

Assessment of crop insect damage using unmanned aerial systems: A machine learning approach

E. Puig^a, F. Gonzalez^a, G. Hamilton^a and P. Grundy^b

^a *Australian Research Centre for Aerospace Automation (ARCAA) and Science and Engineering Faculty,
Queensland University of Technology (QUT), Queensland*

^b *Queensland Department of Agriculture and Fisheries, Queensland
Email: eduard.puiggarcia@qut.edu.au*

Abstract: Agricultural pests are responsible for millions of dollars in crop losses and management costs every year. In order to implement optimal site-specific treatments and reduce control costs, new methods to accurately monitor and assess pest damage need to be investigated. In this paper we explore the combination of unmanned aerial vehicles (UAV), remote sensing and machine learning techniques as a promising technology to address this challenge. The deployment of UAVs as a sensor platform is a rapidly growing field of study for biosecurity and precision agriculture applications. In this experiment, a data collection campaign is performed over a sorghum crop severely damaged by white grubs (Coleoptera: Scarabaeidae). The larvae of these scarab beetles feed on the roots of plants, which in turn impairs root exploration of the soil profile. In the field, crop health status could be classified according to three levels: bare soil where plants were decimated, transition zones of reduced plant density and healthy canopy areas. In this study, we describe the UAV platform deployed to collect high-resolution RGB imagery as well as the image processing pipeline implemented to create an orthoimage. An unsupervised machine learning approach is formulated in order to create a meaningful partition of the image into each of the crop levels. The aim of the approach is to simplify the image analysis step by minimizing user input requirements and avoiding the manual data labeling necessary in supervised learning approaches. The implemented algorithm is based on the K-means clustering algorithm. In order to control high-frequency components present in the feature space, a neighbourhood-oriented parameter is introduced by applying Gaussian convolution kernels prior to K-means. The outcome of this approach is a *soft* K-means algorithm similar to the EM algorithm for Gaussian mixture models. The results show the algorithm delivers decision boundaries that consistently classify the field into three clusters, one for each crop health level. The methodology presented in this paper represents a venue for further research towards automated crop damage assessments and biosecurity surveillance.

Keywords: *Unmanned aerial vehicles (UAV), machine learning, k-means, remote sensing, biosecurity*

1. INTRODUCTION

Invertebrate pests represent a major threat to grain production in Australia. Sorghum is the 5th largest crop in cultivated area and the average annual cost of invertebrate pests to the sorghum industry is \$29.4 million (Murray, 2013). A complex of white grub species (Coleoptera: Scarabaeidae) have been responsible for intermittent but locally severe damage to sorghum crops throughout Australia's eastern grain cropping regions. The complex of endemic species have historically afflicted areas of pasture (Turner et al., 1969) but due to recent changes in farming system practices consisting of grain cropping with standing stubble and minimum tillage, these pests are being increasingly encountered in summer monocot crops such as sorghum and maize (Miles pers. Com., 2015¹). The larvae of these scarab beetles feed on the roots of plants which in turn impairs root exploration of the soil profile and leading to reduced plant growth and possible plant death. As soil borne pests, the larvae are difficult to sample and predict (Buntin, 2012). Because soil pest infestations can have high spatial and temporal variability, comprehensive ground sampling of large crop fields is typically not cost-effective. In order to implement optimal site-specific treatments and reduce control costs, new methods in biosecurity and pest surveillance must be investigated.

Remote sensing applications in agriculture have focused on a wide range of endeavours for the past 25 years (Mulla et al., 2013). Aerial imagery is increasingly being adopted as a management tool in precision agriculture. A variety of sensors have been used for detecting and monitoring pests in agricultural environments (Lan et al., 2009, Lehmann et al., 2015, Gonzalez et al., 2011). Aerial data collection using lightweight, unmanned aerial vehicles (UAV) is one of the most rapidly developing fields in remote sensing technology. UAV-based remote sensing offers a cost-effective sensor platform to acquire near-real time crop data to enhance farm management. Government agencies and research institutions are paying close attention to these emerging technologies (McFadyen et al., 2014), but most recently crop insurance companies have also expressed their interest in adopting UAV platforms for crop damage assessments (Reuters, April 2015).

