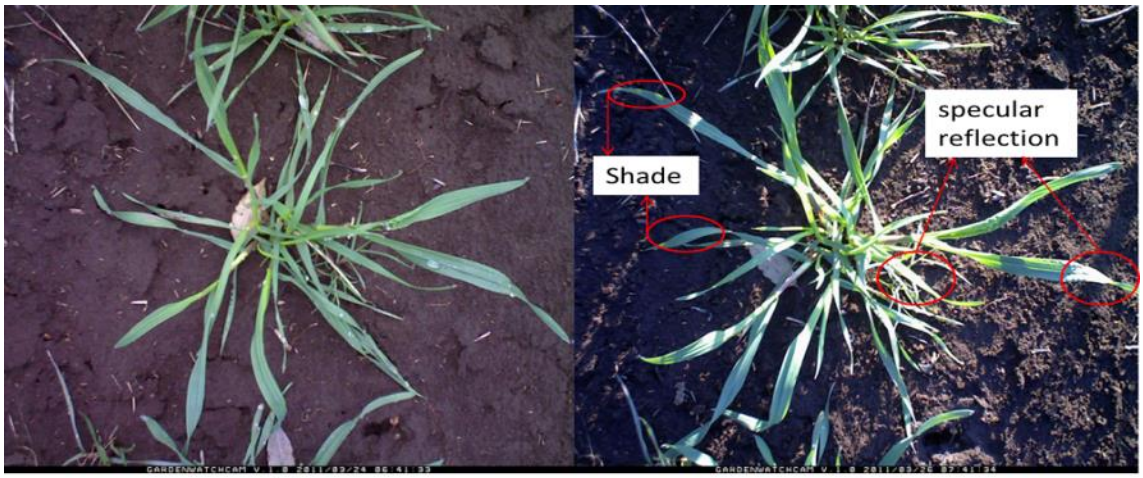


Figure 2-1 Example of wheat images taken under different light conditions.
(a) non-sunny day without direct solar illumination; (b) sunny day.

These color feature methods are based on an implicit assumption that the vegetation and background pixels in each image can be projected into an appropriate plane where they can be distinctly separated by a pre-calculated threshold value. The approaches are, however, not scalable as they may require a different user defined color-threshold for every image except the case of ExG-ExR (Meyer and Neto, 2008), because the values usually depend on the light conditions of the timings when the images are taken. This is rather troublesome when a huge number of the images are taken in a time series. Figure 2-2 demonstrated the color distributions from a piece of time series image set, the images are taken on the same day but different shot time with different illumination conditions. There's no uniformity of those histograms.

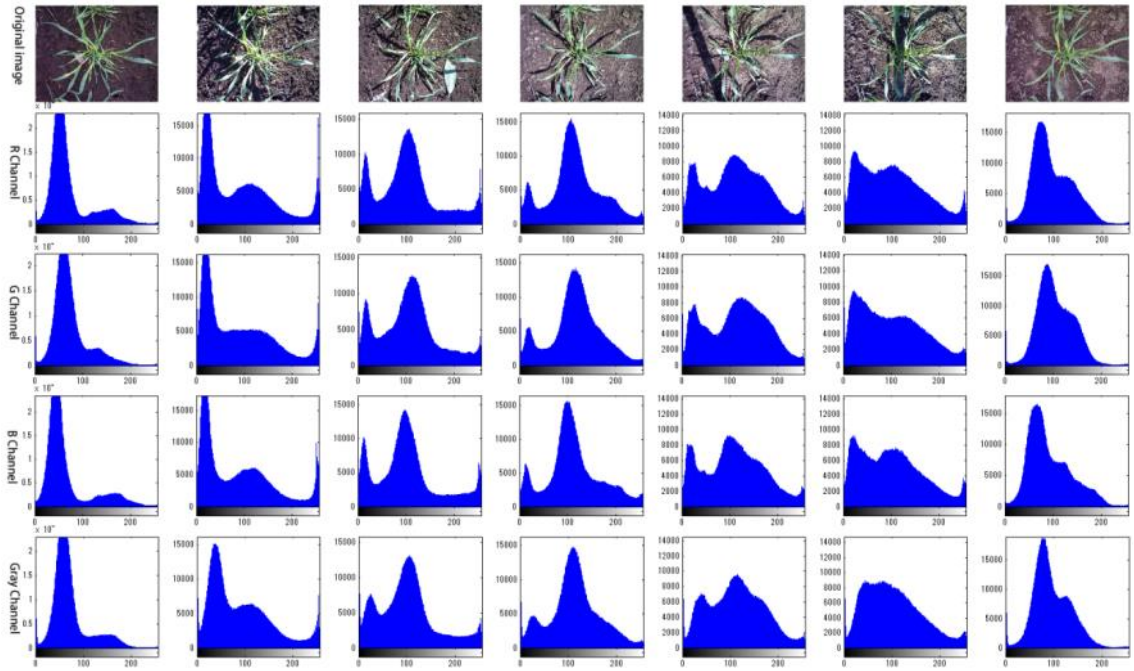


Figure 2-2 Example of a piece of time serious image set (row 1: original image) and the color distributions in histogram (row 2: R channel; row 2: G channel; row 3: B channel; row 4: Gray channel).

Moreover, our preliminary study indicated that these features were not robust to separate vegetation from background particularly when images contain specularly reflected parts and shadowed parts that weaken chromatic features (Figure 2-3). The 1st row of Figure 2-3a, illustrated different color index value (Gray, H, S, V in HSV color space, L,A,B in LAB color space) along the red line that traverse across both background and normal reflected leaf part. While the 2nd row illustrated the different color index value of the red line that cross both background and specularly reflected leaf part which changes the leaf color seriously in image; The 3rd row of Figure 2-3a illustrated the color index values changes very much even on a same leaf. The Figure 2-3b specially showed the line value of 3 widely used color features, we can tell it is not possible to find out an appropriate threshold to separate the leaf part from the background when specular reflection occurs. In fact, (Liu and Pattey, 2010) suggested avoiding taking photographs under strong sunlight conditions. However, in order to utilize a series of images taken automatically with a periodical interval, we can't choose light conditions.

To solve such limitations, (Ruiz-Altisent et al., 2010; Zheng et al., 2010) adopted machine learning methods to utilize the total information of multiple color features. They used clustering methods such as k-means method to automatically build training data sets. They manually assigned a class to each cluster automatically generated and

used the data of each class to train machine learning classification models such as back propagation neural network. However, because each cluster was automatically generated, there was no guarantee on the purity of the classes unless careful human operation is involved. We also preliminary examined the performance of the approach and found that it was not robust in such a case as Figure 2-1.

In this chapter, we propose a new approach to segment vegetation from background in images taken outdoor, aiming at utilization of series of images taken automatically regardless of light conditions. In particular, our proposal is to develop a segmentation model under natural light condition, which is applicable over different images without threshold adjustment for each image.

For the purpose, we also employed a machine learning method, because the nature of the method provides us with highly flexible classification even when values of a set of variables can be totally different for a same class. In comparison to neural network and support vector machine, we used decision tree model (Breiman et al., 1984), which provides the same accuracy as others with no parameter adjustment. For example, Ninomiya & NguyenCong, (1998) showed its high potential in classifying complicated plant shapes where totally different shapes belonged to a same class. As a preliminary study, we evaluated the performance of the proposed approach, using 30 test images and compared it with other methods such as ExG, ExG-ExR, and MExG.

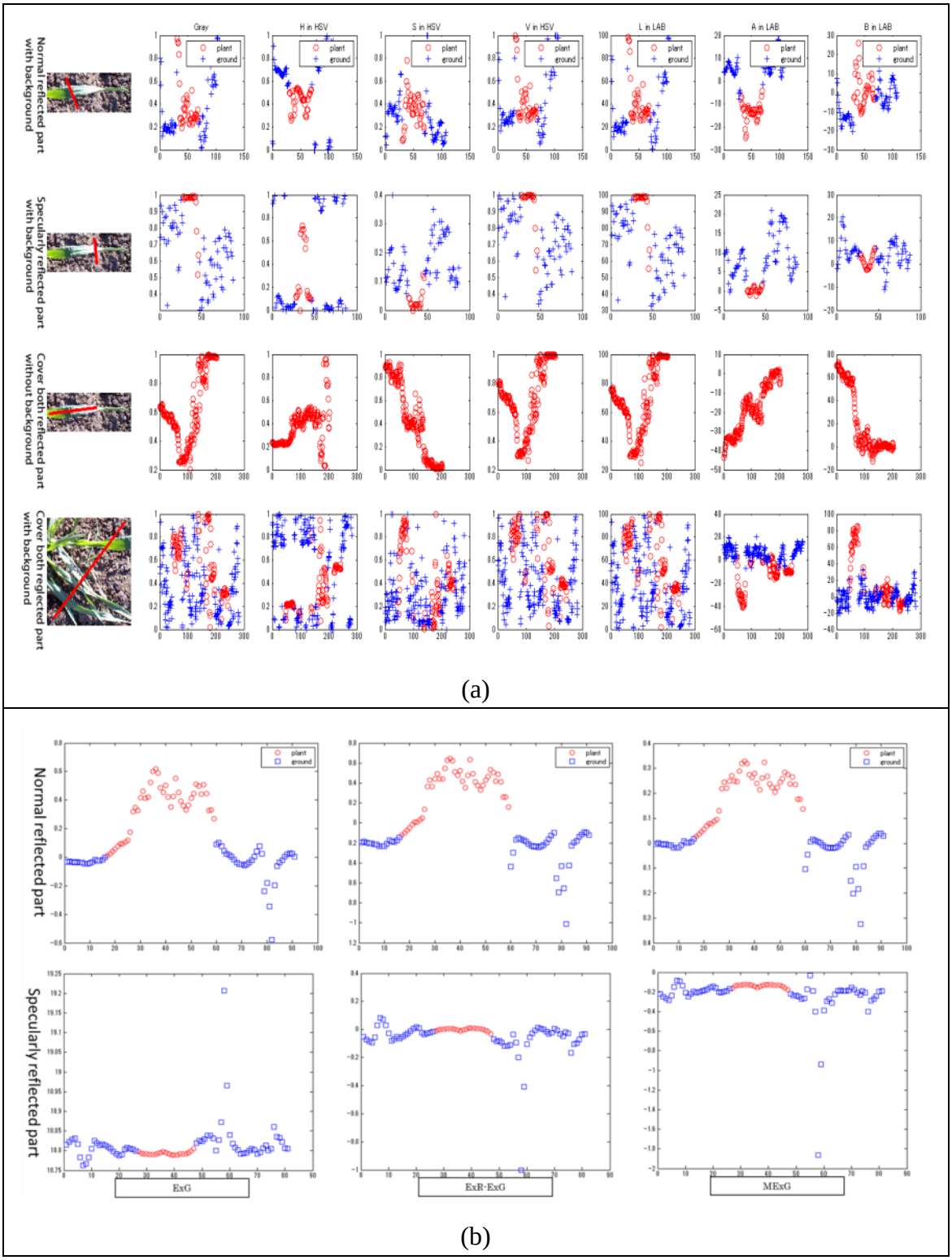


Figure 2-3 Color index values.

(a) Color index value investigation by drawn a line in red.

Column 1: original image and the position of red line;

Column 2~8: color index value of Gray, H, S, V in HSV color space, L,A,B in LAB

color space;
(b) The value investigation for the case of normal reflected part in row 1 of Figure 2-3a and the case of specularly reflected part in row 2 of Figure 2-3a, used three new widely used color features ExG, ExG-ExR, MExG, from column 1~3.

2.2 Materials and Method

2.2.1 Crop materials and image acquisition

In this study, we used wheat (cultivar: *Kinunonami*) as a target crop and grew it in two seasons. It was sown on November 11, 2010 and November 11, 2011, and grown under natural light condition in an outside field of the Institute for Sustainable Agro-ecosystem Services, the University of Tokyo. For the image acquisition, a waterproof digital camera (Garden Watch Cam, Brinno Inc., Taipei) was fixed horizontally in a top-view position, targeting a single wheat plant. Then, a series of the RGB color images (1280×1024) of the wheat were captured with one-hour time interval during the daytime from March 14th to April 1st, 2011 and from April 16th to May 1st, 2012, resulting in a total of 231 images and 213 images taken under various light conditions for each season, respectively. We named the data sets from 2011 and 2012 as Data2011 and Data2012 respectively.

2.2.2 Outline of the proposed model

The segmentation model proposed in this thesis is based on a machine learning model called decision tree. We adopted the CART algorithm to create the tree. The whole process of the proposed method includes the following steps:
1. Acquisition of a training data set from training images to train the model.
2. Training of the model to create a decision tree using the training data set.
3. Vegetation segmentation of test images using the decision tree.
4. Noise reductions on the segmented test images.

Then, the segmentation results were evaluated, using the true values obtained by manually segmented test images.

The proposed model was implemented by a GUI framework that we newly developed in a software package Matlab (MathWorks Inc., Natick) and executed on a Windows 7

PC with 4 core CPU and 16GB memory. The statistical analysis was performed by using a statistical software package R (R Development Core Team, 2012).

2.2.3 Training data

In machine learning based classification models, the model performance strongly depends on the selection of training data. In this paper, firstly we select the training images so that the images should cover heterogeneous natural light conditions considering the wide applicability of the model. Secondly we set the location of the regions of interests (ROIs) for both vegetation and background class, again considering various light conditions.

Figure 2-4 shows an example of the training data acquisition from a training image to extract the training dataset. Red rectangles with 5x5 pixel size in Figure 2-4. Figure 2-4a indicate the locations of the ROIs. Each ROI is selected so that it belongs only to either vegetation or background class. One ROI includes 25 pixels, each of which corresponds to a set of original RGB color information as shown in Figure 2-4b. Using the R, G, B color information of each pixel, 18 color features (r, g, b; Y, Cb, Cr; H, S, L; H, S, V; L*, a*, b*; L*, u*, v*.) defined in 6 ordinarily used color spaces (rgb, YCbCr, HSL, HSV, CIEL*a*b* and CIEL*u*v*) were derived. Then, the training data sets were arranged as Figure 2-4c and Figure 2-4d, concatenating the corresponding color features and classes of the pixels.

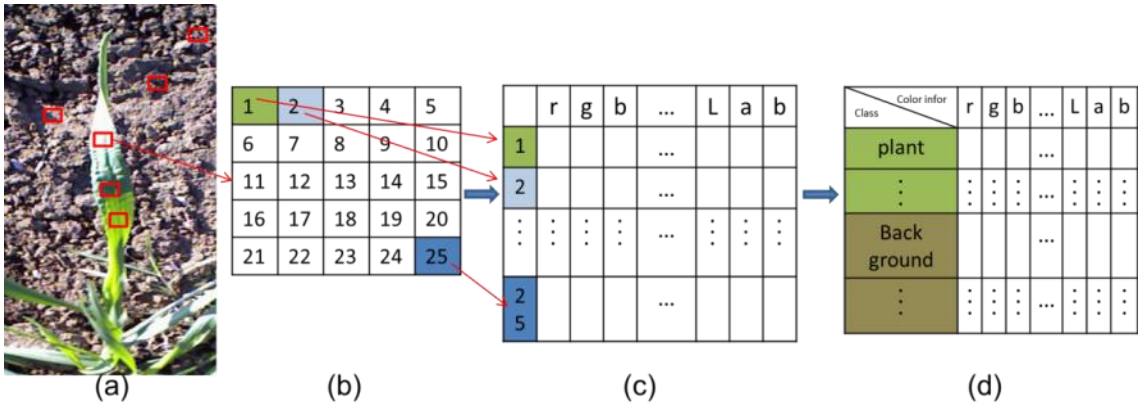


Figure 2-4 Example of training dataset acquisition by setting the ROIs on the training images.
(a) Original image with examples of ROIs; (b) a ROI of 5 x 5 pixels; (c) training data from a ROI; (d) A whole set of training dataset.

2.2.4 Finding key color-features

In the training dataset, we created 18 color features but certain color features could be redundant and/or irrelevant, thereby, decreasing the overall performance of the system. Since there is not a universally acceptable color space to effectively extract vegetation for any real-world application as mentioned above, we addressed the selection of color features for extracting vegetation using feature selection techniques known as “wrapper” (Kohavi and John, 1997). In the wrapper approach, the feature subset selection algorithm exists as a wrapper around the learning algorithm. The feature subset selection algorithm conducts a search for a good subset using the induction algorithm itself as a part of the function evaluating feature subsets. The idea behind the wrapper approach is simple: the learning (or induction) algorithm is considered as a black box. The learning algorithm is run on the dataset, usually partitioned into internal training and holdout sets, with different sets of features removed from the data. The feature subset with the highest evaluation is chosen as the final set on which to run the learning algorithm.

2.2.5 Model construction and generation of vegetation binary images

We generated the decision tree using the selected color features of the training data set by using the CART classifier. Then, the constructed tree model was applied to conduct segmentation on test images, resulting that each pixel belongs to either vegetation or background class, which finally generated binary images of vegetation.

Because some small segments that classified as vegetation were left over the originally background part as the result of misclassification (Figure 2-8b), a noise reduction process using spatial filters was conducted over the binary images generated by the model to obtain the final binary images of vegetation which were used to examine the accuracy of the vegetation segmentation by the model.

2.2.6 Model evaluation

To evaluate the model performance in vegetation segmentation under various natural light conditions, we provided true vegetation binary images by manually segmenting vegetation parts of the test images using an image retouching graphic application

(Adobe Photoshop Elements 9, Adobe Systems Inc., San Jose) and a high-resolution pen tablet (WACOM intuos4 pen tablet, Wacom, Kazo). Because of the illumination complexity of test images taken under sunny conditions, the Photoshop's cutout tools such as "lasso tool" were not usable and we had to very carefully segment those images as a fully manual manipulation. Then, the manually segmented images were used to evaluate the accuracy of the segmentations by the proposed method in this study and other formerly proposed methods such as ExG, ExG-ExR and MExG. The definitions of each method are:

Excess Green Index (Woebbecke et al., 1995):

$ExG = 2g - r - b$	(1)

Excess Green minus Excess Red index (Meyer and Neto, 2008):

$ExG - ExR \quad where: \ ExR = 1.4r - g$	(2)

Modified Excess Green Index (Burgos-Artizzu et al., 2011):

$MExG = 1.262g - 0.884r - 0.311b$	(3)

Where,

$r = \frac{R_N}{R_N+G_N+B_N}, \quad g = \frac{G_N}{R_N+G_N+B_N}, \quad b = \frac{B_N}{R_N+G_N+B_N}$	(4)
$R_N = \frac{R}{R_{max}}, \quad G_N = \frac{B}{G_{max}}, \quad B_N = \frac{B}{B_{max}}$	(5)

Where, $R_{max} = G_{max} = B_{max} = 255$ for 24-bit color images.

The accuracy of the segmentation methods is evaluated by the following performance measurements (Meyer and Neto, 2008):

$Qseg = \frac{\sum_{i=0}^{i=m}\sum_{j=0}^{j=n}(A(v)_{i,j} \cap B(v)_{i,j})}{\sum_{i=0}^{i=m}\sum_{j=0}^{j=n}(A(v)_{i,j} \cup B(v)_{i,j})}$	(6)
$Sr = \frac{\sum_{i=0}^{i=m}\sum_{j=0}^{j=n}(A(v)_{i,j} \cap B(v)_{i,j})}{\sum_{i=0}^{i=m}\sum_{j=0}^{j=n}B(v)_{i,j}}$	(7)

Where A is the set of the vegetation pixels (v=255) or background pixels (v=0) identified by a classification model, B is a reference set of manually segmented vegetation pixels (v=255) or background pixels (v=0), m and n are the image row and column sizes, and i, j are the pixel coordinate indices of the images. The more consistent pixels between A and B, the values become the larger ranging from 0 to 1. Namely, the higher the value, the more accurate the segmentation is. The value of Qseg represents the consistency of both vegetation part and background part while the value of Sr represents the consistency of only vegetation part. Using the values, we examine the

performance of the proposed decision tree based segmentation model (abbreviated to be DTSM in this paper), comparing with the threshold based color-index method ExG, ExG-ExR, MExG.

2.3 Experiments and Results

By using the wheat images taken under natural condition, we examined the performance of the proposed DTSM and compared the results with the presently widely used green color emphasizing formulae based segmentation methods, ExG, ExG-ExR and MExG.

2.3.1 Training images and test images

For the training data set, we selected 5 images from Data2011, considering the variety of light and weather conditions (Figure 2-5). The images include various light conditions such as a sunny condition with shadow, a sunny condition, a sunny condition both with shadow and specular reflection, a cloudy condition and a rainy condition. For the model evaluation, fifteen test images were randomly selected from each of Data2011 (Figure 2-6a) and Data2012 (Figure 2-6b), having a total of 30 test images. The random image selection for Data2011 was conducted after the training images were removed.

The dates of the test images taken varied from March 15th to April 1st and from April 17th to May 1st for Data2011 and Data 2012 respectively resulting in different growth stages of the crop in the images (Appendix A, Table ~ Table 2-5). The timings of the test images taken also varied from early morning to early evening resulting in various illumination conditions over the crop in the images. By visually judging the illumination conditions of the test images, we classified the test images into “sunny” and “non-sunny” (Appendix A, Figure 2-6), having 7 “sunny” images out of 15 images and 6 “sunny” images out of 15 images for Data2011 and Data2012 respectively.

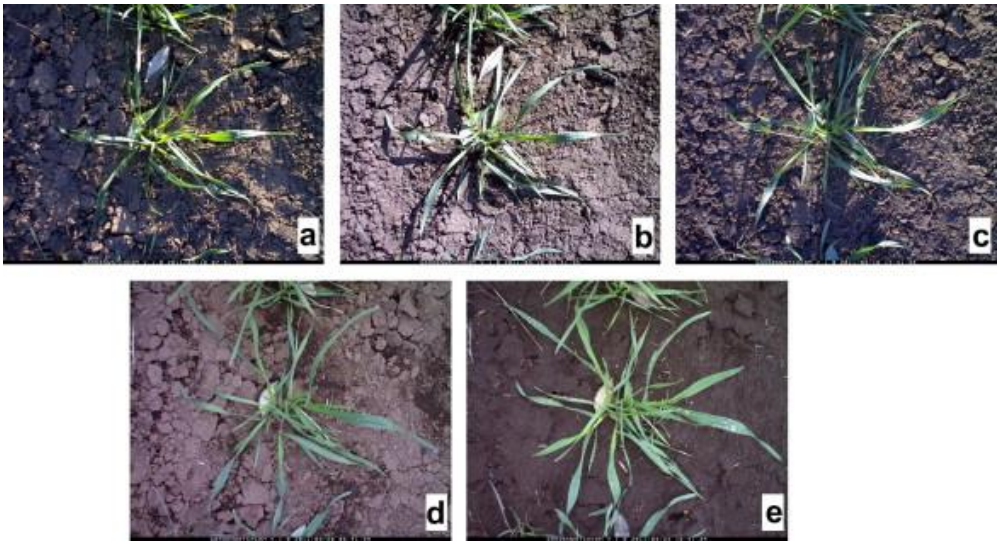
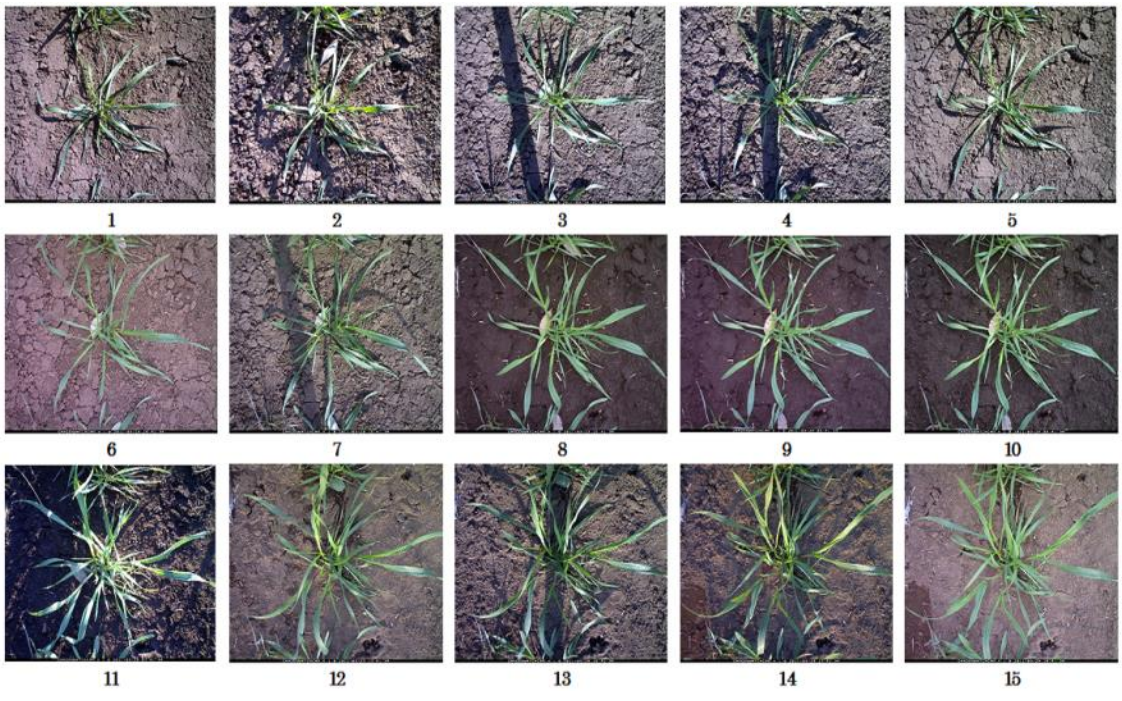
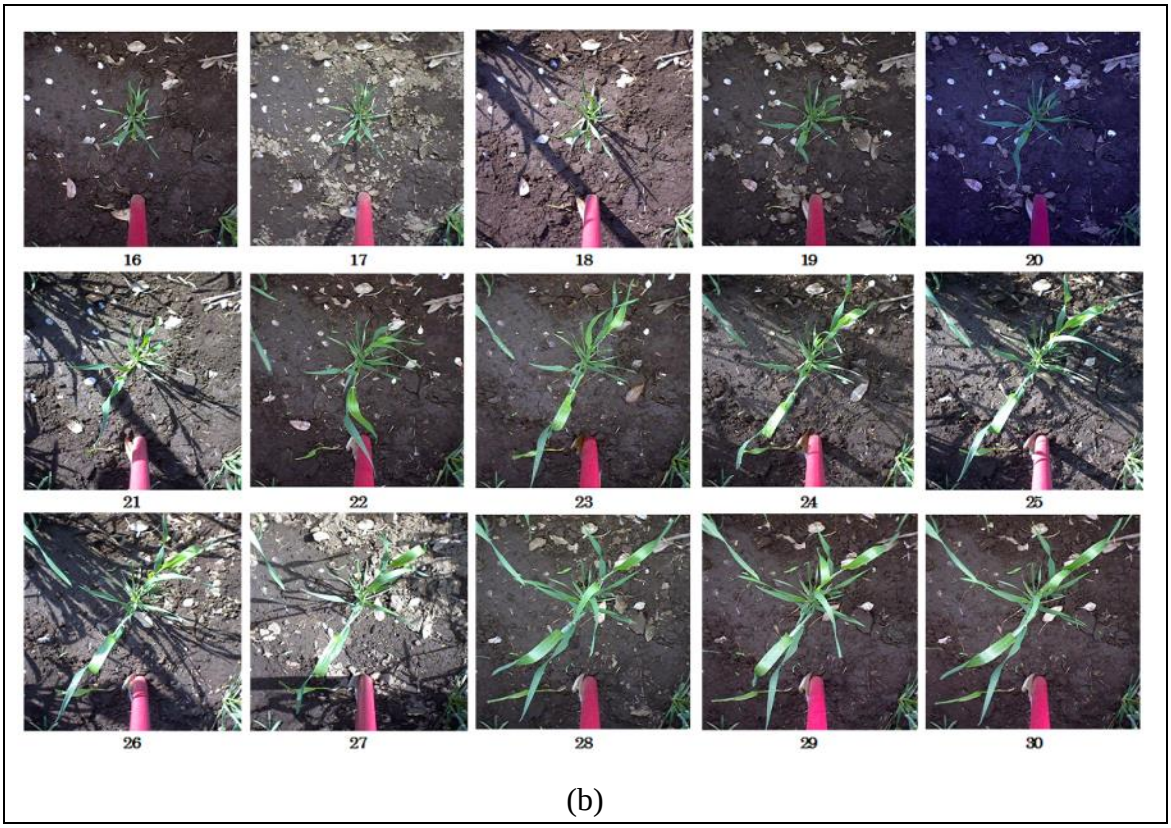


Figure 2-5 Training images used in this study.
The images were taken under: (a) a sunny condition with shadow; (b) a sunny condition; (c) a sunny condition with shadow and specular reflection; (d) a cloudy condition and (e) a rainy condition.



(a)



(b)

Figure 2-6 Test images randomly selected from 2 years image data sets.
(a) test images from Data2011.
(b) test images from Data2012 (the red stick in every image is the camera stander).

2.3.2 Model construction

From the training images, we acquired a total of 45,000 pixels for vegetation class and 90,000 pixels for background class as the training dataset which generated 18 color features (r, g, b; Y, Cb, Cr; H, S, L; H, S, V; CIEL*, a*, b*; CIEL*, u* and v*) for each pixel. The process to acquire a total of 135,000 pixels from the five images for the training data set, took us about 3 minutes using the ROI selecting function in the GUI framework that we developed for this study. Then, by using the feature selection method "wrapper", r, Cb, Cr, S of HSV, S of HIS, a*, u*, v* were selected as the key color features and the DTSM was trained using the key features by the CART algorithm. Figure 2-7 shows the generated tree model and it shows that the color features r, Cb, Cr, S of HSV, S of HSI, a* and u* among the key features, were finally used as the result of the pruning process.

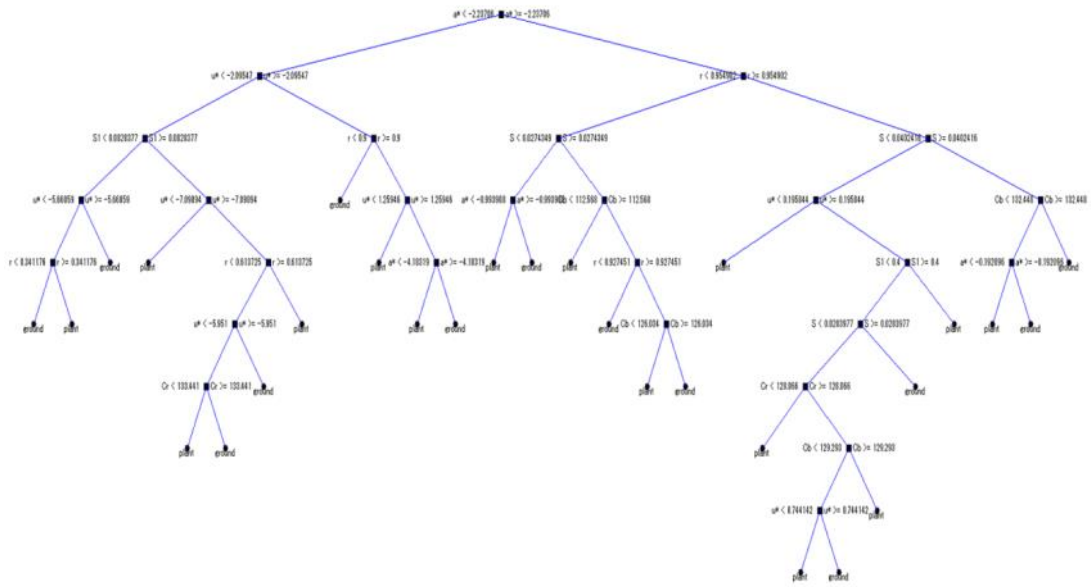


Figure 2-7 Generated decision tree.

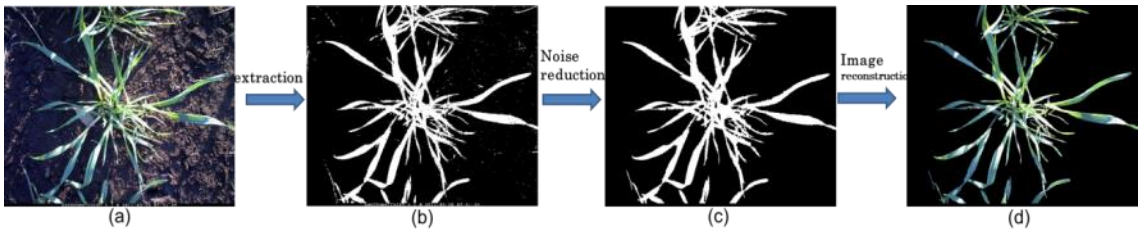


Figure 2-8 Example of the whole processes of vegetation segmentation by DTSM.
(a): original image;
(b): initial vegetation segmentation by the generated tree model where some misclassified pixels are observed on the background;
(c): final vegetation segmentation result after the noise reduction;
(d): Extracted vegetation overlaid on the original image.

