Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper

C.W. Bac*,, J. Hemming*, E.J. van Henten*,****

*Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre, Wageningen, The Netherlands, (e-mail: wouter.bac@wur.nl; jochen.hemming@wur.nl)
**Farm Technology Group, Wageningen University and Research Centre, Wageningen, The Netherlands, (e-mail: eldert.vanhenten@wur.nl)

ABSTRACT

As part of the development of a sweet-pepper harvesting robot, obstacles should be detected. Objectives were to classify sweet-pepper vegetation into five plant parts: stem, top of a leaf (TL), bottom of a leaf (BL), fruit and petiole (Pet); and to improve classification results by post-processing. A multi-spectral imaging set-up with artificial lighting was developed to acquire images of sweet-pepper plants. The background was segmented from the vegetation and vegetation was classified into five plant parts, through a sequence of four two-class classification problems. True-positive detection rate/scaled false-positive rate achieved, on a pixel basis, were 40.0/179% for stem, 78.7/59.2% for top of a leaf (TL), 68.5/54.8% for bottom of a leaf (BL), 54.5/17.2% for fruit and 49.5/176.0% for petiole (Pet), before post-processing. The opening operations applied were unable to remove false stem detections to an acceptable rate. Also, many false detections of TL (>10%), BL (14%) and Pet (>15%) remained after post-processing, but these false detections are not critical for the application because these three plant parts are soft obstacles. Furthermore, results indicate that TL and BL can be distinghuished. Green fruits were post-processed using a sequence of fill-up, opening and area-based segmentation. Several area-based thresholds were tested and the most effective threshold resulted in a true-positive detection rate, on a blob basis, of 56.7 % and a scaled false-positive detection rate of 6.7 % for green fruits (N=60). Such fruit detection rates are a reasonable starting point to detect obstacles for sweet-pepper harvesting. But, additional work is required to complement the obstacle map into a complete representation of the environment.

Keywords: Robot vision, Robustness, Classification, Agriculture, Image analysis, Sensors

1. INTRODUCTION

This research is part of a the EU funded CROPS project, 'Clever Robots for Crops', in which a sweet-pepper harvesting robot will be developed. The manipulator of this harvesting robot should approach a target (fruit or peduncle) while avoiding obstacles. These obstacles should be detected, and eventually localized. Obstacle detection is the scope of this article and obstacles comprise supporting wires, construction elements and plant parts (stem, leaf, fruit and petiole). We separated obstacles in hard obstacles and soft obstacles. Hard obstacles (stems, fruits, supporting wires, construction elements) should be avoided by a manipulator or end-effector, whereas soft obstacles (leaves and petioles) can be touched or pushed aside. In addition, the top of a leaf and the bottom of a leaf were discerned to be able to control the motion of a pushed leaf. Pushing a top of a leaf, namely, will usually result in downward motion of the leaf and pushing a bottom of a leaf will result in an upward motion of the leaf.

Vision-based plant part classification studies under natural lighting conditions are scarce. Two studies describe classification of cucumber plant parts into leafs, stems and fruits: a study on a cucumber leaf picking robot using two near-infrared wavelengths (Van Henten et al., 2006) and a multi-spectral imaging study in which several wavelengths and sensors are compared (Noble and Li, 2012). Unfortunately, both studies lack classification performance values. The article most closely related to the work presented here is classification of grape foliage into leaves, branches and fruits (green or coloured) using RGB cameras under natural lighting conditions. For green grapes, the true-positive rate was 91.9 % with a false-positive rate of 2.7 % (Dey et al., 2012).

Objectives were to 1) classify sweet-pepper vegetation into five plant parts: stem, top of a leaf (TL), bottom of a leaf (BL), fruit and petiole (Pet); 2) improve classification results by post-processing techniques.

2. MATERIALS AND METHODS

A multi-spectral imaging set-up was used to acquire images of sweet-pepper plants (Fig. 1). The crop cultivar was "Viper" and only unripe green fruits occurred in the scenes. A 5 megapixel monochrome camera was used in combination with a filter wheel containing filters with the following wavelengths (bandwidth): 447 (60) nm, 562 (40) nm, 624 (40) nm, 692 (40) nm, 716 (40) nm, 950 (100) nm. For each scene, a set of six 8-bit images was acquired with a resolution of 2082 by 2493 pixels. In total, 12 scenes were recorded under outdoor lighting conditions with additional artificial lighting (Fig. 1).