Remote sensing is an important application field for machine learning and computer vision algorithms. In particular, the research in Geographic Object-Based Image Analysis (GEOBIA) has been sharply increasing for the past 10 years (Blaschke et al. 2014). Both supervised and unsupervised segmentation algorithms are greatly influenced by image quality, spectral bands, spatial resolution and the complexity of the scene. Furthermore most of the available algorithms need to be fine-tuned by the user to extract specific objects of interest (Hay et al., 2008). Thus, image segmentation remains a highly subjective task usually achieved in a trial and error fashion (Belgiu et al., 2014). Cluster analysis, is a major branch of machine learning techniques, aims at organizing the items of a data set into clusters, such that items within a cluster are more "similar" to each other than they are to items in the other clusters. In this study, a neighborhood-oriented parameter is introduced through Gaussian convolution kernel prior to applying K-means clustering, thus providing a single user input to control segmentation quality.

2. METHODS

2.1. Experiment site

The experimental flight campaign was conducted in March 2015 on a private sorghum farm in South East Queensland, Australia. Figure 1a shows the approximate location of the flight area within Australia along with the final orthoimage overlaid on Google Earth. By performing a visual ground sampling, the crop status could be classified in three main groups: healthy canopy, bare soil where plants have been decimated by the pest and transition areas. Figures 1b and 1c show a zoomed-in perspective of one of the severely damaged areas. In Figure 1c it is possible to appreciate the transition from healthy plants in a high density distribution to sparsely distributed unhealthy plants and finally decimated areas.

¹ Dr M. Miles, Principal Entomologist, Qld Dept Agriculture & Fisheries Toowoomba

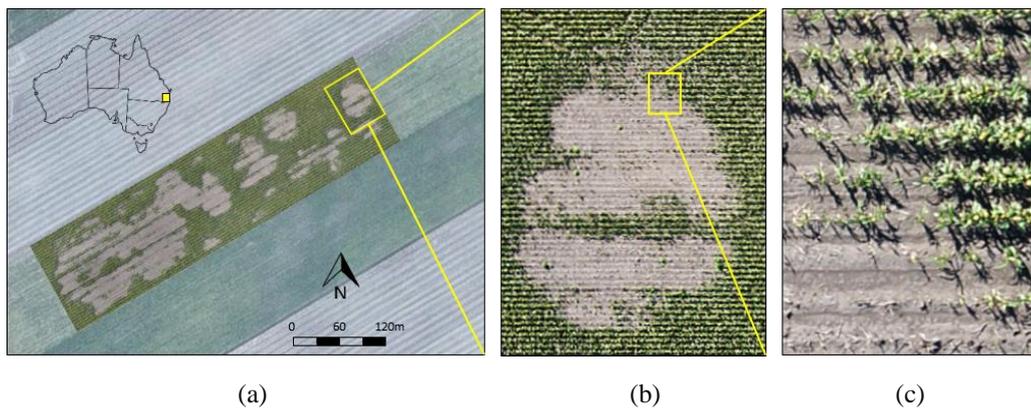


Figure 1. (a) Location of the farm in Queensland, Australia and orthoimage overlaid in Google Earth. (b) Sample damaged zone. (c) High-resolution showing sample image showing levels of crop health.