2.3.3 Model evaluation

The generated model was applied to classify all the pixels of each test image (Figure 2-6) into two classes as vegetation and background. Figure 2-8 shows an example of the whole segmentation process. The performance of ExG, ExG-ExR and MExG were also examined using the same test images. The Otsu method (Otsu, 1979) was used to determine the threshold value for each of the test image one by one in the cases of ExG and MExG, while a fixed threshold zero was adopted in the case of ExG-ExR as the

original study did. After the classifications, the noise reduction processes to eliminate small misclassified pixels on background were applied to all of the results by ExG, ExG-ExR, MExG and DTSM.

To evaluate the results, first we examined the accuracy rates of the segmentation (Qseg and Sr) by ExG, ExG-ExR, MExG and DTSM for each image (see Appendix A). For Qseg, DTSM showed the best performance in 14 images out of 15 images of Data2011 with the mean value of 0.806 and 12 images out of 15 images of Data2012 with the mean value of 0.767. For Sr, DSTM showed the best performance in 11 images out of 15 images of Data2011 with mean value of 0.833 and 9 images out 15 images of Data2012 with the mean value 0.831. Particularly, the performance by DTSM for the images taken under sunny conditions was always best among the segmentation methods we examined in this study.

Because the performance of DTSM was distinctively good under sunny conditions, we conducted the further statistical analyses separately for the images taken under sunny and non-sunny conditions. We also analyzed the results for Data2011 and Data2012 separately because the training images were selected only from Data2011. Table 2-1 and Figure 2-9 shows the mean accuracy rates (Qseg and Sr) by ExG, ExG-ExR, MExG and DTSM, the standard deviation and the results by Tukey's multiple comparison (p=0.05) among the different segmentation methods whose segmentation performances were compared in this study. Because the data were ratio data, the multiple comparisons was conducted after the data were arcsine-transformed.

Under the sunny conditions, the mean accuracy rates of Qseg and Sr by DTSM for both of Data2011 and Data2012 always significantly exceeded the performance of ExG and MExG. Qseg and Sr under the sunny conditions by DTSM were also significantly better than those by ExG-ExR for Data2012 while there were not significant differences between two methods under the sunny conditions of Data2011 although the mean values of Qseg and Sr by DTSM were higher than those by ExG-ExR. Multiple comparison is rather conservative to guarantee the overall type I error of multiple tests and the results of simple Welch's t-test indicate that Qseg and Sr by DTSM for Data2011 are significantly different from those by ExG-ExR (p= 0.0007 for Qseg and p=0.006 for Sr).

On the other hand, under the non-sunny conditions, the performance of DTSM is almost the same as that of ExG, ExG-ExR and MExG except the cases for Data2012 where the accuracy rates by ExG were worse than those by other methods for both Qseg and Sr.

Overall results indicated that the accuracy of both Qseg and Sr by DTSM dominated

ExG, ExG-ExR and MExG under the sunny conditions of Data2011 and Data2012 while its accuracy under the non-sunny conditions is similar to that of ExG-ExR and MExG. The results show that the accuracy of DTSM for Data2012 is even better than that for Data2011 though the decision tree was trained only with the images from Data2011.

Table 2-1

The average accuracy rates of the vegetation segmentations represented by Qseg and Sr for ExG, ExG-ExR, MExG and DTSM. The results for the test images from Datat2011 and Data2012 are separately shown. Note that the model was trained only with the training images from Data2011. See the text for the further details. The "Qty." shows the number of the test images examined

Test image set	**Weather condition**	**Qty.**	**Qseg**			
			ExG	**ExG-ExR**	**MExG**	**DTSM**
Data2011	Sunny	7	0.404 ±0.207	0.680± 0.103	0.578± 0.220	0.783± 0.064
	Non-sunny	8	0.663 ±0.056	0.745± 0.083	0.798± 0.034	0.827± 0.057
Data2012	Sunny	6	0.349 ±0.121	0.588± 0.105	0.472± 0.131	0.783± 0.041
	Non-sunny	9	0.592 ±0.130	0.746± 0.049	0.738± 0.046	0.756± 0.081
Test image set	**Weather condition**	**Qty.**	**Sr**			
			ExG	**ExG-ExR**	**MExG**	**DTSM**
Data2011	Sunny	7	0.414± 0.214	0.719± 0.107	0.607± 0.234	0.811± 0.079
	Non-sunny	8	0.694± 0.075	0.784± 0.105	0.844± 0.059	0.853± 0.066
Data2012	Sunny	6	0.363± 0.131	0.665± 0.137	0.508± 0.152	0.870± 0.056
	Non-sunny	9	0.633± 0.149	0.836± 0.084	0.817± 0.071	0.806± 0.101

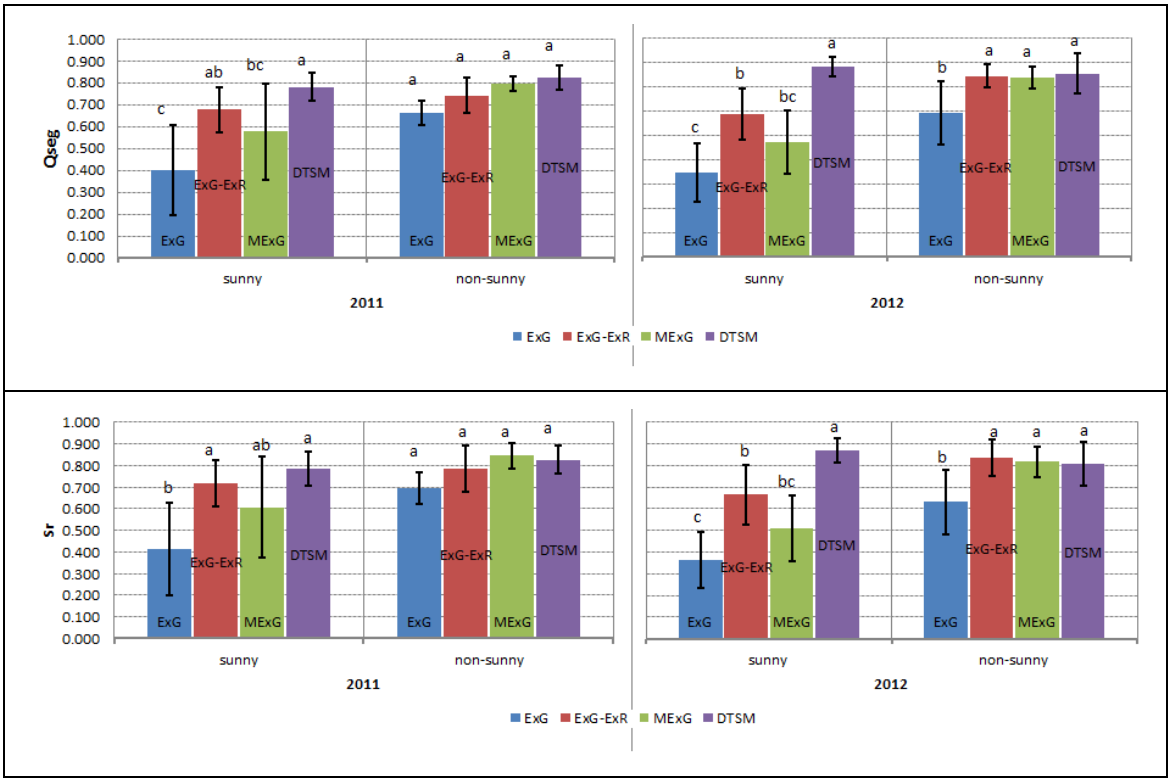


Figure 2-9 Comparison of the mean accuracy rates (Qseg and Sr).

The Qseg and Sr of the segmentations among ExG, ExG-ExR, MExG and DTSM for the test images from Datat2011 and Data2012. The bars indicate the standard deviations. Tukey's multiple comparison (p=0.05) was conducted over the arcsine transformed ratio data and the symbols, a, b or c indicates non-significant differences among the groups with the same symbol.

2.4 Discussion

A model to segment only vegetation from background in digital images was constructed based on decision tree using the CART algorithm and its performance under natural light conditions was examined. The result showed that the accuracy rates of the vegetation extraction for wheat images taken under various natural light conditions by DTSM were better in almost all the cases than those of the previously proposed methods (ExG, ExG-ExR and MExG) which are widely used in crop segmentation.

The DTSM performed outstandingly well on specularly reflected parts of wheat leaves while the other methods could not segment such parts properly (Figure 2-10). This advantage made DTSM perform better under the sunny conditions than ExG-ExR which also performed well compared with ExG and MExG in general. In fact, ExG-ExR could not show as good performance as DTSM when the images include such

specularly reflected parts on wheat (Figure 2-10). Such weakness caused the larger standard errors of the accuracy rates by ExG-ExR under the sunny conditions than those by DTSM (Table 2-1 and Figure 2-9).

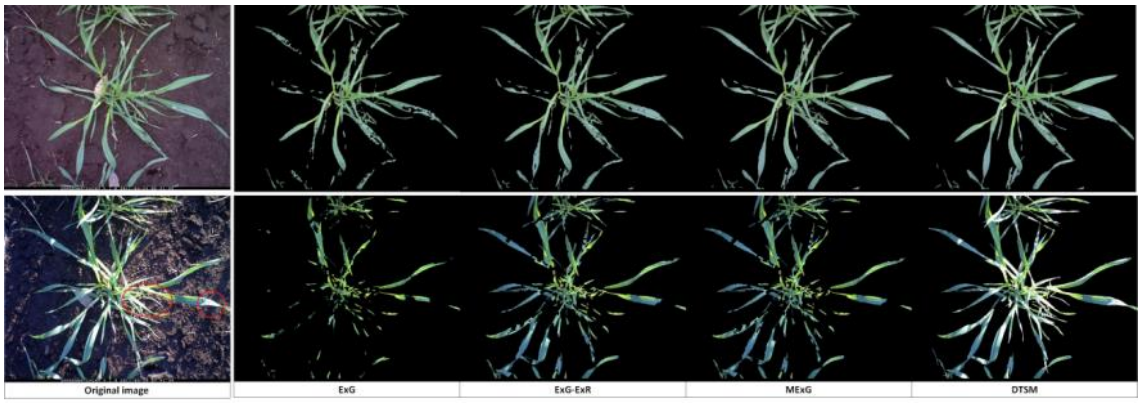


Figure 2-10 An example of the segmentation result.
The segmentation results by ExG, ExG-ExR, MExG and DTSM. The red circles indicate specular reflection parts of the plant.

DTSM was able to be commonly used for the time series data taken under different natural light conditions without any adjustments while most of the previously proposed methods required an adjustment of threshold values for each image. Such thresholding usually needs an interaction of manual operation while DTSM can provide a fully automated tool for vegetation segmentation once the model is trained.

Though DTSM was trained only with the training images from Data2011 in this study, it still performed well for the test images from Data2012. This indicates the generality of the DTSM model to some extent. However, we used only 30 test images in this study and need to examine its real performance using larger test data sets taken under more various conditions. We also wondered how generally DTSM was applicable to other crops. As a preliminary experiment, we simply conducted a classification of a paddy image using the same model trained with the wheat images from Data2011 and found that the model seemed to perform rather well (Figure 2-11) for such an image of a different crop, though we did not evaluate the accuracy based on true values.

Figure 2-11 Result of the segmentation of a paddy image.
Note that no paddy rice images are used in model training.
(a) Original image taken from a paddy rice field; (b) the result of the segmentation.

The performance of the DTSM strongly depends on training data. If a training data set does not cover the information to segment targeted parts, we cannot expect the model to work properly. Therefore, the selection of the training data set is significant and substantial to have a good model. In this study, we selected the training images visually considering the diversity of the light conditions and also manually extracted vegetation and background parts from each training image. Since we used only five training images in this study, the acquisition of the training data was not troublesome. However, because this process can be still subjective and time-consuming when a large number of training images are given, a full or semi-automated system to select training data is ideal. (Ruiz-Ruiz et al., 2009) proposed a use of k-means clustering to obtain training data semi-automatically. We examined it performance using the training images in this study. Following the original study, we conducted 4 group clustering using the same color features (Hue and Saturation). Figure 2-12a shows the results of the clustering of the training images and Figure 2-12b shows the pixels used as vegetation for training. We concluded that this approach was not good enough for semi-automatically obtaining training data sets because the same clusters often include both vegetation and background in our case. Namely, the training data can degrade the performance of the classification model. We also found that the visual assignments of the clusters to vegetation and background were also hard because of such impure clusters.

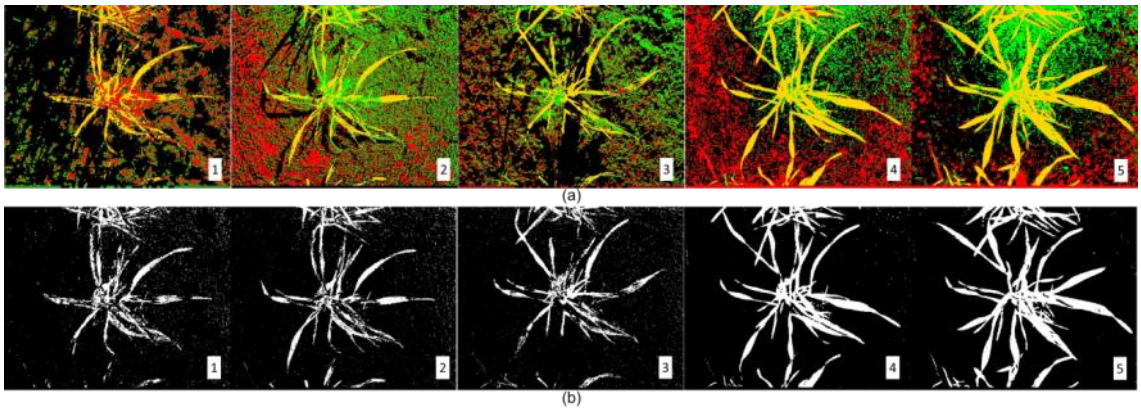


Figure 2-12 Clustering results by using K-means algorithm.
(a) the results of the 4 group clustering for the training images. Gold, red, green and black colors indicate the classes automatically generated; (b) the cluster assigned visually as vegetation for training.

Nowadays, field monitoring systems such as Field Server with high resolution cameras are available. Such systems can acquire series of crop images continuously. The Illumination invariant capability and non-requirement of thresholding of DTSM are definitely useful for those images collected by the field monitoring systems because the number of the images to handle is generally large and the timing of taking images is not selectable considering the light condition. Applying the DTSM proposed in this study which can be used regardless of natural light conditions (see Supplementary data in Appendix B), we would like to expand our study to develop a growth monitoring system of crops.

The accuracy rates represented by Qseg and Sr for DTSM still stay around 0.8 and further improvement is necessary in order to use it practically. In this study, we did not consider the spatial relationships among neighboring pixels for the segmentation except the noise reduction process. It may be, however, useful to improve the accuracy of the segmentation and we also plan to utilize such information in the future study.

2.5 Appendix A.

The segmentation accuracy rates represented by Qseg and Sr for ExG, ExG-ExR, MExG and DTSM for each test image. The image number corresponds to that of Figure 2-6a (Data2011) and Figure 2-6b (Data2012). "*" in the right most column of each image indicates the accuracy rate by DTSM is better than any other methods. The test images were classified into the images taken under sunny conditions and the images under non-sunny conditions, visually judging the illumination conditions of the images.

Table 2-2 Qseg value of Data2011 (Figure 2-4a)

Image No.	Time		Illumination condition	Segmentation Methods			
				ExG	ExG-ExR	MExG	DTSM
1	2011/3/15	12:41	Non-sunny	0.606	0.728	0.786	0.840
2	2011/3/17	10:41	Sunny	0.162	0.513	0.390	0.662
3	2011/3/19	10:41	Sunny	0.164	0.643	0.190	0.732
4	2011/3/19	18:41	Non-sunny	0.570	0.551	0.719	0.694
5	2011/3/20	15:41	Sunny	0.574	0.747	0.729	0.802
6	2011/3/22	8:41	Non-sunny	0.671	0.771	0.812	0.835
7	2011/3/23	15:41	Non-sunny	0.728	0.765	0.798	0.828
8	2011/3/24	6:41	Non-sunny	0.676	0.798	0.818	0.848
9	2011/3/24	9:41	Non-sunny	0.734	0.744	0.814	0.825
10	2011/3/26	7:41	Sunny	0.268	0.579	0.545	0.822
11	2011/3/27	17:41	Sunny	0.640	0.784	0.804	0.852
12	2011/3/28	16:41	Sunny	0.418	0.741	0.652	0.809
13	2011/3/29	17:41	Sunny	0.599	0.751	0.735	0.803
14	2011/3/29	18:41	Non-sunny	0.653	0.808	0.808	0.882
15	2011/4/1	18:41	Non-sunny	0.667	0.793	0.827	0.862
Average				0.542	0.714	0.695	0.806

Table 2-3 Qseg value of Data2012 (Figure 2-4b)

Image No.	Time		Illumination condition	Segmentation Methods			
				ExG	ExG-ExR	MExG	DTSM
16	2012/4/17	8:55	Non-sunny	0.590	0.728	0.709	0.696
17	2012/4/17	11:55	Non-sunny	0.597	0.690	0.675	0.693
18	2012/4/18	8:55	Sunny	0.216	0.403	0.296	0.716
19	2012/4/20	16:55	Non-sunny	0.630	0.740	0.717	0.731
20	2012/4/21	4:55	Non-sunny	0.274	0.648	0.668	0.599
21	2012/4/24	8:55	Sunny	0.207	0.577	0.340	0.786
22	2012/4/26	11:55	Non-sunny	0.668	0.778	0.764	0.796
23	2012/4/27	10:55	Non-sunny	0.686	0.792	0.779	0.813
24	2012/4/28	7:55	Sunny	0.524	0.724	0.651	0.842
25	2012/4/29	7:55	Sunny	0.384	0.581	0.498	0.771
26	2012/4/29	8:55	Sunny	0.361	0.627	0.525	0.801
27	2012/4/29	13:55	Sunny	0.402	0.614	0.519	0.781
28	2012/5/1	10:55	Non-sunny	0.722	0.791	0.788	0.821
29	2012/5/1	13:55	Non-sunny	0.571	0.768	0.770	0.820
30	2012/5/1	16:55	Non-sunny	0.594	0.775	0.768	0.834
Average				0.495	0.682	0.631	0.767

Table 2-4 Sr value of Data2011 (Figure 2-4a)

Image No.	Time		Illumination condition	Segmentation Methods			
				ExG	ExG-ExR	MExG	DTSM
1	2011/3/15	12:41	Non-sunny	0.638	0.773	0.851	0.882
2	2011/3/17	10:41	Sunny	0.163	0.524	0.396	0.666
3	2011/3/19	10:41	Sunny	0.165	0.676	0.191	0.749
4	2011/3/19	18:41	Non-sunny	0.575	0.554	0.728	0.705
5	2011/3/20	15:41	Sunny	0.587	0.781	0.760	0.843
6	2011/3/22	8:41	Non-sunny	0.711	0.817	0.874	0.868
7	2011/3/23	15:41	Non-sunny	0.771	0.780	0.817	0.845
8	2011/3/24	6:41	Non-sunny	0.701	0.850	0.881	0.867
9	2011/3/24	9:41	Non-sunny	0.814	0.772	0.865	0.848
10	2011/3/26	7:41	Sunny	0.278	0.650	0.602	0.895
11	2011/3/27	17:41	Sunny	0.659	0.815	0.840	0.876
12	2011/3/28	16:41	Sunny	0.426	0.792	0.684	0.829
13	2011/3/29	17:41	Sunny	0.617	0.793	0.774	0.819
14	2011/3/29	18:41	Non-sunny	0.665	0.921	0.923	0.936
15	2011/4/1	18:41	Non-sunny	0.677	0.805	0.813	0.873
Average				0.563	0.754	0.733	0.833

Table 2-5 Sr value of Data2012 (Figure 2-4b)

Image No.	Time		Illumination condition	Segmentation Methods			
				ExG	ExG-ExR	MExG	DTSM
16	2012/4/17	8:55	Non-sunny	0.618	0.787	0.762	0.713
17	2012/4/17	11:55	Non-sunny	0.618	0.734	0.715	0.757
18	2012/4/18	8:55	Sunny	0.218	0.417	0.302	0.757
19	2012/4/20	16:55	Non-sunny	0.688	0.868	0.815	0.780
20	2012/4/21	4:55	Non-sunny	0.276	0.682	0.712	0.606
21	2012/4/24	8:55	Sunny	0.209	0.655	0.355	0.889
22	2012/4/26	11:55	Non-sunny	0.724	0.896	0.860	0.848
23	2012/4/27	10:55	Non-sunny	0.743	0.907	0.878	0.876
24	2012/4/28	7:55	Sunny	0.553	0.830	0.711	0.911
25	2012/4/29	7:55	Sunny	0.410	0.706	0.560	0.888
26	2012/4/29	8:55	Sunny	0.373	0.718	0.577	0.896
27	2012/4/29	13:55	Sunny	0.412	0.662	0.545	0.876
28	2012/5/1	10:55	Non-sunny	0.793	0.932	0.905	0.924
29	2012/5/1	13:55	Non-sunny	0.611	0.858	0.862	0.870
30	2012/5/1	16:55	Non-sunny	0.629	0.857	0.844	0.880
Average				0.525	0.767	0.694	0.831

Chapter 3 Applications of DTSM for time series images taken in natural outdoor environments

3.1 Introduction

Given the growing demand for high-throughput phenotyping to support crop breeding, in recent years researchers have been conducting their experiments in fully automated facilities, and have thus been able to successfully assess crop growth and performance with a combination of modern technologies including genetic engineering, robotics and imaging (Bylesjö et al., 2008; Granier et al., 2006; Hartmann et al., 2011). Image analysis techniques are also considered as a powerful tool for use in plant phenomics (Furbank and Tester, 2011; Houle et al., 2010). However, the most useful phenotypic information about crops that are planted in fields is still being obtained by manual sampling that is not ideal because it destroys plants in the process, and it is extremely labor intensive and thus time consuming. The reason so much time is required is that for a given object/region under field conditions, images acquired with digital photography include a wide variety of light intensities, so their analysis involves careful, individual treatment that demands specialized knowledge of observers. Understanding crop biology under field conditions is extremely important, yet it has been impractical for the above reasons, because most phenotyping studies require dealing with very large crop populations. To address those problems, a new method was developed using a bar-code system to help researchers semi-automatically measure and automatically record the phonemic traits of individual crop plants, which is at least three times faster than the traditional approach (Buckler et al., 2009; Yamasaki and Arturo, 2012). Furthermore, when cost is not a limitation, by combining different kinds of imaging equipment such as RGB, Multispectral, Hyperspectral, and Lidar cameras, not only visible traits (color, shape, height, etc.) but also invisible traits such as water content or disease can be detected and measured (Araus and Cairns, 2013; Prashar et al., 2013). Since our objective was to develop a low-cost, easy, high-throughput phenotyping system for use in the field, we focused on assessing the cost, accuracy and feasibility of using only normal RGB images to measure the appropriate phenotypic traits of crops.

In the previous chapter, we proposed an effective and efficient DTSM method to extract vegetation from images taken under diverse light conditions in the field. In this

chapter, we first discuss the ability of DTSM to measure the plant canopy coverage ratio for wheat fields and rice paddies at different growth stages. We also consider other possibilities for evaluation of phenotypic traits from RGB images acquired with a normal consumer digital camera in outdoor settings.

3.2 The ability of DTSM to measure plant canopy coverage ratios

The "plant canopy coverage ratio" is one of the well-known parameters that indicate plant growth. It is usually defined as the percentage of the orthogonal projection area relative to the area of crop foliage in the horizontal plane. This ratio is reported to be highly correlated with Leaf Area Index (LAI), canopy light interception, nitrogen content, and crop yield (Campillo et al., 2008; Fukushima et al., 2003; Takahashi et al., 2012). Since our proposed DTSM provides efficient extraction of crop vegetation regions from images taken under various field conditions, the canopy coverage ratio within a given photograph can be calculated easily by dividing the total number of pixels in the vegetation segments of the image by the total number of pixels in the entire image.

To evaluate the effectiveness of using our DTSM to calculate canopy coverage ratios, we first evaluated the example wheat images used in Chapter 2. For the whole test image data set, the canopy coverage ratios derived from DTSM-segmented images are compared with the value derived from hand-segmented images (considered as true value) in Figure 3-1. Values measured by the DTSM method were positively linearly related to values from the hand segment method (R^2=0.85), with a slight overall underestimation (slope=0.81). Note that variation in canopy coverage was rather narrow (4 to 20%) because we used images from early stages in the growing period.

Furthermore, to confirm the robustness of this new method for natural environments, we also compared the canopy coverage ratio from two separate groups: images taken under sunny (Figure 3-2) and cloudy conditions (Figure 3-3), as in Chapter 2. Again the ratios determined by DTSM and the hand segment method were closely associated, with R^2=0.80 for sunny images and R^2=0.89 for cloudy images; the respective slopes of 0.72 and 0.85 also represent a slight underestimation by the DTSM method under both light conditions. Considering the narrow range (up to 20%) of the canopy coverage ratio during the initial growth stage, the results are acceptably good.

The overall results indicated that DTSM is able to accurately evaluate the canopy coverage ratio of wheat crops, based on only RGB images taken under field conditions.

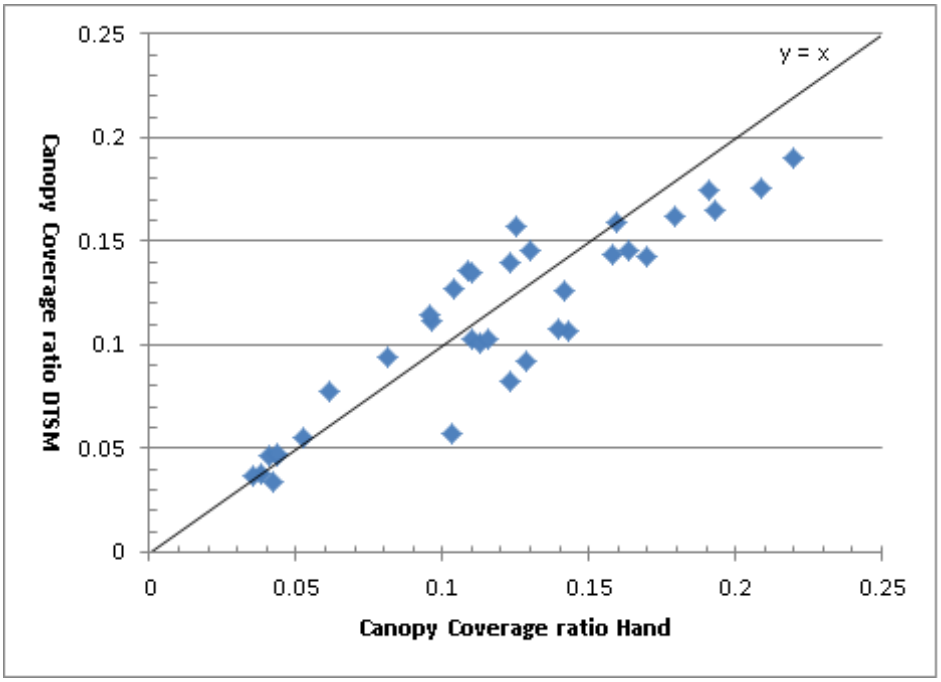


Figure 3-1 The relationship between the canopy coverage ratios by DTSM and manual segmentation. The comparison involves the images taken under all light conditions. The solid line represents a 1:1 relationship.

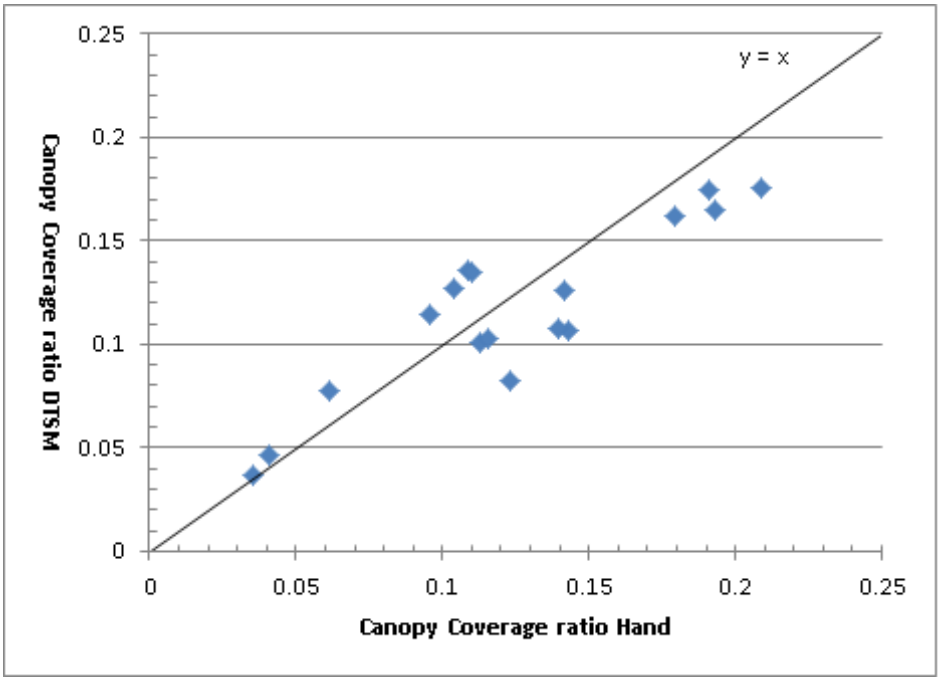


Figure 3-2 The relationship between the canopy coverage ratios by DTSM and manual segmentation. The comparison involves only the images taken under sunny light conditions. The solid line represents a 1:1 relationship.

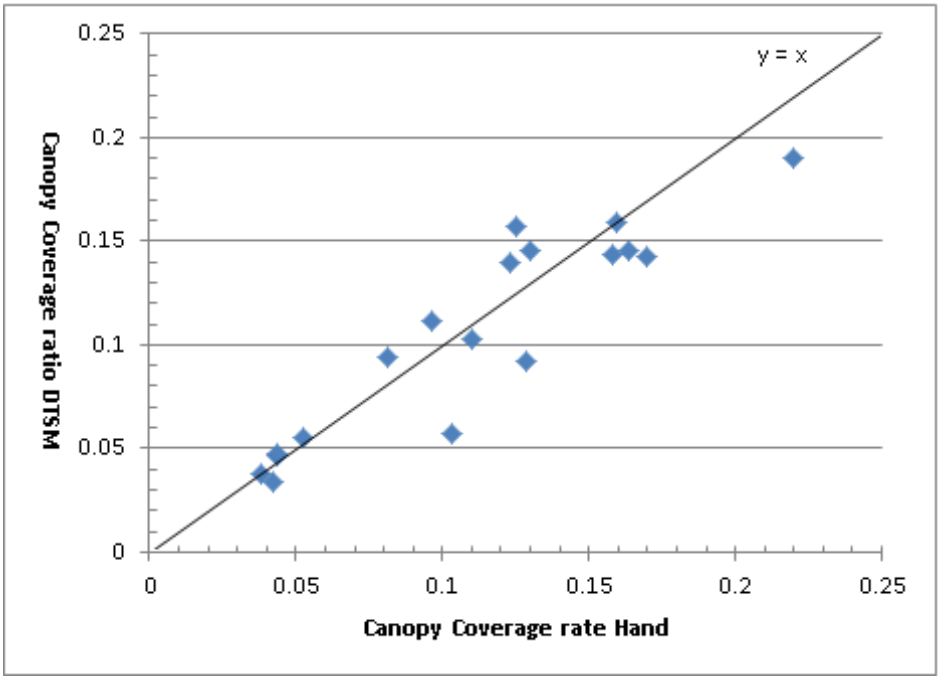


Figure 3-3 The relationship between the canopy coverage ratios by DTSM and manual segmentation. The comparison involves only the images taken under non-sunny light conditions. The solid line represents a 1:1 relationship.

3.3 Evaluation of plant canopy coverage ratios from time series images of rice paddies taken under natural light conditions in the field.

3.3.1 Experimental materials and image acquisition.

For use as the target crop in this experiment, a japonica rice variety *Kinmaze* was sowed on 26 April 2013 and transplanted on 31, May, 2013 at an ordinal outside field of the Institute for Sustainable Agro-ecosystem Services, University of Tokyo, Japan. We used the image acquisition system shown in Figure 3-4. A Canon EOS Kiss x5 digital camera with EF-S18-55mm lens was mounted above rice crops; time-lapse images were taken at 1-h intervals and transmitted to a free webserver (flickr.com) via 3G network (Fukatsu 2011, 2012).

