3D Reconstruction and Feature Extraction for Agricultural Produce Grading

Panitnat Yimyam
Science and Social Sciences,
Burapha University Sakaeo Campus, Sakaeo, Thailand.
e-mail: panitnat@buu.ac.th

Adrian F. Clark
School of Computer Science and Electronic Engineering,
University of Essex, Essex, UK.
e-mail: alien@essex.ac.uk

***Abstract*—This paper examines the grading of agricultural produce from multiple images using colour and texture properties. Some types of agricultural produce need to be inspected from multiple views in order to assess the entire appearance; however, using multiple images may obtain redundant data. Therefore, techniques are presented to reconstruct a 3D object, create new images without duplicated object areas and extract colour and texture features for evaluation. The performance of using multiple view images without duplicated object regions is compared with those of using only top-view images and the original multiple view images. Experiments are performed on apple and guava grading using kNN, NN, SVM and GP for classification. Performance differences from the different image sets are compared using McNemar's test and the Friedman test. It is found that the performance when using multiple view images is superior to that when using single-view images for all experiments. Employing features extracted from multiple view images without object area duplication achieves significantly higher accuracy than employing the original multiple view images for apple grading, but their performances do not differ significantly for guava inspection.**

Keywords*—*Computer vision, 3D reconstruction, colour and texture extraction, genetic programming, performance evaluation, agricultural produce grading.

I. Introduction

Computer vision attempts to emulate a human's ability to understand the content of images, and this has led to its application in a diverse range of tasks, including medicine, robotics and industrial inspection. Some vision systems are intended to replace human labour in dangerous environments but the main driver is that machines can operate more accurately, consistently and speedily than humans for many tasks. Most vision applications are not stand-alone but involve some degree of an artificial intelligence so that decisions can be made autonomously.

Computer vision has attractions for use in the agricultural industry for tasks such as produce grading, replacing manual inspection. A major attraction is improving the consistency of grading because human operators become fatigued or bored when grading over long periods. Indeed, grading machines are in high demand in the agricultural industry, and there is continued interest in improving the effectiveness and flexibility of automatic grading.

Some types of agricultural produce can be inspected completely from a single viewpoint but in general multiple views are needed to assess the appearance of an entire piece of produce. However, multiple images often show redundant data, where the same surface region is observed in several images, and this may result in incorrect classifications. To illustrate this problem, Figure 1 shows top and side views of a box with a different colour on each side; it is easy to see that the red side is visible in all images. If the system evaluates objects using regions in all images, and if the colour of a defect is red, the number of defects would incorrectly be estimated to be greater.

This paper attempts to overcome this problem by reconstructing the 3D shape and appearance of individual products being inspected, using an algorithm that is simple enough to be used in a real-time grading machine. The performance of using features extracted from the 3D object's surfaces is compared with those of using features extracted from top-view images alone and the original multiple view images.

The remainder of this paper is structured as follows. Section II reviews previous work on agricultural produce grading using computer vision, then section III reviews existing techniques for 3D reconstruction. Section IV details the proposed method for 3D reconstruction and feature extraction. Section V then describes the experiments undertaken to assess the use of 3D reconstruction and alternatives and presents the results. Finally, section VI draws conclusions.

II. Assessment of Agricultural Produce by Computer Vision

Many grading techniques based on computer vision have been developed for agricultural produce [1], [2]. Vision systems may acquire one view or multiple views of each object to be assessed. Generally, both colour and texture properties are employed for grading. For example, Xiaobo *et al* [3] developed a grading machine to classify apples according to the proportion of red. In their system, an object was rotated around the stem-calyx axis in 90° increments, so four images of an object were evaluated. Seventeen colour operators, including mean, variance and normalized colour, and eight intervals of hue were extracted for grading. They proposed a classification technique based on a step decision tree algorithm, and its results were compared with those obtained using a back-propagation neural network (BPNN) and a support vector machine (SVM). They found that their technique achieved high accuracy for classification on extra and reject classes but low

978-1-4673-8139-0/16/$31.00 ©2016 IEEE

accuracy for the other classes, and that it was superior to BPNN but inferior to SVM.