2.2. UAV system

The system used in this experiment was the multirotor DJI S800 EVO (DJI, Hong Kong), owned and operated by the Australian Research Centre for Aerospace Automation (ARCAA). The autopilot on board was the DJI Wookong M (WKM) controlled via a 900 MHz datalink by the DJI Ground Station Software (GS), installed in a ground base laptop. The UAV was fitted with a DJI Zenmuse Z15 dampened gimbal providing active 3-axis stabilization of the sensor payload. The gimballed camera mount levelled out any platform movement to ensure the sensor was pointing in nadir direction to the ground at all times during data collection flights. The combined system of the WKM and Zenmuse gimbal allows for autopilot triggering of the camera at user defined intervals. Real-time video was received in the ground base through the DJI 5.8 GHz videolink. A list of relevant features of the UAV platform and sensor payload is presented in Table 1.

2.3. Sensor payload

The imaging sensor deployed in this experiment was the Sony NEX-5R high-resolution camera. The spectral response of the camera is based on the Bayer pattern color filter array. The NEX-5R sensor fulfilled the spatial resolution, image quality and weight requirements for this study. Table 1 lists the relevant features of the camera, including some of the settings selected for the data collection, with “(s)”. Due to the limitations of the WKM autopilot in delivering onboard positioning data after the mission, the Aaronia small-footprint GPS and multi-sensor logger (Aaronia AG, Euscheid, Germany) was installed in the system.

Table 1. UAV platform and sensor specifications

S800 EVO with WKM Autopilot		NEX-5R	
Manufacturer	DJI Innovations	Manufacturer	Sony Corporation
UAV type	Multirotor, hexacopter	Camera type	Digital camera
Added Features	Gimbal, 5.8 GHz videolink	Camera mode (s)	Auto
Max. payload	2 Kg	Sensor type	CMOS sensor
Max. takeoff weight	8 Kg	Sensor size	23.5x15.6mm
Power source	2 x 16Ah 6S Tattu LiPo	Field of view	66° x 46.8°
Endurance	16 min	Spectral range	Red, Green, Blue
WKM max. waypoints	50	Spatial resolution (s)	4592 x 3056
Suitable wind condition	<8m/s (15.6kt, 28.8km/h)	Image format (s)	Jpeg
Max. speed	9m/s vertical, 26m/s hor.	Handling (s)	autopilot triggering
Max. flight height	121m (CASA regulated)	Weight	276g

2.4. Flight planning

Due to the vast size of the farm, a preliminary assessment flight was performed prior to data collection to assess the locations and distribution of pest damaged areas. In this flight, the gimbal was controlled remotely by a gimbal operator while receiving real-time video signal via videolink. A rectangular perimeter of 375m x 750m

containing the area of interest was determined in-situ and roughly referenced on Google Earth for flight path design. The edges of the rectangle were used to define the flight area in the ground station software. A flight path consisting of 19 waypoints in a ‘lawn-mowing’ pattern was implemented and uploaded to the WKM autopilot via datalink. In order to create a high-resolution data set, the flight height was selected at 60m above the ground level, yielding a ground sampling distance between 1.5 and 2 cm/pixel. The horizontal ground speed was selected at 5.5 m/s to achieve a frontal overlap between consecutive images above 85% to guarantee successful orthoimage generation. The distance between parallel transacts was calculated to obtain a lateral overlap above 65% for the same reason. To guarantee each waypoint was reached and the lateral overlap remained constant, ‘stop-and-turn’ waypoint turning mode was implemented.

2.5. Data acquisition

Three consecutive flights were required to complete the flight path. All flights were developed within line of sight, between 3.26pm and 4.14pm (GMT+10) covering the predefined area. Scattered clouds were present during the period of data acquisition. The Camera mode was set to automatic as time and varying light conditions did not allow for an accurate manual configuration of balance, aperture, ISO and shutter speed. The image capture rate was controlled by the WKM autopilot, which was programmed to emit a pulse at desired positions in the flight path to maintain a frontal overlap above 80%. The WKM autopilot was connected to the gimbal system which triggered the remote shutter release of the camera. The Aaronia GPS logger was time-synchronized with the NEX-5R, recording raw GPS, altimeter, compass, gyro and other IMU data.