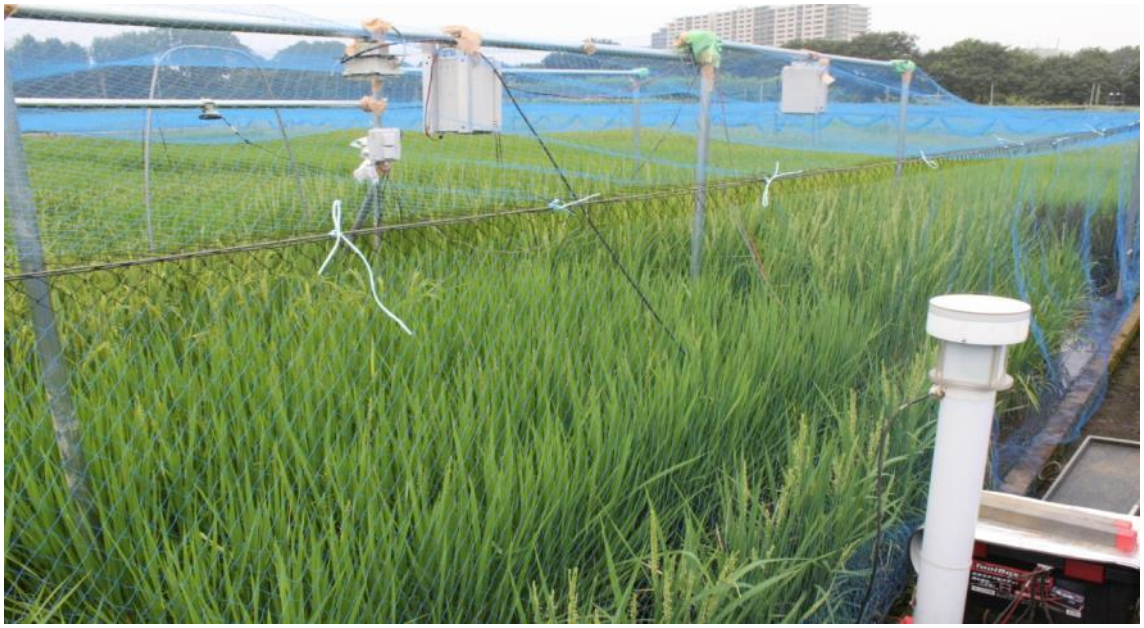


Figure 3-4 Field monitoring system.

The data set of images used to calculate canopy coverage ratios was acquired from 19th June (20 d after transplanting) to 16th August (~1 week before heading), every day from about 8:00AM ~ 16:00PM. To prevent damage by foraging birds, a blue net was placed over the whole experimental field on 30 July. Figure 3-5 shows six different images obtained during the growth stage.

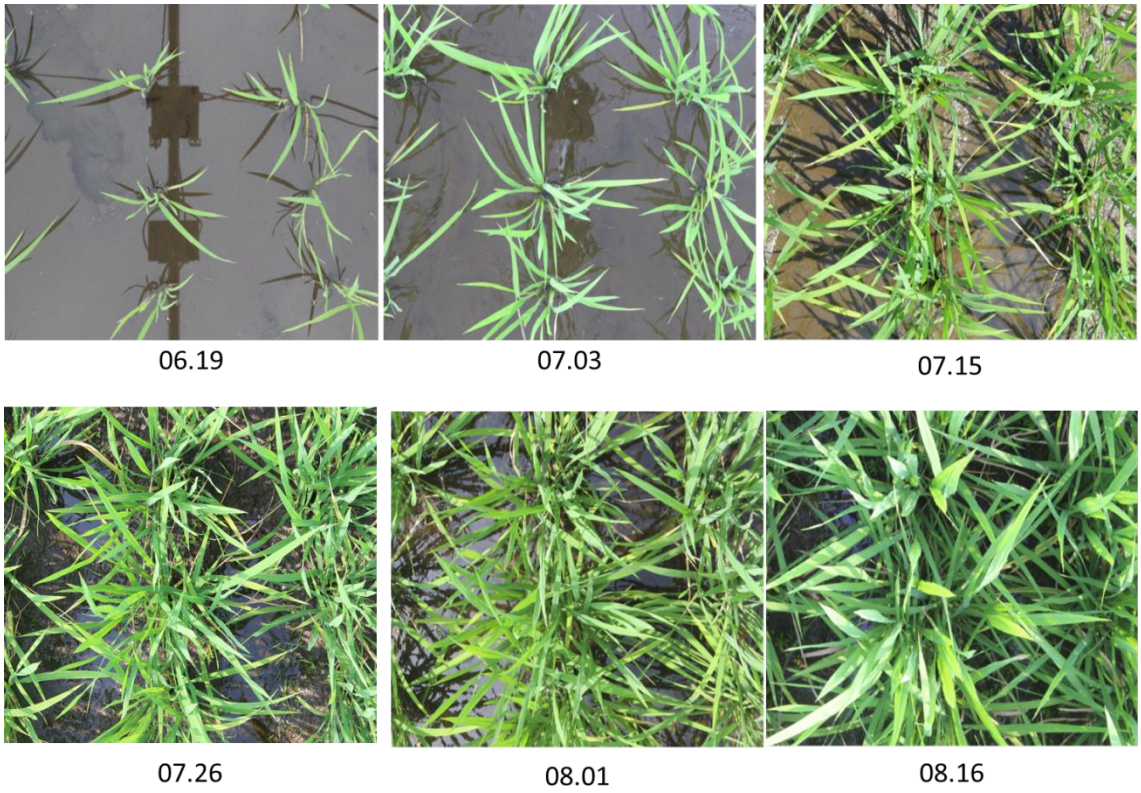


Figure 3-5 Examples of field images obtained throughout the period of observation.

3.3.2 Selection of training data set for DTSM development.

DTSM is a kind of supervised machine learning approach, and deciding the properties of training data that will be used is the most critical step. In our experiments, to enhance the robustness in illumination change, the images used for training data acquisition needed to be preselected and taken under different light conditions. The precise number of images to use is difficult to decide, because too much training data causes overestimation, and too few training data causes underestimation. So, for our DTSM-based approaches, selection of training data needed to be a manual process that required serious attention. Thus, in this chapter, we used the software package Matlab (MathWorks Inc., Natick, U.S.A) to develop a small tool with a graphical user interface that facilitates selection of the training data from six predefined training images. The process of training data selection is as follows:

Start the “Collect training data” function, and follow the notifications to select the point from images that appear automatically by clicking the mouse (Figure 3-6 through Figure 3-10). This is the most important step of the DTSM, directly affecting its ability

to isolate vegetation from the image backgrounds. For all training images, users are required to collect the training data by clicking the mouse: the cursor position of one click is set as the center of a rectangular region, and the color features of all pixels inside the region are automatically extracted and sorted. Users must select the training data separately for the image foreground and background, which are considered as positive and negative classes in the machine learning model. It is highly recommended that all possibly different parts of each training image should be selected (for example, the shadowed part and spectral reflected part of the same image).

In this experiment, 10 training images were manually selected from the images taken during the initial and early middle growth stages (between the 20-th day and the 78-th day after the transplanting), considering the variation of weather and light conditions. Then, training data for two classes were carefully selected from each of 10 images.

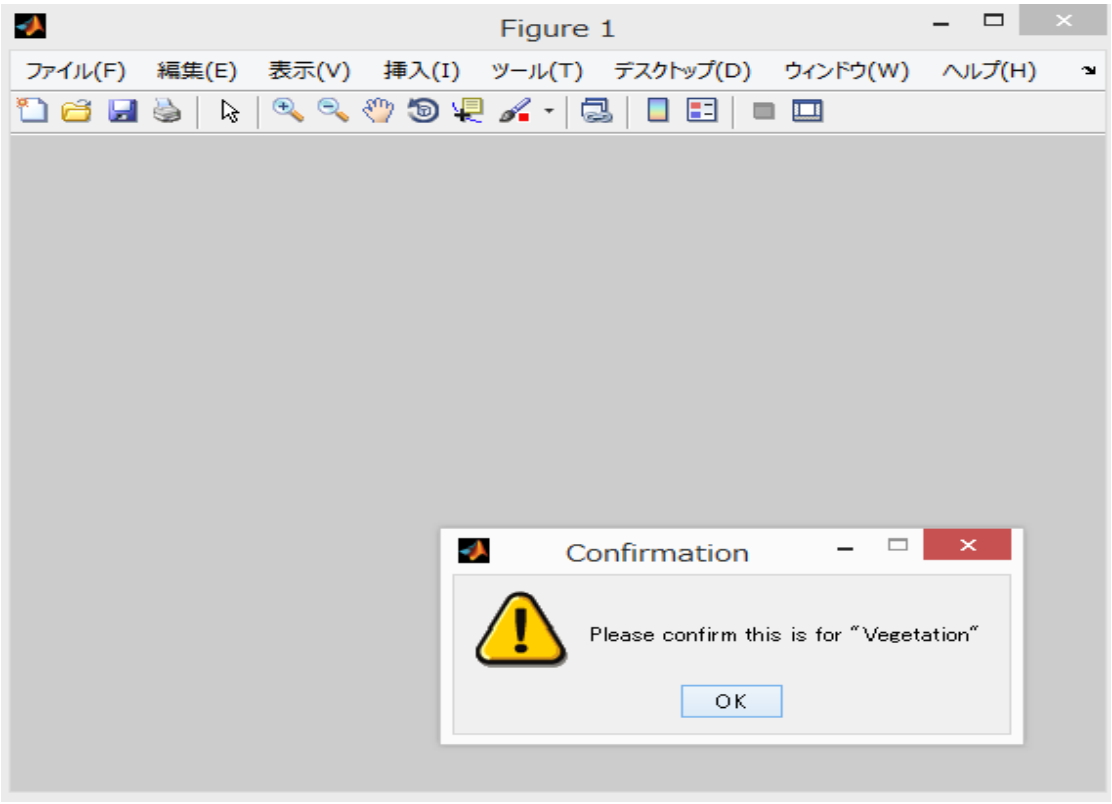


Figure 3-6 Notification of foreground (vegetation) training data collection.

Figure 3-7 Collection of foreground (vegetation) training data.

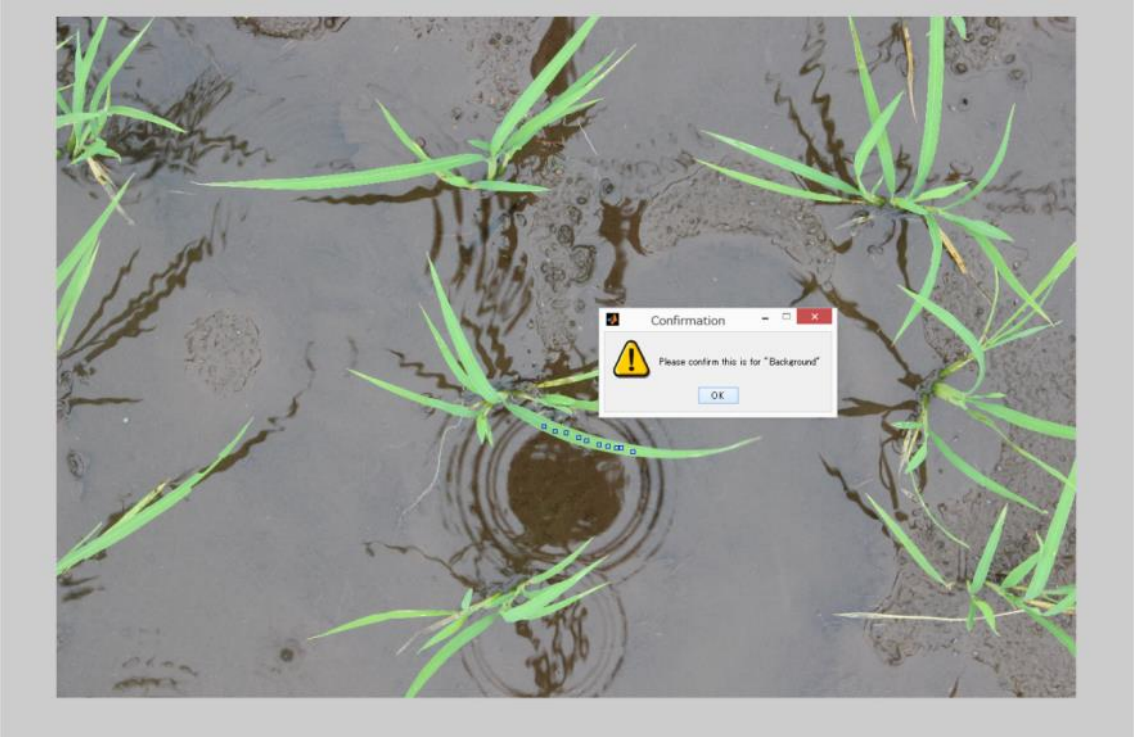


Figure 3-8 Notification of background training data collection.

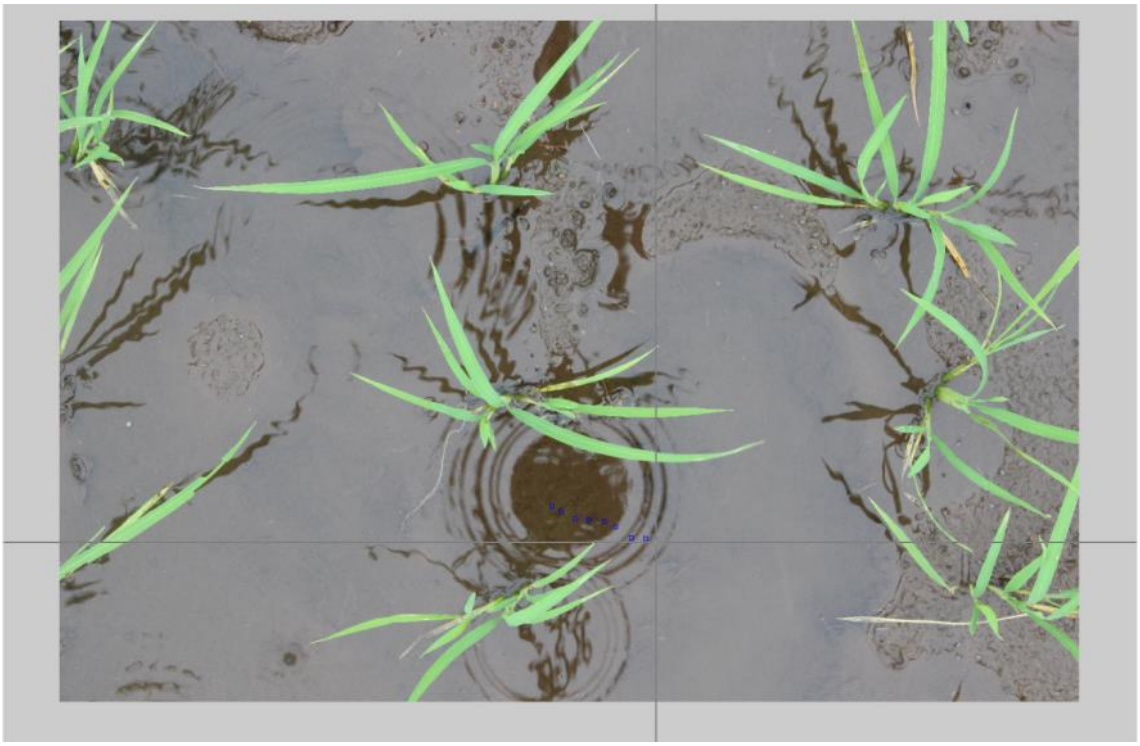


Figure 3-9 Collection of background training data.

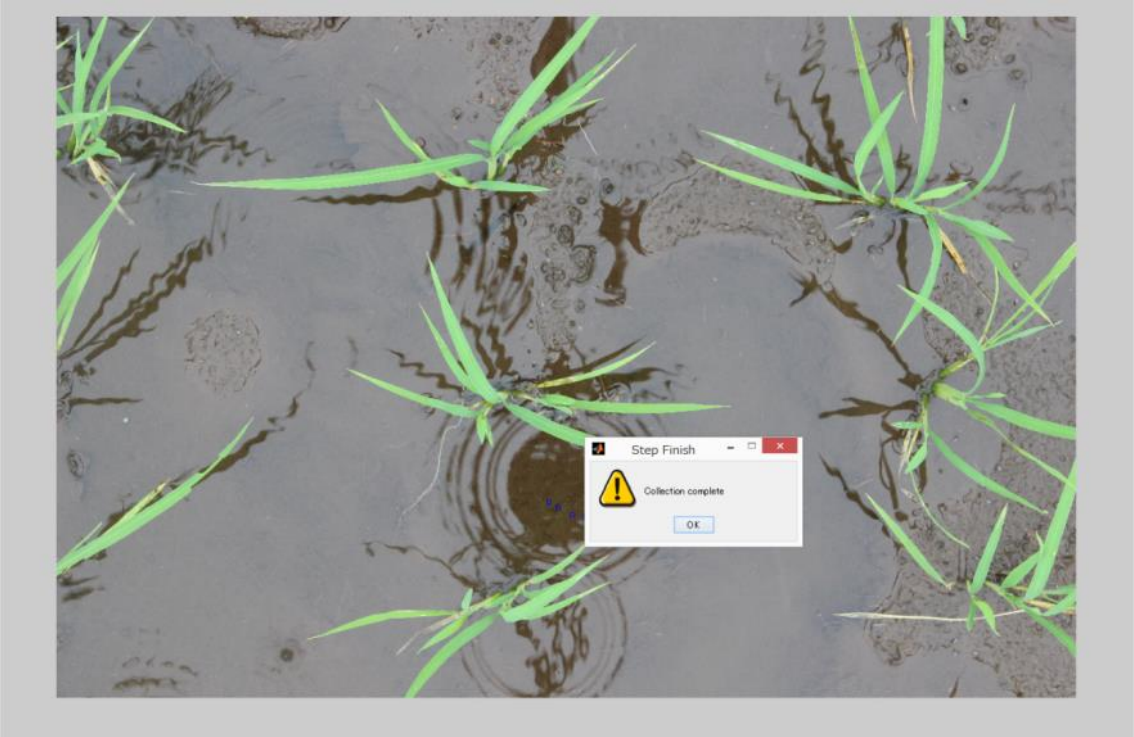


Figure 3-10 Notification that collection of training data is complete.

3.3.3 Experimental results.

Canopy coverage ratios are derived using the proposed DTSM (Figure 3-11); the x axis indicates the days after transplant, and the y axis indicates the canopy coverage ratio. The black plus marks represent the calculated value based on DTSM and the red dots represent values derived from the same images by careful hand-isolation of the vegetation (i.e., true values). The graph shows that canopy coverage ratio increases as time since transplant progresses. There was a high positive correlation between the DTSM-derived values and true values for canopy coverage ratios, with R^2=0.99 and a slope of 0.96 (in Figure 3-12).

Three suspicious data points resulted from the evaluation of vegetation cover by DTMS (circled red in Figure 3-11). The first dubious point suggested that the canopy coverage ratio on day 20 was greater than the value on day 21, a result that was caused by strong wind (Figure 3-13). Another dubious point, in the middle of the graph, suggested that the coverage ratio of day 63 (the image '201307311601') suddenly increased, and declined two days later at day 65; this temporary flattening of several stems was caused by an unknown source (possibly raccoon dogs) and lasted for two

days (Figure 3-14). The last dubious point suggested that the canopy coverage ratio suddenly dropped from 73% to 58%. The reason is that the raw images were taken at the 77-th day after transplanting, when the rice plants grew larger and had many overlapping leaves. Due to the direction of sunlight, some of leaves near the ground are covered with strong shadows, which greatly weakens the color features in digital images therefore also weakens the ability of DTSM to accurately isolate the vegetation. Evaluation results of image '201308151502' and another taken 1 h previously ('201308151402') are shown in Figure 3-15. The shadows in both of them caused vegetation coverage to be significantly underestimated, but image '201308151502' loses more vegetation pixels because of the larger dark areas within the image. This error is acceptable because we did not use any vegetation pixels from shadowed parts of those images as training data — i.e., the model classified those dark pixels as those belonging to the background class because the color values are close to the training data of background elements. This is the weakness of DTSM which is strongly affected by the selection of training data. The training data are selected manually, which makes it difficult to include all the possible dubious cases to address the aforementioned underestimation problem. As a preliminary experiment, we added 12000 pixels selected from the dark regions of the crop image to the training data set for the vegetation class, and trained the DTSM model again with the new training data. Figure 3-16 shows the comparison of the segmentation and the coverage ratio by the newly constructed DTSM model and the former model without the training data from the dark regions, . The new model improved the segmentation result for the image '201308151502',with the true coverage ratio of 76%, raising the estimated value from 58% to 79%, whereas the result for '201308151402' which does not have dark shadowed region keeps almost the same accuracy. And the values obtained at the early growth stage also remained almost the same as before (Figure 3-17), and the R^2 between the DTSM derived values and true values is as high as 0.99 (Figure 3-18).

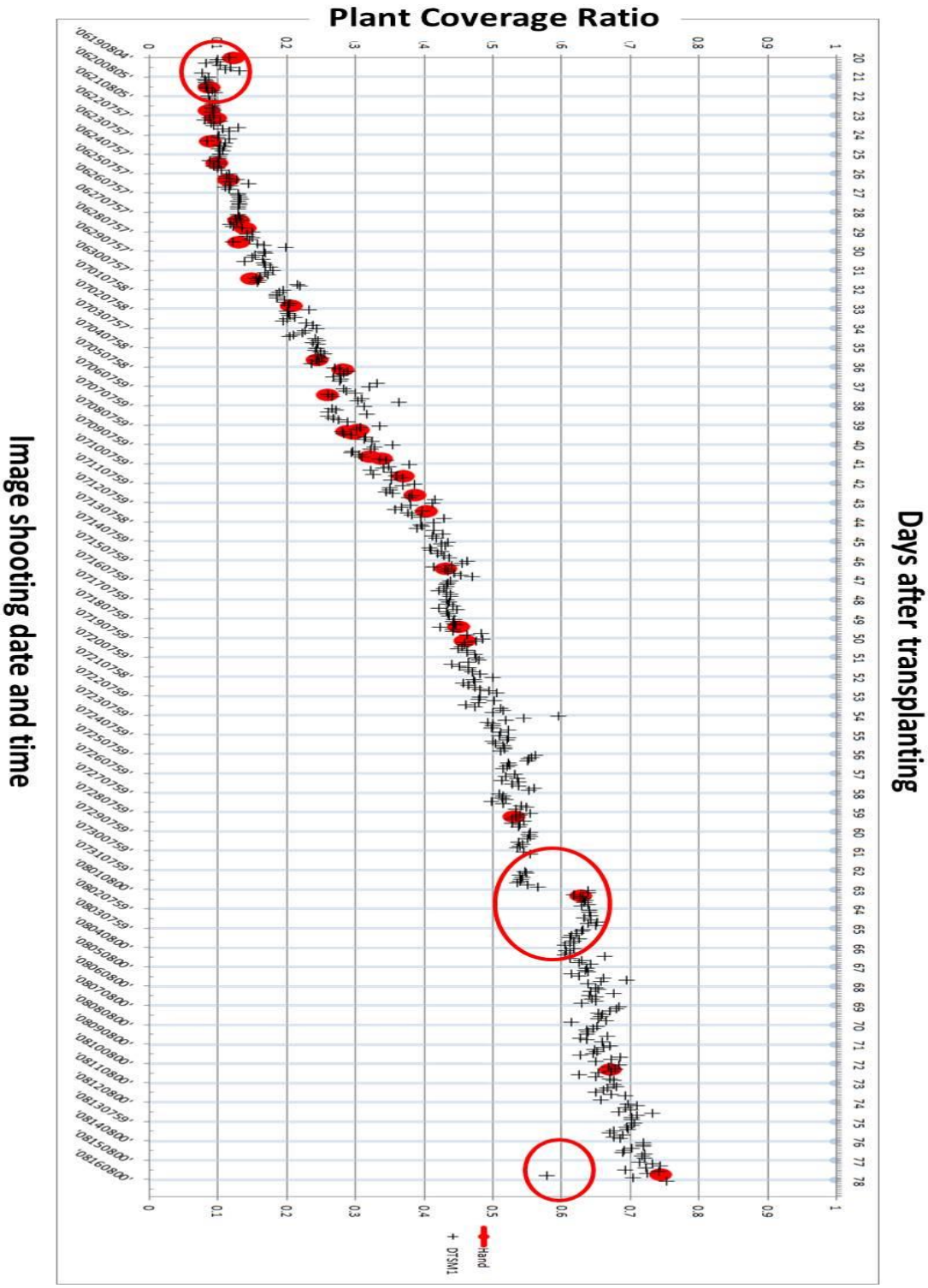


Figure 3-11 Canopy coverage ratio of rice paddies. Crosses and red dots indicated the estimated coverage ratios by DTSM and the manually derived true values respectively. See the text for the further details.

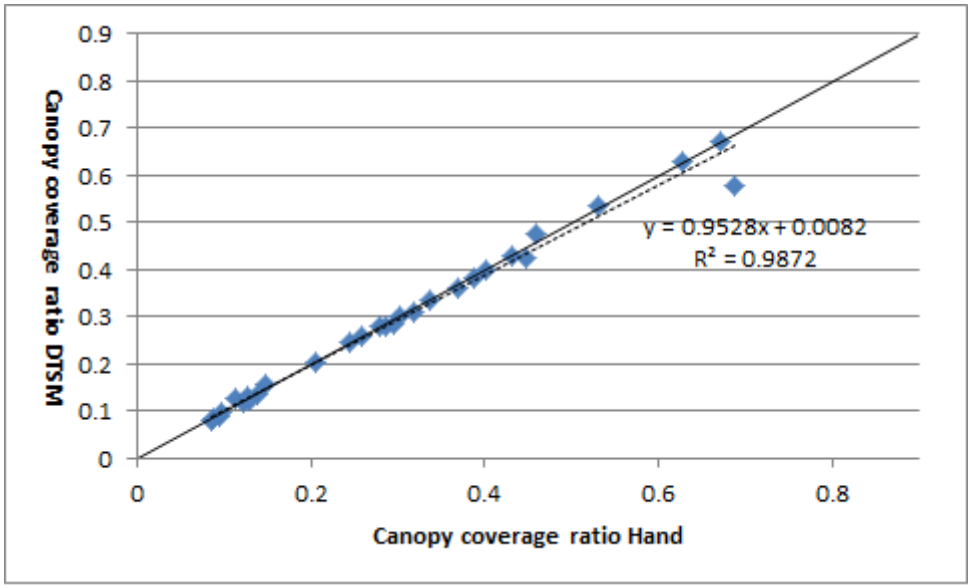


Figure 3-12 Relationship between the canopy coverage ratios by DTSM and the manually derived true values. The solid line and the dashed line represent a 1:1 relationship and the linear regression respectively.

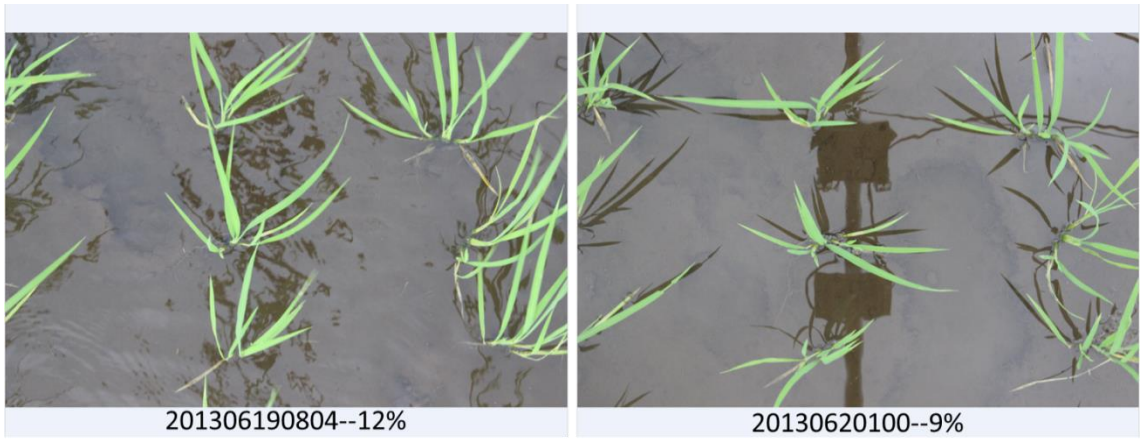


Figure 3-13 Wind strongly affects evaluation of the canopy coverage ratio from images. The image on the left side is taken in windy condition. The number before "--" such as "201306190804"indicates that the image shooting timing is 19th, Jun, 2013 at 08:04AM, and the number behind such as "12%" indicates that the evaluated canopy coverage ratio is 12%.

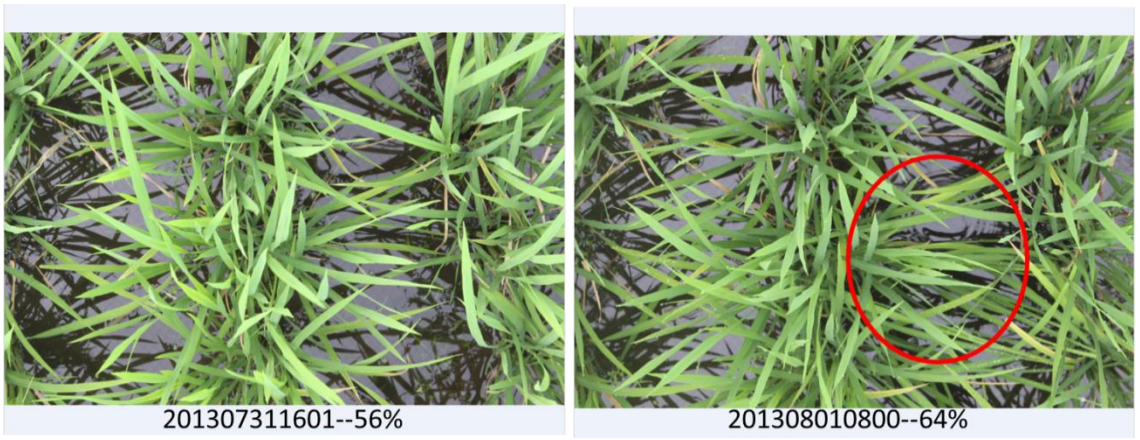


Figure 3-14 Physical disturbances of vegetation can affect the evaluation of canopy coverage ratio from images; e.g., the vegetation in the right image was slightly flattened by an unknown cause.

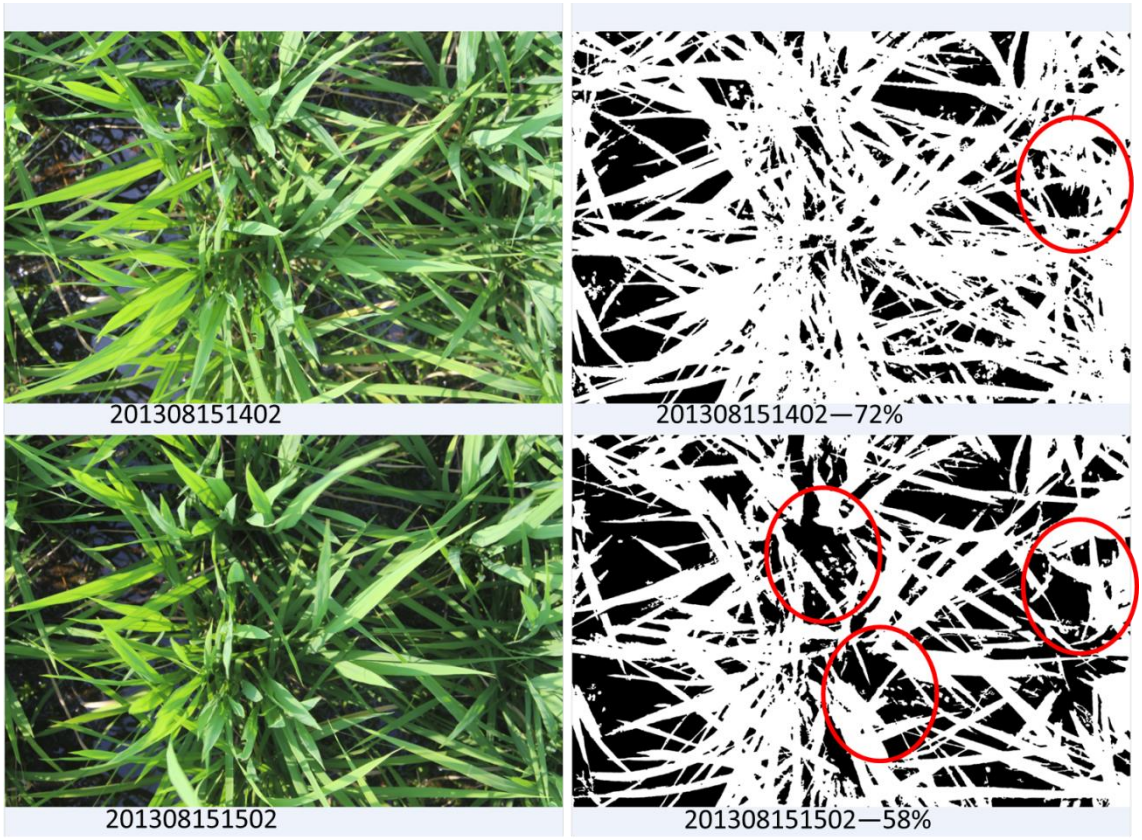


Figure 3-15 Dark shadows such as those that appear on the lower photo at left result in underestimation of plant canopy coverage.