Xu *et al* [4] presented a vision system for apple blemish discrimination. Their system employed three cameras, capturing images while the object was rolled; three images were taken per camera. As the intensity of stem-end, calyx and defect areas are similar to each other, they remarked that stem-end and calyx regions could not be visible in the same image. If two or more doubtful blobs (stem-end, calyx or defect areas) appeared in an image, the fruit appeared to contain a defect. Their experiments were done in three groups: using nine images from three cameras, six images from two cameras and three images from one camera. It was found that the first case was able to cover the entire object surface while the second case lost 5–10% of surface; 15–20% was lost for the last case. The system achieved 94.5% accuracy when using nine images, 83.7% with six images and 63.3% with three images. They suggested using more images results in better accuracy for defective apple detection because there is more opportunity to find blemishes. However, as more spot blush areas were also found, apples with good skins could be misclassified. In the authors' opinion, this problem occurs because using multiple view images obtains redundant data, and the problem can be solved if object areas are seen in only one image. This case inspired the authors to develop a solution to the problem, and the resulting technique will be described later.

Feng and Qixin [5] proposed an apple grading system. Three cameras were employed for capturing an object from top, top-right and top-left positions. As the amount of red area on the fruit surface is related to the sugar content, the ratio of red pixels to the surface area of three-view images was calculated and used for classification using a Bayes classifier. The samples were categorized into three groups according to the percentages of red area. The classification result reached 90% mean accuracy. Rao *et al* [6] presented a method to grade apples based on colour properties. Each apple was rotated around its stem-calyx axis, and six images were taken per fruit. They found that creating colour grading indices using RGB components was difficult but converting the colour representation to one with a hue channel overcame this. The system was approximately 98% accurate.

Colour is not the only criterion that has been used to inspect apples; texture features have also been employed. For instance, Kavdir and Guyer [7] applied a grey-level co-occurrence matrix method to evaluate texture properties from object images taken from a monochrome camera with a sensitivity range of 400–2000 nm. One image was captured per object. Texture and histogram features were used for classification, using BPNN, decision tree (DT), k-nearest-neighbour (kNN) and Bayesian approaches. The samples were classified into good skin, defect and stem-calyx categories. They found that BPNN with texture features achieved the best results in most cases.

Leemans and Destain [8] proposed an apple grading machine capturing several pictures per fruit while it was rolled on a belt, capturing the appearance of the whole fruit surface. Five colour (in RGB), five texture (in RGB), four shape and one position features were extracted to discriminate defective apples from healthy ones. The system performance reached 73% accuracy. It was found that the main causes of errors were fruit with a few small defects being classified into an accepted group, and that some defect appearance was similar to healthy tissue.

Li *et al* [9] developed a sorting machine to detect defective apples. The system involved two monochromatic cameras set above and below the object. An object was laid in a fruit cup without a bottom. Two mirrors were fixed beside the object to reflect side views into the upper camera; hence, one top view, one bottom view and two side views were employed for inspection. Defect areas were segmented based on reference images. As true defects and stem-calyx concave areas were similar, five fractal dimensions were extracted and used for classification by BPNN. They found that stem-calyx areas were often bigger than defects. The classifier achieved about 93% accuracy.

Multiple view analysis was also been employed for other types of fruit. For example, Noordam *et al* [10] developed a high speed potato-grading machine. They presented a technique that used two mirrors attached so as to obtain a full view of the fruit to inspect the quality. Meanwhile Khojastehnazhand *et al* [11] captured two images of lemons, and a hue component was extracted and used for classification. Texture properties are also important for grapefruit inspection. For instance, Kim *et al* [12] extracted texture features based on co-occurrence matrices in HSI space. Some 39 texture features were generated and used for classification on normal and five diseased peel conditions of grapefruit. The SAS procedure DISCRIM was applied for classification using first selected and then all texture features. It was found that employing 13 selected operators from HSI texture features was able to achieve the best accuracy of 96.7%.

