DETECTION OF FRUIT IN CANOPY NIGHT-TIME IMAGES: TWO CASE STUDIES WITH APPLE AND MANGO

O. Cohen, and R. Linker

Faculty of Civil and Environmental Engineering, Technion
Israel

A. Payne, and K. Walsh

School of Medical and Applied Sciences, Central Queensland University
Queensland, Australia

ABSTRACT

Reliable estimation of the expected yield remains a major challenge in orchards. In a recent work we reported the development of an algorithm for estimating the number of fruits in images of apple trees acquired in natural daylight conditions. In the present work we tested this approach with nighttime images of similar apple trees and further adapted this approach to nighttime images of mango trees.

Working with the apple images required only minimal re-parameterization of the algorithm and did not require changes of the algorithm itself. Twenty images were used to calibrate the algorithm, and after re-parameterization the number of objects detected by the algorithm corresponded to 79.9% of the number of apples visible in the images. The procedure was tested with 144 images containing close to 7000 apples and the number of objects detected by the algorithm corresponded to 80.2% of the number of apples visible in the images.

The analysis of the mango images was more challenging and some modifications of the algorithm were required, mostly to handle the elliptical shape of mangoes and avoid false positive detection on trunks. Twenty images were used to perform these changes and thus calibrate the procedure, which was then tested on three datasets containing 164 images with over 6000 fruit. The number of objects detected by the algorithm in the calibration set corresponded to 72% of the number of fruit counted by visual inspection. In the validation sets the number of objects detected ranged from 66% to 69% of the number of fruit identified by visual inspection.

Considering the differences between the two datasets in terms of shape and size of the fruit and leaves, these results are very encouraging and show that, after site-specific calibration using a small number of images, the proposed approach could provide good yield estimates.

Keywords: Yield estimation, artificial vision, imaging processing

INTRODUCTION

Estimation of the expected yield is still a major challenge in orchards. To date, yield forecasts are based on visual inspection of some trees by the farmer. This method is extremely time-consuming and, due to the high variability that exists in orchards, the result is far from reliable. Since the early days of artificial vision systems, numerous studies have been devoted to the use of this approach for fruit localization in tree canopies, not only for yield estimation but also for robotic harvesting (e.g. Jimenez et al., 2000; Bulanon et al., 2001; Annamalai and Lee, 2003; Zhao et al., 2005; Moorinta et al., 2010; Kapach et al., 2012; Payne et al., 2013, 2014). Linker et al. (2012) recently reported a first step toward the development of a vision system for yield estimation in apple orchards. Although the results were encouraging, this study, which was conducted with daylight images, underscored the dependence of the final results on illumination, which is very difficult to standardize in outdoor daytime applications. As an alternative, the present study focused on the use of nighttime images acquired with artificial lighting. Furthermore, whereas the study of Linker et al. (2012) was limited to apples, the present study considered both apple and mango trees.

DATASETS

Apple dataset

Images were recorded in a Golden Delicious orchard in the Matityahu Research Station located in Northern Israel. In order to create a dataset that included trees with both high and low fruit loads, images were acquired in two areas of orchards known to produce very different yields. The planting pattern in both area was 2.0m by 4.5m. Altogether, the images of 26 trees were taken during a single night in August 2013, when the apple diameter was approximately 6cm. Each side of the tree was photographed using three identical cameras (Canon PowerShot G7) mounted on a vertical pole at approximately 110cm, 180cm and 240cm from the ground. The pole was mounted on a trailer which was pulled by an electric cart. Two additional poles were mounted on the cart at 60cm from the camera poles and LED lights were mounted on each pole at heights 130cm and 250cm from the ground. The cart was driven in the middle of the row so that the distance between the cameras and the tree trunk was approximately 360cm. The cameras were operated in full automatic mode but without flash. The images were not taken "on-the-fly" but the cart stopped in front of each tree to prevent image blurring.

Twelve images (six images from each area) were selected randomly as calibration images which were used to tune the parameters of the image processing algorithm. In order to determine the performance of the automatic procedure, the number of fruit in each image was determined by visual inspection.

In addition, the actual yield of each tree was determined by individual picking at the end of the season.