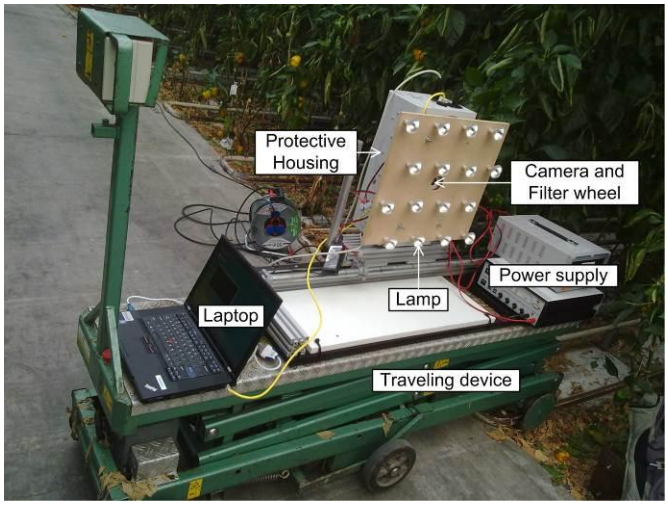

Fig. 1.Overview of the experimental set-up in the greenhouse.

Images were recorded from 11:00 am until 11:30 am under a clear sunny sky. Outdoor solar irradiance was measured during image recording and varied between 374 and 435 W/m^2. Camera-stem distance was on average 80 cm and varied in a range of 63 cm to 109 cm among scenes.

Ground truth data was obtained through manual labelling of pixels in recorded images. In total, $14.6 \cdot 10^6$ pixels were labelled and this number comprised 29.4% of the vegetation present in the 12 scenes. Labelled pixels were mostly leaves – TL (54.6%), BL (22.4%). Other labelled pixels were fruits (15.6%), stems (3.7%) and petioles (3.7%). A large part of the vegetation pixels, mostly leaves, was not labelled (70.6%) because we assumed labelled leaves already represented the majority of the leaf variation occurring in the scene. In addition, labelling all samples would increase the computational load dramatically during training of the classifier. An example of a labelled scene is in Fig. 2.

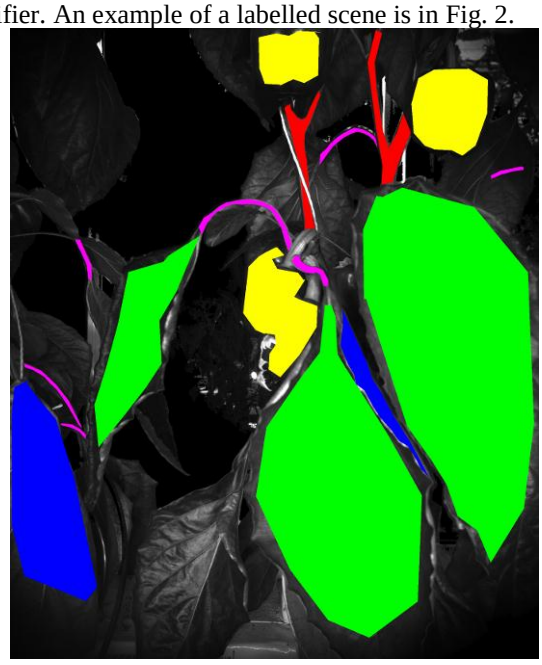

Fig. 2. Labelled image comprising five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), green fruit (yellow) and petiole (purple).

2.1. Performance assessment

Performance was assessed as a binary classification problem. Hence, we compared the detection rates for one class versus the union of the other four classes. As a result, we calculated five 2 by 2 sized confusion matrixes. The elements in each matrix describe true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) detected pixels (Bradley, 1997). Based on these elements, a true-positive detection rate *TPR* (equation (1)) and a scaled false-positive detection rate *SFPR* (equation (2)) were calculated.

$$TPR = \frac{100 \cdot TP}{TP + FN} \quad [\%] \quad (1)$$

$$SFPR = \frac{100 \cdot FP}{TP + FN} \quad [\%] \quad (2)$$

Note that this measure *SFPR* is identical to how authors of previous and recent fruit detection literature refer to 'false-positive rate' (Jiménez et al., 2000, Bulanon et al., 2010, Linker et al., 2012). Such a definition of false-positive rate is, however, confusing because in other research disciplines false-positive rate is calculated as FP/(FP+TN) (Mackinnon, 2000, Gu et al., 2009). Hence, we use scaled false-positive rate instead of false-positive rate to avoid ambiguous definitions of false-positive rate.