III. Previous 3D Reconstruction Techniques

The 3D volume of produce has been constructed to represent the object structure without surface details [13], [14]; subsequently, various techniques have been devised to generate better results. Seitz and Dyer [15] proposed a voxel-colouring technique and initiated a photo-consistency idea to model a colour 3D object based on colour consistency estimation across many images. An object was put on a turntable, and a camera was placed higher than the object position so as to capture top and side views of the object simultaneously. Pictures were taken as the turntable was rotated every few degrees. A virtual volume was carved in the near-to-far direction relative to the camera position. In other words, the closet voxel layer of the camera was traversed before the further layers. Voxel occlusions were managed by occlusion-compatible order and marking pixel (footprint) techniques [16]. When a voxel was visited, its projected pixels were checked. If they did not overlap any footprints of previously visited voxels, it means the voxels were not occluded by others. It assigned colour in the next step, and its projected pixels were marked as a footprint. If its projected pixels overlapped any footprints, the voxel was occluded; therefore, it did not need colour identification. After that, the voxel was checked again by a voxel consistency technique. Several images were employed to evaluate the colour of the voxel. If the colour was accepted using a colour consistency threshold, the voxel would be assigned by this colour; otherwise, the voxel would be removed from the 3D space. As this technique traverses voxels by sweeping a single plane in depth order relative to the camera position, it is

a limitation on using arbitrary camera placements. In order to reduce the constraint, several alternative approaches have been explored; for example, Kutulakos and Seitz [17] proposed space carving to generate a 3D object from multiple images taken from arbitrary but known locations, sweeping in six directions, the positive and negative directions of the X, Y and Z axes. The idea of visiting occluders before the voxels that they occluded was still used. However, only a set of images that the voxel was seen were employed for the consistency checking in each sweep [18].

Culbertson and Malzbender [19] presented the notion of generalized voxel colouring (GVC). The idea of colour consistency was employed for voxel removal. They claimed that the space carving technique used only a subset of images from which the voxel was seen. Their technique employed the entire images from which the voxel was visible for the colour consistency check. They proposed two algorithms, employing different data structures to manage voxel visibility: item buffers (GVC-IB) and layered depth images (GVC-LDI). For GVC-IB, an item buffer of each image was generated to store each pixel and its closet voxel; while for GVC-LDI, layered depth images were applied to store a list of all surface voxel projected to each pixel and the voxels were sorted according to the distance from the camera. They found that GVC-IB required less memory but GVC-LDI obtained higher performance.

Leung et al [20] presented embedded voxel colouring to reconstruct a 3D model with unconstrained camera positions. They proposed the idea of 'water-tightness,' presenting a visibility relationship between voxels and their corresponding pixels on the images. A voxel belongs to a pixel if the voxel was projected to the pixel and it was the closest voxel from the pixel; moreover, a voxel was occluded with respect to a pixel if there was no clear line of sight between them. This idea is useful for dealing with voxel visibility based not only on arbitrary camera location setting but also using low-resolution voxel space. They demonstrated a leaking projection problem when employing a low-resolution volume. An adaptive thresholding algorithm was proposed for the colour consistency check [21].

As the above-mentioned techniques aim to render a virtual 3D model with high quality for the representation of an inanimate object, many pictures have been employed for the reconstruction, such as 21 images in [16], 16 images in [18] and 36 images in [21]. Thus, photo-consistency was appropriate to apply when model-building.

Conversely, much research has been done to reconstruct a 3D model of a moving human, often trying to render the result in real-time, and some techniques generated a 3D object based on voxel-based reconstruction [22], [23], [24]. Similar to the reconstruction of an inanimate object, a very large cube was initialized in the 3D space. Each voxel was projected to all images. Because of the use of a low voxel resolution for voxel-carving, Balan [22] projected eight vertices of a voxel to all images. If a proportion of a convex hull formed by the vertices fell outside a silhouette, the voxel would be removed from the space. Cheung [23] proposed a sparse pixel occupancy test (SPOT) for volumetric reconstruction: if selected pixels in a convex hull of a voxel were inside the silhouettes, the voxel was assigned to the object volume. The remaining voxels in the space were formed as a reconstructed 3D object. Kehl *et al* [24] developed an application for human tracking. Because they used video streams as input, they created a fixed look-up table (LUT) for each camera to store a list of each pixel with pointers to all projected voxels. The following step was to assign colour to the reconstructed object. As non-surface voxels were not seen in any image, only surface voxels were assigned a colour. Surface voxels were detected by considering six-connected neighbours: if a voxel did not contain one or more nearest neighbours in the six directions, it acted as a surface voxel. Since multiple voxels were able to be projected onto the same area on an image plane, the closest voxel was supposed to belong to the pixel colour. The depth of voxels was able to be applied to deal with voxel occlusions. For example, in [24] a depth buffer was used to indicate the depth of voxels. If an occlusion was not detected from the depth buffer, the colour of the corresponding pixel was assigned to the closest surface voxel and the voxel became the new visible voxel in the depth buffer.