Mango dataset

Mango images were collected in a tropical Australian orchard at 'stone hardening' stage at night under artificial lighting. The images were divided into four sets (20, 80, 10 and 74 images respectively) collected over two evenings in October 2011. All images were recorded with the same camera (Canon 50D SLR) and with the same illumination platform that consisted of four 6*3 W LED spotlights (SCA P/L). The trees in this orchard were relatively small and widely spaced so that each image included a full view of a single tree. The fruit in these images did not yet have significant or consistent red coloration (or 'blush'), and included a high number of split, commercially unviable fruit, making the yield estimation process more difficult. The fruit in all images were counted by visual inspection. Additional details about the dataset can be found in Payne et al. (2014). The first dataset (20 images) was used to calibrate the image processing algorithm.

IMAGE ANALYSIS OUTLINE

The image processing algorithm has been described in Linker et al. (2012). Basically, it includes four main stages: Pixel classification based on color and smoothness; Detection of blobs of connected pixels (seed areas) and extension of these blobs to neighboring pixels with similar properties; Segmentation of the contours of the blobs into arcs; Grouping of arcs into objects and comparison with a reference object. Since this procedure was originally developed for detection of apples in daylight images, its use for detection of apples in nighttime images required only minimal re-parameterization of the procedure and the algorithm itself did not need to be modified. The parameters which were updated are listed in Appendix 1.

The analysis of the mango images was more challenging on three accounts: (1) the fruit was elliptical rather than spherical, (2) the fruit color was not uniform and ranged from reddish to green, and (3) there were a large number of very small branches that could cause erroneous segmentation of the fruit. In order to deal with the elliptic shape of the mango it was necessary to modify the function that quantifies the discrepancy between the object built from the arcs and the reference object. The function originally used by Linker et al. (2012) contained four terms:

$$\Delta T = \sqrt{\left(\frac{\Delta R}{\delta_{50}^{R}}\right)^2 + \left(\frac{\Delta V}{\delta_{50}^{V}}\right)^2 + \frac{1}{\delta_{50}^{r}}\left(\sum_{i=1}^{n}\frac{(\theta_{2i}-\theta_{1i})\Delta r_i}{P_V}\right)^2 + \frac{1}{\delta_{50}^{a}}\left(\sum_{i=1}^{n}\frac{(\theta_{2i}-\theta_{1i})\Delta a_i}{P_V}\right)^2}$$

The first term is a measure of the difference between the radius of the detected object and the radius of the model apple; the second term is a measure of the difference between the length of the detected contour and the contour of a fully visible apple; and the third and fourth terms are a measure of the radial and angular deviation of each arc, respectively. In order to handle elliptical objects, a

fifth term which is a measure of the difference between the aspect ratio of the object and the aspect ratio of the reference model, was added:

$$\Delta T = \sqrt{\left(\frac{\Delta R}{\delta_{50}^{R}}\right)^2 + \left(\frac{\Delta V}{\delta_{50}^{V}}\right)^2 + \frac{1}{\delta_{50}^{r}}\left(\sum_{i=1}^{n}\frac{(\theta_{2i}-\theta_{1i})\Delta r_i}{P_V}\right)^2 + \frac{1}{\delta_{50}^{a}}\left(\sum_{i=1}^{n}\frac{(\theta_{2i}-\theta_{1i})\Delta a_i}{P_V}\right)^2 + \left(\frac{\Delta A}{\delta_{50}^{A}}\right)^2}$$

where ΔA is the difference between the actual and model aspect ratio. Also, the average of the lengths of the ellipse axes was used instead of the circle radius R in the first term. The other issues were solved by adjusting some of the parameters of the algorithm (values listed in Appendix 2). However, the preliminary results showed a relatively large number of "false positive" detections on tree trunks, which under the artificial illumination had a grey-pink color which was also typical of some mangoes. In order to reduce the number of these "false positive" detections, trunk areas were labeled in eight of the calibration images and an additional color classifier was built to detect trunk color. For each blob of pixels originally recognized as mango, the following index was calculated:

$$\alpha = \frac{1}{n}\sum_{i=1}^{n}\left(p_i^{mango} - p_i^{trunk}\right)$$

where n is the number of pixels in the blob, p_i^{mango} is the probability of pixel i to belong to a fruit according to the first color classifier, and p_i^{trunk} is the probability of pixel i to belong to a trunk according to the second color classifier. A blob was retained as mango only if this index was positive, meaning that its probability of being a fruit was larger than its probability of being a trunk.