Similar to authors of fruit detection literature, we consider scaled false-positive rate to be a more useful measure to report than false-positive rate because false-positive detections are expressed in terms of the class to be detected. False detection rate is therefore not biased by unbalanced class sizes, as is the case with false-positive rate. A drawback of scaled false-positive rate is, however, that rates can exceed 100 %.

2.2. Pixel classification of vegetation

Pixels were classified using Classification And Regression Trees (CART) in MATLAB® 2007b in combination with the Sequential Floating Forward Selection (SFFS) feature selection algorithm (Pudil et al., 1994). Vegetation classification was performed on a computer with an Intel Core i5 CPU 2.4 GHz Quad core processor including 4 GB memory.

The first step in the image processing sequence was to remove the background to obtain remaining pixels of interest. To remove the background, a threshold operation was applied on the 900 nm image (gray-value threshold: $\leq$ 27) and holes in the background region were filled by a fill-up operation. Subsequently, overexposed regions, mostly construction elements and supporting wires, were removed by a threshold (>139) on the 447 nm image. As a result, only vegetation remained. The vegetation was classified into five plant parts: stem, top of a leaf, bottom of a leaf, green fruits and petiole. Pixel-based features were used, i.e. Normalized Difference Index (NDI) (Davies, 2009) and raw gray-values. In total 15 NDI features were calculated from the six wavelengths. Consequently, 15+6 = 21 features were used.

Classification of the five plant parts was split into four two-class classification problems instead of one five-class classification problems because this approach resulted in greater accuracy in previous research (Kavdır and Guyer,

2004). Fig. 3 shows which plant part classes were separated in each classification problem.

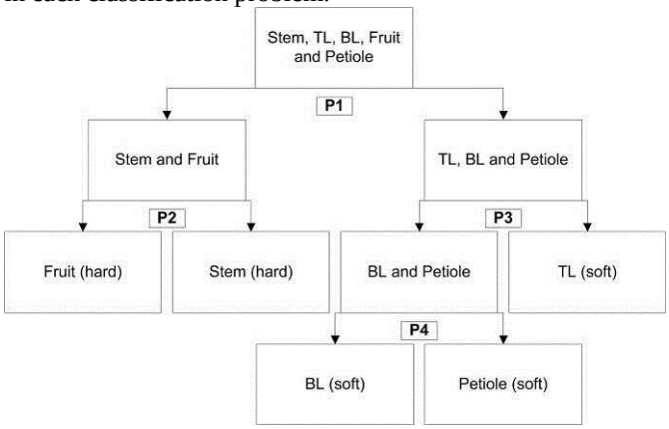

Fig. 3. Approach taken to reduce a five-class classification problem into four two-class problems: P1, P2, P3 and P4.

The decision trees for each two-class problem were trained and pruned in MATLAB®. Two scenes were used for training and ten scenes were used for testing. Before training, class sample sizes were balanced such that the classifier would equally favour both classes. After pruning, trees were exported to Halcon (MVTEC, 2012). As a result, both classification and post-processing techniques were applied in Halcon.

2.3. Post-processing applied to each class

For post-processing, morphological image processing was first applied to improve results of pixel classification. We applied an opening operation, with rectangular mask sizes of 3x3, 5x5 and 7x7 pixels, to each of the five plant parts. Subsequently, results were compared with unprocessed data, in terms of true-positive detection rate and scaled false-positive detection rate. Note that only labelled regions in the image were assessed for performance of classification and post-processing on a pixel basis.

2.4. Green fruit detection

Detection of the fruits, turned out to be more successful than the other plant parts and therefore additional post-processing was applied to improve fruit detection. We applied a sequence of fill-up, opening (circular mask with a radius 2.5 of pixels), connection and area-based segmentation. The circular mask was chosen because fruits mimic a circular shape more than a rectangular shape. The value for area-based segmentation was tested in a range of 1000 to 13000 pixels, with steps of 2000 pixels, to determine the effect on true-positive fruit detection rate and scaled false-positive detection rate.