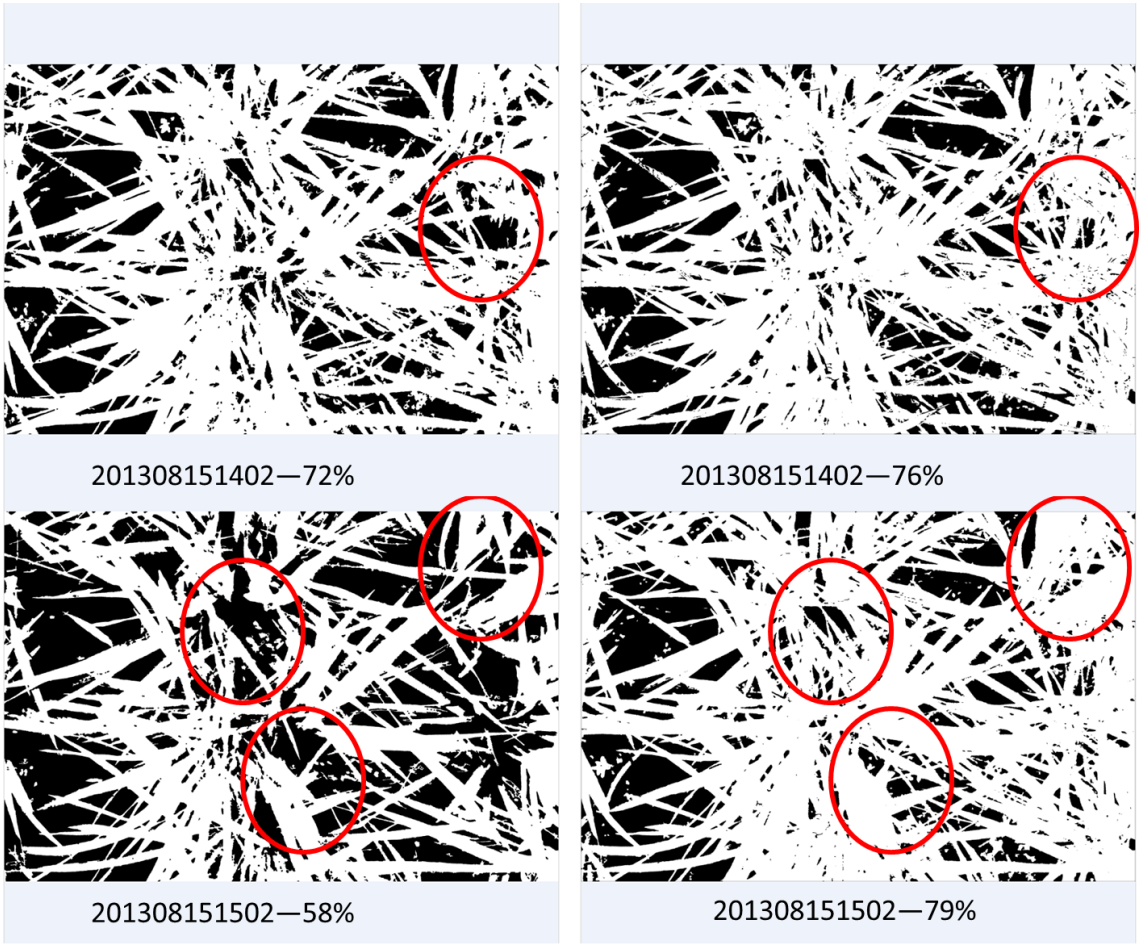


Figure 3-16 Comparison of the isolation of vegetation via DTSM, between the original model versus the model that includes information from the dark regions. Left column = results without the training data from dark regions; right column = results with the training data from the dark regions.

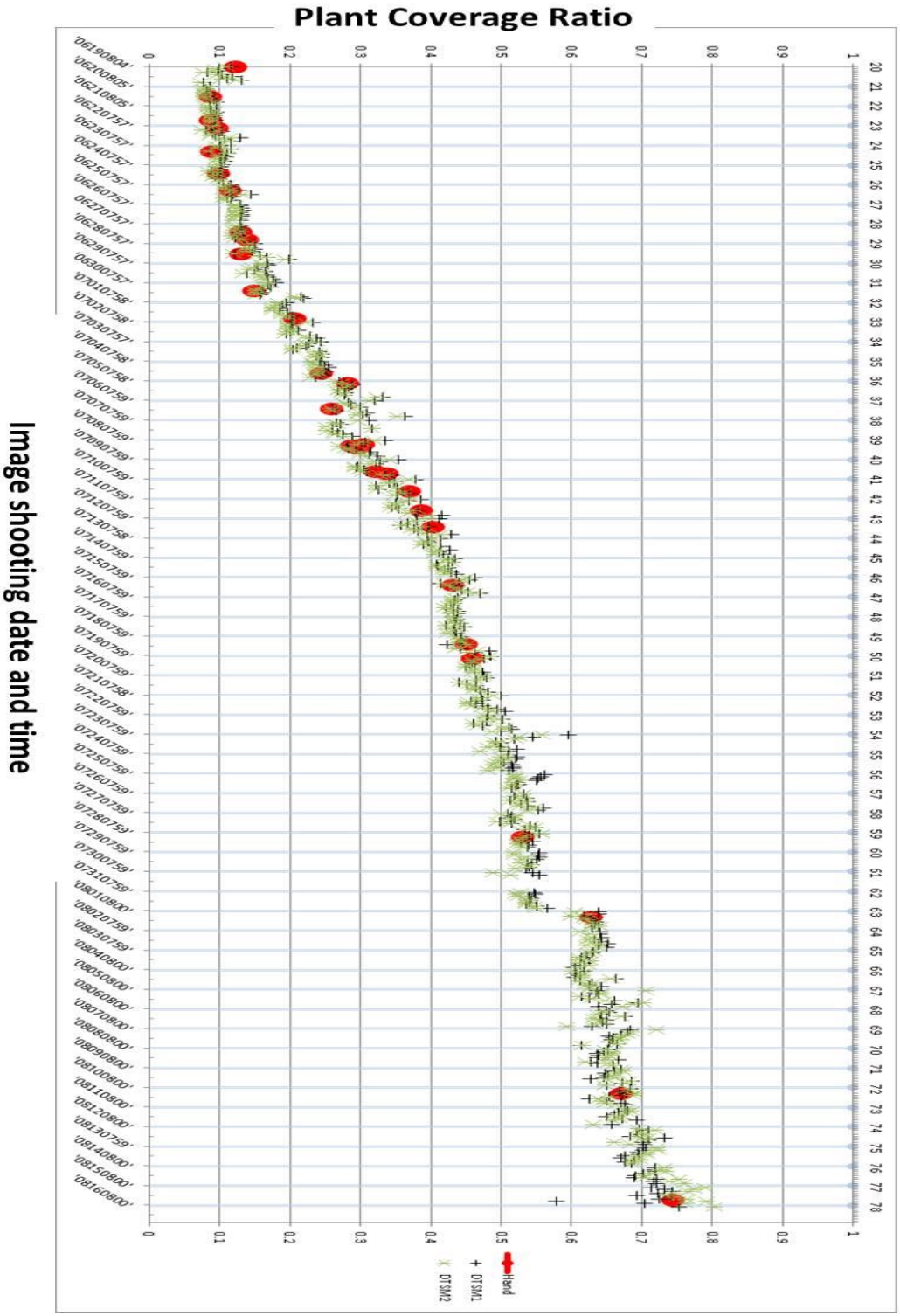


Figure 3-17 Comparison of rice crop canopy coverage ratios calculated by the DTSM trained without (black crosses) and with (green asterisks) the newly added training data from dark crop regions and manual segmentation (red dots).

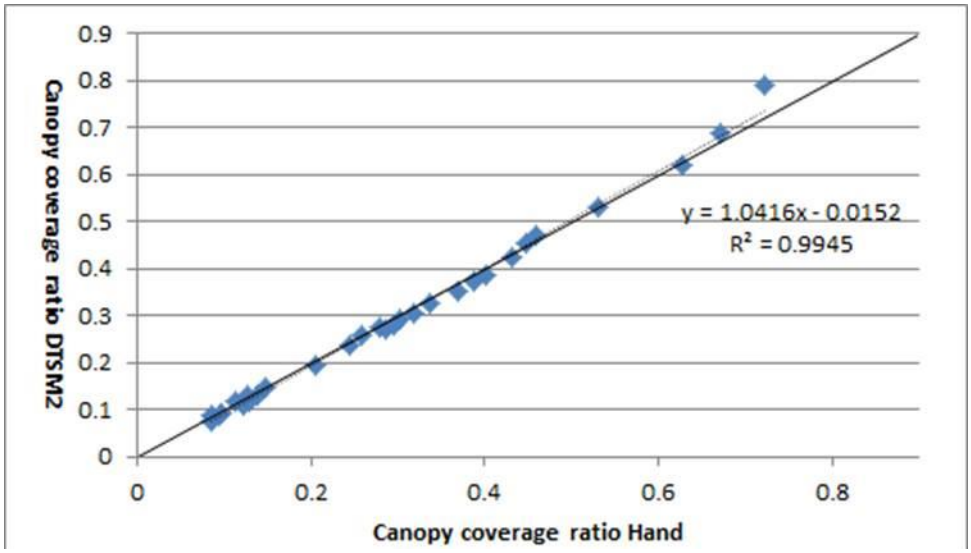


Figure 3-18 Relationship between the canopy coverage ratios by DTSM trained with the newly added training data and the manually derived true values. The solid line and the dashed line represent a 1:1 relationship and the linear regression respectively.

3.4 Discussion.

In this Chapter, we apply the DTSM algorithm to evaluating sets of images constituting a time series taken in wheat and paddy rice field. We derived one important phenotypic trait, i.e. canopy coverage ratio. The results demonstrated the high accuracy of DTSM to extract vegetation segments from other background elements in the images, though the robustness of this approach was influenced by various environmental factors such as wind, animal destruction. For future applications, we plan to combine weather information into DTSM in order to easily identify outliers caused by particular environmental conditions such as strong wind which affects the shape of rice paddies, especially during the initial growth stage. Such a possibility is shown in Figure 3-21 which shows that the fluctuation of the canopy coverage is caused by wind.

Because of equipment problems, the images for the canopy coverage estimation of the variety *Kinmaze could* not fully acquired for the whole growth stage — ideally we would like to evaluate the ability of our DTSM to measure plant canopy coverage ratios throughout the whole growth stage. As a preliminary experiment, we used images of another rice variety *Kamenoo,* taken from the initial growth stage to the mature growth stage though the manually derived true canopy ratio values are not available for the images. The estimated canopy coverage ratios (Figure 3-19) clearly simulated the transition from the initial growth stage to the maximum LAI stage, and the senescence during maturing stage. The crop grew quickly and plant sizes became stable in the later

middle growth stage, and then dropped down due to the growth of panicles and changes in leaf color. Although there are some values that stood out as being anomalous, they were able to be treated as outliers from the entire data set and could be removed by moving average. However, some other values were also noticeably different from the majority of results, which will affect overall accuracy of evaluation, such as the value circled red in Figure 3-19. We checked the relevant raw image data linked to those values and determined the cause: artificial objects appear in the image because of the field work (Figure 3-20). We realized that under field conditions, for an automatic image acquisition system, some "bad" images are inevitably included in the huge resulting datasets for plant canopy coverage ratio evaluation. To correct this problem, we can first apply the same DTSM to the entire image data set, in order to identify and automatically remove those poor images (whose evaluated plant canopy coverage ratios are extremely different from others).

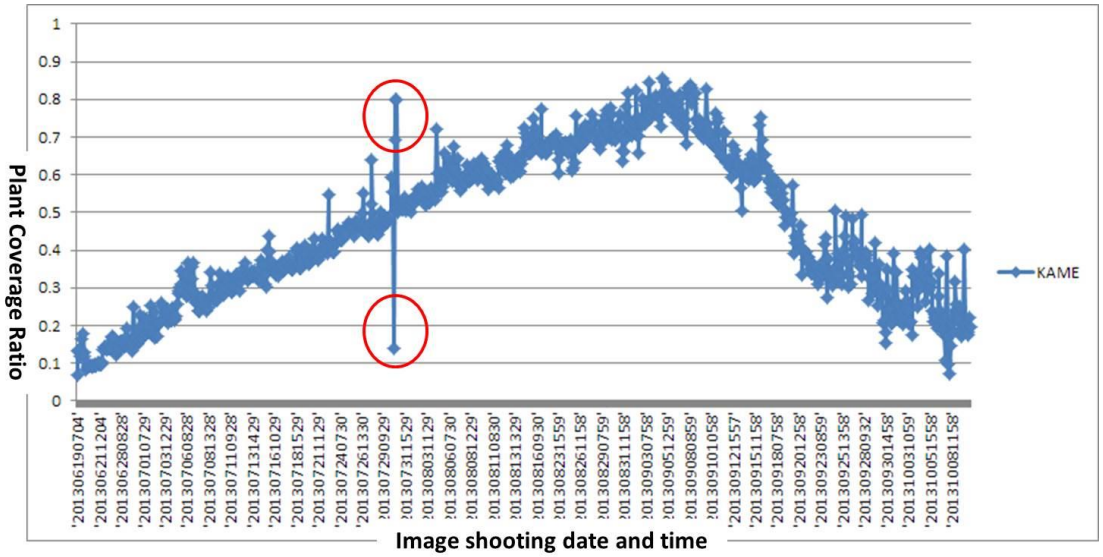


Figure 3-19 The canopy coverage ratios of *Kamenoo* crops throughout the whole growth stage.

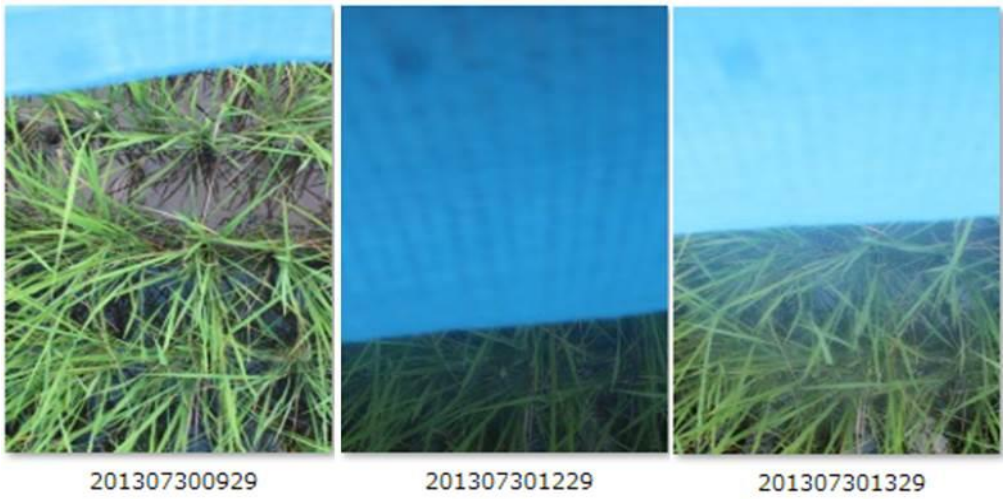


Figure 3-20 The blue sheet is included in the experimental image set.
The blue sheet is used for protect the net while cover it on field.

As mentioned several times before, the accuracy of the proposed machine learning based model is strongly affected by the selection of training data. Training data are currently selected manually and the quality of those data totally depends on the experience of selectors. In the future, we would like to develop a method to select optimal training data automatically.

Additionally, because the DTSM precisely isolated the vegetation from crop images taken under various light conditions, other phenotypic traits could also be extracted from the images. For example, a preliminary experiment confirmed that it is possible to count the number of leaves in a rice paddy during the initial growth stage (Figure 3-22).

However, the use of images taken by just a single camera from the top view over crops has its limitations. As a next step, we plan to add at least one or more cameras pointed in different directions to enable the accurate measurement of more crop phenotypic characteristics.

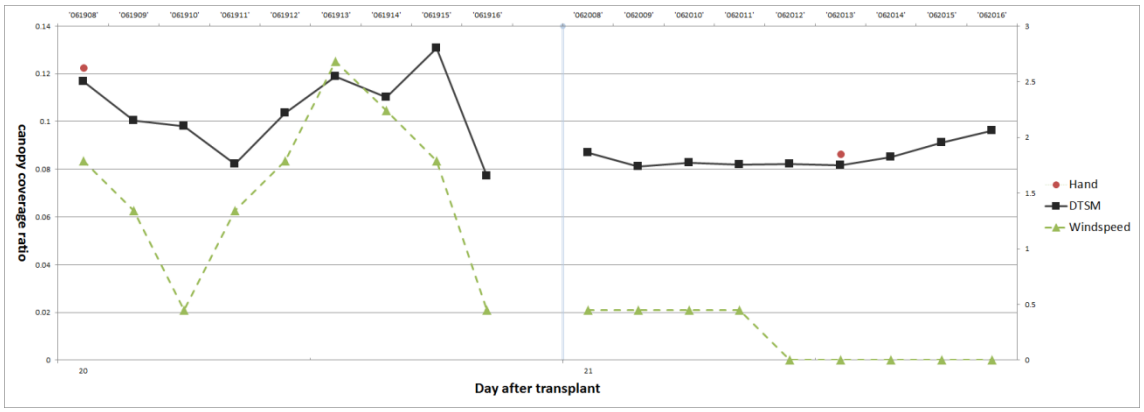


Figure 3-21 Results of combining wind information into the evaluation of canopy coverage ratios. The black line indicates the evaluated plant canopy coverage ratios at each observation time; the green dash line indicates the wind speed at the same time. The plant canopy coverage ratio changes if the observed wind speed fluctuates and keep stable if no wind blows.

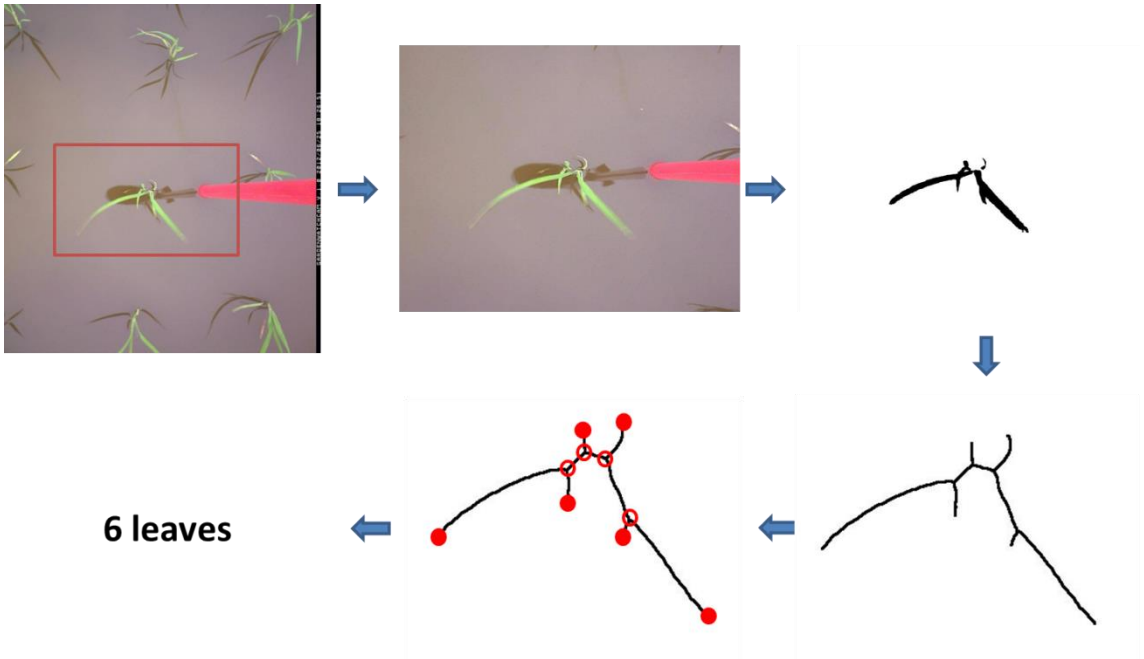


Figure 3-22 Experimental trial of the use of DTSM to count the number of leaves on a rice plant at the initial growth stage.

Chapter 4 Modification of DTSM to solve the problem of segmenting vegetation regions from complicated images taken under practical crop fields

4.1 Introduction

The previous two chapters demonstrated that the DTSM we developed is able to extract regions of crop vegetation from photographic images of both wheat and rice fields that are taken under diverse outdoor lighting conditions. As the next step, we targeted the use of the model to assess images taken in organic/natural farming fields. Herbicides are not or less used on such fields, so numerous varieties of weeds coexist with crops, such as the green algae and other floating weeds that are found in rice paddies (Figure 4-1). The color and shape of those weeds are highly similar to those of rice plants, creating a complicated background that makes segmentation of the crop vegetation regions from the photo extremely difficult. Until now, no published research has addressed this problem, to the best of our knowledge.

In this chapter, we detail our efforts aimed at finding a solution to extract only rice vegetation regions from images photographed under natural/organic farming field environments, whose numerous weeds lead to very complicated backgrounds.

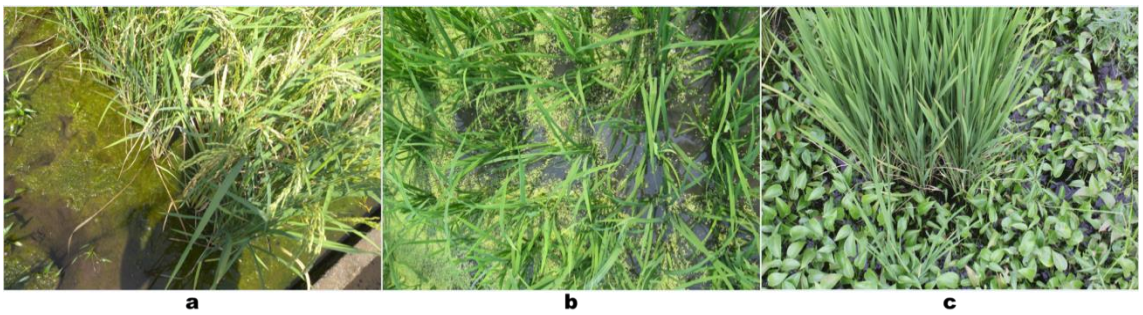


Figure 4-1 Weeds that commonly grow in organic/natural rice paddies.
(a). green algae; (b) floating weeds (c) other weeds. The tall, thin leaves are the rice plants.

4.2 Experiment and method

4.2.1 Materials and field experimental design

In this study, we used two varieties of paddy rice (cultivar: *Nihonbare* and *Takanari*) as the target crop. They were grown in an ordinal field at the Institute for Sustainable Agro-ecosystem Services, University of Tokyo, Japan. Both varieties were sown on 26 April 2012 and transplanted to the field on 8 Jun 2012. In order to simulate the growth conditions of an organic/natural farming field, we transplanted the individual rice plants into seeding boxes, and put each box on top of a platform to ensure that water remained under them. The simulated environment is shown in Figure 4-2.

For image acquisition, we used a waterproof digital camera (Garden Watch Cam, Brinno Inc., Taipei) that was fixed horizontally in a top-view position, targeting the rice paddies. Then, two series of RGB color images of the two varieties of rice were captured at one-hour intervals during the daytime throughout all stages of the growing season.

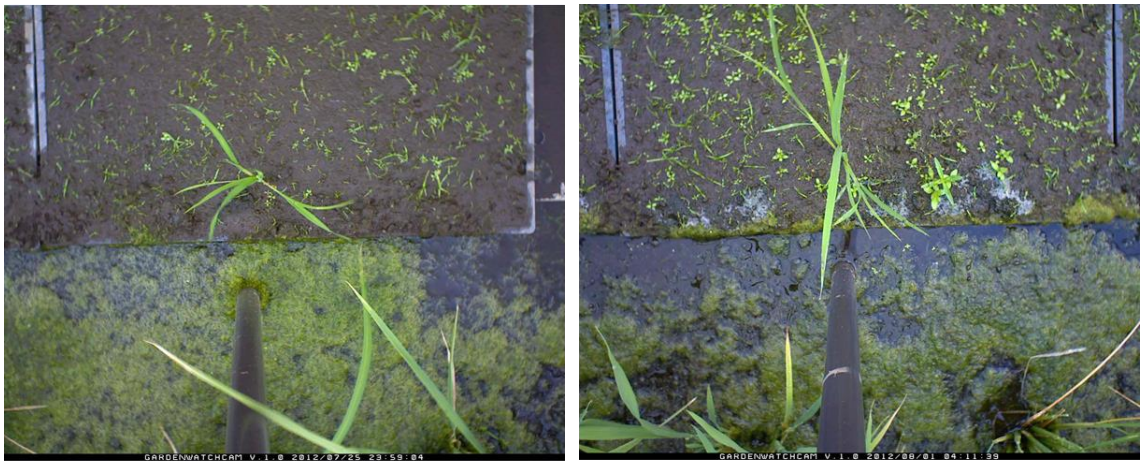


Figure 4-2 Experimental design of rice growth conditions.
Two varieties were planted: *Takanari* (left) and *Nihonbare* (right). Half of each image is covered with green algae floating on the water, and the other half is full of other weeds on the soil.

4.2.2 Results

As outlined in Chapters 2 and 3, the original DTSM algorithm includes the following

four main steps. (1) Generate training data sets by manually selecting pixels from training images; each pixel includes 18 dimensional color information from six different color spaces. (2) Generate the decision tree from training data. (3) Separate all pixels within a single test image into two classes, i.e., the vegetation class and the background class (non-vegetation: the soil or water). (4) Eliminate the misclassified pixels as "noise" to obtain the final resulting image that includes only the regions of crop vegetation.

As DTSM is a machine learning approach, its segmentation capability depends strongly on the selection of the training dataset — "good" training data results in a good segmentation performance. Because the background color of the images used in this experiment is almost the same as the color of the crop vegetation regions, it is difficult to separate them based solely on color features. Instead, we first constructed the DTSM by using the training data extracted from images that included green algae and weeds. Then, through assessment of the test image with the newly constructed model, every pixel of the image was classified into either the vegetation class or the background class, as described in Chapter 2. As an example of the results obtained by using two different DTSMs, Figure 4-3a shows an test image, and Figure 4-3b is the result acquired with the model that was constructed in Chapter 2; almost all the pixels of algae/weeds were misclassified into the vegetation class. The reason is that no color features of the weedy region were included in the background class for the model training. The result acquired by using the reconstructed DTSM is shown in Figure 4-3c. Although some pixels from weedy regions were added into the background class, most of them were still misclassified as belonging to the vegetation class. So in this case, only the addition of new training data to reconstruct the model enabled accurate extraction of the crop vegetation region. Moreover, the regions consisting of connected, misclassified pixels were not small enough to be treated as image noise in order to be eliminated by the noise reduction filter. To solve this issue, we first generated the edge image of the original image using the "Canny" method (Canny, 1986). By then combining it with the DTSM-segmented image, we were able to separate the image into several small regions. Finally, using spatial information from each small region and the noise reduction filter, the desired result was acquired. The whole treatment consists of the following steps, as detailed in Figure 4-4.

1. Segment the test image to obtain a binary result (image s1).
2. Use "Canny" method to acquire the edge image s2 from the gray level image of the original one.
3. Combine the eroded s1 and s2 images using logical ".AND" and ".XOR" commands to separate s1 into several regions and eliminate the small regions as the image

noises, shown in s3. The reason we erode s1 is to keep its original structure.

4. Check for regions that are constructed by connected pixels; each such region is surrounded by a bounding box, as shown in s4. Because of the different morphological features of rice plants versus weeds, the spatial penetration rate between them is different. We predefined a threshold for each region: if the spatial penetration rate was less than 70%, the region was classified as rice vegetation. The final result is shown in the last image of Figure 4-4.

Several examples of the segmented results are shown in Figure 4-5. The performance of the proposed method seems good, in terms of extracting the crop vegetation from the background. However, some unsuccessful segmentation also resulted, such as the region surrounded by a red rectangle in Figure 4-6. The likely cause is shown in Figure 4-7; the edge of that part was not detected because of the low resolution of original image. This also can be demonstrated by plotting the grayscale values in three dimensions, as shown in the right side of Figure 4-7. The black points are the manually placed markers of the position of the leaf pixels, while the other colors demonstrate the grayscale value of each pixel, and there is little difference between crop vegetation and the background, especially at the tip of the leaf.

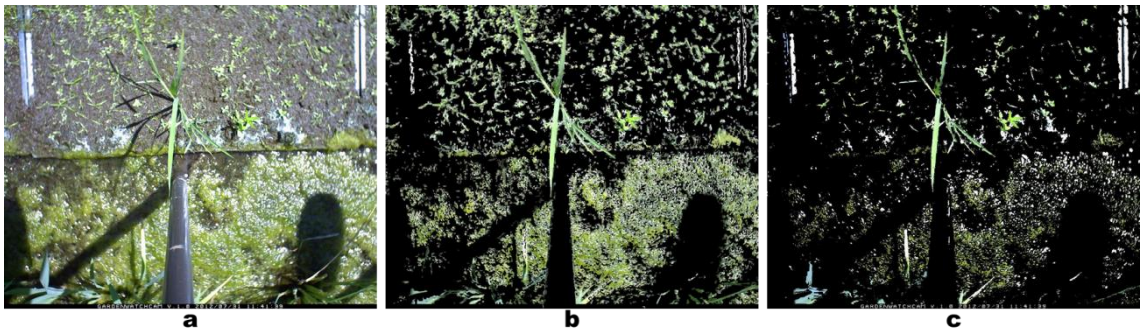


Figure 4-3 Segmentation based solely on color features: (a) the original image; (b) the segmented result of using the model from Chapter 2; (c) the segmented result of using the model conducted in this chapter. The difference between the two models is the inputted training data.

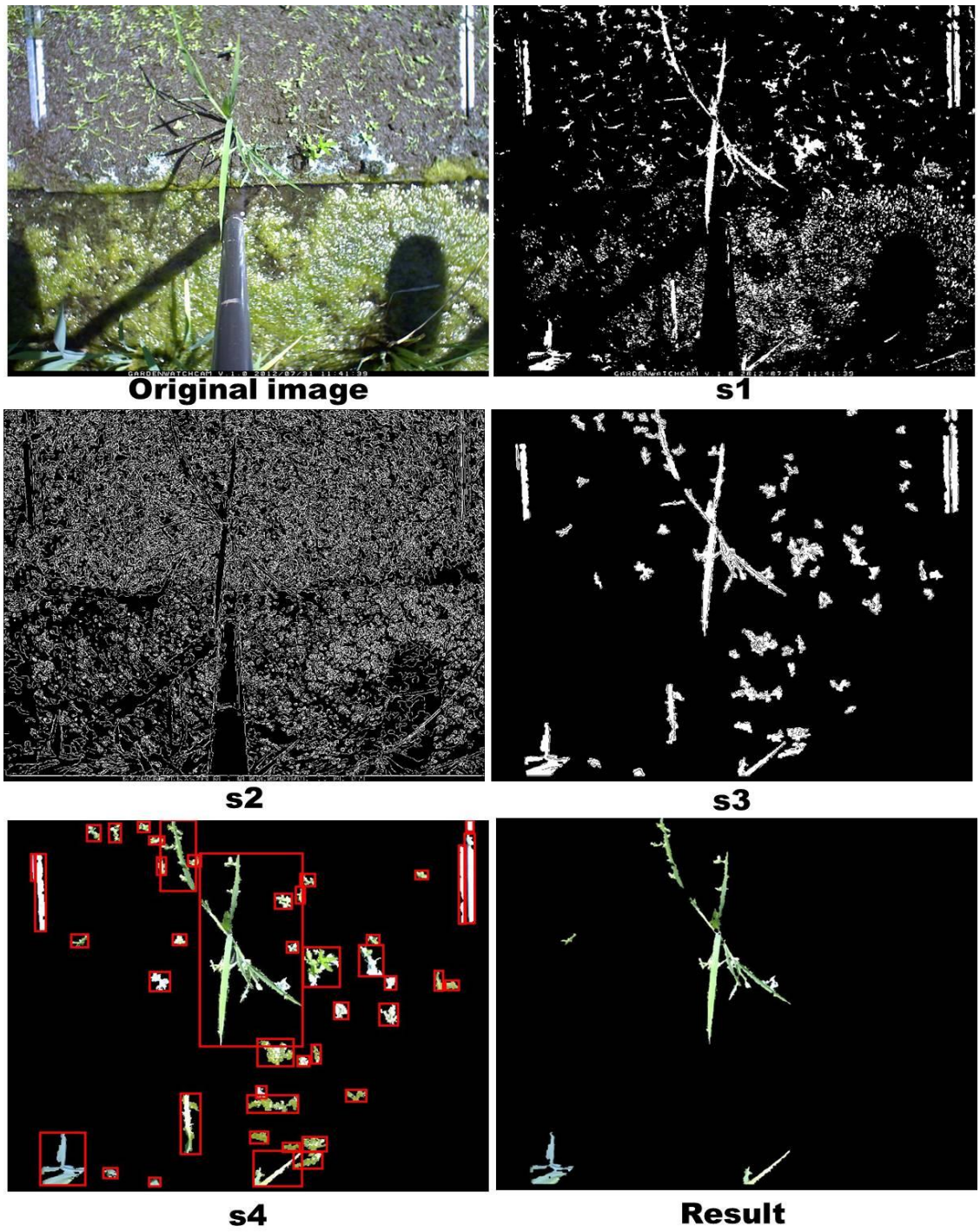


Figure 4-4 The steps involved in the proposed segmentation method. (S1) Binary image generated by DTSM. (S2) Edge image acquired using the "Canny" method.
(S3) Result acquired by combining the eroded s1 and s2 using logical ".AND" and ".XOR" commands/operators; all connected pixels are separated into small regions.
S4. Checking the spatial penetration rate of each bounding box (red rectangles).

