THE UNIVERSITY OF
WARWICK

Image analysis and statistical modelling for measurement and quality assessment of ornamental horticulture crops in glasshouses

N. R. Parsons[*,c], R. N. Edmondson[a] and Y. Song[a,b]

[a]Warwick HRI, [b]Computer Science and [c]Statistics Department, University of Warwick, Coventry, CV4 7AL,UK

* Corresponding author. Address for correspondence:

Warwick Medical School, University of Warwick, Coventry CV2 2DX.

Tel.: +44 2476968627; Fax: +44 2476968617; E-mail address: nick.parsons@warwick.ac.uk

Abstract

Image analysis for ornamental crops is discussed with examples from the bedding plant industry. Feed-forward artificial neural networks are used to segment top and side view images of three contrasting species of bedding plants. The segmented images provide objective measurements of leaf and flower cover, colour, uniformity and leaf canopy height. On each imaging occasion, each pack was scored for quality by an assessor panel and it is shown that image analysis can explain 88.5%, 81.7% and 70.4% of the panel quality scores for the three species respectively. Stereoscopy for crop height and uniformity is outlined briefly. The methods discussed here could be used for crop grading at marketing or for monitoring and assessment of growing crops within a glasshouse during all stages of production.

Keywords: Image analysis; Artificial neural networks; Ornamental crops; Crop scanning

1.   Introduction

Monitoring the growth and development of crops in commercial glasshouses requires accurate quantification of a wide range of plant characteristics from plant height and flower cover to indications of the presence of pests and diseases. Ehret *et al.* (2001) reviewed automated monitoring for glasshouse crops and Caponetto *et al.* (2000) discussed the use of sensors to monitor the crop environment. Langton *et al.* (2004) discussed direct measurement of the crop response to environment while Koumpouros *et al.* (2004) discussed image processing for distance diagnosis in pest management.

Modern digital imaging technology is used widely for routine monitoring in industry, and particularly in food production (Brosnan and Sun, 2004) and software methods developed for industrial applications can be readily adapted for grading, on-line sorting and management of individual high value ornamental plants (Brons *et al.*, 1993; Timmermans and Hulzebosch, 1996). Image grading systems are used for individual plants such as pot-plants on moving tracks or belts and machine graders for trays of seedlings where individual seedlings can be clearly distinguished from their neighbours. However, the general problem of assessing an area of tightly packed plants or a batch or pack of plants is much less well developed. Image analysis methods for complex structures such as beds or packs of flowering plants are much more challenging and present many difficulties both in the collection and the processing of the images.

Ehret *et al.* (2001) discuss applications of digital imaging for monitoring glasshouse crops, but do not present any practical results. For many producers, grading and monitoring is tackled almost exclusively by a combination of manual measurement and visual inspection

which is highly labour-intensive and subjective. Computer vision systems and image analysis offer the potential to automate detailed measurement and assessment of plants as crops, not just individual plants and as a first step towards this goal, we develop reliable methods for routinely extracting important attributes from images of multiple ornamental plants in packs. Plant height and height variation is an important quality attribute and we discuss methods for determining leaf canopy heights using single fixed side images. Recent research (Song *et al*. 2007) has shown the feasibility of stereoscopic methods for height measurement using overhead images and future development of stereoscopic methods should allow height information to be collected for tightly spaced crops on benches or growing floors where side-imaging is not feasible.

Commercial protected glasshouse crop environments allow many important environmental factors (e.g. lighting and temperature) to be partly or totally controlled during growth. Currently, most crop husbandry management decisions during production are based on experience, frequent visual inspection of the crop and time-consuming manual methods such as tracking plant heights using hand measurement of pre-selected plants in a growing crop (e.g. poinsettia tracker, Harwood and Hadley, 2004). The aim of this paper is to relate overall plant quality to key plant characteristics measured from digital crop images using automated image analysis methods. The methodology is applied to bedding plants in packs and could be used directly for imaging individual packs on a grading line at marketing. Stereoscopy is not discussed in detail here but we do show how height information from a stereoscopic image can be used to improve the fit of a plant quality model.