Performance of fruit detection was not only determined on a pixel basis, but also on a blob (or region) basis. To compare results of blob analysis to the literature, the full image was classified instead of only labelled regions. Separation of fruit clusters into individual fruits is a challenging task in fruit detection (Linker et al., 2012) and was not performed in this research. To calculate the number of individual fruits detected we manually counted the number of fruits present in a cluster, before and after detection. If, for instance, a blob covered three fruits, we counted such a blob as three successfully detected fruits. We counted a fruit as successfully detected if at least some part (>0 %) of the visible fruit surface was detected. In total, 60 distinct fruits were visible (partially or completely) in the ten test scenes.

3. RESULTS

3.1. Pixel classification of vegetation

NDI features turned out to be the strongest features to classify plant parts. True-positive detections with NDI features were about 4-6 % greater than for raw spectral features. The result of plant part classification, using NDI features, is shown in Fig. 4.

Fig. 4. Classification of vegetation into five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), fruit (yellow) and petiole (magenta). Black parts are either background regions or overexposed regions.

Fig. 4 shows that performance of classification based on only pixel information is limited because many false-positive detections occur in the scene. At pixel level, average true-positive detection rates (standard deviations) among scenes (N=10) are: 59.2 (7.1)% for hard obstacles and 91.5 (4.0)% for soft obstacles. Furthermore, detection rates per class are: 40.0 (12.4)% for stem, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole.

Total execution time for one scene in Halcon was 2.3 s: 1.5 s for calculation of 12 NDI features and 0.8 s for decision tree classification. Calculation time of post-processing methods reported in the following sections was in the order of 1-20 ms and therefore negligible to the time required for classification.

3.2. Post-processing applied to each class

To assess the effect of post-processing on true-positive detection rates and scaled false-positive detection rates, results are shown for three post-processing operations and for unprocessed data (Fig. 5).

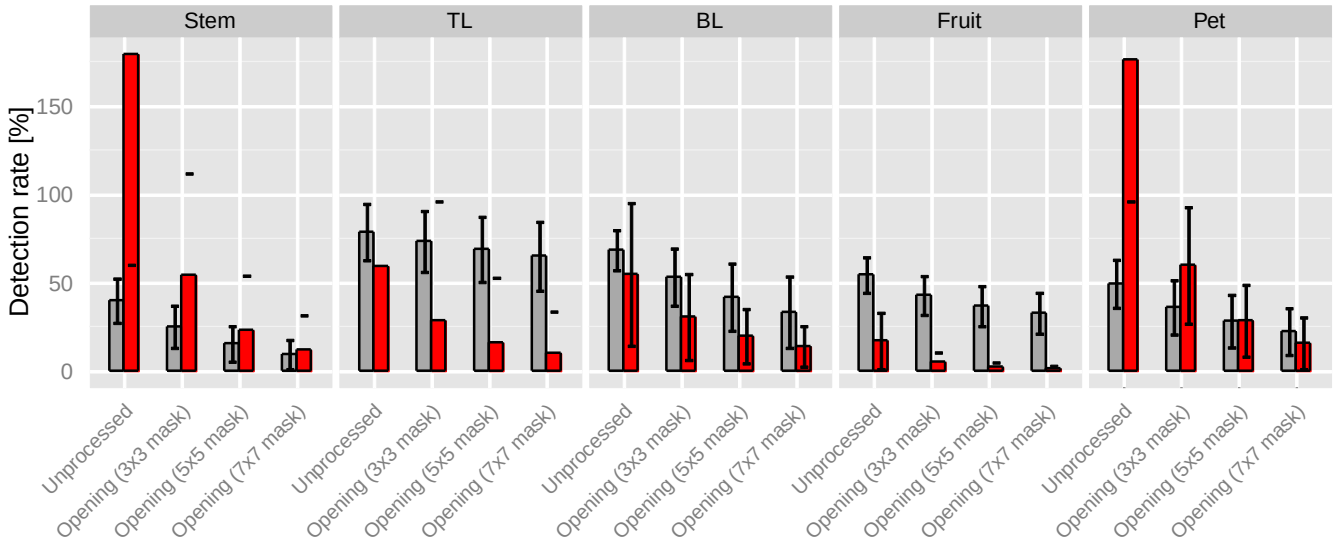

Fig. 5. Mean (N=10 scenes) and SD of true-positive detection rate (■) and scaled false-positive detection rate (■), on a pixel basis, for unprocessed pixels and for three post-processing operations: opening with mask sizes 3x3, 5x5 or 7x7 pixels. These results are shown for each of the five plant part classes: Stem, Top of a Leaf (TL), Bottom of a Leaf (BL), Fruit and Petiole (Pet). Post-processing of fruits resulted in the lowest scaled false-positive rate compared with other classes: 5.2 % for a 3x3 mask, 2.5 % for a 5x5 mask and 1.4 % for a 7x7 mask.