RESULTS

Apple

Two typical calibration images are shown in Figures 1 and 2. Although perfectly homogeneous illumination can not be achieved within complex 3D tree canopies, these images are much more balanced than images acquired on daylight conditions.

Calibration of the image processing algorithm yielded the relationship shown in Figure 3. This Figure shows that the number of objects detected corresponded to ~80% of the fruit identified by visual inspection of the images (478 objects detected compared to 598 apples). In other words, according to the calibration results an estimation of the actual number of fruit in an image can be obtained by multiplying the number of identified objects by a correcting factor equal to 1.25.

The calibrated procedure (including the correction factor) was applied to the remaining 144 images and the results are shown in Figure 4 and Table 1. Regardless of the area of the tree covered by the image (camera height), there is a very good agreement between the number of apples estimated by the automatic procedure and the number of apples counted by visual inspection (overall estimate error less than 1% of actual number of apples).

Fig. 1. Typical image of tree with high fruit load (middle camera)

Fig. 2. Typical image of tree with low fruit load (middle camera)

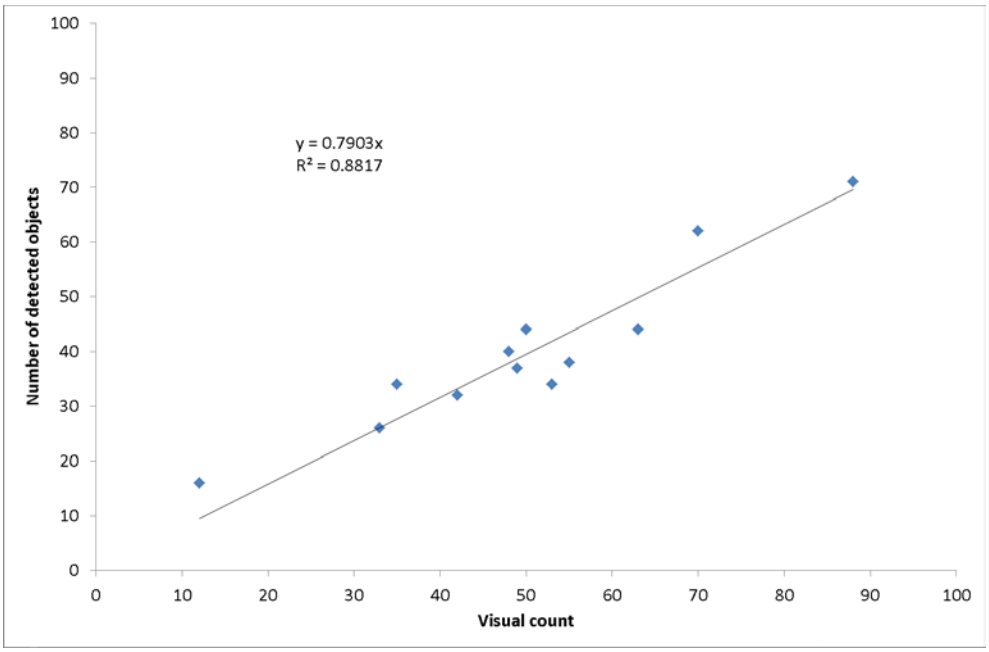

Fig. 3. Calibration: Relationship between the number of objects detected by the automatic procedure and the number of apples identified by visual inspection of the 20 calibration images

Since for practical applications the information of interest is the actual yield per tree rather than the number of fruit visible in an image, each tree was picked individually at the end of the season and an attempt was made to correlate the yield with the combined results of the six images covering the tree (three from each side). The result images (in which the detected objects are indicated by red circles) were registered manually and the number of detected objects in this mosaic image was counted manually (multiple detections at the same location in different images were counted only once). The registration procedure is illustrated in Figure 5. In six cases the image registration was too poor to provide reliable results (the irregular 3D shape of the canopy caused too great a misalignment) and the corresponding trees were excluded from the analysis. The results for the remaining 20 trees are shown in Figure 6. Although these results are based on a small dataset and must be considered with caution, they are encouraging in the sense that they show a clear linear relationship between the actual yield per tree and the estimated tree load (sum of objects detected in the two mosaic images).