IV. Feature Extraction

As mentioned in section II, some types of agricultural produce can be inspected from one view but images from several views that cover the whole surface are needed to inspect others. However, feature extraction from multiple view images may obtain redundant data. This paper attempts to extract features of objects from multiple view images without duplicated object areas, and the easiest way to do that is to reconstruct the 3D shape.

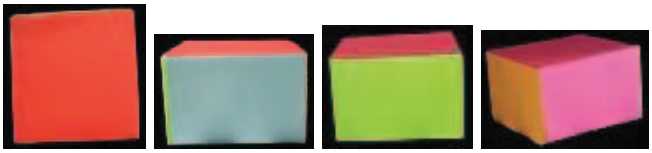
A. Approach used for shape reconstruction

The technique for 3D volume reconstruction presented in [25] is the basis for this paper. A large cube is initialized that encloses the object. This virtual volume is divided into equally-sized voxels, all initially opaque according to the camera parameters derived from camera calibration. If the projection falls outside any silhouette, it will be deleted from the space by setting it to be transparent. The remaining voxels represents a reconstructed 3D volume, which will be approximately equal to or greater than the real object volume. The more images are used, the closer the reconstructed shape is to the real object. The generated 3D volume can then be employed for shape analysis.

B. Object shown in each image without duplicated areas

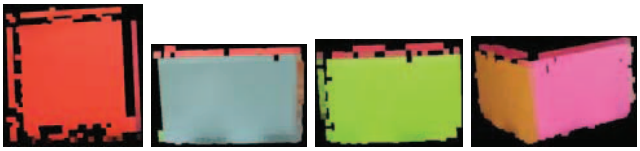
The camera configuration used here is the same as that presented in [25]: one camera captures a top view and three cameras laid nearly in the corners of triangle capture side views, covering most of the surface of an object. An example of object images taken from the proposed camera setting is presented in Figure 1.

Analogously to the above-mentioned research for 3D reconstruction of a moving human, after the 3D volume has been reconstructed, the colour is evaluated and assigned to surface voxels. This paper does not aim to represent a reconstructed 3D object but instead to evaluate the colour and texture properties for inspection. As features can be directly extracted from images, instead of assigning colour to surface voxels, a viewpoint ID is associated with a surface voxel. These IDs can be shown in only one viewpoint so as to eliminate duplicated areas in the original images. An algorithm to evaluate which



(a) Viewpoint 1 (b) Viewpoint 2 (c) Viewpoint 3 (d) Viewpoint 4

Fig. 1: Segmented images of a box taken from different viewpoints



(a) Viewpoint 1 (b) Viewpoint 2 (c) Viewpoint 3 (d) Viewpoint 4

Fig. 2: Object areas of the box are shown in only one viewpoint

viewpoints that surface voxels can be seen in is presented in Algorithm 1.

Algorithm 1: Evaluation on viewpoints that surface voxels can be visible

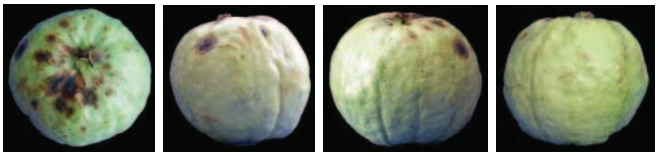
```
1 foreach i ∈ {images} do                                   // iterate images
2     foreach v ∈ {surface voxels} do                       // iterate surface voxels
3         i[v] = depth(i, v);                               // measure voxel depth
4     ascendingSort(i[v]);                                  // ascending sort voxels
5     foreach v ∈ {i[v]} do                                 // iterate thro' voxels
6         project voxel v to image i;
7         if the footprint area does not completely overlap the previous ones
          then
8             v[imageID] = v[imageID] ∪ iID;                // store image ID
9             mark a footprint area of voxel v;             // mark projected pixels of
              the voxel
```

where $depth(i, v)$ is depth measured between the voxel and the camera centre and *ascendingSort(i(v))* sorts the set of voxels of the viewpoint in ascending order according to depth.