2.6. Photogrammetry

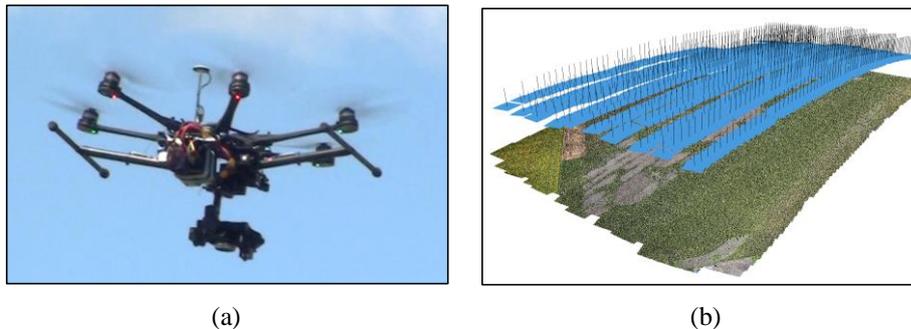


Figure 2. (a) S800 EVO platform with Z15 gimbal and NEX-5R imaging sensor pointing in nadir direction. (b) Screenshot of the image orthorectification and 3D reconstruction result using Agisoft Photoscan.

An image processing pipeline was implemented in order to create a single orthoimage of the area of interest for further analysis. First, the synchronized Aaronia positioning data was used to stamp each image with an estimate of the 3D position from which it was triggered. The resulting data set was imported from the Agisoft PhotoScan Pro software (Agisoft LLC, 2013) for further processing. This software’s strength lies in its ability to orthorectify each individual image and accurately generate a 3D reconstruction (Figure 2b) based on a structure-from-motion (SfM) algorithm. The reconstruction was georeferenced with 5 ground control points (GCPs) spread in the field during the mission. The high-resolution orthoimage exported from Photoscan covered an area of 28.3ha, had dimensions of 40784 x 33177 pixels and a volume of 2.5 Gb. A sample is shown in Figure 1c. From the original orthoimage, a target area of approximately 130m x 472m (6.1ha) was selected. The resulting orthoimage was lowered to 2322 x 640 pixels to reduce the computational cost in further image analysis while maintaining the relevant features of the image. The Google Earth overlay is shown in Figure 1a and a zoomed-in sample is shown in Figure 1b.

2.7. Image analysis

Cluster analysis or clustering refers to a broad set of machine learning techniques used for finding natural clusters in a data set given a measure of similarity. Clusters can be defined as high-density regions in the feature space separated by lower density regions. When the data set is an image, the matrix of feature vectors has an $N \times M$ size, where N is the total number of pixels and M represents the M -dimensional feature space. In order to achieve the objective of classifying the orthoimage F into three levels of crop health, a soft K-means

algorithm is formulated and implemented in Python programming language. K-means is one of the most widely used clustering algorithms in pattern recognition and computer vision (Jain, 2010). Prior to applying the algorithm, a neighbourhood-oriented parameter σ is introduced in order to control high-frequency components in the image. For that purpose the product of two Gaussian convolution kernels is applied to each channel $c = r, g, b$ of the image $F = [F_r, F_g, F_b]$:

$$h_\sigma(u, v) = h_\sigma(u) \cdot h_\sigma(v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{2\sigma^2}} \quad (2.1)$$

The operation is described in (2.2), where (α, β) is the pixel position, (u, v) the position of the kernel coefficients and G is the resulting image. There are a number of advantages in using a Gaussian kernel, namely it is isotropic and assigns higher weights to neighbours closer to the central pixel. In addition, the coefficients fall almost to zero at the kernel's edges and being separable leads to a faster computational implementation.

$$G_c(\alpha, \beta) = h_\sigma(u, v) \otimes F_c(\alpha, \beta) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} h_\sigma(u, v) \cdot F_c(\alpha - u, \beta - v) \quad (2.2)$$