2.   Materials and Methods

2.1.   Plants and image acquisition

Bedding plant crops are grown in packs containing typically 12, 24 or 30 plants and must meet tight marketing specification often with a tight marketing schedule. A range of bedding plant species including pansy (Turbo yellow) in spring 2006 and dianthus (Festival Cherry Picotee), viola (Bicolour-Magnifico) and cyclamen (Silverado Scarlett) in autumn 2006 were grown simultaneously at a UK commercial grower and at a large research glasshouse facility located at WHRI (University of Warwick, UK). The various species were typical examples of commercial plant crops selected to represent dense and sparse leaf canopy cover and flower distributions, and also wide height variations (see `www.warwick.ac.uk/go/cropimaging` for example images). All plants were subjected to a range of treatments, reported in detail elsewhere (Edmondson *et al.*, 2007). At the end of production, 72 packs of pansy, 24 packs each of cyclamen and dianthus, and 48 packs of viola (24 packs of 24 small plants and 24 packs of 12 large plants) were selected for imaging and quality assessment. Only the cyclamen, dianthus and viola results are discussed in detail here but the pansy is used to illustrate the application of stereo imaging for bedding plants.

Images were collected both at the end of production and after two weeks of simulated shelf-life. A single fixed image captured from the side of each pack allowed estimates to be made of mean leaf canopy height and flower height. Red-green-blue (RGB) colour digital images were captured, at a resolution of 1136×852 pixels, from all packs routinely under controlled conditions in a purpose-built light-box from a fixed height (0.5 m) and orientation directly above and to one side (0.4 m) of each pack, using a pair of standard commercially available digital cameras.

4

Although three-dimensional representations of objects are used widely in medicine and biology, applications in agriculture are rare (Takizawa *et al.*, 2005), with few developments in commercial horticulture. In situations where it is difficult or impossible to collect side view images (e.g. during plant production), the three-dimensional representation proves to be an important tool for routine plant quality assessment and crop scheduling. A detailed discussion of the recovery of height information from stereo images is beyond the scope of this paper, so will not be discussed further; however a detailed description of the methods used here for determining individual pixels heights is given by Song *et al.* (2007) and Song (2008). The stereo algorithm used here provided a matrix of heights corresponding to the original image file, allowing a map of the canopy surface to be constructed.

2.2. Image segmentation

The first step in analysis was segmentation into multiple regions of interest (sets of pixels), corresponding to different objects or parts of objects. A comprehensive discussion of the principles of image processing is provided by Jain (1989) and applications in the biological sciences are discussed by Glasbey and Horgan (1994). Our main aim was to identify important plant features (e.g. leaves and flowers) from individual images, which was achieved by classifying individual pixels into disjoint groups based on the similarity of the colour characteristics and vertical positioning (height), relative to a fixed base point, of each pixel. A wide range of possible approaches to image segmentation have been described in the research literature (see for example Pal and Pal, 1993 and Cheng *et al.*, 2001) but in general, the selection of an appropriate segmentation algorithm depends largely on the type of image and the application. For our purposes, unsupervised approaches such as k-means and fuzzy c-means (MacQueen, 1967; Bezdek, 1981) were found to be slow and unreliable so a

supervised learning method was developed. This used a training data set, consisting of pairs of input and target outputs, to develop a model that allocated all valid inputs to only one of a small number of output classes.

Artificial neural networks are widely applied to problems in image processing (Egmont-Petersen *et al.*, 2002) and have been found to be particularly useful for classification problems generally in agriculture (Jayas *et al.*, 2000; Cerna and Chytry, 2005). More specifically they have been used both in computer vision applications for classification of potted plants (Brons *at al.*, 1993; Timmermans and Hulzebosch, 1996; van Kaam, 2001) and for complex knowledge acquisition, manipulation and extraction of data for other horticulture applications (Woodford, 2008). Feed-forward neural networks provide a powerful and general framework for representing non-linear relationships between input and output variables; for a detailed description see for instance Bishop (1995). For classification problems involving mutually exclusive classes, the most widely used error function (Ripley, 1994; Bishop, 1995) for training feed-forward neural networks is:

$$E = -\sum_r \sum_k t_{rk} \log_e (y_{rk}) \,.$$

This represents the cross-entropy between the logarithm of the predicted class membership $\log_e y_{rk}$ and the observed indicator of class membership, t_{rk} for K mutually exclusive classes and R training pixels. To avoid over-fitting, it is common to use the sum of squares of the weights C as a penalty on the fit criterion E to give $E + \lambda C$. This penalty is called weight decay and it reduces the 'noisiness' of the fit by reducing the magnitude of weights at each step in the fitting algorithm; the degree and smoothness of fit is controlled by the weight decay term λ.