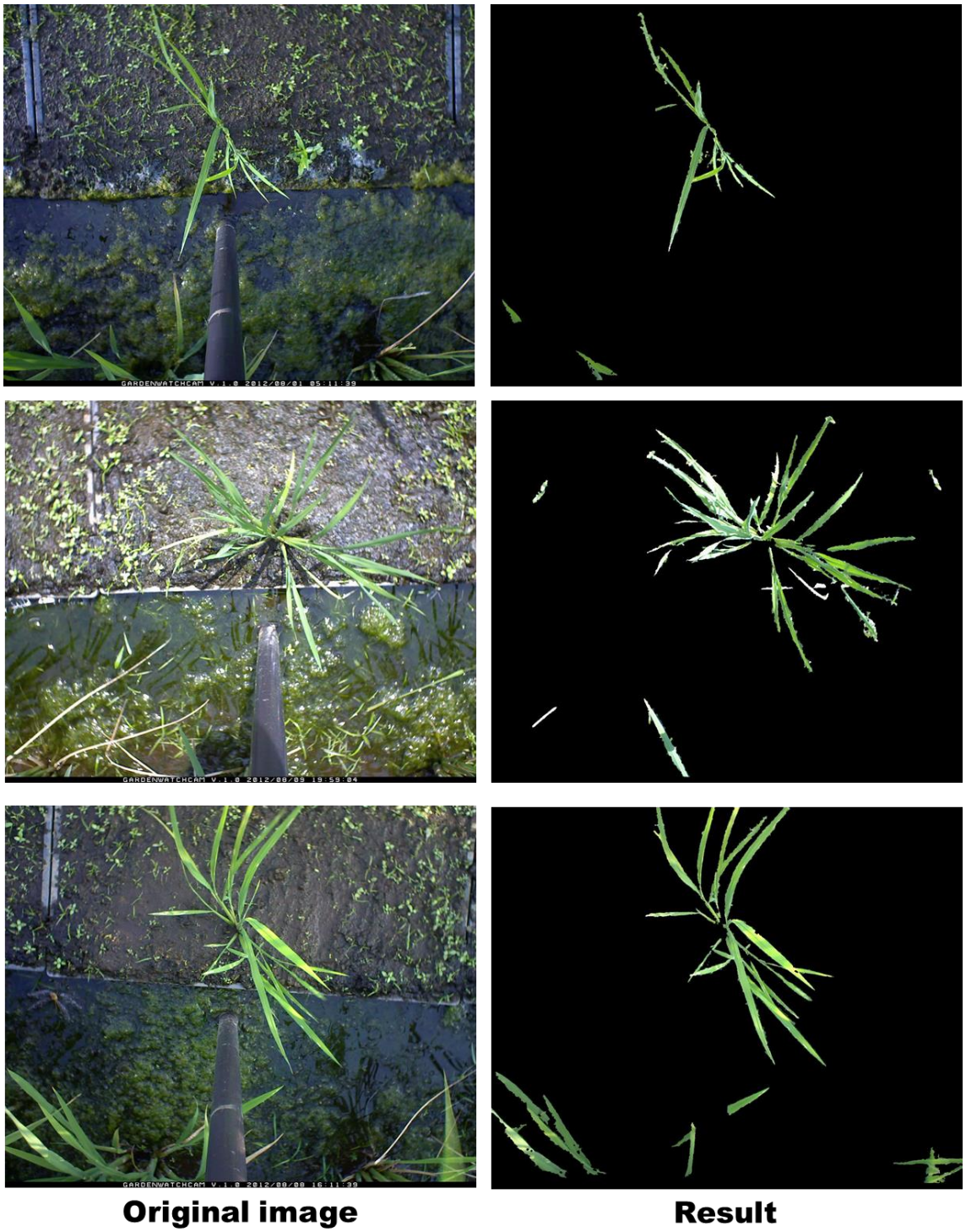


Figure 4-5 Results of vegetation segmentation from a complicated background.

Figure 4-6 Unsuccessful segmentation. An example where the background cannot be separated from the vegetation is enclosed in the red rectangle.

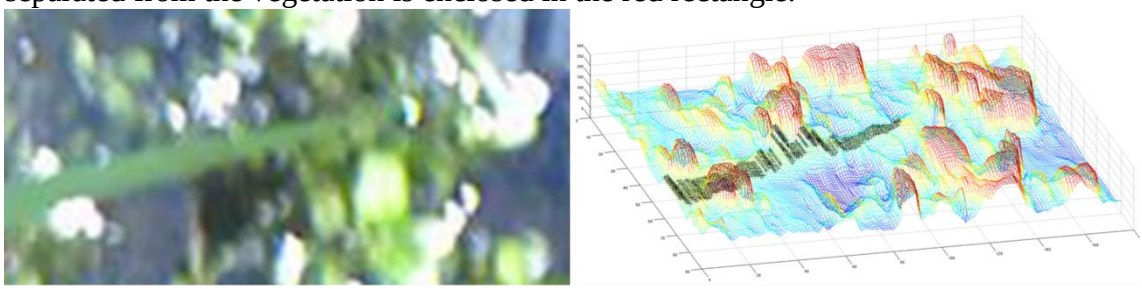


Figure 4-7 The reason for the unsuccessful segmentation: because of the image's low resolution, the edge of the vegetation cannot be distinguished from the background in the original image.

4.3 Evaluation of accuracy and discussion

In order to evaluate the accuracy of segmentation by the model, we prepared the true/reference binary images. Using an image retouching graphics application (Adobe Photoshop Elements 9, Adobe Systems Inc., San Jose, CA, USA) and a high-resolution pen tablet (WACOM intuos4 pen tablet, Wacom, Kazo), 28 reference images were carefully segmented by hand in order to derive the true values. The Qseg and Sr were then calculated as in Chapter 2, with a mean Qseg of 0.47 and mean Sr of 0.48 (Table 4-1),

The causes of low accuracy have been discussed. During the training data generation as shown in Figure 4-8, we targeted only one individual rice plant, while the leaves of other rice plants were not considered. For example, the smaller plants in the lower edge

of the images in Figure 4-8 had been also considered as belonging to the background class for the proposed method, while the manually made true images include them. If the test images only include the individual target plant, accuracy was about 0.63–0.70 for the example in Figure 4-9. As a preliminary study, this is acceptable even it cannot work for all situations.

Table 4-1. The accuracy of segmentation using the proposed method

Img. Num.	Qseg	Sr	Img. Num.	Qseg	Sr
1	0.283528	0.302799	15	0.660957	0.689729
2	0.160811	0.179341	16	0.524031	0.529018
3	0.301017	0.305933	17	0.575955	0.582744
4	0.278381	0.286358	18	0.587798	0.594648
5	0.261063	0.274193	19	0.550431	0.553332
6	0.235781	0.266977	20	0.599659	0.609274
7	0.31868	0.329862	21	0.611365	0.63578
8	0.483235	0.512159	22	0.570117	0.643238
9	0.527827	0.54214	23	0.561436	0.582178
10	0.343215	0.35717	24	0.530312	0.573349
11	0.490478	0.497875	25	0.53054	0.565237
12	0.391095	0.39235	26	0.535408	0.57529
13	0.42705	0.531629	27	0.642056	0.714542
14	0.549626	0.551658	28	0.499888	0.505789

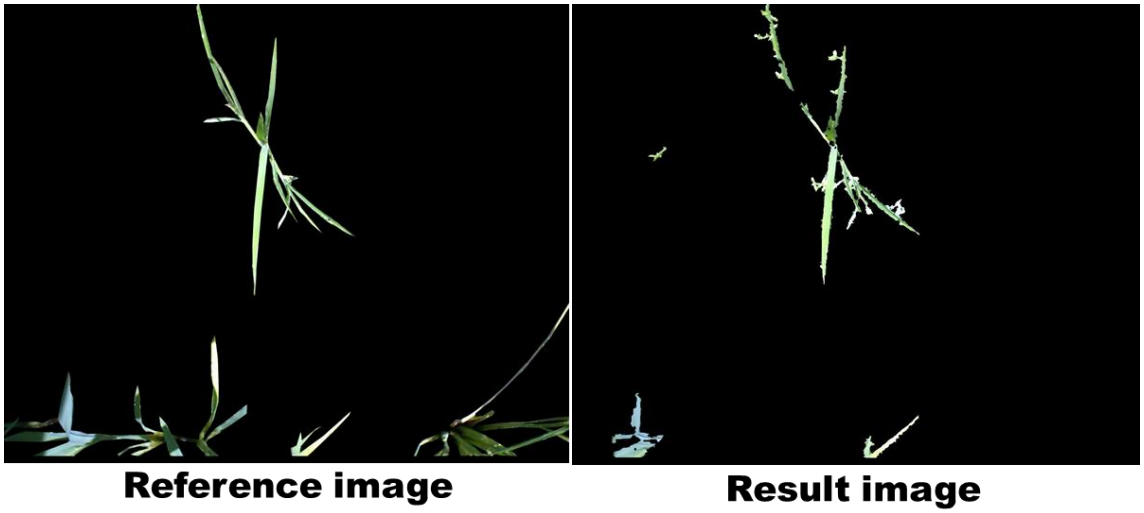


Figure 4-8 The test image that caused low accuracy with the proposed method.
The leaves of other rice plants were left in the reference image (left), which caused a low accuracy of segmentation.

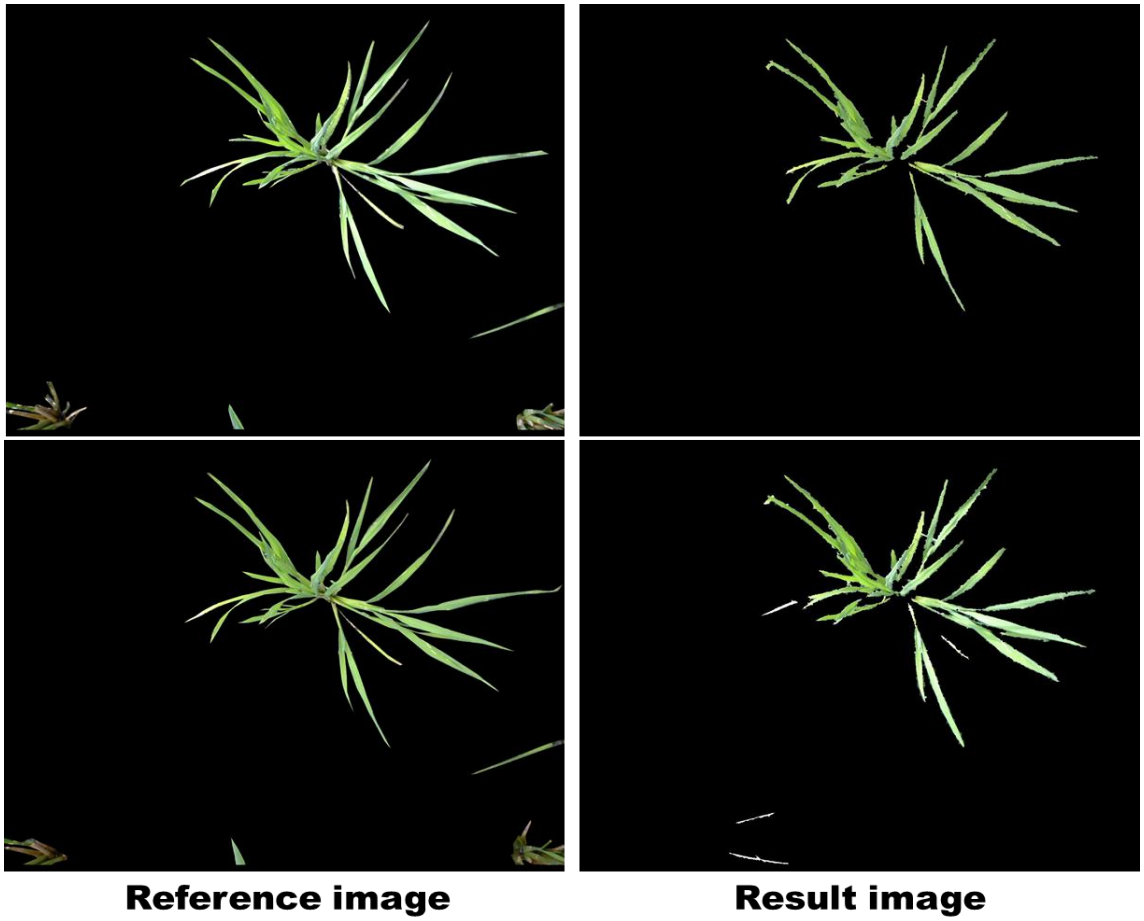


Figure 4-9 Test images that lead to higher segmentation accuracy.
There are no other rice plants left in the reference (left).

Chapter 5 Automatic flowering detection in paddy rice

5.1 Introduction

Knowing, understanding, and identifying the growth stages of rice are critical for managing the rice crop. Most cultural management practices after planting are based on crop growth stages. The entire rice life-cycle (3–6 months) is normally completed within three distinct growth phases: (Yoshida, 1981)

1. Vegetative phase: During the vegetative phase, the plant goes through germination-emergence, seedling growth, tillering, and internode elongation stages of development.
2. Reproductive phase: Includes panicle initiation, culm elongation, heading, and flowering stages.
3. Ripening phase: Includes the milky, dough, and final maturity stages

Machine learning and computer vision technology have become easily applicable to different research fields with the exponential growth of PC performance according to Moore's Law. These technologies have often been used in agronomic applications to analyze canopy/leaf construction, color, shape, etc., during the vegetative phase (Evers et al., 2009; Guo et al., 2013; Liu et al., 2013; Mielewczik et al., 2013; Panneton and Brouillard, 2009; Royo and Villegas, 2011; Sakamoto et al., 2011; Welles and Cohen, 1996; Yu et al., 2013). However, almost no image-based studies have surveyed the phenotypic characteristics during the reproductive phase. For example, the flowering stage, which is one of the most important phenotypic characteristics of the rice plant (Kobayasi, 2012; Matsui et al., 1997; Yoshida and Nagato, 2011), is still surveyed by humans, which is very labor and time intensive.

In this chapter, we seek a solution to automatically detect paddy rice flowering using RGB images taken under natural field conditions. We surveyed the biological and physical properties of paddy rice flowering, established a strategy to detect the flowering spikelet, and discuss the accuracy and applications of the proposed method.

5.1.1 Basic paddy rice flowering knowledge

Flowering (also called "Spikelet anthesis") begins with panicle exsertion (heading) or the day following panicle exsertion. Flowering typically starts at approximately 9:00

AM with the sun rise, regardless of whether the panicles are fully out or not. There are four basic steps to rice flowering (Figure 5-1).

Palea begin to open. Then, about 5 min later it opens to approximately 20–30 ° and the anther filaments begin elongating and exserting. About 30 min later, the anther filaments elongate and then the anthers completely exsert, scattering the pollen. Soon after, the anther filaments wither and droop. At the same time, the spikelets begin to close. Finally, the anthers are left outside to die, and the closed spikelet never opens again.

Heading is a synonym for flowering of paddy rice. The entire course of events occurs in 1–2.5 h of a day during the flowering stage and is very sensitive to the outside environment. Temperature and weather conditions directly affect the timing of flowering. Figure 5-2 shows images of flowering taken from our experimental field.

Besides the basic biological properties of rice flowering, the physical properties also need to be discussed. First, rice plants are rather flexible, as the shape of their organs can be easily deformed by external environmental parameters such as wind, rain, etc. Second, none of the organs maintains the same construction or shape during growth because of the physiological complexity of the plant. Moreover, because of overlapping organs, changing light conditions also cause differences when representing the plant using two-dimensional RGB digital images. Figure 5-3 shows various scenes of rice flowering spikelets, and Figure 5-4 demonstrates how they change according to growth and the external environment. Figure 5-4a shows images of two panicles taken over three days and indicates changes in the physical size and shape due to growth. Figure 5-4b andFigure 5-4c show that images taken 5 min before and after are totally different because of changes in the natural light conditions and the overlap caused by wind. Thus, detecting flowering panicles from a large number of time series images without using color or shape becomes very difficult. The first approach is to identify stable and common features of flowering panicles and use such features as training data to establish a model to recognize the same object from other images. This is similar to the concept of generic object recognition technology, which is still a very challenging task in machine vision.

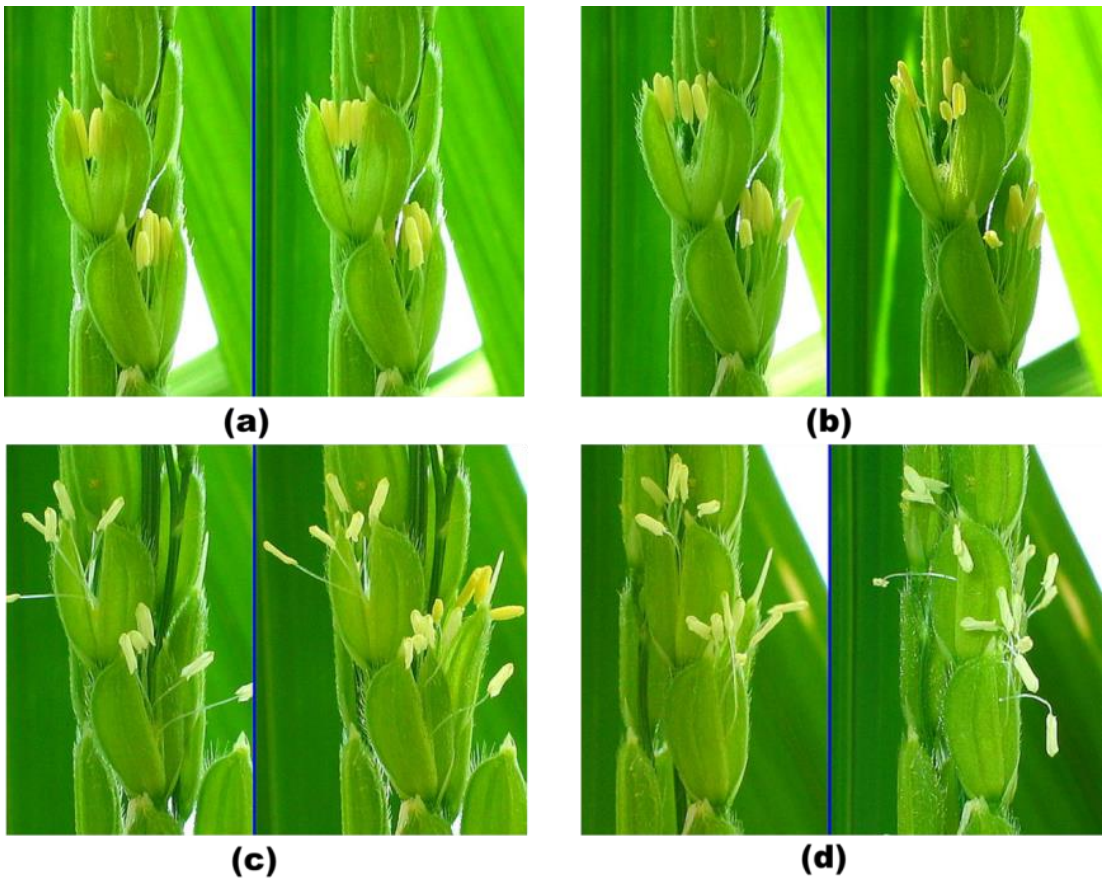


Figure 5-1. The four steps of rice flowering.
(a) Initial step of flowering; the palea begins to open.
(b) Maximum opening. The palea open to about 20–30°, the anther filaments start to elongate, and the anthers exsert.
(c) Closing of the spikelet. After the anther filaments elongate and the anthers have exserted completely, the pollen scatters. Then, the anther filaments wither and droop, and the spikelets start closing simultaneously.
(d) Last stage of flowering. The anthers are left outside to die, and the closed spikelet never opens again.
Images downloaded from:
http://homepage3.nifty.com/knmn/ine/ine109.htm

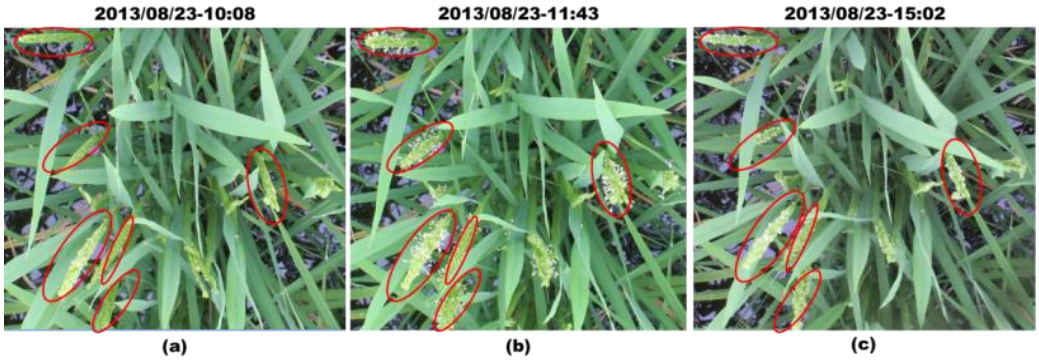


Figure 5-2. Rice flowering in one day.
The flowering period is short, and is only visible from images taken around 12:00 AM in (b), but not in (a) or (c).

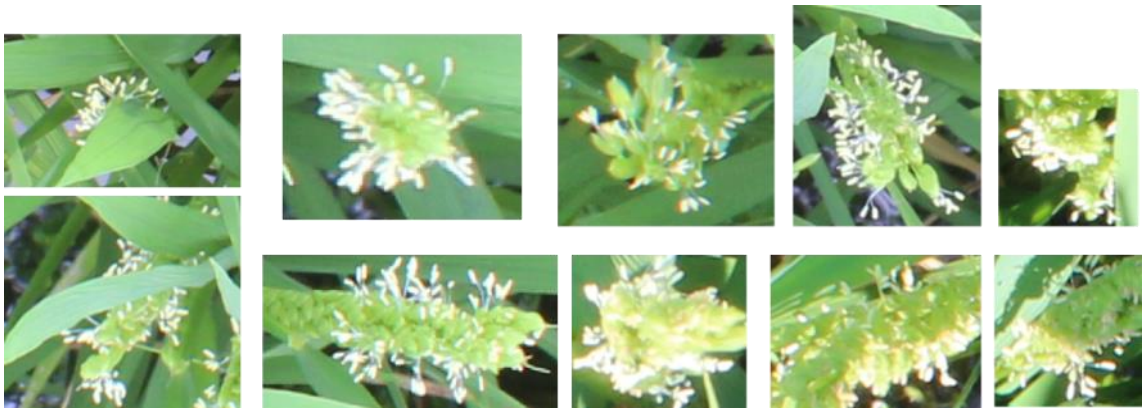


Figure 5-3. Different rice flowering panicle images.

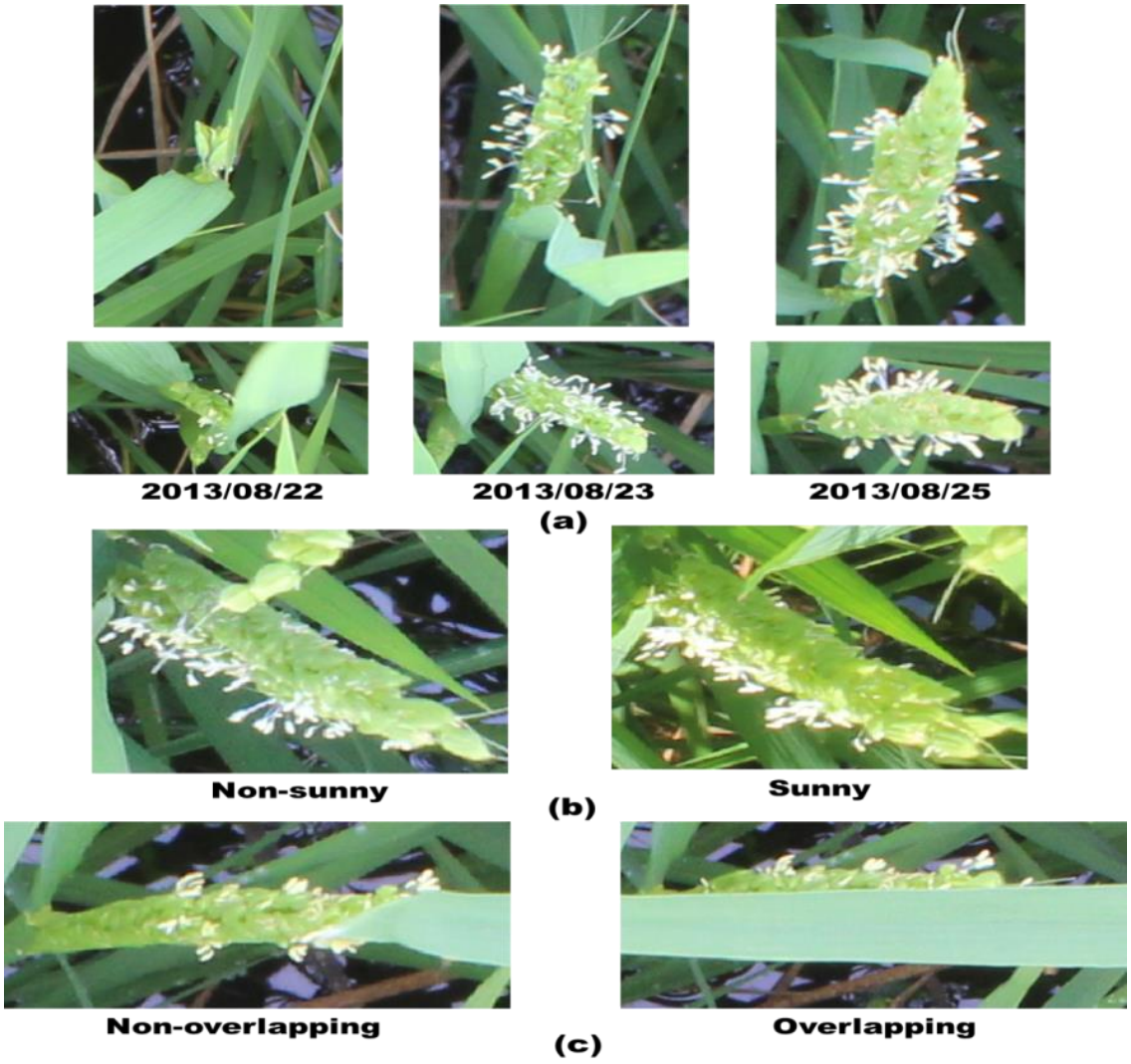


Figure 5-4. Flowering panicles appearance changes depending on the plant growth stage and the external environment.

(a) Images of two flowering panicles taken over three days. Physical size and shape have changed due to growth.

(b) Images of a flowering panicles taken under different light conditions.

(c) Images of a flowering panicles that is overlapped or not near a leaf. Note that the overlap is caused by wind.

5.1.2 Generic object recognition technology

Recognizing a generic object from natural scenes has been a long standing goal of computer vision studies. Due to the constant progression of computer and network technology, generic object recognition has rapidly progressed in the past five years. The

main purpose of this technology is to recognize an object from real-world scenes and classify it into a generic category (Li et al., 2007). For example, Figure 5-5 shows images taken under a natural environment. Although none of the images are the same in background, color, or shape, they are still classified into generic categories of "butterflies" and "cougars."

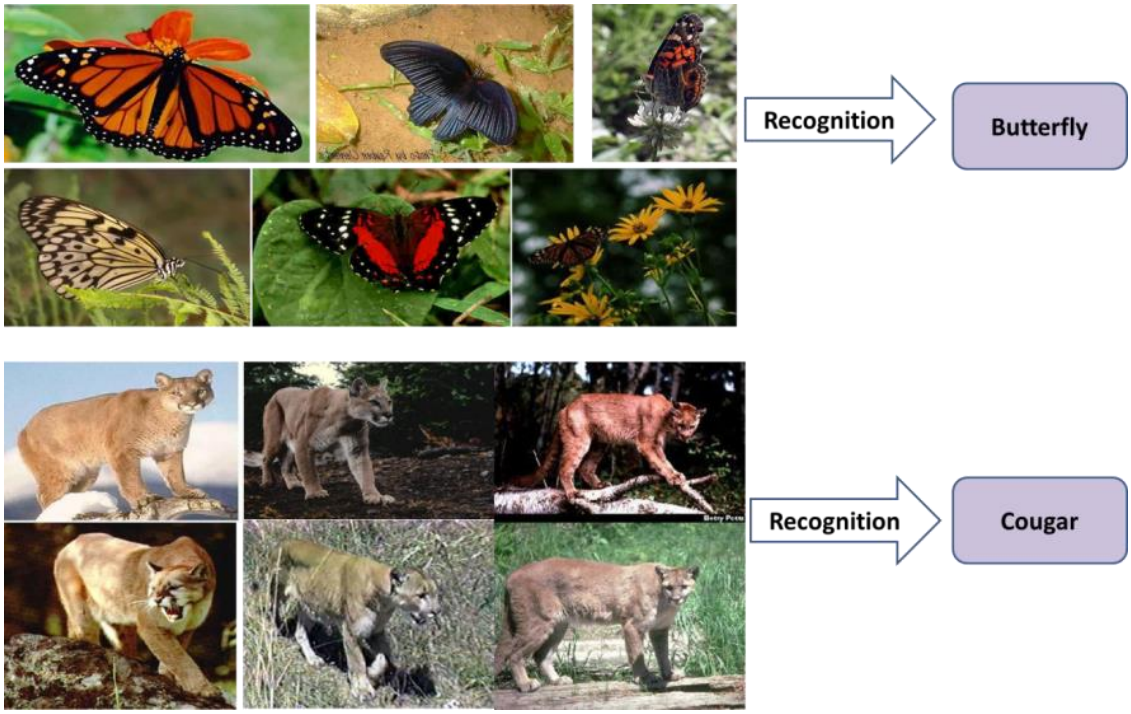


Figure 5-5. Examples of generic object recognition.
Images were downloaded from the Caltech-256 object reorganization database:
http://www.vision.caltech.edu/Image_Datasets/Caltech256/256_ObjectCategories.tar

The most commonly used method to build a generic object recognition model is the local-feature based method, which can be divided into four steps:

1. Extract the local features from a training image data set. The features should have sufficient stability to describe the target object, which is unaffected by nearby clutter or partial occlusion, illumination, three-dimensional projective transforms, and common object variations. Two problems are solved in this step: Where to extract the feature? and what kind of feature should be extracted? For the first problem, some studies have used a "feature point detector" to detect the location of the interest point of an image such as "the corner." Schmid et al., (2000) described several famous feature detectors, whereas others are used for "dense sampling" to extract features from a regular grid or from randomly selected locations. Nowak et

al., (2006) showed that the density of extracted features is the dominant factor in determining recognition performance. For the second problem, many local features have been proposed, such as the Gabor wavelet (Mutch and Lowe, 2006a), maximum entropy (Mutch and Lowe, 2006b), scale-invariant feature transform (SIFT) (Lowe, 2004), image patch (Agarwal et al., 2004), speeded-up robust features (Bay et al., 2008), and Texton (Malik et al., 2001, 1999; Shotton et al., 2006). The invariance of illumination, deformation, rotation, and scale are considered the most important points to choose which features to use for a model (Vedaldi et al., 2010).