Post-processing improved ratio between true-positive detection rate and scaled false-positive detection rate for all classes. Yet this ratio does not exceed a value of one for the stem and petiole, which indicates that stem and petiole detection was difficult. Furthermore, standard deviation (SD) of detection rate among scenes does not decrease significantly, except for false-positive fruit detections. Apparently, these opening operations do not decrease the variability of detection among scenes.

Fig. 6 demonstrates that many false-positive stem detections remain after post-processing. Such false stem detections are unacceptable because the motion planning algorithm of the robot arm considers these false 'hard obstacle' detections as forbidden areas during calculation of a collision-free path, whereas in reality a path can be planned through these soft obstacles (e.g. leaf).

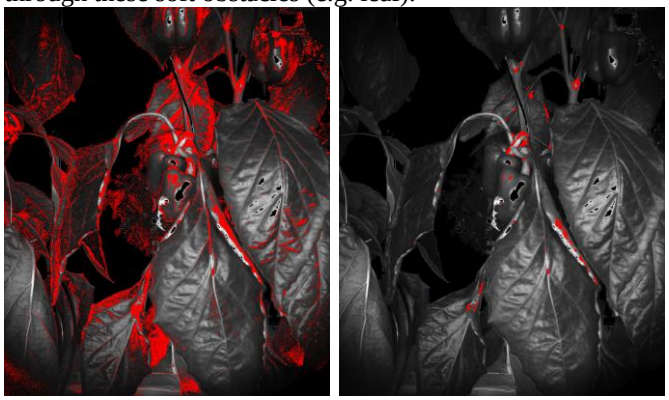

Fig. 6. Stem detections after classification and unprocessed (left). Stem detections after an opening operation (7x7) mask (right). Despite the opening operation, many false detections occur on the leaf and fruit, and many true detections disappear.

Although ratio between true-positive detection rate and scaled false-positive detection rate increased after post-processing (Fig. 5), scaled false-detection rates remain greater than 10 % for TL and greater than 14 % for BL. An example of an opening operation applied to TL is in Fig. 7.

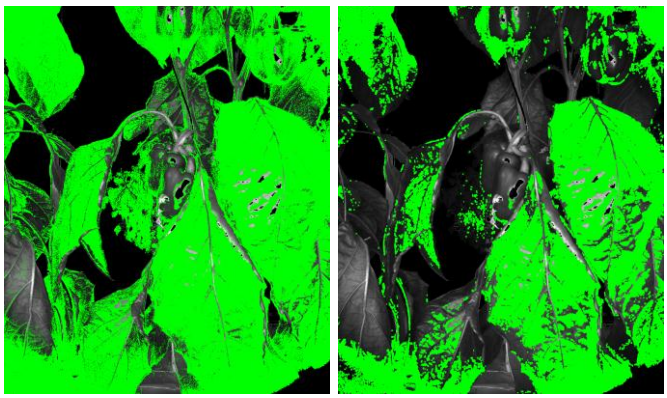

Fig. 7. Top of a Leaf (TL) detections after classification and unprocessed (left). TL detections after an opening operation (7x7) mask (right). Despite the opening operation, many false detections occur on the stem, bottom of a leaf, fruit and petiole.

Fig. 7 (right) demonstrates that few false BL detections remain after an opening operation (in bottom left part of image). Such remaining false detection may be removed through area-based segmentation. Similarly, few false BL detections occurred in the TL (not shown). Hence, these results indicate that TL and BL can be distinguished.

3.3. Green fruit detection

In contrast to detection of other plant parts, a scaled false-positive detection rate of <5.2%, on a pixel basis, was achieved for fruit detection (Fig. 5). The remaining blobs were further processed and Fig. 8 demonstrates that many false-positive blob detections were removed, whereas most true-positive fruit detections remained.