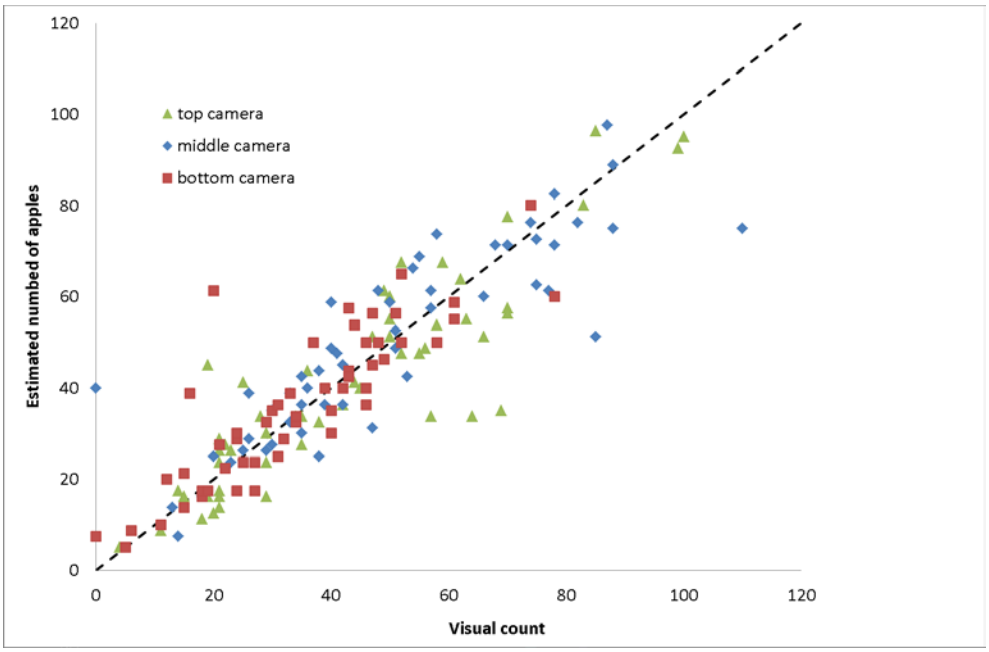

Fig. 4. Validation: Relationship between the number of apples identified by visual inspection and the number of apples estimated by the automatic procedure of the 144 images.

Table 1. Summary of the results for the apple validation images.

	Camera position			Total
	Top	Middle	Bottom	
Visual count	2276	2605	1832	6713
Automated estimate	2187	2632	1924	6743
Ratio	0.961	1.010	1.050	1.004