2. Describe the object by the many features extracted in step 1. A new method called "bag-of-visual-words" (BoVW) (also called "bag-of-feature-points" or "bag-of-features") is used to represent the distribution of the features. The BoVW is inspired by models used for natural language processing. It aims to translate an image into a set of visual words, and a the histogram of visual word frequency is calculated to determine the contribution of visual words as well as to estimate relevance (Gabriella et al., 2004; Perronnin et al., 2006; Sivic and Zisserman, 2003a). A bag-of-words corresponds to a histogram of the number of occurrences of a particular image pattern in a given image. The main advantages of this method are its simplicity, its computational efficiency, and its invariance to affine transformations, as well as occlusion, lighting, and intra-class variations.
3. Select the most appropriate classifier to establish the object recognition model (modeling). Several machine learning-based approaches have been applied such as support vector machines (SVMs) (Ahmed et al., 2012; Joze and Drew, 2010; Joly et al., 2013a; Pereira et al., 2012; Perronnin et al., 2006; Rumpf et al., 2012; Vedaldi and Fulkerson, 2010a; Yusof et al., 2013; Zhang et al., 2010), random forests (Bosch et al., 2007), artificial neural networks (Arribas et al., 2011; Jafari et al., 2014; Joly et al., 2013b; Pereira et al., 2012), and boosting (Liu et al., 2007; Mikolajczyk and Schmid, 2005; Perronnin, 2008). Generally, there is a trade-off between the speed of object detection and accuracy. Many studies have discussed the difference between those classifiers (Buddhiraju and Rizvi, 2010; Mollazade et al., 2012; Morra et al., 2010; Pinto et al., 2011).
4. Use the model to recognize the object from unknown test images. Normally there are four different tasks of generic object recognition, as shown in Figure 5-6: a. Classify the entire image into one category; b. Classify the image into a combination of categories; c. Segment the image into several regions, and then classify those regions into different categories; d. Detect the window that includes

the entire object (Yanai, 2010).

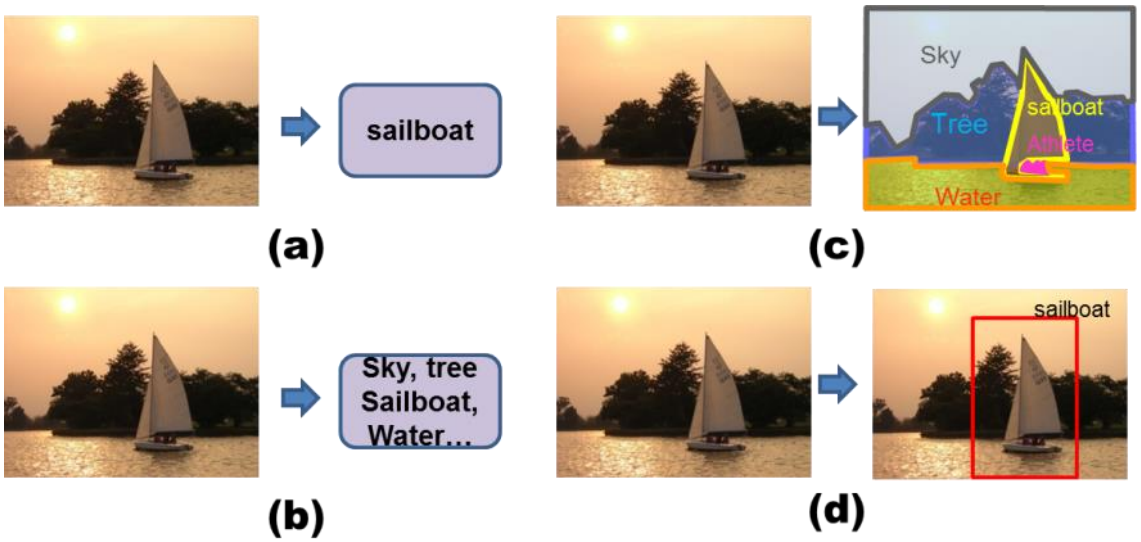


Figure 5-6. Tasks of generic object recognition.
(a) Classify the image into one category: "a sailboat."
(b) Classify the image into a combination of categories: "sky, tree, sailboat, and water."
(c) Segment the image into specified regions and classify each region into different categories: "sky, tree, sailboat, and water."
(d) Detect the window that includes the interested object. In this image, the object "sailboat" was detected in the red rectangle.
The image "sailboat" was downloaded from:
http://people.csail.mit.edu/torralba/shortCourseRLOC

We would like to introduce more details of the key technologies of generic object recognition that will be used in the experiment: the BoWV, which is used to describe the images.

5.1.3 Bag-of-visual-words (BoVW).

The BoVW is a method that represents images as orderless collections of local features (SIFT descriptor in this study). The name is derived from the Bag-of-Words representation used in textual information retrieval. A visual word is constructed to represent the dictionary using clustering features (SIFT in this study) extracted from a set of training images. The image features represent local areas of the image, just as words are local features of a document. Clustering is required, so that a discrete word can be generated from millions of local features sampled from the training images. Each

feature cluster is a visual word. Given a novel image, features are detected and assigned to their nearest matching terms (cluster centers) from the visual words. The term vector is the normalized histogram of the quantized features detected in the image. The BoVW term vector is a compact representation of an image that discards large-scale spatial information and the relative locations, scales, and orientations of the features. Figure 5-7 is an example of how visual words are generated. The SIFT features are extracted from each image and clustered into eight visual words. Figure 5-8 shows an example of the histogram, and Figure 5-9 shows how the features are assigned to the generated visual words.

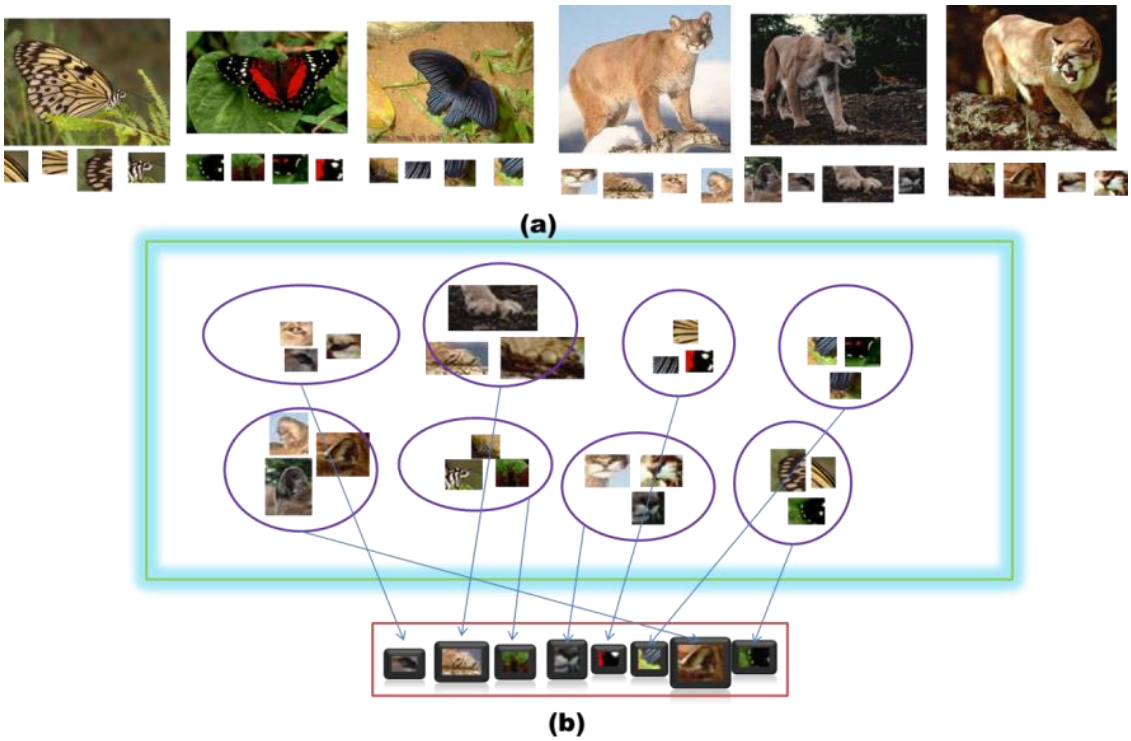


Figure 5-7. Generation of the visual words.
(a) Local features are extracted from each training image.
(b) Visual words are created by clustering the local features extracted from (a). Each visual word indicates the cluster center.

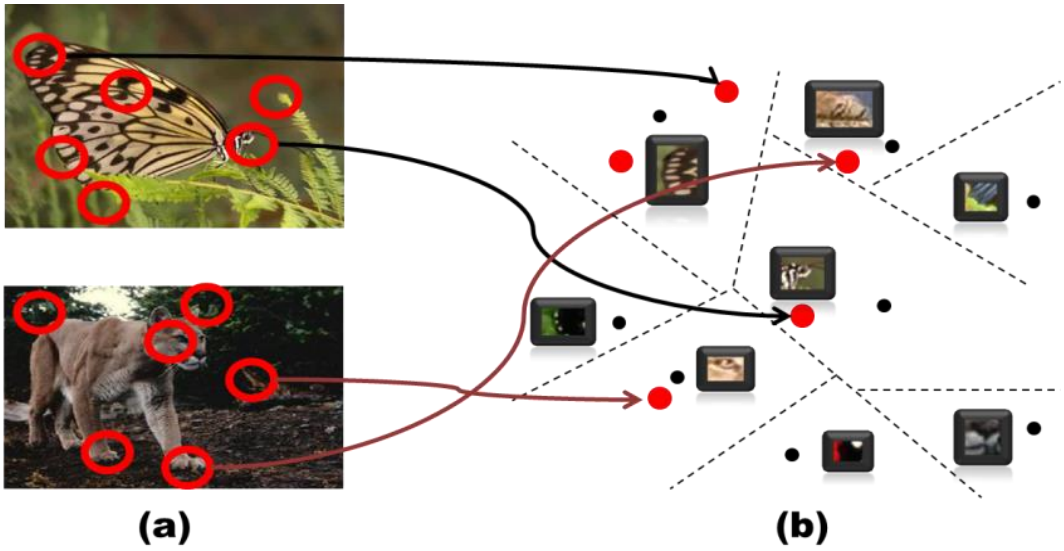


Figure 5-8. Assign scale-invariant feature transform (SIFT) feature descriptors.
(a) The SIFT features are extracted and described for each image.
(b) The extracted SIFT descriptors are assigned to each visual word using the nearest neighbor search.

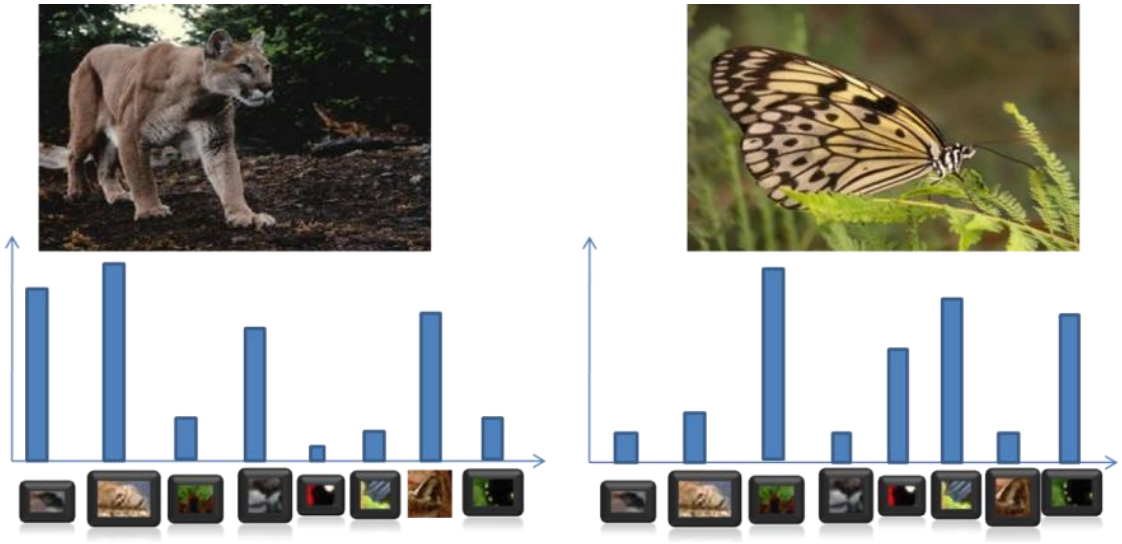


Figure 5-9. A histogram of visual words is used to describe the image.

5.2 Detecting rice flowering from field images

5.2.1 Experimental materials and image acquisition

Seed of the japonica variety *Kinmaze*, which were sown on April 26 and transplanted on May 31, 2013 to an ordinal external field at the Institute for Sustainable

Agro-ecosystem Services, the University of Tokyo, were used as the target crop in this experiment. A DSLR Canon EOS Kiss X5 camera with an EF-S18-55 mm lens was set on the vertical top of the rice crop at a 2 m height, and time-lapse images were taken and transmitted to a free webserver (www.flickr.com) via a 3G network from a field server (Fukatsu et al., 2012, 2011). The time interval was 1 h during days 81–83 and 5 min during days 84–97 after transplanting, resulting in 857 images to construct the experimental image data set. We cropped only the middle region of the original image to avoid the distortion caused by camera lens. The actual size was 30 × 45 cm, and each region contained three rice plants. Figure 5-10 shows the generation of an experimental image. The size of the image on the right side is 2001 × 1301 pixels, and the resolution is 43 pixels/cm approximately.

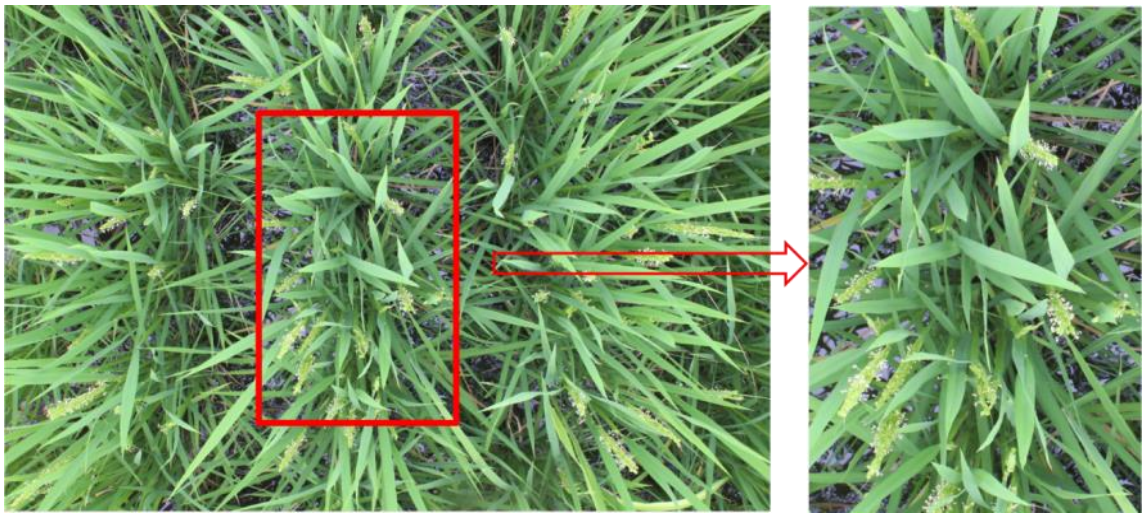


Figure 5-10. Generation of the experimental image (right).
The experimental image was cropped from the middle region of original image. The actual size is 30 × 45 cm, and each region contained three rice crops.

5.2.2 Experiment

Our objective was to detect only flowering panicle parts from the rice images taken under field conditions, which is a simple task when compared to normal generic object recognition from a large-scale image database that requires discriminating hundreds of objects. We simplified our study to discriminate only flowering and non-flowering parts from images. The proposed model was implemented using a GUI framework that we newly developed in the software package Matlab (MathWorks Inc., Natick) with the "VLFeat" open source library (Vedaldi and Fulkerson, 2010b) and that was executed on a Windows 7 PC with 4 core CPU and 16 GB memory. The statistical analysis was

performed using the statistical software package R (R Development Core Team, 2012).

The whole process is demonstrated in Figure 5-11 and can be separated into two parts: the training part (upper line) and the testing part (lower line). The training part is represented by the following steps:

1. Create the training database by manually cropping the rectangular regions from an experimental image data set that include flowering panicles. We created the training database including about 300 images of rice flowering panicle parts as positive data and about 400 images of non-flowering parts as negative data. The positive data are shown partly in Figure 5-3 and Figure 5-4, and they include almost all different situations: weather conditions such as sunny, rainy, and cloudy; growth stage such as initial, middle, and final flowering stages; background such as water, flat leaf, interacted leaves; and location such as with occlusion or not, overlapped or not. The sizes of the images were not uniform. However, it was not appropriate to use all of the training images to generate BoVWs that can be used to build the classifier, because too much training data for a machine-learning approach always leads to a over-fitting problem, which strongly affects the general capability of the training model (Bishop, 2006). Therefore, we used the training images with strong/robust features that describe the kind of image generically such as the flowering panicles of paddy rice. In this study, we discuss how to identify the most appropriate training images, which are the most efficient to recognize flowering panicles from images.
2. Extract the local feature points and the descriptor of those points from training images. We used SIFT in our experiment. However, instead of a SIFT detector to detect the location of the feature points, we used a dense sampling method to extract the points, because it performs better in object generic recognition approaches. Therefore, the SIFT descriptors were computed at points on a regular grid with spacing of M pixels. Here, M = 15, as shown in Figure 5-12a. The SIFT descriptors were computed at each grid point over four circular support patches with radii r = 4, 6, 8, and 10 pixels. Consequently, each point was represented by four SIFT descriptors, and each is a 128 dimensional vector. Figure 5-12b–e demonstrates how to describe a point with a different radius (scale). The point used for demonstration is the one surrounded by the black rectangle shown in Figure 5-12a. The descriptor included 16 patches (red rectangles in Figure 5-12b–e), which are rotated to the dominant orientation of the feature point, and each patch is described in eight different directions and gradient magnitudes (red bins inside the red rectangles). Note that the descriptors are rotation invariant.
3. Generate the visual words and represent the training images using the histograms of

the generated visual words. In this study, we used K-means, because they have good performance in genetic generic object implementations (Lazebnik et al., 2006; Sivic and Zisserman, 2003b). The choice of initial centroid position and the number of clusters (k) affects the resulting vocabulary in the K-means clustering method. In this study, we predefined k = 600 (number of visual words). Then, we ran K-means several times with random initial assignments of points as cluster centers, and the best result was used to select the best performing vocabulary.

4. Use the histogram BoVW as the training data to train the classifier. SVM is one of the most popular classifiers for BoVW. In this study, we used the SVM with a χ^2 kernel, which is particularly suited to data in the histogram format (Jiang et al., 2007; Zhang et al., 2007). Furthermore, a homogeneous kernel map was used to approximate the χ^2 kernel to quicken the process. The map transforms the data into a compact linear representation that reproduces the desired kernel to a very good level of approximation. This representation enables very fast linear SVM solvers (Vedaldi and Zisserman, 2012, 2010). The source code is available from the "VLFeat" open source library (Vedaldi and Fulkerson, 2010b).

We used the sliding window approach in the testing part, as demonstrated in Figure 5-6d. The basic concept of a sliding window is to first scan the whole test image with a predefined size window to segment the test image into several blocks. Then, generate a classifier for each block to judge if it is the target object (flowering panicle parts). Mark all detected blocks on the test image. The marked blocks are considered as the recognized flowering panicle parts of the test image. The size of the block strongly affects the capability of our recognition algorithm. We discuss how to identify the most appropriate block (sliding window) size.

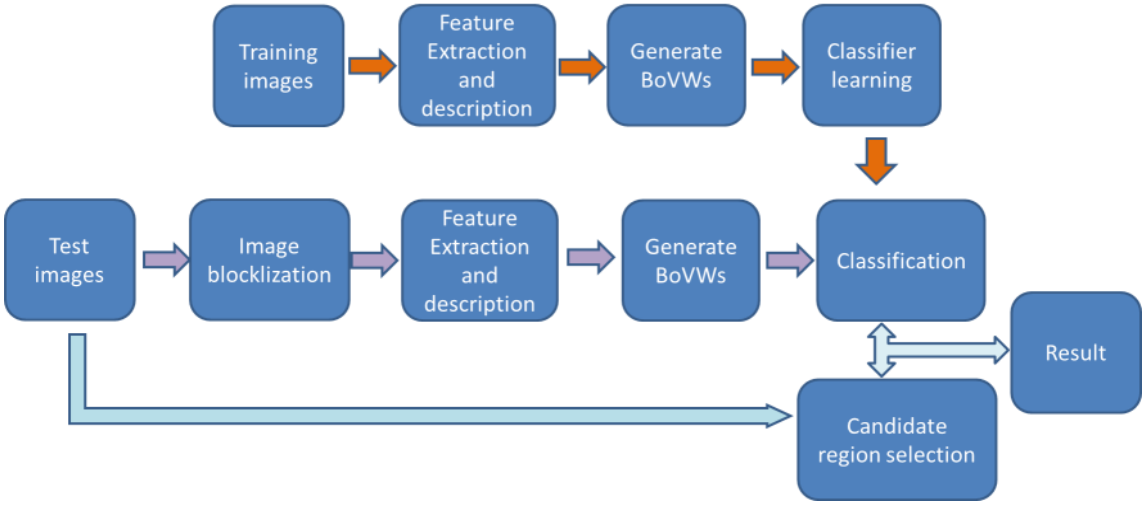


Figure 5-11. A flowchart of the proposed flowering detection model.

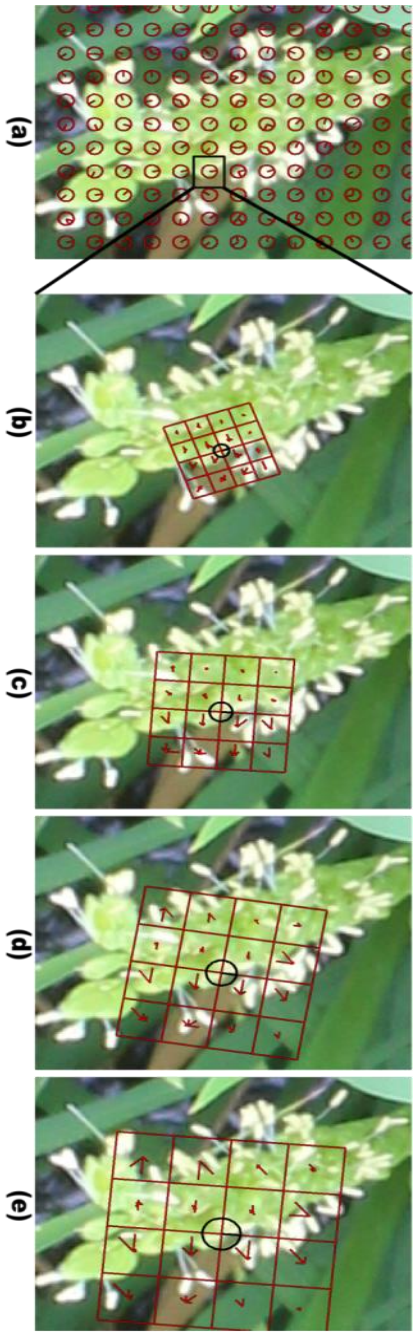


Figure 5-12. An example of dense sampling and SIFT feature point description.
(a) The SIFT descriptors are computed at points on a regular grid with spacing of 15 pixels, as represented by the red circle.
(b–e) Pickup one point as an example, and use SIFT to descript the point at four different scales, as shown in the black circle with radii r = 4, 6, 8, and 10. A descriptor included 16 patches, represented by the red rectangles, which are rotated to the dominant orientation of the feature point. Each patch is descripted in eight different directions and gradient magnitudes (red bins inside the red rectangles).

5.2.2.1 Training image selection

We conducted a preliminary experiment to select the most appropriate training images, as follows: First, we defined 5, 15, 30, 50, and 100 as the number of training images. Then, for each number, we randomly selected the training image from the database with 10 replications to construct the SVM classifier, named as methods M_5_1 to M_5_10, M_15_1 to M_15_10, M_30_1 to M_30_10, M_50_1 to M_50_10, and M_100_1 to M_100_10. Finally, we applied each classifier to all the test images.

In order to evaluate the performance of each method, the flowering panicles were carefully counted visually by human eyes and the number was considered as the true value of the flowering panicles.

In this study, we use a sliding window method which does not detect a whole flowering panicle but detect a region (window/block) of a flowering panicle. We call such detected regions as blocks in this study.

We compared the correlation coefficient between the visually counted flowering panicle number and the number of the blocks detected by each method in order to determine the most appropriate number of training images. The correlation coefficient values are plotted in Figure 5-13. The x axis represents the number of replications, and the y axis represents the R value. The five different lines are related with each method. However, the only observation from this graph is that the training data selection strongly affects the capability of the model.

Due to the R value of each method, we used the methods M_15_6, M_30_6, M_30_8, M_30_9, M_50_3, and M_50_8 whose R values were >0.8. Finally, we checked the mean value and standard deviation (SD) of each method, as shown in Figure 5-14. The R values increased with the number of training images until it reached M_30, and then it declined significantly after M_30. Thus, we chose 30 as the training image number and the images used in M_30_6, which hold the highest correlation coefficient of R = 0.82 in method M_30, as the training images.

5.2.2.2 Determining block size (sliding window size)

The sliding window size for the test images strongly affects the recognition capability of our proposed algorithm. In this section, we sought the most appropriate size for method M_30_6, which was selected in the last section. We tested seven different block sizes of 30, 50, 70, 90, 110, 130, and 150. We also calculated the correlation coefficient of the M_30_6-based method to check the performance of each block size, which is shown in Figure 5-15. Finally, three block sizes whose R values were >0.8 were selected; they are M_30_6_50 (R = 0.83), M_30_6_70 (R = 0.83), and M_30_6_90 (R =

0.82). However, it was difficult to tell the best block size by only considering the R values.

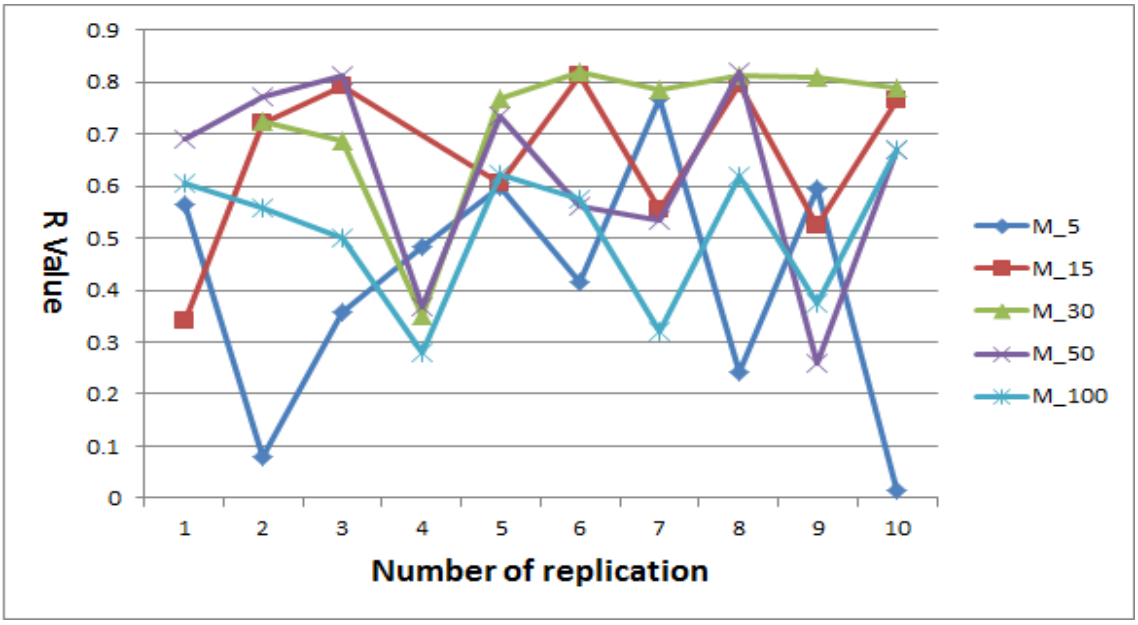


Figure 5-13. Correlation coefficients between the detected number of blocks that contained flowering panicle parts and the visually counted number of flowering panicles.

The number after 'M_' indicates the number of training images for model construction. Ten replications were used for each number in the randomly selected training images.

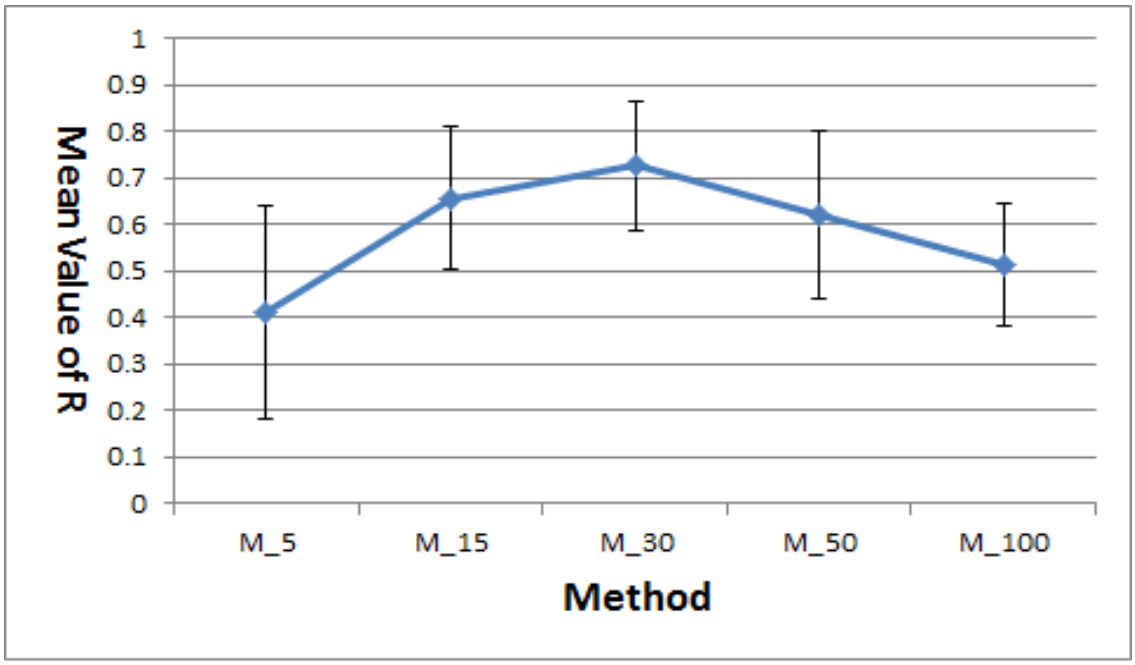


Figure 5-14. Mean values of the correlation coefficients for each proposed method. Method M_30 had the highest correlation coefficient for all 10 replications of each method with an R = 0.82 and lowest SD = 0.13.

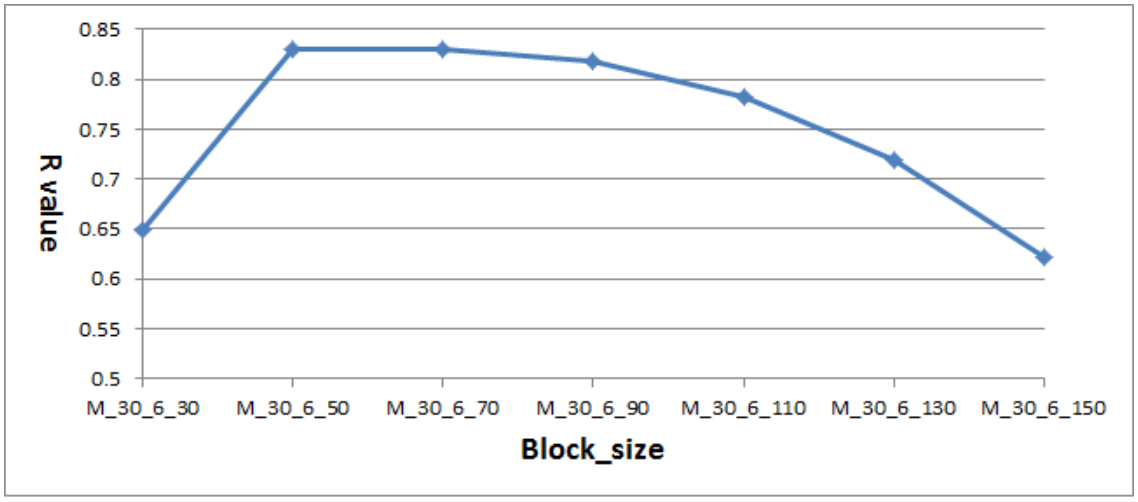


Figure 5-15 Correlation coefficients between the detected number of blocks that contain flowering panicle parts using seven different block sizes and the visually counted number of flowering panicles.

5.3 Experimental results

The flowering parts of each test image were evaluated using the M_30_6 method with three different block sizes of 50, 70, and 90. Figure 5-16 shows examples of the

detected flowering panicles by proposed method.

We detected the flowering parts of each test image during the entire observation period. Besides evaluating the number of blocks detected, we also evaluated the collected regions of the blocks, to determine the relationship between block number, connected region number, and visual judged flowering panicles number. The correlation coefficient values are shown in Table 1.