In practice, the extent of the kernel edges is defined by the parameter σ , which is the same for each channel $\sigma = \sigma_r = \sigma_g = \sigma_b$ in this study and defines the shape of the Gaussian curve. The influence of pixels located further apart grows with higher values of σ . The new image G can be rearranged into a $N \times M$ matrix of feature vectors. In this case $M = 3$, one column for each channel. To introduce the K-means algorithm, let $X = \{x_1, \dots, x_N\}$ be the data set in an M -dimensional Euclidean space and $\{X_1, \dots, X_K\}$ a partition of X with corresponding K cluster centroids $\{a_1, \dots, a_K\}$. Once the number of clusters K is defined, the purpose of the K-means algorithm is to minimize the objective function φ , a sum of squared within-cluster errors:

$$\varphi = \sum_{i=1}^K \sum_{x_j \in X_i} d(x_j - a_i)^2 = \sum_{i=1}^K \sum_{x_j \in X_i} \|x_j - a_i\|^2 = \sum_{i=1}^K \sum_{j=1}^N w_{ij} \|x_j - a_i\|^2 \quad (2.3)$$

Where $w_{ij} = w_i(x_j) = 1$ if $x_j \in X_i$ and $w_{ij} = w_i(x_j) = 0$ if $x_j \notin X_i$. This optimization problem computationally difficult. However, an iterative descent can solve the problem in most cases by searching minimizers $\{w_1, \dots, w_K\}$ and $\{a_1, \dots, a_K\}$ of φ using the following conditions:

$$a_i = \frac{\sum_{j=1}^N w_{ij} x_j}{\sum_{j=1}^N w_{ij}} \quad \text{and} \quad w_{ij} = w_i(x_j) = \begin{cases} 1, & \text{if } \|x_j - a_i\|^2 = \min_{1 \leq k \leq K} \|x_j - a_k\|^2 \\ 0, & \text{otherwise.} \end{cases} \quad (2.4)$$

The algorithm initializes by randomly assigning a number, from 1 to K , to each observation. These allocations serve as the initial cluster for further iteration. The first step consists of computing the centroid for each of the K clusters (2.4). In the second step each observation is assigned to the cluster whose centroid is closest as measured by the Euclidean distance (2.4). As steps one and two are repeated, φ will decrease the value of the objective function. When the result no longer changes, a global or a local optimum has been reached depending on the initial random assignment of clusters. It is common practice to run the algorithm multiple times with different initial membership assignments and choose the solution with the smallest value of the objective function. As an overview of this formulation, when K-means is applied to F , the algorithm is performing a *hard* assignment of data points to clusters. Whereas applying it K-means to G leads to a *soft* assignment based on a Gaussian convolution kernel. In fact, this *soft* K-means formulation is similar to the Expectation-Maximization (EM) algorithm for Gaussian mixture models (Hastie et al., 2009).

3. RESULTS

In previous sections, a set of techniques to acquire and process high-resolution imagery from a UAV platform have been described. Furthermore, a *soft* K-means clustering has been introduced to perform image analysis. The objective is classifying the area represented in the orthoimage in three crop health levels: healthy, bare soil and transition. For this purpose the algorithm is initialized with $K = 3$ clusters. The results are presented in Figure 3 as a membership map (a) and a decision boundary diagram (b). Each image is complemented by a zoomed-in sample to visualize the result in higher detail.