An unbiased estimate of the misclassification rate is found by using *n*-fold cross-validation (Stone, 1974; Bishop, 1995), with cross-validated misclassification rates (CVMR) determined

by counting zero if the predicted class (class with the largest predicted probability y_{rk}) is correct and one otherwise. A more sensitive method is to use logarithmic scoring, where $-\log_e y_{rk}$ is the score for the true class k and zero is the score for the other classes, giving cross-validated logarithmic scores (CVLS); mean logarithmic scores, across the R training cases, provide a useful measure of separation for discrimination between classes (Titterington *et al.*, 1981).

2.3. Image colour characteristics

Each pixel in an $M \times N$ image can be readily allocated to an appropriate class (e.g. leaf or flower) by a simple visual inspection of the image and, by replicating this process for sets of randomly selected pixels a large training database of pixel characteristics can be created. The colour values for each pixel, in each RGB digital image file, are given by r_{mn}, g_{mn} and b_{mn} where, for the data used here, the red, green and blue colour intensities are measured on an 8-bit scale from 0 to 255. The colour properties of each pixel are best expressed in the normalised colour space using the red R_{mn} and green G_{mn} chromatics and the overall intensity I_{mn} where:

$$R_{mn} = \frac{r_{mn}}{r_{mn} + g_{mn} + b_{mn}}, \; G_{mn} = \frac{g_{mn}}{r_{mn} + g_{mn} + b_{mn}} \text{ and } I_{mn} = \tfrac{1}{3 \times 255}(r_{mn} + g_{mn} + b_{mn}),$$

It follows from the definitions that $0 \leq R_{mn} \leq 1$, $0 \leq G_{mn} \leq 1$ and $0 \leq I_{mn} \leq 1$.

The pixels neighbouring each pixel are also important for determining its classification (see for example Foucher *et al.*, 2004), therefore the local region properties based on the 8-connected neighbourhood region for each pixel were also calculated. The mean red $\overline{R}_{mn}$ and

green $\overline{G}_{mn}$ chromatics and intensity $\bar{I}_{mn}$ within the 8-connected region for each pixel are given by the following expressions:

$$\overline{R}_{mn} = \frac{1}{9} \sum_{\substack{p=m-1,m,m+1 \\ q=n-1,n,n+1}} R_{pq}, \ \ \overline{G}_{mn} = \frac{1}{9} \sum_{\substack{p=m-1,m,m+1 \\ q=n-1,n,n+1}} G_{pq} \ \ \text{and} \ \ \bar{I}_{mn} = \frac{1}{9} \sum_{\substack{p=m-1,m,m+1 \\ q=n-1,n,n+1}} I_{pq},$$

Standard deviations of the red $sd(R_{mn})$ and green $sd(G_{mn})$ chromatics and the intensity $sd(I_{mn})$ within the 8-connected spatial region can be defined, and constrained to lie in the interval $[0,1]$, in an analogous manner to conventional standard deviations for each pixel.

2.4. Image analysis

The extraction of quantitative information from images, such as counts of objects, areas, distances and boundaries, is the aim of much image analysis (Glaseby and Horgan, 1994). Previous work has investigated the relationship between hand measured plant attributes and overall quality for other ornamental plant species (Parsons, 2004; Parsons *et al*., 2006) and has shown that a small number of key factors usually account for a large proportion of the variability in recorded quality assessments. The four most important of these characteristics are defined here and methods are described for their estimation from segmented images;

(i)   *Leaf and flower cover*. The percentage leaf and flower cover within an image were estimated by calculating the percentage class membership based on counts of the number of pixels in each class in the segmented image.

(ii)   *Leaf and flower cover uniformity*. The variability in leaf and flower cover across each pack image was used to estimate uniformity. The image was sub-divided into regions covering the individual cells of a pack and the coefficient of variation of the percentage cover across these regions provided an estimate of pack uniformity.