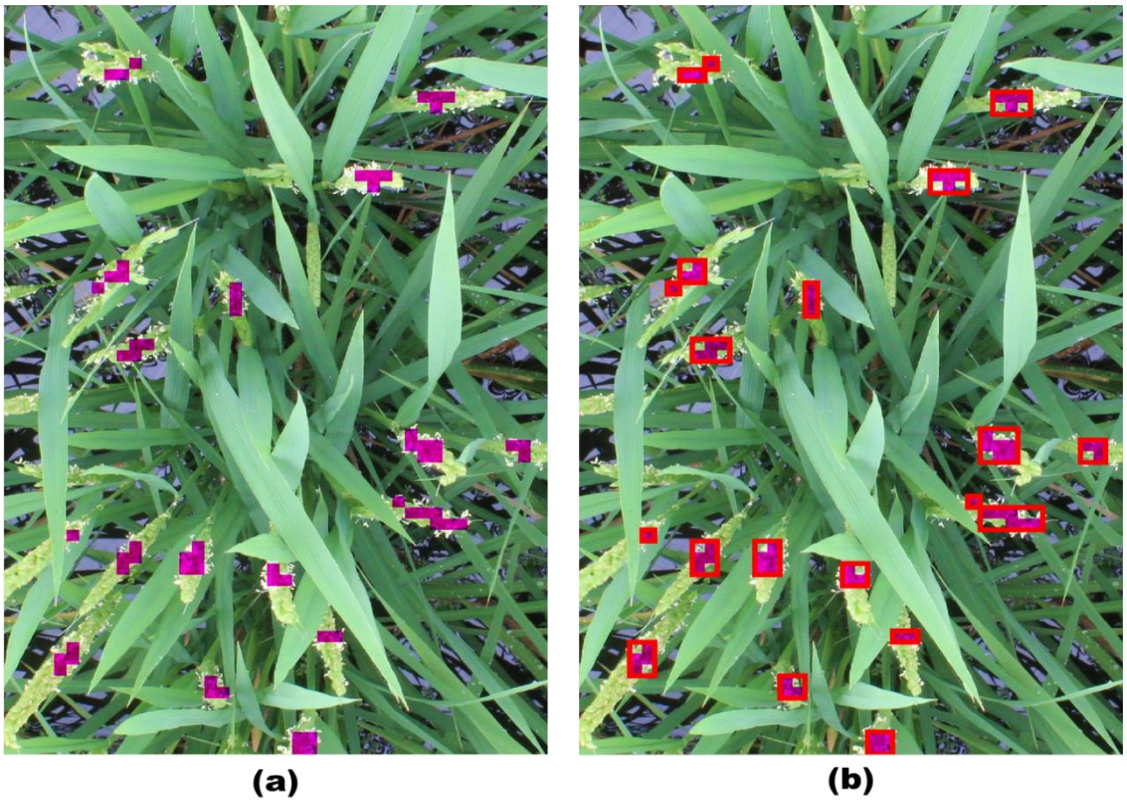


Figure 5-16. An example of flowering panicles detection by proposed method.

Table 1. The correlation coefficients for the block number and the connected region

	M_30_6_50	M_30_6_70	M_30_6_90
R of Block number	0.83	0.83	0.82
R of Connected region	0.83	0.84	0.83

As an example, the result for method “M_30_6_90” is plotted in Figure 5-16 in time series order. The lower x axis represents the image shooting date and time, the upper x axis represents days after transplant, the left y axis represents the visually counted number of flowering panicles, and the right y axis represents the number of blocks that were detected as flowering panicle parts using the proposed method. Note that the scales of the two y axes are the same. The whole observation period was from days 84–97 after transplant. However, because of defects in the image acquisition system, most of

the image data on days 86, 92, and 93 were missing. The green line with the green triangle indicates the number of blocks that were detected as flowering panicle parts named "M_30_6_90." The black line with the black rectangle indicates the number of flowering panicles, which was derived by counting the number of connected regions of the detected blocks named as "M_30_6_90_connected." The blue grid line with the red rhombus indicates the visually counted number of flowering panicles called "Visually counted," which is considered the true value for this experiment. Although the scales of the result values (block number detected as flowering panicle parts) are different from those of the true values (the number of visually counted flowering panicles), the trends are very similar, and the correlation coefficient was 0.82, indicating good detection capability for the proposed method. Additionally, the number of connected regions of the detected blocks (black line) was much closer to the true value with a correlation coefficient of 0.83, which approximated the real number of flowering panicles.

Figure 5-18 and Figure 5-19 show the relationship between weather conditions and flowering time within one day. Figure 5-18 plots the number of flowering panicles counted at every five min interval for seven days of flowering. The different line colors represent days 85–91. The number of flowering panicles peaked from 11:00 AM to 13:00 PM except on day 87 (green line). According to the hourly precipitation data, which are shown in Figure 5-19, rain occurred on the morning of day 87. Because rice does not start flowering under a rainy weather condition, the number of evaluated flowering panicles reached a peak at approximately 14:00–15:00 PM after the rain stopped.

In addition, as an application of proposed method, we tried to automatically predict the flowering/heading date of rice by using only the evaluated number of flowering panicles. The rice flowering/heading date is defined as the day that 40–50% of all panicles observed flowered in a unit area of the rice field. Thus, we predicted the flowering date by counting the detected flowering panicle number, as shown in Figure 5-20. Figure 5-20 shows the cumulative distribution of the overall number of flowering panicles for every 5 minutes. The detected flowering number grew quickly in the beginning and became stable at the end of flowering stage. The red line represents the 40% of the total cumulative amount. Both the manually counted number and the machine detected number reached the 40% level at the 88-th day. In older to evaluate the predicted result, we also asked an expert to observe the test image data set and gave his judgment about the flowering date (normally it is judged qualitatively by experts). The flowering data given by the experts was a day between the 85-th to the 88-th day which was coincide with the date predicted by proposed method.

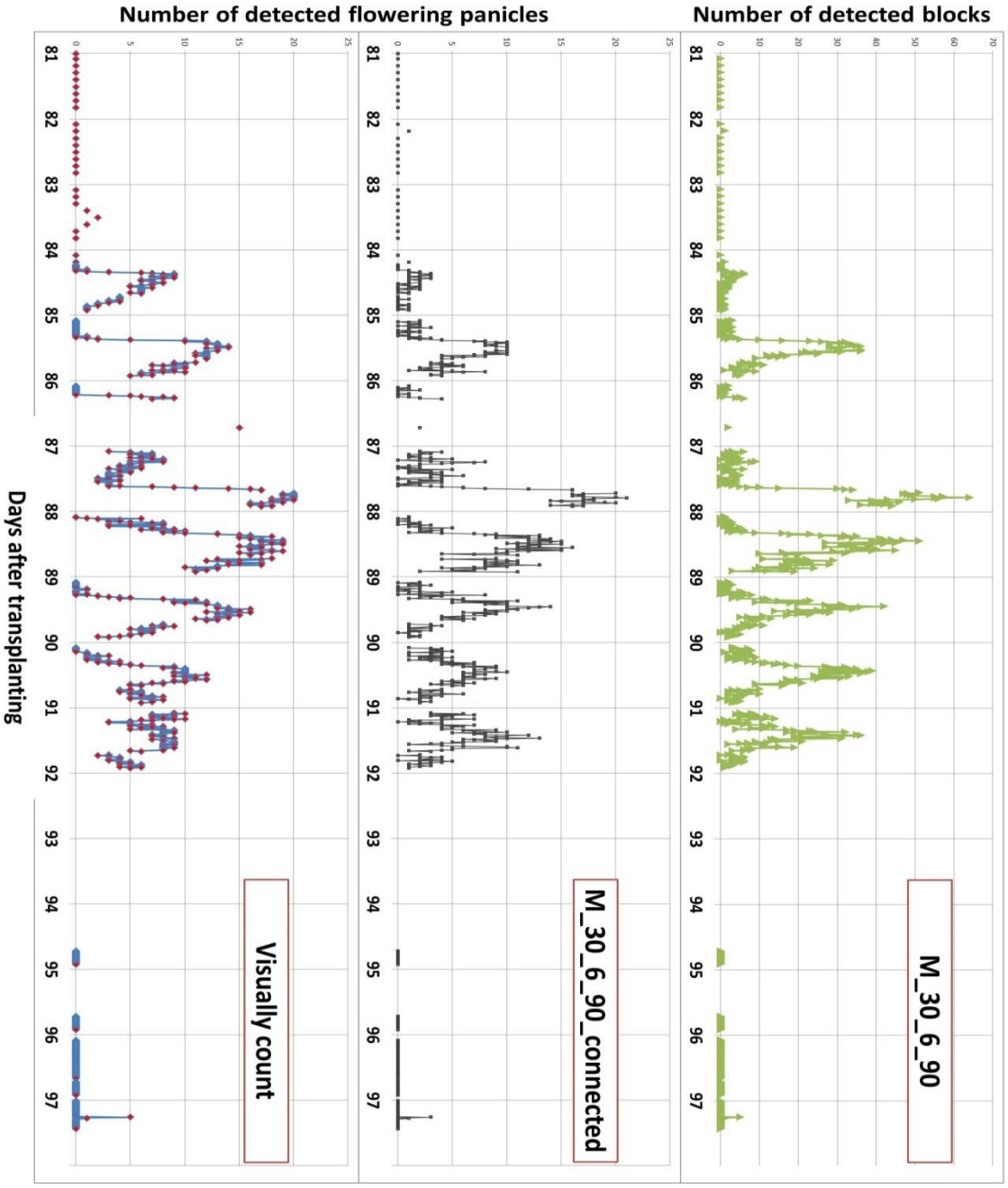


Figure 5-17. Result of M_30_6_90 method during days 81–97 after transplant.
The green line with the green triangle named “M_30_6_90” indicates the number of blocks detected as flowering panicle parts. The black line with the black rectangle named “M_30_6_90_connected” indicates the number of flowering panicles, derived by counting the number of connected regions of the detected blocks. The blue grid line with the red rhombus named “Visually count” indicates the visually counted number of flowering panicles which is considered as the true value.

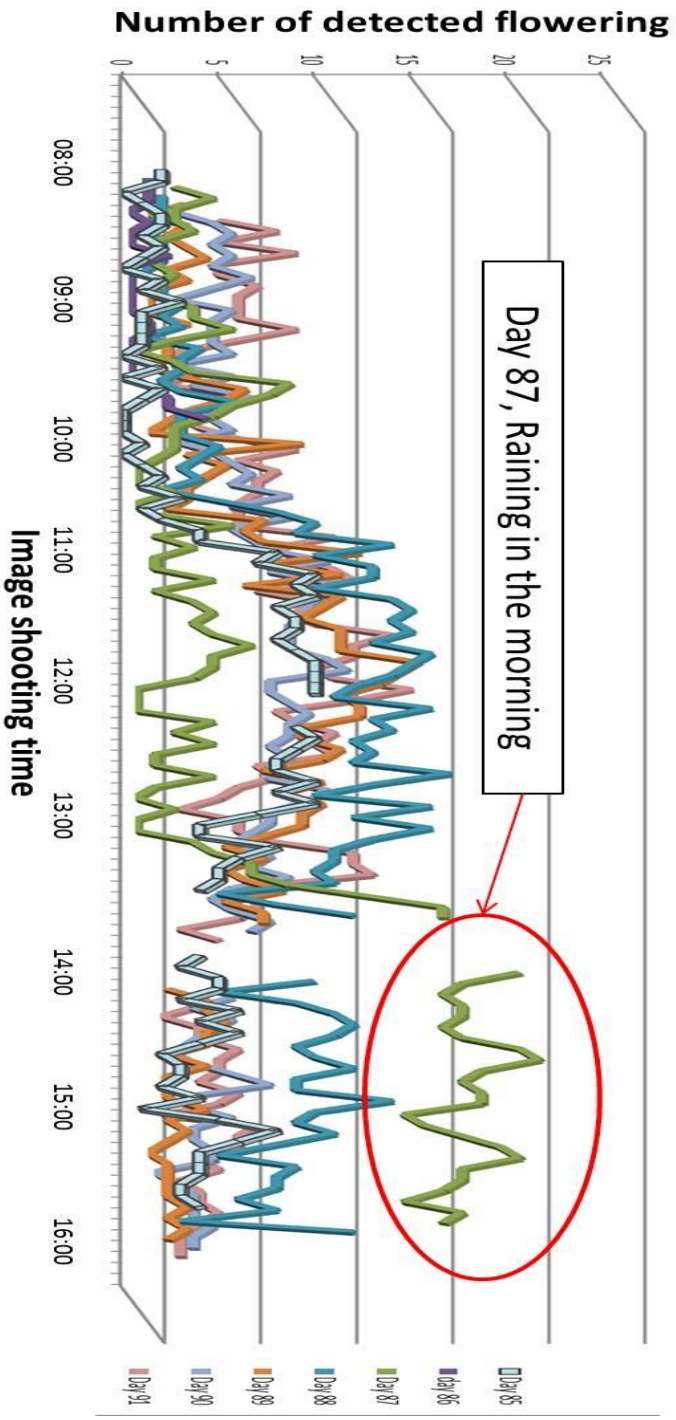


Figure 5-18 Differences in flowering timing among seven different days.

Each line indicates the number of detected flowering panicles number (connected block number) from 8:00AM to16:00PM of different days in 5 minute time interval. Seven different color of lines indicate different days of the flowering period. The number normally reached to peak during 12 AM. But on the 87-th day (the green line) after transplanting reached the peak around 15PM.

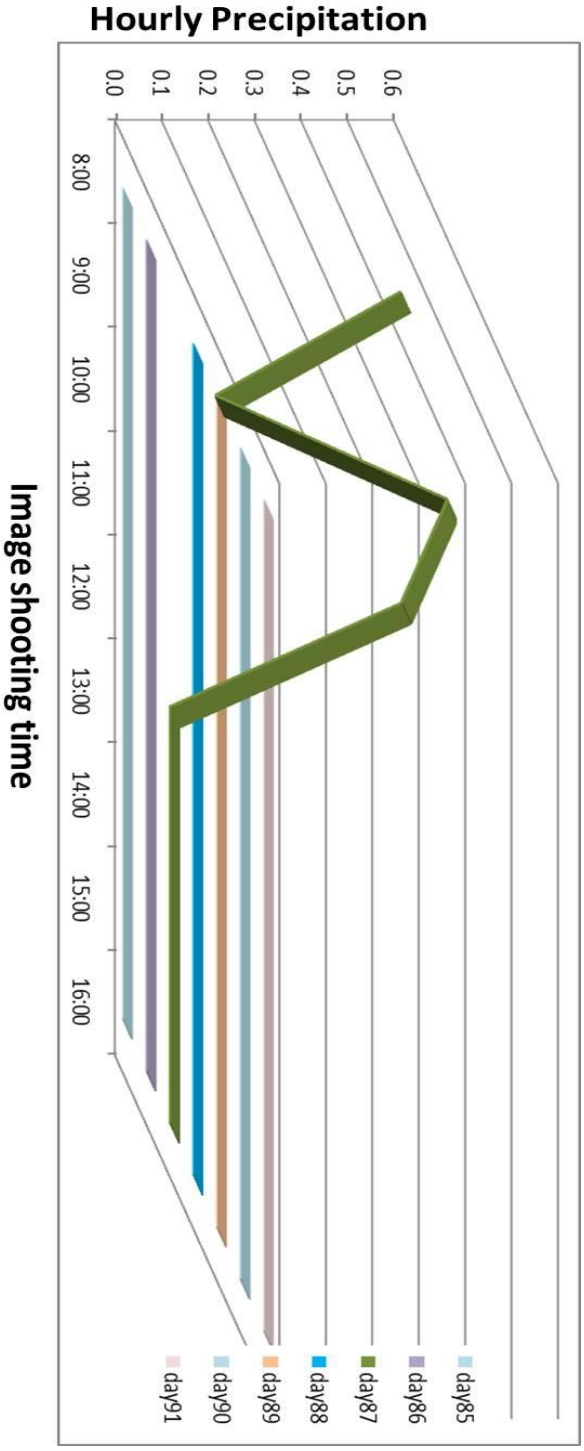


Figure 5-19. Hourly precipitation for seven different days.

Seven different color of lines indicate the hourly precipitations from 8:00AM to16:00PM of different days of the flowering period. Only the 87-th day (the green line) after transplanting was rainy in the morning.

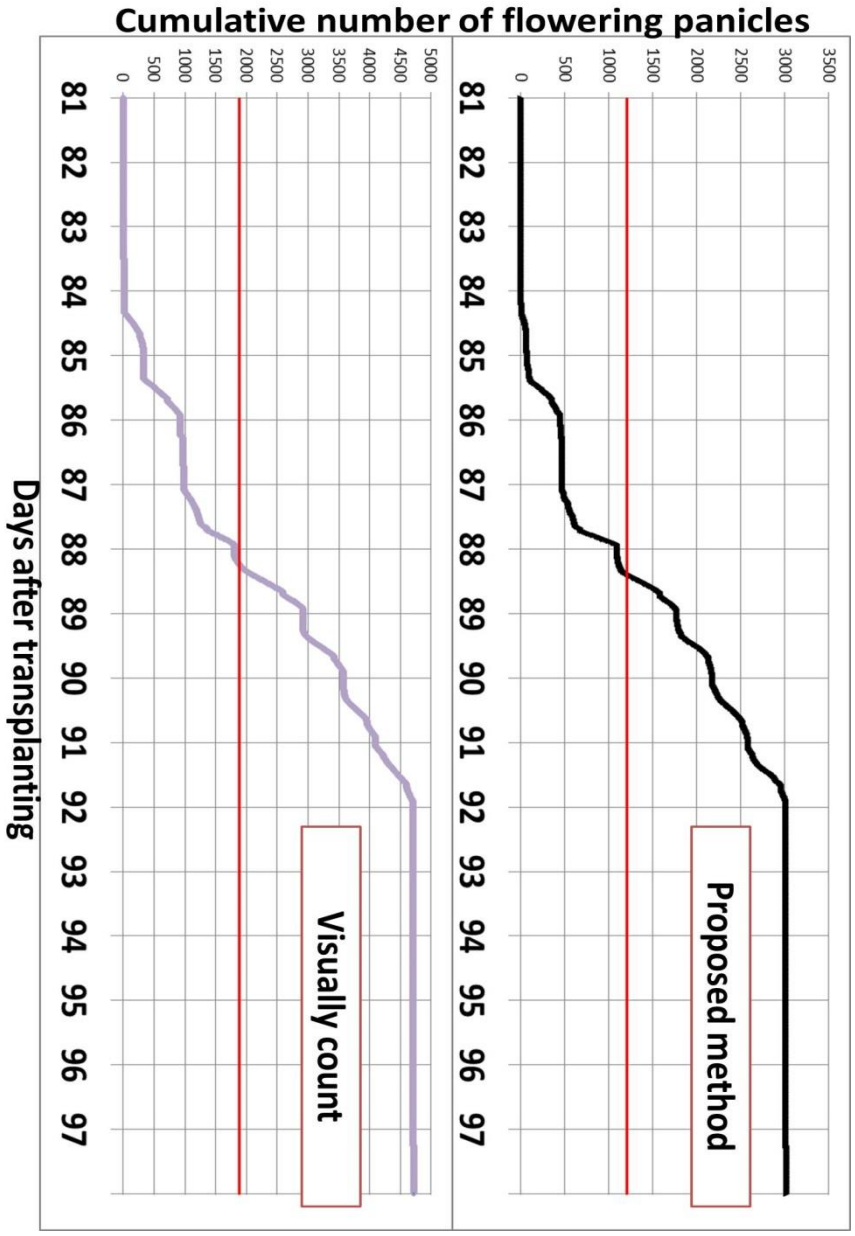


Figure 5-20. Prediction of flower/heading date.

The x axis represents days after transplanting, and the y axis represents the cumulative number of flowering panicles visually counted and the number of the connected blocks counted by the proposed method. The purple line represents the number of visually counted flowering panicles, and the black line represents the number of the connected blocks counted by the proposed method. The red lines represent the 40% value of the total cumulative amounts. Both reached the 40% level at the 88-th day.

5.4 Discussion

We proposed a method to detect rice flowering panicles from a time series of RGB images taken under natural conditions in a real field. The idea for our method comes from generic object recognition technique. Several experiments were conducted to determine the most appropriate training images from a training image database to enhance the SVM classifier during the training stage. Then, we identified three block sizes to detect flowering panicles during the testing stage. The results showed a reasonably good ability to automatically detect the flowering stage during the entire growth stage and to estimate flowering time during the flowering stage. The proposed method also showed stable detecting ability even under various light conditions. We also conducted three preliminary experiments to confirm generalization ability of the method:

1. We applied the flowering detection model to two different set of images that were resized to 75% and 50% of the resolution of original test images.
2. We applied the flowering detection model to a larger scale (image that occupies a larger area of the field than the test image).
3. We applied the same model to images taken from a different field with a different rice variety, *Kinmaze*, in addition to the three different image sizes tested.

The results of experiment 1 are shown in Figure 5-21. The size change did not affect the detection capability of our proposed method. The results of experiment 2 are shown in Figure 5-22. The picture was taken from a higher position on top of the field, but almost all flowering panicles were detected. Thus, the proposed method shows the possibility to be applied to images taken by UAV, which can handle a huge field area in a very short time. The results of experiment 3 are shown in Figure 5-23. Most of the flowering panicles were also well detected, even though the images of this variety were not used for the training data in the model. Although small misdetections were not fully avoided, we believe the results of our proposed method are acceptable for automatic rice flowering detection approach.

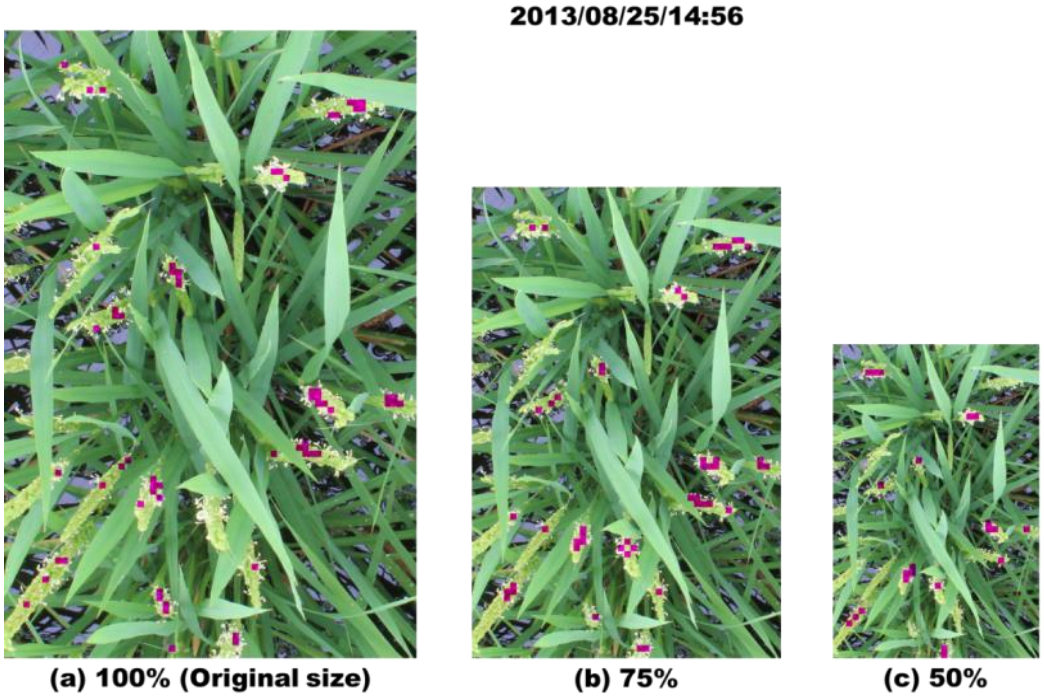


Figure 5-21. Flowering detection results with three different image sizes.
The variety is *Kinmaze*, which is the same one used for collecting the training image data.

Figure 5-22. Detecting flowering panicles on a large scale image.

Figure 5-23. Flowering detection results for "*kamenoo*."
Images of the variety "*kamenoo.*" No training images were used with this variety.

However, several disadvantages exist for our proposed method. First, we only used the gray level information of the images for the feature description step during the training stage, which may have affected accuracy. Second, although we tested several different block sizes (sliding window size) for each experiment, size was fixed; therefore, if the objectives scale of the object change, or the resolution of images change, our method may not detect the object properly.

The generic object recognition technique is still continuously developing. For example, the BoVW only counts the number of occurrences of visual words conducted by local features, and loses the location information of each features or other important information, which might help train the classifier. Therefore, studies are focusing on increasing the dimensions of BoVW by adding more statistical information such as the vector of locally aggregated descriptors (Jegou et al., 2010), the super vector coding (Zhou et al., 2010), the fisher vector (Perronnin et al., 2010), and the vector of locally aggregated tensors (Picard and Gosselin, 2011). These new concepts have been proposed to accurately recognize/classify large scale images in the real world. This is different from our objective in agriculture, but the ideas may help us develop the same kind of feature information to detect different parts of crops for extracting phenotypic characteristics of agricultural products.

We would like to improve the accuracy and general versatility of the flowering detection model described in this study. This goal will require adding more efficient training images and testing more machine learning-based classifiers at the training stage. We will improve the test strategy to be more flexible such as with the dynamic sliding window approach.

Chapter 6 General discussion

6.1 Overall conclusions

The extraction/segmentation of target crop vegetation from the background of digital photographs is a critical step in crop phenotyping based on the processing of images. A machine learning-based image segmentation method, “DTSM”, was proposed. Based on the use of only the color features of pixels, the DTSM proved capable of overcoming diverse natural light conditions that can strongly affect the profile of crop images taken outdoors. The single model can therefore be easily used to segment crop vegetation regions from vast amounts of time-series images.

Based on the DTSM, some crop canopy phenotypic characteristics can be evaluated from time-series images with high accuracy, such as the dynamic change in crop canopy coverage ratio throughout the growth cycle. Moreover, local weather data collected by a field server was combined with the whole evaluation process, and proved to be a very important impact factor for evaluating the crop canopy coverage ratio from huge amounts of time-series images taken in the field during all growth stages.

A preliminary study was conducted to address problems with image interpretation posed by weeds and algae, which complicate the background of crop images and thus make it very difficult to isolate crop vegetation, particularly under organic and natural paddy farming conditions. The proposed method proved able to extract individual crop vegetation from such complicated backgrounds, but it was not useful for images taken of the whole plant community. More experimentation will be necessary to devise methods to solve that problem.

Finally, a new method was developed to detect the flowering panicles in time-series images of rice paddies. It was then used to evaluate two of the most important phenotypic characteristics of rice crops, flowering time and the relative amount of flowers. Currently, such detection relies fully on qualitative human judgments. The “bag of visual words” technology was used to describe the flowering panicle images, which are difficult to describe with shape or color features alone. Then, a support vector machine approach was used to generate the flower detection model. A new training image selection method was also discussed, whose detection accuracy was high.

Overall, this series of experiments shows that the machine learning technologies are a very powerful tool for extracting information useful for evaluating the visible

phenotypic characteristics of crops. Owing to the advantage of general applicability, the machine learning based approaches were also considered very suitable for field-based phenotyping systems. However, the accuracy of such approaches depends strongly on the selection of training data. In this dissertation, the training data were selected manually or semi-manually, so the use of "good" training data still depends on human experience.

6.2 Future research prospects

In this dissertation, several new methods were proposed to evaluate visible phenotypic characteristics of plants. In the future, we would like to turn each method into publishable tools, which will be freely available for the use of agronomics researchers. We also plan to share all the source codes with any computer science researchers who wish to improve our proposed algorithms. Figure 6-1 shows an example of a tool with GUI interface that could be used easily for vegetation segmentation and evaluation of plant canopy coverage ratios.

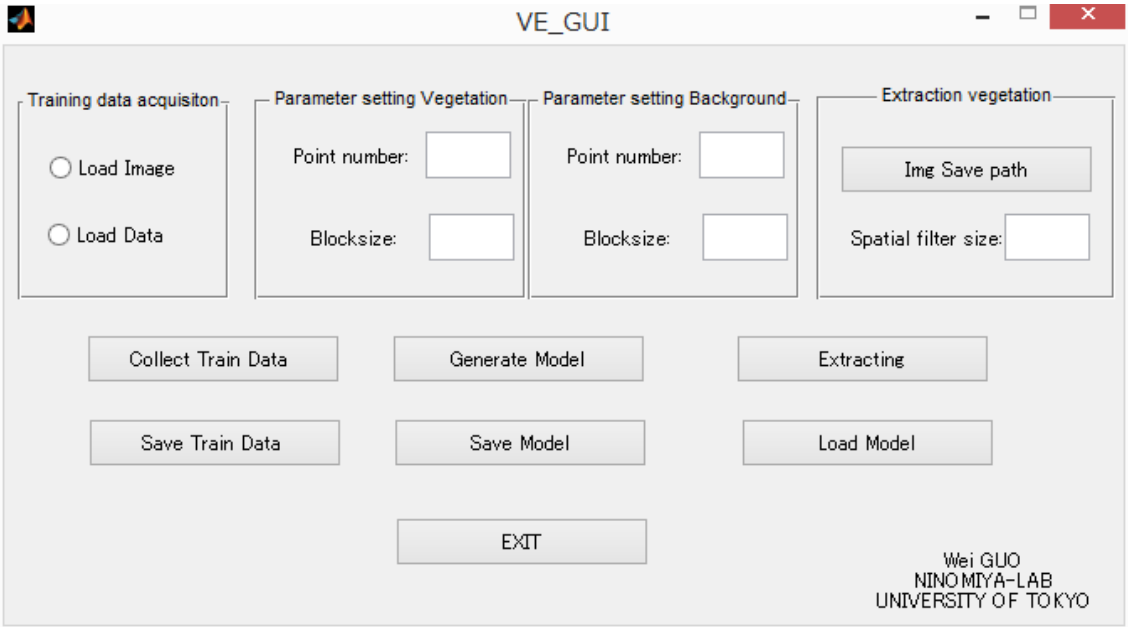


Figure 6-1 An example of a tool for vegetation segmentation and evaluation of plant canopy coverage ratios.

As mentioned before, the ability and effectiveness of machine learning approaches depend strongly on the appropriate selection of training data. The training data acquired from this experiment cannot cover all situations found in different experiments, so researchers still need to add or renew the training data to satisfy their own particular

needs. Development of an easy-to-use training data selection tool will be considered in the future.

In this study we used only images taken by cameras set above crops. The use of images taken from only one direction has its limitations. As the next step, we will add at least one more set of images taken from a different direction, to try to evaluate more phenotypic information such as leaf area index.

Also, the cameras were all set at a fixed observation position to acquire the images, so only very small area of the crop field was included in the image. This does not provides enough information for a high-throughput phenotyping system. We are planning to design an UAV (unmanned aerial vehicle)-based platform to acquire crop field images with a wide area, and to then evaluate crop phenotypic information from those images.

Our proposed method to segment crop vegetation from very complex backgrounds needs to be discussed further. More images should be collected for further algorithm development. How to best design the field experiment and the system to acquire useful images are challenging tasks, so in the future we will pay more attention to experimental design.

The application of image analysis technologies to field-based plant phenotyping is still an emerging research topic. We anticipate that more and more new algorithms will be developed in this field in the near future. However, evaluating the image segmentation accuracy of an algorithm is not easy. For example, true values are needed. This means that for a given test image, a correct, completely segmented image should be provided. This is normally done by hand labelling or drawing, processes that are labour intensive and costly. Moreover, because that process is manual, different observers would give different true values that add error to the evaluation of segmentation algorithms. To avoid these issues, many databases already exist that provide true values of images, such as “The Berkeley Segmentation Dataset and Benchmark”,
(http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/),
and the “segmentation evaluation database”
(http://www.wisdom.weizmann.ac.il/~vision/Seg_Evaluation_DB/).

Researchers can easily download the test images and check them with their own algorithm. As we know, until now there was no such image databases created specially for use in crop phenotyping. We have already created over 150 marked images for the evaluation of crop vegetation segmentation algorithms, and 300 rice flower images for detection of rice flowering — we would like to share those images to start a database

server for other researchers who wish to evaluate their own proposed algorithms. We are also expecting to attract more computer science researchers to this effort.

In recent years, the newest and hottest topic for object recognition has turned to the Large-Scale Generic Object/Scene recognition. Researchers are aiming to develop algorithms that help computers to recognize more than 30,000 generic objects in the world using only digital images. This challenging task is progressing rapidly, and the ideas generated by such approaches are powerful for object recognition. In the case of crop phenotyping in agricultural research, because we focused only on several major plants/organs, detecting them from RGB images could be made more efficient by integrating the latest object recognition technologies. This will better enable the extraction of useful phenotypic crop characteristics with high speed and accuracy, which will facilitate the generation of low cost, field-based, high throughput phenotyping.

Acknowledgements.

I would like to offer special thanks to Dr. Seishi Ninomiya of the Graduate School of Agricultural and Life Sciences, who not only served as my adviser but as a professor and friend who gave me the opportunity to work with him during the past three and half years and provided support and encouragement to complete my degree.