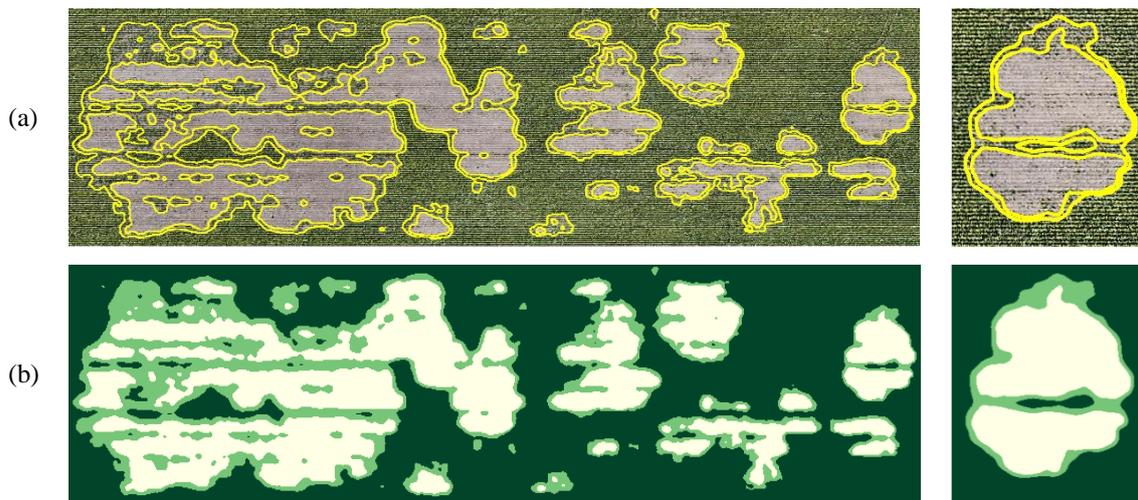


Figure 3. Membership map and decision boundaries after applying a *soft* K-means clustering with $K = 3$ clusters and Gaussian parameter $\sigma = 8$.

The decision boundary diagram (Figure 3a) demonstrates how the algorithm can create accurate clusters for each crop health level. The only user input required to achieve this result is tuning the parameter σ which defines the shape of the Gaussian kernels. In this case a value of $\sigma = 8$ delivers cluster boundaries that properly represent the spatial distribution of each health level. Higher values of σ lead to blob-like shapes, while lower values will tend to classify small gaps in between plants as bare soil. The Gaussian kernels smooth high-frequency components in the orthoimage and providing a measure of plant density based on the spectral signature and texture of groups of pixels. The membership map (Figure 3b) shows each of the clusters in a different color code to visualize the geospatial distribution of each cluster. By knowing the membership of each individual pixel, we can compute an estimate of the area of each crop health level. Thus the area represented in the image can be classified as 3.25 ha of healthy crops, 1.71 ha of decimated crops and 1.13 ha of transition areas. As an observation, the pest movement appears to be highly anisotropic. Some crop rows seem remarkably resilient, given that they're flanked by bare soil on both sides over relatively long distances. This behavior, as well as the location and distribution of transition areas is relevant information in order to design a site-specific control strategy.

4. CONCLUSIONS

This study shows how UAV platforms are evolving into aerial sensor platforms that can be deployed cost-effectively to acquire vegetation data, cover farm scale areas and deliver near real-time assessments. A high-resolution RGB sensor collected imagery over a sorghum crop in South East Queensland, where certain areas were severely damaged by white grub pest. An image processing pipeline has been presented in order to convert individual high-resolution images into a single orthoimage prior to image analysis. In this experiment, the objective of image analysis was classifying the field in three crop health levels: healthy canopy, decimated crop and transition areas with lower plant density. Keeping simplicity and minimal user input requirements in mind, an unsupervised machine learning approach has been selected to approach the challenge. A *soft* K-means clustering algorithm similar to Gaussian Mixture Models and the EM algorithm has been formulated and implemented by combining Gaussian convolution kernels with the K-means algorithm. The advantage of this approach is the introduction of a neighborhood-oriented parameter to control high-frequency components in each band of the image, thus requiring a single parameter to be tuned in order to adjust the smoothness of the cluster edges. Decision boundaries overlaid on the original image demonstrate accurate cluster separations. Membership maps are used to measure the areas of the field assigned to each health level. The methodology presented in this paper is likely to be successful in situations of severe crop damage where bare soil, healthy canopy and transition areas are visible in the image. In many pest management applications it is the transition zones that will be harder to quantify and classify and this is where performance of the algorithm could show the clearest benefits for management. The accurate spatial mapping of pest complexes within fields in a non-invasive way enables the study of spatial population dynamics over time. This approach could have applications in collecting evidence for better understanding the ecology of the beetle species within agricultural