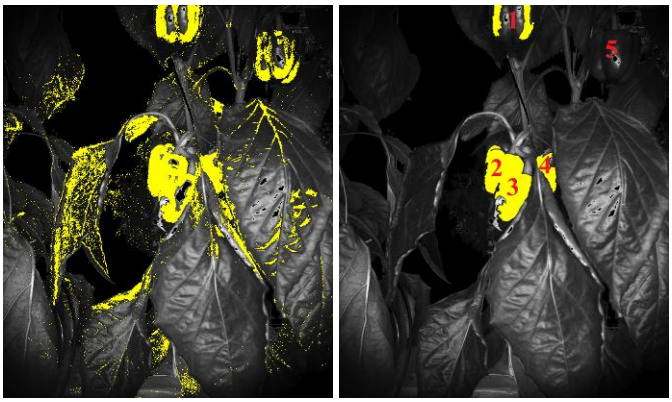

Fig. 8. Classified fruit pixels (left) and the result after fill-up, opening, connection and area-based segmentation (right). The fruit visible in the top right (no. 5) is unfortunately not detected after a sequence of fill-up, opening and area-based (>5000 pixels) segmentation. Hence, four (no. 1-4) of the five fruits (no. 1-5) were detected.

Results of different thresholds, for area of blobs, are in Table 1.

Table 1. True-positive detection rate (TPR) and scaled false-positive detection rate (SFPR) of green fruits, on a blob basis, for different area-based thresholds.

Area-based threshold [> pixels]	True-positive detection rate [%]	Scaled false-positive detection rate [%]	Ratio of TPR/SFPR [-]
1000	88.3	126.7	0.70
3000	83.3	55.0	1.51
5000	78.3	30.0	2.61
7000	68.3	18.3	3.73
9000	68.3	13.3	5.14
11000	61.7	11.7	5.27
13000	61.7	10.0	6.17
15000	56.7	8.3	6.83
17000	56.7	6.7	8.46
19000	55.0	6.7	8.20
21000	51.7	6.7	7.72

We chose the threshold with greatest ratio of true-positive detection rate vs. scaled false-positive detection rate because here the post-processing approach is most effective in removing false detections while preserving true fruit detections. Greatest ratio (8.46) of true-positive vs. false-positive detections is achieved at an area-based threshold of 17000 pixels: TPR = 56.7% and SFPR = 6.7%.

4. DISCUSSION

Accurate hard obstacle detection is more critical for the application than accurate soft obstacle detection because false hard obstacle detection limit the free workspace of a robot manipulator, whereas false soft obstacle detections do not. A motion without hitting either soft or hard obstacles is ideal, but, if required, a motion can be planned through soft obstacles. A motion through hard obstacles is, however, unacceptable because a damage to the stem affects growth of the plant and a damage to the fruit causes yield loss.

Detection rate achieved for soft obstacles (top of a leaf, bottom of a leaf and petiole) is limited, but acceptable for this application.

Regarding hard obstacles, stem detection rate is too small for a useful obstacle map because, after an opening operation (7x7 mask), a TPR of 9.5% and a SFPR of 12.2% was achieved. For green fruit detection, a TPR of 56.7% and a SFPR of 6.7% was achieved. This performance is far from perfect, but probably a reasonable starting point to find a collision-free path to a ripe fruit. However, with such fruit detection rates, the obstacle map will not be complete and additional detection is therefore required. Such detection may be obtained from additional sensors on the end-effector. Also, the manipulator should be able to adapt its path during motion if the sensor detects additional obstacles, or if a false-positive obstacle detection is re-detected as a true negative detection.

Fruit detection rate is worse than related work regarding green apple detection. Linker et al. (2012) achieved a TPR of 88% and a SFPR of 25% under intense natural lighting and a TPR of 95% and a SFPR of 4% under diffuse natural lighting. The methods used in this article are rather basic compared with methods used for apple detection, which may elucidate why those authors achieved a better performance. However, for green citrus detection, Kurtulmus et al. (2011) achieved a TPR of 75% and a SFPR of 27%, which is comparable to results achieved in this research.

The artificial lighting used mitigates the effect of outdoor lighting variations. Nevertheless, classification performance is rather limited and, in addition to outdoor lighting variations, varying plant-camera distances may elucidate why classification performance was rather limited. In related work, geranium cuttings were classified based on RGB images recorded indoor and, seemingly, camera-object distances were more constant (Humphries and Simonton, 1993). These authors achieved a TPR of 85% for leaf, 21% for petiole and 74% for stem, on a pixel basis. They did not report false-positive detection rate, which renders it hard to compare results with this research. Nevertheless, their true-positive rates are slightly higher than results reported here, except for petiole classification. Yet, our study probably benefitted from the near-infrared wavelengths used and the performance gap may have been larger if we would have recorded RGB images combined with such varying lighting conditions and varying camera-object distances. In summary, one can observe that pixel-based classification is limited and addition of object-based features can improve performance because, after addition of object-based features, Humphries and Simonton (1993) achieved a TPR of 97% for leaf, 95% for petiole and 94% for stem.