8

(iii) *Leaf canopy height and shape*. Leaf canopy heights were estimated from side view images by sub-dividing each image into a grid which, for our images were 24 x 24 pixels representing a square of area 1 cm^2, and then determining the proportion of leaf material within each grid square. The centre of the lowermost (and uppermost) square in each column within the grid designated as leaf material, that is a square containing $\geq 50\%$ leaf material, directly below (above) a square designated as non-leaf material ($< 50\%$ leaf) was used as an estimate of the lower (upper) extent of the canopy. Given upper and lower limits within each column, a quadratic curve was fitted across the width of the image grid and estimates for the mean height and linear and quadratic coefficients were determined using regression analysis (Sen and Srivastava, 1997).

(iv) *Leaf and flower colour.* The leaf greenness index was estimated as the normalized colour characteristic given by the mean value of $100 \times G_{mn}$ (Section 2.3) for all pixels identified as belonging to class leaf. Similar definitions, lead to estimates for flower blue and red normalized colour characteristics being given by $100 \times B_{mn}$ and $100 \times R_{mn}$.

Nine characteristics, denoted as z_{rk} where $r = 1,2 \ldots 9$, were estimated for each of the k packs for each species; (i) leaf cover (%), (ii) flower cover (%), (iii) leaf uniformity (%), (iv) flower uniformity (%), (v) leaf canopy height (cm), (vi) leaf canopy height squared (cm^2), (vii) leaf greenness (%), (viii) flower blueness (%) and (ix) flower redness (%).

2.5. Plant quality scores

An assessment panel comprising 20 members of the WHRI (University of Warwick) staff were randomly selected from individuals with an interest in horticulture. Assessors from both a horticultural (specialist) and non-horticultural (non-specialist) background were used. Packs

were scored for quality by each member of the panel at two separate occasions, two weeks apart. On each occasion, two assessments were made on consecutive days using a different randomized ordering of the packs on each assessment day. Packs were assessed visually and each panel member gave a score indicating their own independent assessment of the quality of the plant material using a scale where 1 = very poor, 2 = poor, 3 = average, 4 = good and 5= very good quality. The ability of individual panel members to repeat their quality scores was assessed by calculating the Spearman rank correlation coefficient after the randomized change in pack number. The mean quality score for each pack was calculated as the weighted mean of the individual scores from the 20 panel members, using the individual correlations as weights (all correlations were positive). This resulted in a mean score $\bar{q}_k$ for each of the K packs for each species that gave little or no weight to those panel members who were unable to consistently replicate their scores on the same material at each occasion.

2.6. Statistical models

The effect of the pack characteristics, described in Section 2.4, on the pack mean quality score (Section 2.5) for each species were investigated using the multiple regression model

$$E(\bar{q}_k) = \beta_0 + \sum_{r=1}^{9} \beta_r z_{rk} \ ,$$

where $E(\bar{q}_k)$ denotes the expected quality score for pack k, and β_0 and β_r are regression coefficients that are estimated from the data. Assuming a Normal distribution for the response variable $\bar{q}_k$, the regression coefficients were estimated using generalized estimating equations (Hardin and Hilbe, 2002), that accommodated a first order autoregressive correlation model for the repeat scores on the same packs made at two occasions separated by two weeks. Model fitting proceeded using the open source statistical software R (R Development Core

Team, 2007) and the `gee` procedure library. The significance of individual regression coefficients was assessed using Z-statistics based on the robust variance estimator, the goodness of fit for each species model was assessed using an adjusted R^2_{adj} value and diagnostic plots were used to check the model assumptions.

3. Results

3.1. Image processing

Fifty pixels were randomly selected from each of the 48 available cyclamen and dianthus side images, and visually classified into one of three mutually exclusive classes (flower, leaf and other) using the *Image Processing Toolbox* (version 5.1) available in Matlab (The MathWorks, 2007). Similarly, 25 pixels were randomly selected from each of the 96 available cyclamen side images and visually classified. The number of pixels selected within each image was based largely on pragmatic considerations and for the most complex networks the ratio of data records to network weights was never less than 30:1; the overall performance of the networks were not unduly sensitive to the number of selected training pixels. The nine colour characteristics, described in Section 2.4, were calculated for each of the sampled pixels and neural networks, with 10 input layer nodes and 3 output layer nodes, were fitted using function libraries available in the open source statistical package R (R Development Core Team, 2007). All data, model fitting code and example processed side and top view images for dianthus, cyclamen and viola are available from the authors at
`www.warwick.ac.uk/go/cropimaging.`