environments. Although this is an important first step in the use of non-intrusive sensing to estimate damage levels in crops, in a more complete study further steps could be taken to validate results. Metrics such as ROC analysis (Swets, 1988) could be used to assess the relative accuracy of classification. In conclusion, this paper demonstrates how UAV-based remote sensing and machine learning techniques can have a major contribution to biosecurity surveillance and pest management. Further research will include testing the accuracy of the algorithm with new data sets and investigating robust approaches to further automate biosecurity assessments with unmanned aerial systems.

ACKNOWLEDGMENT

The authors thank Dr. Jonathan Kok and the ARCAA operations team for their excellent contribution to the data collection campaign.

REFERENCES

- Belgiu, M. and Dragut, L. (2012). Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 96 (2014) 67-75.
- Blaschke, T., Hay, G., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz, R., van der Meer, F., Werff, H., van Collie, F. and Teide, D. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing* 2014, 87, 180-191.
- Buntin, G. D. (2012). Grain Sorghum Insect Pests and Their Management. University of Georgia Cooperative Extension. Bulletin 1283, revised March 2013.
- Candiago S., Remondino, F., De Giglio, M., Dubbini, M. and Gatelli, M. (2015). Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sensing* 2015, 7, 4026-4047.
- Gonzalez, F., Narayan, P., Castro, M. and Zeller, P.G. (2011). Development of an autonomous unmanned aerial system to collect time-stamped sampled from the atmosphere and localize potential pathogen sources. *Journal of Field Robotics*. P961-976.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). Elements of statistical learning. Data mining, inference and prediction. Springer 2009.
- Hay, G., J. and Castilla, G. (2008). Geographic object-based image analysis (GEOBIA): a new name for a new discipline. *Lecture Notes in Geoinformation and Cartography*. Springer, 2008, Pages 75-89.
- Jain, A., K. (2010). Data Clustering: 50 years beyond K-means. *Pattern Recognition Letters*, Volume 31, Issue 8, 2010, pages 651-666.
- Lan Y., Huang, Y, Martin, D.E. and Hoffman, W.C. (2009). Development of an airborne remote sensing system for crop pest management. *American Society of Agricultural and Biological Engineers Vol. 25(4)*: 607-615
- Lehman, J. R., Nieberding, F., Prinz, T., and Knoth, C. (2015). Analysis of unmanned aerial system-based CIR images in forestry – A new perspective to monitor infestation levels. *Forests* 2015, 6, 594-612.
- McFadyen, A. L., Gonzalez, Campbell, D. and Eagling, D. (2014). Evaluating unmanned aircraft systems for deployment in plant biosecurity. Plant Biosecurity Cooperative Research Centre, Bruce, A.C.T.
- Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. *Biosystems Engineering*, 2013. Volume 114, Issue 4, April 2013, Pages 358–371
- Murray, D. A. H., Clarke, M.B. and Ronning, D.A. (2013). The current potential costs of invertebrate pests in Australia. Grains Research and Development Corporation.
- Reuters (April, 2015). ADM could use drones from 2016 to get crop insurance claims data. <http://www.reuters.com/article/2015/04/22/archer-daniels-insurance-dronesidUSL1N0XJ37U20150422>.
- Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 240, 1285–1293.
- Turner, J.W., and Shaw G.D. (1969). Black soil scarab in south Queensland. *Queensland Agricultural Journal* Volume: 95, Issue: 3, Pages: 204-207