5. CONCLUDING REMARKS

This study is one of the first multi-spectral imaging studies, under varying lighting conditions, in which detection rates are reported. True-positive detection rate/scaled false-positive rate achieved, on a pixel basis, are: 40.0/179% for stem, 78.7/59.2% for top of a leaf (TL), 68.5/54.8% for bottom of a leaf (BL), 54.5/17.2% for fruit and 49.5/176.0% for petiole (Pet). The opening operations applied were unable to remove false-positive stem detections to an acceptable rate. An

improved stem detection algorithm is therefore a task for future work. Also, many false detections of TL, BL and Pet remained after post-processing, but these false detections are not critical for the application because these three plant parts are soft obstacles. Furthermore, results indicate that TL and BL can be distinguished.

Green fruits were post-processed using a sequence of fill-up, opening and area-based segmentation. Several area-based thresholds were tested and the most effective threshold resulted in a true-positive detection rate of 56.7 % and a scaled false-positive detection rate of 6.7 % for green fruits (N=60). Such fruit detection rates are a reasonable starting point for an obstacle map in an application regarding sweet-pepper harvesting. But, additional sensor information and detection is required to complement the obstacle map into a complete representation of the environment.

ACKNOWLEDGEMENTS

We thank Anja Dieleman for providing and maintaining the sweet-pepper plants in the experimental greenhouse in Wageningen. This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555).

REFERENCES

BRADLEY, A. P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition,* 30**,** 1145-1159.

BULANON, D. M., BURKS, T. F. & ALCHANATIS, V. 2010. A multispectral imaging analysis for enhancing citrus fruit detection. *Environmental Control in Biology,* 48**,** 81-91.

DAVIES, E. R. 2009. The application of machine vision to food and agriculture: A review. *Imaging Science Journal,* 57**,** 197-217.

DEY, D., MUMMERT, L. & SUKTHANKAR, R. Year. Classification of plant structures from uncalibrated image sequences. *In:* Proceedings of IEEE Workshop on Applications of Computer Vision, 2012. Grape, 329-336.

GU, Q., ZHU, L. & CAI, Z. 2009. Evaluation measures of the classification performance of imbalanced data sets. *Communications in Computer and Information Science.*

HUMPHRIES, S. & SIMONTON, W. 1993. Identification of Plant Parts Using Color and Geometric Image Data. *Transaction of the ASABE,* 36**,** 1493-1500.

JIMÉNEZ, A. R., CERES, R. & PONS, J. L. 2000. A survey of computer vision methods for locating fruit on trees. *Transactions of the American Society of Agricultural Engineers,* 43**,** 1911-1920.

KAVDıR, İ. & GUYER, D. E. 2004. Comparison of Artificial Neural Networks and Statistical Classifiers in Apple Sorting using Textural Features. *Biosystems Engineering,* 89**,** 331-344.

KURTULMUS, F., LEE, W. S. & VARDAR, A. 2011. Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture,* 78**,** 140-149.

LINKER, R., COHEN, O. & NAOR, A. 2012. Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture,* 81**,** 45-57.

MACKINNON, A. 2000. A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. *Computers in Biology and Medicine,* 30**,** 127-134.

MVTEC 2012. HALCON image processing library developed by Machine Vision Technologies (MVTEC). HALCON version: 11.0.0.1. www.mvtec.com.

NOBLE, S. & LI, D. 2012. Segmentation of Greenhouse Cucumber Plants in Multi-spectral Imagery. *International Conference of Agricultural Engineering, CIGR-Ageng.* Valencia, Spain.

PUDIL, P., NOVOVIČOVÁ, J. & KITTLER, J. 1994. Floating search methods in feature selection. *Pattern Recognition Letters,* 15**,** 1119-1125.

VAN HENTEN, E. J., VAN TUIJL, B. A. J., HOOGAKKER, G. J., VAN DER WEERD, M. J., HEMMING, J., KORNET, J. G. & BONTSEMA, J. 2006. An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering,* 94**,** 317-323.