3.2. Stereoscopy

The image stereoscopy work is reported elsewhere (Song *et al.* 2007 and Song 2008) and is not discussed in detail here. However, Figure 1 illustrates the use of stereoscopy to generate a false colour relative surface height map for a pansy pack in centimetres. The false colour contour map accurately represents the relative height of the flowers and leaves in the image and can be used for capturing information about the uniformity and habit of a crop. With proper calibration, the algorithm can also be used to calculate absolute crop height for crop growth monitoring (Harwood and Hadley, 2004) but this application is beyond the scope of this article and will not be discussed further here.

3.3. Segmentation

CVLS, based on 10 replications of ten-fold leave-one-pixel-out cross-validation, were not unduly sensitive to the weight decay parameter used for values < 0.1, so the largest of the tested values (0.01) was selected to give the smoothest fit to the data without loss of predictive power. There was little substantive improvement in fit for more than 3 hidden layer nodes for any of the species used here, so the best performing network for segmenting side images consisted of 10 input nodes, 3 hidden layer nodes and 3 output nodes; 30 weights connecting the input nodes to the three hidden layer nodes, 9 weights connecting the hidden layer nodes to the outputs and three weights connecting the bias node to each output – 42 weights in total. No attempt is made here to fully optimize the networks for robust performance, for instance by pruning redundant nodes and weights; this would be required for practical implementation in a commercial system, but is not discussed here for reasons of brevity. The observed mean CVLS for dianthus, viola and cyclamen (0.070, 0.175 and 0.036) indicated that on average each pixel was classified correctly with a probability of 0.932, 0.840 and 0.965; considerably better than would be achieved by purely random allocation of pixels

to the three designated classes (0.333). The misclassification tables (cross-tabulation of true versus predicted classes) for these networks gave no indications that the networks performed differentially better at classifying pixels into any of the individual classes and the final trained networks for dianthus, viola and cyclamen gave good segmentation of images into flowers and leaves, misclassifying 2.25%, 6.83% and 1.38% of pixels respectively.

Segmentation for the dianthus crop was challenging due to lack of visual distinction between white flower parts and the white polystyrene pack and for this example, height information was obtained using stereoscopic methods. Data were collected from top view images using random pixel selection in exactly the manner and number described for the side view images and visually classified into one of four mutually exclusive classes (flower, leaf, pack and compost). In addition to the nine colour characteristics, described in Section 2.4, pairs of stereo images for dianthus were used to calculate heights at each point in the image. Figure 2 shows measured heights from single fixed side view images for individual packs of dianthus plants and mean pack heights measured from matched stereo pairs of the corresponding packs. The plot error bars indicate the pixel to pixel height variability for each pack image, showing how fixed side view heights simplify the true within and between plant height variability within a crop. The individual estimated absolute and relative pixel heights, determined by calibration based on either the distance between camera and pack or on the maximum height recorded within a pack respectively were also included as inputs to the neural networks, as a means of discriminating between pixels located on flowers and packs possessing similar colour characteristics, but located at differing heights within the crop canopy.

13

As for the side view image networks, a weight decay parameter of 0.01 was used for the ANN for the top view images that comprised 10 input nodes and 4 output nodes for cyclamen and viola, and 12 input nodes (2 additional nodes for absolute and relative crop heights) and 4 output nodes for dianthus. The observed mean CVLS for dianthus, viola and cyclamen (0.103, 0.109 and 0.125) indicated that on average each pixel was classified correctly with a probability of 0.902, 0.897 and 0.882. The trained networks for dianthus, viola and cyclamen gave good segmentation of images into flowers and leaves, misclassifying 4.08%, 3.62% and 4.12% of pixels respectively. Misclassification rates were generally not as good for the top view images as for the side view images, because of the greater depth and complexity of the images, due mainly to the shading of underlying material from leaves and flowers further up the plant canopy. The results of the segmentation for dianthus are illustrated in the right-hand pane of Figure 3. The flowers and the polystyrene pack are sufficiently well distinguished to make good estimates of class cover; leaf = 66.8%, flower = 19.4%, pack = 5.1% and compost = 8.7%.