I would like to thank Prof. Kazuhiko Kobayashi, Prof. Masaru Mizoguchi, Prof. Junko Yamagishi, Prof. Hiroyoshi Iwata, Mr. Daisuke Tanada, Mr. Kyosuke Yamamoto, Mr. Kazuhiro Nishioka, Dr. Sudharsan of the University of Tokyo, Dr. Tokihiro Fukatsu, Dr. Hiroshi Nakagawa, Dr. Hiroe Yoshida, Dr. Masayuki Hirafuji, Dr. Takuji Kiura of the National Agriculture and Food Research Organization, Dr. Yoichiro Kato of the International Rice Research Institute (IRRI) for providing valuable guidance and suggestions. Special thanks to Prof. Kazuhiko Kobayashi, Prof. Masaru Mizoguchi, Prof. Junko Yamagishi, Prof. Hiroyoshi Iwata for serving on my dissertation committee and their careful review and many suggestions that were provided on the original manuscripts. Special thanks to Dr. Tokihiro Fukatsu for providing field monitoring system (Field sever).

I would like to thank All the Tech. Support Staffs of Institute for Sustainable Agro-ecosystem Services of The University of Tokyo for providing helps of my field experiment.

I would like to thank MEXT (Ministry of Education, Culture, Sports, Science & Technology in Japan) for funding my assistantship through RECCA (Research Program on Climate Change Adaptation) project and two years scholarship (Monbukagakusho Honors Scholarship for Privately Financed International Students).

Finally I would like to thank my parents, my wife Lake and my son Huaiyuan, who gave me support throughout these years and the encouragement to complete this degree. To them, I gratefully dedicate this dissertation.

References

Agarwal, S., Awan, A., & Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. doi:10.1109/TPAMI.2004.108

Agarwala, A., Agrawala, M., Cohen, M., Salesin, D., & Szeliski, R. (2006). Photographing Long Scenes with Multi-viewpoint Panoramas. *ACM Trans. Graph.*, *25*(3), 853–861. doi:10.1145/1141911.1141966

Ahmed, F., Al-Mamun, H. A., Bari, A. S. M. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, *40*(0), 98–104. Retrieved from http://www.sciencedirect.com/science/article/pii/S026121941200124X

Ambuel, J., Ninomiya, S., & Takahashi, N. (1997). Fuzzy Logic Evaluation of Soybean Plant Shape. *Breeding science*, *47*(3), 253–257. Retrieved from http://ci.nii.ac.jp/naid/110001815381/en/

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., … White, J. W. (2013). Development and evaluation of a field-based high-throughput phenotyping platform. *Functional Plant Biology*. doi:10.1071/FP13126

Araus, J. L., & Cairns, J. E. (2013). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*. Retrieved from http://www.sciencedirect.com/science/article/pii/S1360138513001994

Arribas, J. I., Sánchez-Ferrero, G. V., Ruiz-Ruiz, G., & Gómez-Gil, J. (2011). Leaf classification in sunflower crops by computer vision and neural networks. *Computers and Electronics in Agriculture*, *78*(1), 9–18. doi:10.1016/j.compag.2011.05.007

Bay, H., Ess, A., Tuytelaars, T., & Gool, L. Van. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, (September). Retrieved from http://www.sciencedirect.com/science/article/pii/S1077314207001555

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning* (p. 738). Springer. Retrieved from http://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738

Bosch, A., Zisserman, A., & Muoz, X. (2007). Image Classification using Random Forests and Ferns. *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. doi:10.1109/ICCV.2007.4409066

Breiman, L., Friedman, J. H., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees* (p. 358). Wadsworth International Group.

Brown, M., & Lowe, D. (2002). Invariant Features from Interest Point Groups. *BMVC*, 23.1–23.10. doi:10.5244/C.16.23

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., … McMullen, M. D. (2009). c. *Science (New York, N.Y.)*, *325*(5941), 714–8. doi:10.1126/science.1174276

Buddhiraju, K. M., & Rizvi, I. A. (2010). Comparison of CBF, ANN and SVM classifiers for object based classification of high resolution satellite images. *2010 IEEE International Geoscience and Remote Sensing Symposium*, 40–43. doi:10.1109/IGARSS.2010.5652033

Burgos-Artizzu, X. P., Ribeiro, A., Guijarro, M., & Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, *75*(2), 337–346. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168169910002620

Burgos-Artizzu, X. P., Ribeiro, A., Tellaeche, A., Pajares, G., & Fernández-Quintanilla, C. (2010). Analysis of natural images processing for the extraction of agricultural elements. *Image and Vision Computing*, *28*(1), 138–149. doi:10.1016/j.imavis.2009.05.009

Bylesjö, M., Segura, V., Soolanayakanahally, R. Y., Rae, A. M., Trygg, J., Gustafsson, P., … Street, N. R. (2008). LAMINA: a tool for rapid quantification of leaf size and shape parameters. *BMC plant biology*, *8*(1), 82. doi:10.1186/1471-2229-8-82

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M. D., & Araus, J. L. (2012). High-throughput phenotyping and genomic selection: the frontiers of crop breeding converge. *Journal of integrative plant biology*, *54*(5), 312–20. doi:10.1111/j.1744-7909.2012.01116.x

Cairns, J. E., Sanchez, C., Vargas, M., Ordoñez, R., & Araus, J. L. (2012). Dissecting maize productivity: ideotypes associated with grain yield under drought stress and well-watered conditions. *Journal of integrative plant biology*, *54*(12), 1007–20. doi:10.1111/j.1744-7909.2012.01156.x

Campillo, C., Prieto, M., & Daza, C. (2008). Using digital images to characterize canopy coverage and light interception in a processing tomato crop. *HORTSCIENCE*, *43*(6), 1780–1786. Retrieved from http://hortsci.ashspublications.org/content/43/6/1780.short

Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-8*(6), 679–698. doi:10.1109/TPAMI.1986.4767851

Cobb, J., DeClerck, G., Greenberg, A., Clark, R., & McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, *126*(4), 867–887. doi:10.1007/s00122-013-2066-0

Duan, L., Yang, W., Huang, C., & Liu, Q. (2011). A novel machine-vision-based facility for the automatic evaluation of yield-related traits in rice. *Plant methods*, *7*(1), 44. doi:10.1186/1746-4811-7-44

Evers, J. B., Huth, N. I., & Renton, M. (2009). Light Extinction in Spring Wheat Canopies in Relation to Crop Configuration and Solar Angle. In *2009 Third International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications* (pp. 107–110). IEEE. doi:10.1109/PMA.2009.20

Fei-Fei, L., & Perona, P. (2005). A Bayesian hierarchical model for learning natural scene categories. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. doi:10.1109/CVPR.2005.16

Fiorani, F., & Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual review of plant biology*, *64*, 267–91. doi:10.1146/annurev-arplant-050312-120137

Fritz, G., Seifert, C., Kumar, M., & Paletta, L. (2005). Building Detection from Mobile Imagery Using Informative SIFT Descriptors. In *Proceedings of the 14th Scandinavian Conference on Image Analysis* (pp. 629–638). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/11499145_64

Fukatsu, T., & Hirafuji, M. (2005). Field monitoring using sensor-nodes with a Web server. *Journal of Robotics and Mechatronics*, *17*(2), 164–172. Retrieved from http://model.job.affrc.go.jp/FieldServer/paper/field_server02.pdf

Fukatsu, T., Kiura, T., & Hirafuji, M. (2011). A web-based sensor network system with distributed data processing approach via web application. *Computer Standards & Interfaces*, *33*(6), 565–573. Retrieved from http://www.sciencedirect.com/science/article/pii/S092054891100033X

Fukatsu, T., Watanabe, T., Hu, H., Yoichi, H., & Hirafuji, M. (2012). Field monitoring support system for the occurrence of Leptocorisa chinensis Dallas (Hemiptera: Alydidae) using synthetic attractants, Field Servers, and image analysis. *Computers and Electronics in Agriculture*, *80*, 8–16. Retrieved from http://www.sciencedirect.com/science/article/pii/S016816991100233X

Fukushima, A., Kusuda, O., & Furuhata, M. (2003). Relationship of vegetation cover ratio to trowth and yield in wheat. *Report of the Kyushu Branch of the Crop Science Society of Japan*, (69), 33–35. Retrieved from http://ci.nii.ac.jp/naid/110001784847/en/

Furbank, R. T. (2009). Plant phenomics: from gene to form and function. *Functional Plant Biology*, *36*(11), v–vi. Retrieved from http://www.publish.csiro.au/paper/FPv36n11_FO

Furbank, R. T., & Tester, M. (2011). Phenomics--technologies to relieve the phenotyping bottleneck. *Trends in plant science*, *16*(12), 635–44. doi:10.1016/j.tplants.2011.09.005

Gabriella, C., Dance, C. R., Fan, L., Willamowski, J., & Bray, C. (2004). Visual categorization with bags of keypoints. *Workshop on Statistical Learning in*

Computer Vision, ECCV, 1–22. Retrieved from http://217.109.185.161/layout/set/print/content/download/20785/148346/file/2004_010.pdf

Garan, S. A. (2003). Phenomics: a new direction for the study of neuroendocrine aging. *Experimental Gerontology*, *38*(1), 218. Retrieved from http://www.sciencedirect.com/science/article/pii/S0531556502001845

Gleadow, R., Johnson, A., & Tausz, M. (2013). Crops for a future climate. *Functional Plant Biology*, *40*(2), iii. doi:10.1071/FPv40n2_FO

Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S. J., Dauzat, M., Hamard, P., … Tardieu, F. (2006). PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *The New phytologist*, *169*(3), 623–35. doi:10.1111/j.1469-8137.2005.01609.x

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P. J., Burgos-Artizzu, X. ., & Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, *75*(1), 75–83. doi:10.1016/j.compag.2010.09.013

Guo, W., Rage, U. K., & Ninomiya, S. (2013). Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Computers and Electronics in Agriculture*, *96*, 58–66. doi:10.1016/j.compag.2013.04.010

Hamid Reza Vaezi Joze and Mark S. Drew. (2010). Improved machine learning for image category recognition by local color constancy. In *International Conference on Image Processing* (pp. 1–4). HongKong. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651069

Harris, C., & Stephens, M. (1988). A Combined Corner and Edge Detector. *Procedings of the Alvey Vision Conference 1988*, 23.1–23.6. doi:10.5244/C.2.23

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

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: the next challenge. *Nature reviews. Genetics*, *11*(12), 855–66. doi:10.1038/nrg2897

Ishizuka, T., Tanabata, T., Takano, M., & Tomoko, S. (2005). Kinetic measuring method of rice growth in tillering stage using automatic digital imaging system. *Environ. Control Biol.*,, *43*(2), 83–96.

Iwata, H., Nesumi, H., Ninomiya, S., Takano, Y., & Ukai, Y. (2002). The Evaluation of Genotype × Environment Interactions of Citrus Leaf Morphology Using Image Analysis and Elliptic Fourier Descriptors. *Breeding science*, *52*(4), 243–251. Retrieved from http://ci.nii.ac.jp/naid/110001803200/en/

Iyer-Pascuzzi, A. S., Symonova, O., Mileyko, Y., Hao, Y., Belcher, H., Harer, J., … Benfey, P. N. (2010). Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant physiology*, *152*(3), 1148–57. doi:10.1104/pp.109.150748

Jack, A., Ninomiya, S., & Takahashi, N. (1997). Fuzzy Logic Evaluation of Soybean Plant Shape. *Breeding science*, *47*(3), 253–257. Retrieved from http://ci.nii.ac.jp/naid/110001815381/

Jafari, A., Mohtasebi, S. S., Jahromi, H. E., & Omid, M. (2006). Weed detection in sugar beet fields using machine vision. *International Journal of Agriculture & Biology*, *8*(5), 602–605. Retrieved from http://www.fspublishers.org/ijab/past-issues/IJABVOL_8_NO_5/8.pdf

Jafari, Abdolabbas, Fazayeli, A., & Zarezadeh, M. R. (2014). Estimation of orange skin thickness based on visual texture coarseness. *Biosystems Engineering*, *117*(0), 73–82. doi:http://dx.doi.org/10.1016/j.biosystemseng.2013.08.010

Jegou, H., Douze, M., Schmid, C., & Perez, P. (2010). Aggregating local descriptors into a compact image representation. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 3304–3311). IEEE. doi:10.1109/CVPR.2010.5540039

Jeon, H. Y., Tian, L. F., & Zhu, H. (2011). Robust Crop and Weed Segmentation under Uncontrolled Outdoor Illumination. *Sensors (Basel, Switzerland)*, *11*(6), 6270–6283. doi:10.3390/s110606270

Jiang, Y.-G., Ngo, C.-W., & Yang, J. (2007). Towards Optimal Bag-of-features for Object Categorization and Semantic Video Retrieval. In *Proceedings of the 6th ACM International Conference on Image and Video Retrieval* (pp. 494–501). New York, NY, USA: ACM. doi:10.1145/1282280.1282352

Joly, A., Goëau, H., Bonnet, P., Bakić, V., Barbe, J., Selmi, S., … Barthélémy, D. (2013). Interactive plant identification based on social image data. *Ecological Informatics*. Retrieved from http://www.sciencedirect.com/science/article/pii/S157495411300071X

Kobayasi, K. (2012). Effects of Solar Radiation on Fertility and the Flower Opening Time in Rice Under Heat Stress Conditions. Retrieved from http://cdn.intechopen.com/pdfs/33354/InTech-Effects_of_solar_radiation_on_fertility_and_the_flower_opening_time_in_rice_under_heat_stress_conditions.pdf

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1-2), 273–324. doi:10.1016/S0004-3702(97)00043-X

Kuhl, F., & Giardina, C. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, *18*(3), 236–258. doi:10.1016/0146-664X(82)90034-X

Lati, R. N., Filin, S., & Eizenberg, H. (2011). Robust Methods for Measurement of Leaf-Cover Area and Biomass from Image Data. *Weed Science*, *59*(2), 276–284. doi:10.1614/WS-D-10-00054.1

Lazebnik, S., Schmid, C., & Ponce, J. (2006). Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. doi:10.1109/CVPR.2006.68

Li, F.-F., Rob, F., & Pietro, P. (2007). Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding*, *106*(1), 59–70. Retrieved from http://www.sciencedirect.com/science/article/pii/S1077314206001688

Liu, J., & Pattey, E. (2010). Retrieval of leaf area index from top-of-canopy digital photography over agricultural crops. *Agricultural and Forest Meteorology*,

150(11), 1485–1490. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168192310002029

Liu, J., Pattey, E., & Admiral, S. (2013). Assessment of in situ crop LAI measurement using unidirectional view digital photography. *Agricultural and Forest Meteorology*, *169*, 25–34. doi:10.1016/j.agrformet.2012.10.009

Liu, Yaokai, Mu, X., Wang, H., & Yan, G. (2011). A novel method for extracting green fractional vegetation cover from digital images. (G. Henebry, Ed.)*Journal of Vegetation Science*, *23*(3), 406–418. doi:10.1111/j.1654-1103.2011.01373.x

Liu, Ying, Zhang, D., Lu, G., & Ma, W.-Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.*, *40*(1), 262–282. doi:10.1016/j.patcog.2006.04.045

Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant physiology*, *160*(4), 1686–97. doi:10.1104/pp.112.208298

Louhaichi, M., Johnson, M. D., Woerz, A. L., Jasra, A. W., & Johnson, D. E. (2010). Digital Charting Technique for Monitoring Rangeland Vegetation Cover at Local Scale. *International Journal of Agriculture & Biology*, (VOL_12_NO_3), 406–410. Retrieved from http://fspublishers.org/ijab/past-issues/IJABVOL_12_NO_3/17.pdf

Lowe, D. (1999). Object recognition from local scale-invariant features. *Computer Vision, 1999. The Proceedings of the*, 1150–1157. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=790410

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*. Retrieved from http://www.springerlink.com/index/H4L02691327PX768.pdf

Malik, J., Belongie, S., Leung, T., & Shi, J. (2001). Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, *43*(1), 7–27. Retrieved from http://link.springer.com/article/10.1023/A%3A1011174803800

Malik, J., Belongie, S., Shi, J., & Leung, T. (1999). Textons, contours and regions: Cue integration in image segmentation. In *IEEE International Conference on Computer*

Vision. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=790346

Matsui, T., Omasa, K., & Horie, T. (1997). High temperature-induced spikelet sterility of Japonica rice at flowering in relation to air temperature, humidity and wind velocity conditions. *Japanese Journal of Crop …*. Retrieved from http://park.itc.u-tokyo.ac.jp/joho/Omasa/255.pdf

Meyer, G. E., & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, *63*(2), 282–293. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168169908001063

Mielewczik, M., Friedli, M., Kirchgessner, N., & Walter, A. (2013). Diel leaf growth of soybean: a novel method to analyze two-dimensional leaf expansion in high temporal resolution based on a marker tracking approach (Martrack Leaf). *Plant methods*, *9*(1), 30. doi:10.1186/1746-4811-9-30

Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. doi:10.1109/TPAMI.2005.188

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., … Gool, L. V. (2005). A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, *65*(1-2), 43–72. doi:10.1007/s11263-005-3848-x

Mizoguchi, M., Ito, T., & Mitsuishi, S. (2010). Ubiquitous Monitoring of Agricultural Fields in Asia using Sensor Network. In *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World* (Vol. 24, pp. 1–4). Retrieved from http://www.iuss.org/19th WCSS/Symposium/pdf/0850.pdf

Mollazade, K., Omid, M., & Arefi, A. (2012). Comparing data mining classifiers for grading raisins based on visual features. *Computers and Electronics in Agriculture*, *84*, 124–131. doi:10.1016/j.compag.2012.03.004

Montes, J. M., Melchinger, A. E., & Reif, J. C. (2007). Novel throughput phenotyping platforms in plant genetic studies. *Trends in plant science*, *12*(10), 433–6. doi:10.1016/j.tplants.2007.08.006

Morra, J., Tu, Z., & Apostolova, L. (2010). Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. *Medical Imaging,*, 1–15. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4957035

Mutch, J., & Lowe, D. G. (2006). Multiclass Object Recognition with Sparse, Localized Features. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. doi:10.1109/CVPR.2006.200

Ninomiya, S. (2009). Evaluation of Crop Shape. *Kansei Engineering International Journal*, *9*(1), 26–30. Retrieved from http://www.jske.org/muinmglj4-127/?block_id=127&active_action=multidatabase_view_main_detail&multidatabase_id=6&content_id=237

Ninomiya, S., & NguyenCong, V. (1998). Evaluation of soybean plant shape based on tree-based models. *Breeding Science*, *48*(3), 251–255.

Ninomiya, S., Oide, M., & Takahashi, N. (1998). Fuzzy-Logic Search Engine for Soybean Shape Data. *Japanese Journal of Crop Science*, *67*(1), 104–108. Retrieved from http://ci.nii.ac.jp/naid/110001742062/en/

Ninomiya, S., Sasaki, A., & Takemura, K. (1992). Evaluation of fineness of wrinkles on husks of malting barley (Hordeum vulgare L.) by texture analysis of digital image data. *Euphytica*, *64*(1-2), 113–121. doi:10.1007/BF00023544

Ninomiya, S., & Shigemori, I. (1991). Quantitative Evaluation of Soybean (Glycine max L. MERRILL) Plant Shape by Image Analysis. *Japanese journal of breeding*, *41*(3), 485–497. Retrieved from http://ci.nii.ac.jp/naid/110001826208/en/

Nowak, E., Jurie, F., & Triggs, B. (2006). Sampling strategies for bag-of-features image classification. *Computer Vision–ECCV 2006*, *3954*, 490–503. Retrieved from http://link.springer.com/chapter/10.1007/11744085_38

Ohta, Y., Kanade, T., & Sakai, T. (1980). Color information for region segmentation. *Computer Graphics and Image Processing*, *13*(3), 222–241. doi:10.1016/0146-664X(80)90047-7

Oide, M., & Ninomiya, S. (1998). Evaluation of Soybean Plant Shape by Multilayer Perceptron with Direct Image Input. *Breeding science*, *48*(3), 257–262. Retrieved from http://ci.nii.ac.jp/naid/110001807910/en/

Oide, M., & Ninomiya, S. (2000). Discrimination of soybean leaflet shape by neural networks with image input. *Computers and Electronics in Agriculture*, *29*(1), 59–72. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168169900001368

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, *9*(1), 62–66. doi:10.1109/TSMC.1979.4310076

Panneton, B., & Brouillard, M. (2009). Colour representation methods for segmentation of vegetation in photographs. *Biosystems Engineering*, *102*(4), 365–378. doi:10.1016/j.biosystemseng.2009.01.003

Pereira, L. A. M., Nakamura, R. Y. M., de Souza, G. F. S., Martins, D., & Papa, J. P. (2012). Aquatic weed automatic classification using machine learning techniques. *Computers and Electronics in Agriculture*, *87*, 56–63. doi:10.1016/j.compag.2012.05.015

Pereyra-Irujo, G. a., Gasco, E. D., Peirone, L. S., & Aguirrezábal, L. a. N. (2012). GlyPh: a low-cost platform for phenotyping plant growth and water use. *Functional Plant Biology*, *39*(11), 905. doi:10.1071/FP12052

Pérez, A. J., López, F., Benlloch, J. V., & Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture*, *25*(3), 197–212. Retrieved from http://dx.doi.org/10.1016/S0168-1699(99)00068-X

Perronnin, F, Liu, Y., Sanchez, J., & Poirier, H. (2010). Large-scale image retrieval with compressed Fisher vectors. *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. doi:10.1109/CVPR.2010.5540009

Perronnin, Florent. (2008). Universal and adapted vocabularies for generic visual categorization. *IEEE transactions on pattern analysis and machine intelligence*, *30*(7), 1243–56. doi:10.1109/TPAMI.2007.70755

Perronnin, Florent, Dance, C., Csurka, G., & Bressan, M. (2006). Adapted vocabularies for generic visual categorization. In *European Conference on Computer Vision–ECCV* (pp. 464–475). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4359362

Philipp, I., & Rath, T. (2002). Improving plant discrimination in image processing by use of different colour space transformations. *Computers and Electronics in Agriculture*, *35*(1), 1–15. doi:10.1016/S0168-1699(02)00050-9

Picard, D., & Gosselin, P.-H. (2011). Improving image similarity with vectors of locally aggregated tensors. *Image Processing (ICIP), 2011 18th IEEE International Conference on*. doi:10.1109/ICIP.2011.6116641

Pinto, N., Barhomi, Y., Cox, D. D., & DiCarlo, J. J. (2011). Comparing state-of-the-art visual features on invariant object recognition tasks. *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*. doi:10.1109/WACV.2011.5711540

Prashar, A., Yildiz, J., McNicol, J. W., Bryan, G. J., & Jones, H. G. (2013). Infra-red thermography for high throughput field phenotyping in Solanum tuberosum. *PloS one*, *8*(6), e65816. doi:10.1371/journal.pone.0065816

R Development Core Team. (2012). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from http://www.r-project.org/

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. (J. P. Hart, Ed.)*PloS one*, *8*(6), e66428. doi:10.1371/journal.pone.0066428

Reuzeau, C., Pen, J., Frankard, V., Wolf, J. de, Peerbolte, R., Broekaert, W., & Camp, W. van. (2010, September 6). TraitMill: a Discovery Engine for Identifying Yield-enhancement Genes in Cereals. *Plant Gene and Trait*. Retrieved from http://bio.sophiapublisher.com/index.php/pgt/article/view/53

Riomoros, I., Guijarro, M., Pajares, G., Herrera, P. J., Burgos-Artizzu, X. P., & Ribeiro, A. (2010). Automatic image segmentation of greenness in crop fields. In *2010 International Conference of Soft Computing and Pattern Recognition (SoCPaR)* (pp. 462–467).

Royo, C., & Villegas, D. (2011). Field Measurements of Canopy Spectra for Biomass Assessment of Small-Grain Cereals. In D. Matovic (Ed.), *Biomass - Detection, Production and Usage* (p. 496). InTech. Retrieved from http://www.intechopen.com/source/pdfs/19066/InTech-Field_measurements_of_canopy_spectra_for_biomass_assessment_of_small_grain_cereals.pdf

Ruiz-Altisent, M., Ruiz-Garcia, L., Moreda, G. P., Lu, R., Hernandez-Sanchez, N., Correa, E. C., … García-Ramos, J. (2010). Sensors for product characterization and quality of specialty crops—A review. *Computers and Electronics in Agriculture*, *74*(2), 176–194. doi:10.1016/j.compag.2010.07.002

Ruiz-Ruiz, G., Gómez-Gil, J., & Navas-Gracia, L. M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture*, *68*(1), 88–96. doi:10.1016/j.compag.2009.04.009

Rumpf, T., Römer, C., Weis, M., Sökefeld, M., Gerhards, R., & Plümer, L. (2012). Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsium arvense and Galium aparine. *Computers and Electronics in Agriculture*, *80*, 89–96. doi:10.1016/j.compag.2011.10.018

Sakamoto, T., Shibayama, M., Kimura, A., & Takada, E. (2011). Assessment of digital camera-derived vegetation indices in quantitative monitoring of seasonal rice growth. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 872–882. Retrieved from http://www.sciencedirect.com/science/article/pii/S0924271611000931

Schmid, C., Mohr, R., & Bauckhage, C. (2000). Evaluation of Interest Point Detectors. *International Journal of Computer Vision*, *37*(2), 151–172. doi:10.1023/A:1008199403446

Sequencing ProjectInternational Rice Genome. (2005). The map-based sequence of the rice genome. *Nature*, *436*(7052), 793–800. doi:10.1038/nature03895

Shotton, J., Winn, J., Rother, C., & Criminisi, A. (2006). TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-class Object Recognition and Segmentation. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision* –

ECCV 2006 SE - 1 (Vol. 3951, pp. 1–15). Springer Berlin Heidelberg. doi:10.1007/11744023_1

Sivic, J., & Zisserman, A. (2003). Video Google: a text retrieval approach to object matching in videos. *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. doi:10.1109/ICCV.2003.1238663

Sparks, D. L., Cairns, J. E., Sonder, K., Zaidi, P. H., Verhulst, N., Mahuku, G., … Prasanna, B. M. (2012). Maize Production in a Changing Climate. In *Advances in Agronomy* (Vol. 114, pp. 1–58). Retrieved from http://www.sciencedirect.com/science/article/pii/B9780123942753000067

Sudharsan, D., Adinarayana, J., Tripathy, a. K., Ninomiya, S., Hirafuji, M., Kiura, T., … Sreenivas, G. (2012). GeoSense: A Multimode Information and Communication System. *ISRN Sensor Networks*, *2012*, 1–13. doi:10.5402/2012/215103

Takahashi, K., Rikimaru, A., Sakata, K., & Endou, S. (2012). A Study of the characteristic of the observation angle on the terrestrial image measurement of paddy vegetation cover (In Japanese). *Japan Society of Photogrammetry and Remote Sensing*, *50*(6), 367–371. Retrieved from http://ci.nii.ac.jp/naid/10030287221/en/

Tang, X., Liu, M., Zhao, H., & Tao, W. (2009). Leaf Extraction from Complicated Background. In *2nd International Congress on Image and Signal Processing, 2009. CISP ’09.* (pp. 1–5).

Tellaeche, A., Burgos-Artizzu, X. P., Pajares, G., & Ribeiro, A. (2008). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition*, *41*(2), 521–530. doi:10.1016/j.patcog.2007.07.007

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20260–4. doi:10.1073/pnas.1116437108

Tsuduki, Y., & Fujiyoshi, H. (2008). A Method for Visualizing Pedestrian Traffic Flow Using SIFT Feature Point Tracking. In *Proceedings of the 3rd Pacific Rim*

Symposium on Advances in Image and Video Technology (pp. 25–36). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-540-92957-4_3

Vedaldi, A., & Fulkerson, B. (2010). Vlfeat: An Open and Portable Library of Computer Vision Algorithms. In *Proceedings of the International Conference on Multimedia* (pp. 1469–1472). New York, NY, USA: ACM. doi:10.1145/1873951.1874249

Vedaldi, A., Ling, H., & Soatto, S. (2010). Knowing a good feature when you see it: Ground truth and methodology to evaluate local features for recognition. *Computer Vision: Detection, Recognition and Reconstruction*, *285*, 27–49. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-12848-6_2

Vedaldi, A., & Zisserman, A. (2010). Efficient additive kernels via explicit feature maps. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. doi:10.1109/TPAMI.2011.153

Vedaldi, A., & Zisserman, A. (2012). Efficient Additive Kernels via Explicit Feature Maps. (A. Zisserman, Ed.)*IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(3), 480–492. Retrieved from http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.153

Vergauwen, M., & Van Gool, L. (2006). Web-based 3D Reconstruction Service. *Mach. Vision Appl.*, *17*(6), 411–426. doi:10.1007/s00138-006-0027-1

Welles, J. M., & Cohen, S. (1996). Canopy structure measurement by gap fraction analysis using commercial instrumentation. *Journal of Experimental Botany*, *47*(9), 1335–1342. doi:10.1093/jxb/47.9.1335

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., … Wang, G. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, *133*, 101–112. Retrieved from http://www.sciencedirect.com/science/article/pii/S037842901200130X

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil residue and lighting conditions. *Transactions of the ASAE*, *38*, 259–269.

Yamanaka, N. (2001). An Informative Linkage Map of Soybean Reveals QTLs for Flowering Time, Leaflet Morphology and Regions of Segregation Distortion. *DNA Research*, *8*(2), 61–72. doi:10.1093/dnares/8.2.61

Yamasaki, M., & Arturo, G. (2012). FieldBook : development of the bar-coded phenotyping system and the integration into plant breeding and genetics(Proceedings of the 40th Symposium). *Journal of crop research*, (57), 55–59. Retrieved from http://ci.nii.ac.jp/naid/110009553892/en/

Yanai, K. (2010). Recent Progresses in Object Recognition. *Journal of the Robotics Society of Japan*, *28*(3), 257–260. doi:10.7210/jrsj.28.257

Yoshida, H., & Nagato, Y. (2011). Flower development in rice. *Journal of experimental botany*, *62*(14), 4719–30. doi:10.1093/jxb/err272

Yoshida, S. (1981). *Fundamentals of Rice Crop Science* (p. 269). Int. Rice Res. Inst. Retrieved from http://books.google.com/books?id=323XxM076SsC&pgis=1

Yoshioka, Y, Iwata, H., Ohsawa, R., & Ninomiya, S. (2005). Quantitative evaluation of the petal shape variation in Primula sieboldii caused by breeding process in the last 300 years. *Heredity*, *94*(6), 657–63. doi:10.1038/sj.hdy.6800678

Yoshioka, Yosuke, Honjo, M., Iwata, H., Ninomiya, S., & Ohsawa, R. (2007). Pattern of geographical variation in petal shape in wild populations of Primula sieboldii E. Morren. *Plant Species Biology*, *22*(2), 87–93. doi:10.1111/j.1442-1984.2007.00180.x

Yoshioka, Yosuke, Iwata, H., Tabata, M., Ninomiya, S., & Ohsawa, R. (2007). Chalkiness in Rice: Potential for Evaluation with Image Analysis. *Crop Science*, *47*(5), 2113. doi:10.2135/cropsci2006.10.0631sc

Yoshioka, Yosuke, Ohsawa, R., Iwata, H., Ninomiya, S., & Fukuta, N. (2006). Quantitative Evaluation of Petal Shape and Picotee Color Pattern in Lisianthus by Image Analysis. *J. Amer. Soc. Hort. Sci.*, *131*(2), 261–266. Retrieved from http://journal.ashspublications.org/content/131/2/261.abstract

Yu, Z., Cao, Z., Wu, X., Bai, X., Qin, Y., Zhuo, W., … Xue, H. (2013). Automatic image-based detection technology for two critical growth stages of maize:

Emergence and three-leaf stage. *Agricultural and Forest Meteorology*, *174–175*(0), 65–84. doi:http://dx.doi.org/10.1016/j.agrformet.2013.02.011

Yusof, R., Khalid, M., & M. Khairuddin, A. S. (2013). Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system. *Computers and Electronics in Agriculture*, *93*(null), 68–77. doi:10.1016/j.compag.2013.01.007

Zhang, J., Marszałek, M., Lazebnik, S., & Schmid, C. (2007). Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study. *International Journal of Computer Vision*, *73*(2), 213–238. doi:10.1007/s11263-006-9794-4

Zhang, Z., Li, Z. N., & Drew, M. S. (2010). Learning image similarities via probabilistic feature matching. In *Image Processing (ICIP), 2010 17th IEEE International Conference on* (pp. 1857–1860). HongKong: IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5653990

Zheng, L., Shi, D., & Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminant. *Pattern Recognition Letters*, *31*(9), 920–925. doi:10.1016/j.patrec.2010.01.016

Zhou, X., Yu, K., Zhang, T., & Huang, T. T. (2010). Image Classification Using Super-Vector Coding of Local Image Descriptors. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010 SE - 11* (Vol. 6315, pp. 141–154). Springer Berlin Heidelberg. doi:10.1007/978-3-642-15555-0_11