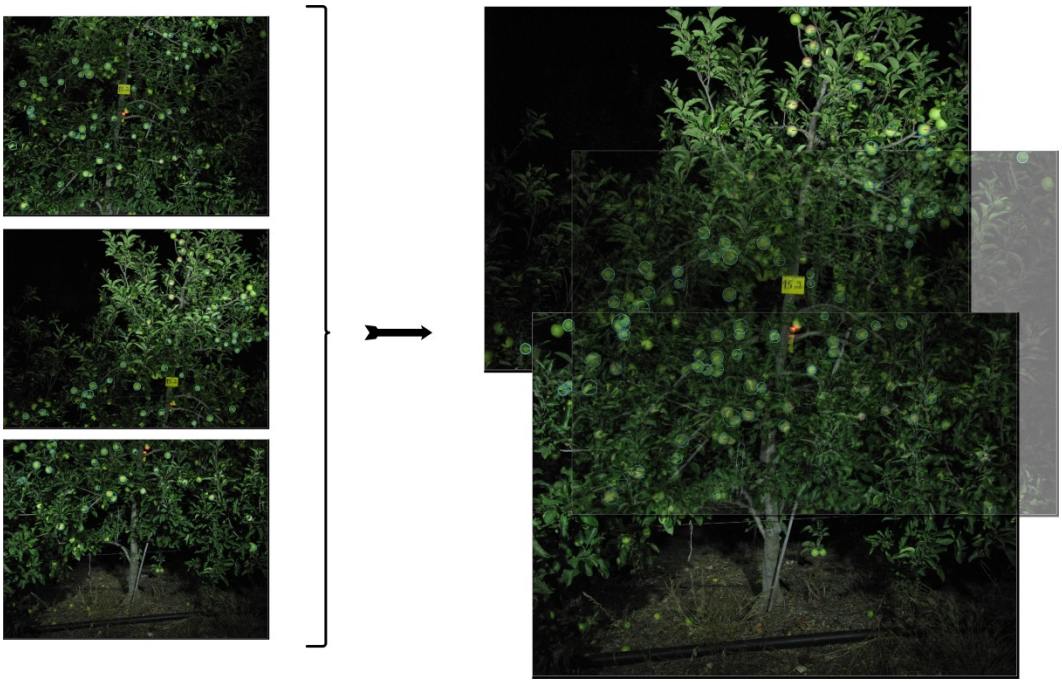

Fig. 5. Tree reconstruction from the images acquired by the three cameras. The original images are shown on the left and the mosaic image obtained by manual registration of the result images is shown on the right.

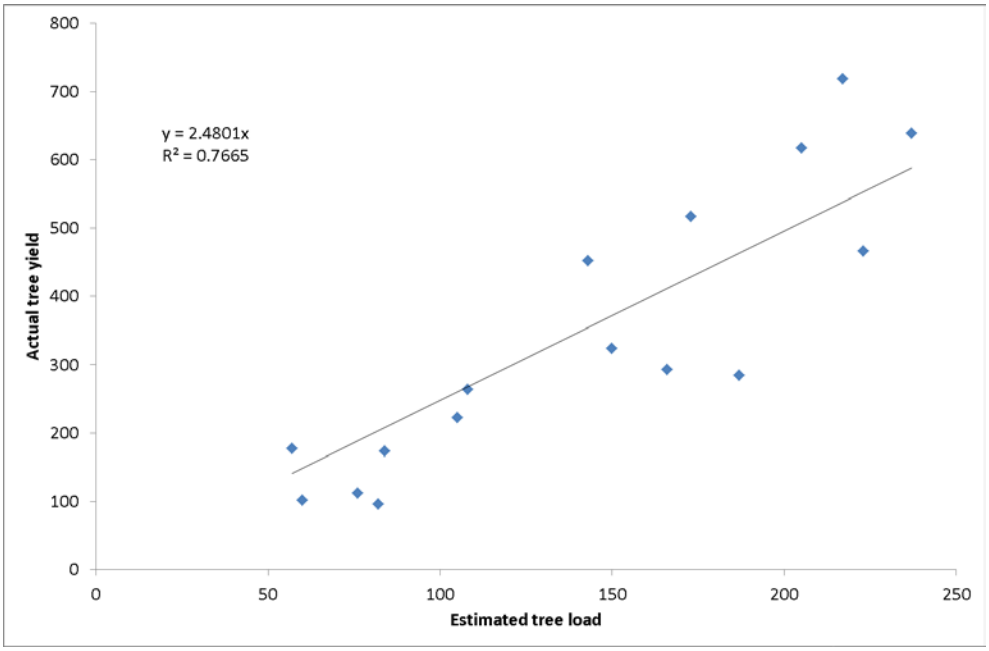

Fig. 6. Relationship between the actual tree yield and the number of objects detected in the two mosaic images of each tree.

Mango

A typical calibration image is shown in Figure 7. Comparison with Figures 1 and 2 clearly shows that the mangoes color is much less uniform than that of apples. On the other hand, mangoes are larger and tend to be less grouped in clusters than apples. Also, there are fewer occlusions by leaves but inflorescence stalks often cause initial segmentation of the fruit into several objects.

The calibration stage yielded the relationship shown in Figure 8. This result was poorer than the one achieved for the apple images, both in terms of R^2 and in terms of the ratio between the number of detected objects and the actual number of fruit (507 detected objects vs. 704 counted mangoes). Nonetheless, as demonstrated by Figure 9 and Table 2, applying this calibrated procedure to the validation sets led to a very accurate estimation of the number of fruits in all three sets of images (estimate within 10% of visual count).

Fig. 7. Typical image.