3.4. Plant quality models

The quality scores from the panel of 20 assessors were weighted and averaged, using the methods described in Section 2.5, to give a mean score $\bar{q}_k$ for each of the K packs for each species. Models of the form described in Section 2.6 were fitted to data for each species using the model selection methods described by Parsons *et al*. (2006). Estimated regression parameters and standard errors are shown for each species in Table 1. Diagnostic plots of model residuals showed no significant departures from the model assumptions.

The parameter estimates in Table 1 describe the affect of each attribute on overall plant quality. In general, increasing leaf and flower cover increased assessments of pack quality and the significant negative coefficients for leaf greenness indicated that quality assessments improved for packs that had darker green coloured leaves than those packs that had paler leaves. For dianthus and cyclamen, flower colour was also important, with coefficients indicating a preference amongst the assessment panel for darker red coloured flowers. A more even distribution of flowers across the visible pack surface improved quality for cyclamen, whereas for viola the opposite effect was observed and quality improved when flower uniformity indicated that flower cover was more variable across the visible pack surface although the reason is unclear. Leaf canopy height showed a clear maximum in quality at approximately 11cm for dianthus, but for the range of pack heights available for viola quality increased as pack heights decreased. Models accounted for 88.5%, 81.7% and 70.4% of the variation in quality scores for dianthus, viola and cyclamen respectively. The relationship between observed and predicted quality scores for each pack, based on the fitted models, is shown in Figure 4.

4.   Discussion

Artificial neural networks based on colour characteristics from individual pixels and surrounding pixels together with information on height obtained from stereo pairs of images, performed well in segmenting all the images used here. Where segmentation based on pixel data alone is poor, information about the typical shape of objects or knowledge of the spatial arrangement and textures of objects could also be incorporated into the networks (Egmont-Petersen *et al.*, 2002). Random pixel selection worked well for the well-populated classes

discussed here but for detecting rarer classes such as disease systems, a more structured sampling procedure could be appropriate.

Our image segmentation work has been based mainly on ANNs but an alternative method can be based on support vector machines (Vapnik, 2000; Karatzoglou *et al.*, 2006), a supervised machine learning method that is becoming increasing popular (Karimi *et al.*, 2006). We have little direct experience of support vector machines but the literature suggests that support vector machines are at least as powerful as ANN methods therefore a comparison of the two methods for classification of plant quality could be of value.

Our models accounted for 88.5%, 81.7% and 70.4% of the variation in quality scores for dianthus, viola and cyclamen respectively but it should be remembered that the panel quality scores were themselves estimates and were subject to considerable variation. Our view is that the image analysis scores are in fact more objective and more reliable than panel scores and the high level of correlation between the two sets of scores confirms that image analysis methods can routinely capture the same information that a human assessor can reliably capture. For accurate estimates of heights and more detailed descriptions of the canopy shape and variability of a growing crop in a commercial setting, three-dimensional images can be captured from cameras directly above the crop. We have shown that this can be achieved using a single camera collecting images at two fixed positions, but the similar results could be obtained using a pair of cameras or from a single moving camera collecting a series of images over a growing crop. These are currently areas of active research.

Although we have concentrated on image analysis for grading and crop monitoring, we would also expect, more generally, that digital imaging would be capable of detecting localised pest

and disease or other types of plant damage. The imaging sensors could also be easily combined with additional sensors to monitor automatically for stress using infra-red temperature (Langton *et al.*, 2004), chlorophyll fluorescence (Van Kooten *et al.*, 1991; DeEll *et al.*, 1999), spectral information (Zwiggelaar, 1998) and potentially other sensors at the same time that the crop was being visually imaged.

5.   Conclusion

The development of methods for image processing and statistical modelling of quality scores for routine quality control of ornamental bedding plant crops has been reported. Artificial neural networks have been used to segment digital images into regions of interest and key morphological and chromatic characteristics of the crops have been incorporated into statistical models that explain much of the variation in overall quality scores made by a panel of assessors. The morphological and chromatic characteristics for each ornamental plant pack provide a rich source of information that have the potential to assist in making complex decisions regarding the scheduling and quality management of a growing plant crop. In the near future, we would expect machine vision and intelligent image analysis software to become of key importance in the monitoring and control of growing crops especially in protected crop environments in glasshouses.

Acknowledgements

imaging for quality assessment' and PhD studentship CP37 on 'Decision support for glasshouse crop production using digital imaging and artificial neural networks'.

References

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford.

Brons, A., Rabatel, G., Ros, F., Sevila, F., Touzet, C., 1993. Plant grading by vision using neural networks and statistics. Computers and Electronics in Agriculture 9(1), 25-39.

Brosnan, T., Sun, D.W., (2004). Improving quality inspection of food products by computer vision — a review. Journal of Food Engineering 61(1), 3-16.

Caponetto, R., Fortuna, L., Nunnari, G, Occhipinti, L. Xibilia, M.G., 2000. Soft computing for greenhouse climate control. IEEE Transactions on Fuzzy Systems 8(6), 753-760.

Cerna, L., Chytry, M., 2005. Supervised classification of plant communities with artificial neural networks. Journal of Vegetation Science 16(4), 407-414.

Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J., 2001. Color image segmentation: advances and prospects. Pattern Recognition 34(12), 2259-2281.

DeEll, J., van Kooten, O., Prange, R., Murr, D., 1999. Applications of chlorophyll fluorescence techniques in postharvest physiology. Horticultural Reviews 23(1), 69-107.

Edmondson, R.N., Parsons, N., Adams, S., Song, Y., 2007. The measurement and improvement of robust bedding plant quality and the use of digital imaging for quality assessment: So you think you know plant quality? HDC News, 139(1), 20-21.

Egmont-Petersen, M., de Ridder, D., Handels, H., 2002. Image processing with neural networks – a review. Pattern Recognition 35(10), 2279-2301.

Ehret, D.L., Lau, A., Bittman, S., Lin, W., Shelford, T., 2001. Automated monitoring of greenhouse crops. Agronomie 24(4), 403-414.

Foucher, P., Revollon, P., Vigouroux, Chassériaux, B., 2004. Morphological image analysis for the detection of water stress in potted forsythia. Biosystems Engineering 89(2), 131-138.

Glasbey, C.A., Horgan, G.W., 1994. Image Analysis for the Biological Sciences. John Wiley & Sons, Chichester.

Hardin, J.W., Hilbe, J.M., 2002. Generalized estimating equations. Chapman and Hall, New York.

Harwood, T.D., Hadley, P., 2004. Graphical tracking systems revisited: a practical approach to computer scheduling in horticulture. Acta Horticulturae 654(1), 179-186.

Jain, A.K., 1989. Fundamentals of Digital Image Processing. Prentice-Hall International, New Jersey.

Jayas, D.S., Paliwal, J., Visen, N.S., 2000. Multi-layer neural networks for image analysis of agricultural products. Journal of Agricultural Engineering Research 77(2), 119-128.

Karatzoglou, A., Meyer, D., Hornik, K., 2006. Support vector machines in R. Journal of Statistical Software 15(9).

Karimi, Y., Prasher, S.O., Patel, R.M., Kim, S.H., 2006. Application of support vector machine technology for weed and nitrogen stress detection in corn. Computers and Electronics in Agriculture 51(2), 99-109.

Koumpouros, Y., Mahaman, B.D., Maliappis, M., Passam, H.C., Sideridis, A.B., Zorkadis, V., 2004. Image processing for distance diagnosis in pest management. Computers and Electronics in Agriculture 44(2), 121-131.

Langton, F.A., Horridge, J.S., Holdsworth, M.D., Hamer, P.J.C., 2004. Control and optimization of the greenhouse environment using infra-red sensors. Acta Horticulturae 633(1), 145-152.

MacQueen, J.B., 1967. Some Methods for classification and analysis of multivariate observations. In: Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281-297.

Pal, N.R., Pal, S.K., 1993. A review of image segmentation techniques. Pattern Recognition 26(9), 1277-1294.

Parsons, N.R., Edmondson, R.N., Gilmour, S.G., 2006. A generalized estimating equation method for fitting autocorrelated ordinal score data with an application in horticultural research. Journal of the Royal Statistical Society, Series C 55(4), 507-524.