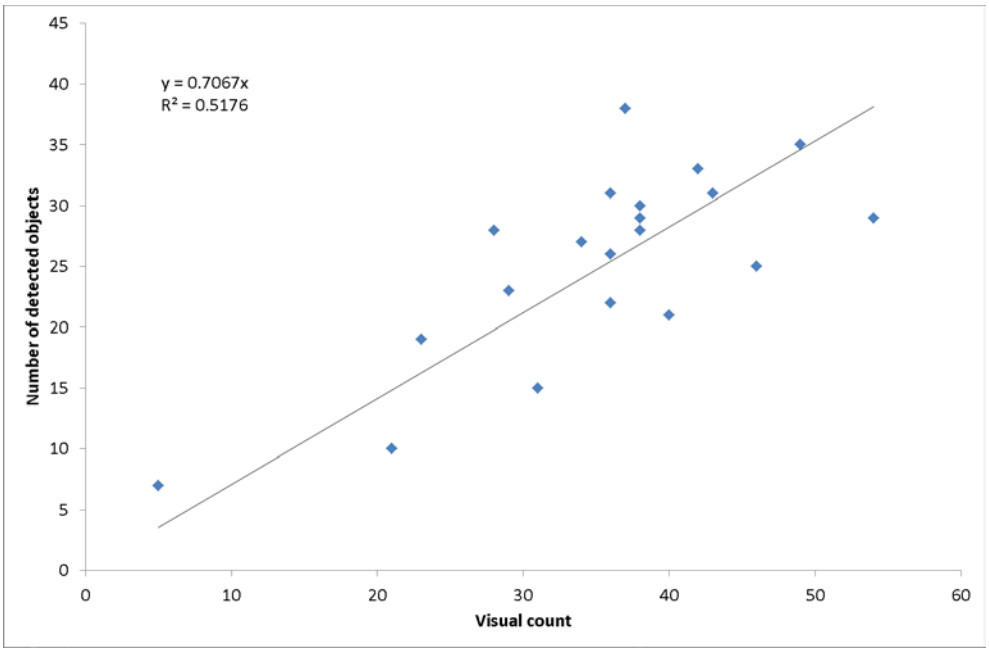

Fig. 8. Calibration: Relationship between the number of objects detected by the automatic procedure and the number of mangoes identified by visual inspection of the 20 calibration images.

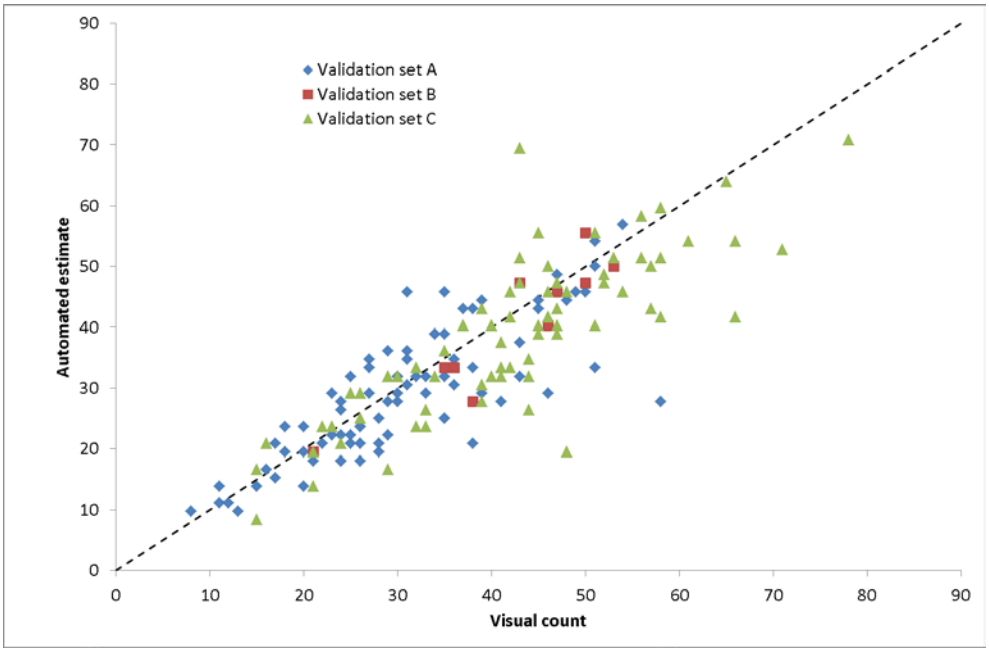

Fig. 9. Validation: Relationship between the number of mango fruit identified by visual inspection and the number of fruit estimated by the automatic procedure for 164 images.