Parsons, N.R., 2004. Statistical methods for improving pot-plant quality and robustness. Unpublished PhD thesis. University of London, UK.

R Development Core Team 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ripley, B.D., 1994. Neural networks and related methods for classification. Journal of the Royal Statistical Society, Series B 56(3), 409-456.

Sen, A., Srivastava, M., 1997. Regression analysis: Theory, Methods and Applications. Springer-Verlag, New York.

Song, Y., Wilson, R.G., Edmondson, R.N, Parsons, N.R., 2007. Surface modelling of plants from stereo images. In: Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM 2007), pp. 312-319.

Song, Y. 2008. Modelling and analysis of plant image data for crop growth and monitoring in horticulture. Unpublished PhD thesis. University of Warwick, UK.

Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society, Series B 36(2), 111-147.

Takizawa, H., Ezaki, N., Mizuno, S., Yamamoto, S., 2005. Measurement of plants by stereo vision for agricultural applications. In: Proceedings of the Seventh IASTED International Conference, Signal and Image Processing, Hawaii. p238.

The MathWorks (2007). Matlab. http://www.mathworks.com/.

Timmermans, A.J.M., Hulzebosch, A.A., 1996. Computer vision system for on-line sorting of pot plants using an artificial neural network classifier. Computers and Electronics in Agriculture 15(1), 41-55.

Titterington, D.M., Murray, G.D., Murray, L.S., Spiegelhalter, D.J., Skene, A.M., Habbema, J.D.F., Gelpke, G.J., 1981. Comparison of discrimination techniques applied to a complex data set of head injured patients. Journal of the Royal Statistical Society, Series A 144(2), 145-175.

Van Kaam, C.J.H.M., 2001. Neural networks used for classification of potted plants. Acta Horticulturae 562(1), 109-115.

Van Kooten, O., Mensink, M., Otma, E, van Doorn, W., 1991. Determination of the physiological state of potted plants and cut flowers by modulated chlorophyll fluorescence. Acta Horticulturae 298(1), 83-91.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Springer, New York.

Woodford, B.J., 2008. Evolving neurocomputing systems for horticulture applications. Applied Soft Computing 8(1), 564-578.

Zwiggelaar, R., 1998. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. Crop Protection 17(3), 189-206.

Figure 1. A false colour image, showing crop canopy height (cm) for an individual pack of bedding plants, and the original pack image with the area selected for height determination.

Figure 2. Measured heights from single fixed side view images of individual packs of plants and mean heights measured from matched stereo pairs of the corresponding packs for dianthus. Bars represent one standard deviation.

Figure 3. Images showing dianthus plants at marketing (left) and segmented into four classes (right); red pixels represent flowers, green pixels leaves, white pixels the polystyrene pack and black pixels the growth medium.

Figure 4. Observed and predicted quality scores, based on the fitted statistical models, for each species.

Table 1. Parameter estimates, standard errors (s.e.) and Z statistics for crop quality models for each species

Parameter	Species								
	Dianthus			Viola			Cyclamen		
	Est.	s.e.	Z	Est.	s.e.	Z	Est.	s.e.	Z
Constant	0.224	1.205	0.19	4.921	3.095	1.59	-2.408	2.440	-0.99
Leaf cover (%)	0.032	0.007	4.54	0.129	0.031	4.12	0.057	0.025	2.25
Flower cover (%)	0.015	0.006	2.43	0.091	0.030	3.03	0.061	0.024	2.55
Leaf uniformity (%)	-	-	-	-	-	-	-	-	-
Flower uniformity (%)	-	-	-	-0.007	0.002	-3.91	0.002	0.001	2.86
Leaf height (cm)	0.345	0.060	5.76	-0.191	0.026	-7.43	-	-	-
Leaf height^2 (cm^2)	-0.016	0.004	-4.52	-	-	-	-	-	-
Leaf greenness (%)	-0.091	0.024	-3.77	-0.246	0.021	-11.59	-0.078	0.034	-2.28
Flower blueness (%)	-	-	-	-	-	-	-	-	-
Flower redness (%)	0.083	0.011	7.19	-	-	-	0.051	0.009	5.79