Table 2. Summary of the results for the three mango validation sets.

	Number of images	Manual count	Automated procedure	Ratio
Validation set A	80	2464	2340	0.950
Validation set B	10	419	400	0.954
Validation set C	74	3148	2873	0.912
All	164	6031	5613	0.931

CONCLUSION

This work has shown the robustness of the image processing procedure originally reported in Linker et al. (2012). Applying this procedure, which was developed using daytime images of apple trees, to nighttime images of similar trees required only minimal re-parameterization of the algorithm. After calibration, the difference between the number of fruit determined by the automatic procedure and by visual count was less than one percent. Encouraging preliminary results showing a relationship between the actual tree yield and the fruit load estimated from the images were also obtained. Analysis of the mango trees images required some modifications of the algorithm itself in order to handle the elliptical shape of the mangoes. Once calibrated, the procedure yielded estimates which were within ten percent of the number of fruit as determined by visual inspection of the images. In both studies the required adjustments were performed using only a small number of images (20), and such site-specific calibrations could be realistically performed in practical applications.

REFERENCES

Annamalai, P., Lee, W.S. 2003. Citrus yield mapping system using machine vision. ASAE Paper 03-1002.

Bulanon, D.M., Kataoka, T., Zhang, S., Ota, Y., Hiroma, T. 2001. Optimal thresholding for the automatic recognition of apple fruits. ASAE Paper 01-3133.

Jimenez, A.R., Ceres, R., Pons, J.L. 2000. A survey of computer vision methods for locating fruit on trees. Trans of the ASAE 43, 1911–1920.

Kapach, K., Barnea, E., Mairon, R., Edan, Y., Ben-Shahar, O. 2012. Computer vision for fruit harvesting robots–state of the art and challenges ahead. International Journal of Computational Vision and Robotics 3, 4–34.

Moonrinta, J., Chaivivatrakul, S., Dailey, M.N., Ekpanyapong, M. 2010. Fruit detection, tracking, and 3d reconstruction for crop mapping and yield estimation, in: IEEE International Conference on Control, Automation, Robotics and Vision.

Payne, A., Walsh, K.B., Subedi, P., Jarvis, D. 2013. Estimation of mango crop yield using image analysis–segmentation method. Computers and Electronics in Agriculture 91, 57–64.

Payne, A., Walsh, K., Subedi, P., Jarvis, D. 2014. Estimating mango crop yield using image analysis using fruit at stone hardening stage and night time imaging. Computers and Electronics in Agriculture 100, 160–167.

Pla, F., Juste, F., Ferri, F., Vicens, M. 1993. Colour segmentation based on a light

reflection model to locate citrus fruits for robotic harvesting. Computers and Electronics in Agriculture 9, 53–70.
Tabb, A.L., Peterson, D.L., Park, J. 2006. Segmentation of apple fruit from video via background modeling. ASABE Paper 06-3060.
Zhao, J., Tow, J., Katupitiya, J. 2005. On-tree fruit recognition using texture properties and color data. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 263–268.

APPENDIX A – PARAMETERS FOR APPLE DATASETS

L_{int} : 40
R_{m} : 40
R_{min} : 25
R_{max} : 100
δ_{50}^{r} : 0.4
Ω_{max} : 0.3

The value of parameters not included in the list was as Dataset #1 in Linker et al. (2012).

APPENDIX B – PARAMETERS FOR MANGO DATASETS

A_m : 1.3 (aspect ratio of model)
L_{int} : 40
R_{m} : 75
R_{min} : 20
R_{max} : 250
δ_{50}^{A} : 1.9
δ_{50}^{V} : 0.8 * 2π
δ_{50}^{r} : 0.4
Ω_{max} : 0.3
τ : 0 (not used)

The value of parameters not included in the list was as Dataset #1 in Linker et al. (2012).