In lines 1–3 in order to visit surface voxels in near-to-far ordering relative to the camera position, the depth between the surface voxels and the images are measured. In line 4, the voxels are sorted in ascending order according to depth. In lines 5–8, the voxels are projected to the image in order. If a footprint area completely overlaps previous ones, it means the surface voxel is occluded and not visible in the image; otherwise, the voxel is seen in the image, so the image ID is assigned to the surface voxel. Note that as a surface voxel can be visible in multiple images, it can contain a set of image IDs. In line 9, the footprint area of the voxel is marked. A footprint area in this paper indicates an region of pixels in an enclosing rectangle of the projected vertices of a voxel.

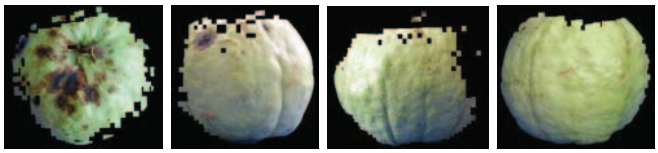
This algorithm can be applied to identify the viewpoints from which surface voxels are visible without concern about camera locations. As a surface voxel can include multiple image IDs, only one ID will be selected to assign to the voxel. A couple of rules are made to choose a viewpoint for surface voxels seen in multiple images.

1) If a voxel is visible in all images, the ID of the top view is assigned to the voxels.



(a) Viewpoint 1 (b) Viewpoint 2 (c) Viewpoint 3 (d) Viewpoint 4

Fig. 3: Segmented images of a guava taken from different viewpoints



(a) Viewpoint 1 (b) Viewpoint 2 (c) Viewpoint 3 (d) Viewpoint 4

Fig. 4: Object areas of the guava are shown in only one viewpoint

2) For other cases, the image IDs of neighbours in a defined area are considered. The most common ID will be assigned to the voxel.

For other designs of camera setup, using only rule 2 would be appropriate. An example of results generated from the proposed technique to create new images of the box (as shown in Figure 1) without duplicated object areas is presented in Figure 2. Duplicated object areas such as the red region are roughly visible in only one image. The proposed method is examined for agricultural produce using a defective guava. Figure 3 shows a segmented image of the sample from each viewpoint, while Figure 4 shows the result from the proposed technique.

C. Colour and texture extraction

Colour and texture were extracted from images of the object. Colour operations are analyzed in binary, RGB and HSI modes. Sample colour operators contain normalization, mean, standard deviation, $c_1c_2c_3$ and $l_1l_2l_3$. Texture evaluation is based on co-occurrence matrices and run-length approaches, and texture features are extracted in RGB and HSI spaces. The details of colour and texture operations used in this paper are described in detail in [26], [27]. There are 29 colour and 123 texture operators in total.

V. Experiments

Experiments were performed to compare the performances of using only top view images (ImageSet1), multiple view images (ImageSet2), and multiple view images without duplicated object areas (ImageSet3) for agricultural produce grading based on the colour and texture features listed above. The experiments consisted of defect inspection on apples and guavas. Each experiment included twenty sub-experiments with different training data. The vision system used involves genetic programming (GP) for classification and is described in [27]. The results of using different image sets were compared by McNemar's test (using $\alpha = 0.05$ and critical value

$Z_{\text{crit}} = 1.96$) and the Friedman test (using α = 0.05 and the critical value $\chi^2_{0.05} = 5.99$ for $df = 2$). The GP results were compared with other machine learning approaches: BPNN, kNN and SVM. These classifications performed using the same extracted features training sets and testing sets, using Weka.

A. Apple defect inspection

Bruising of the skin of apples can be caused by dropping during harvesting or subsequent knocks. Bruised areas are generally darker than normal skin. Royal gala apples were employed for this experiment: 76 fruits with good skin and 136 fruits with defective skin. 30% of the samples were randomly selected for training and 70% for testing.

Table I presents the accuracy results from all three input sets as classified by kNN, BPNN, SVM and GP. The kNN classifier was the least accurate of the group for all input sets, though the classification performances of all learning approaches do not vary greatly. However, the performance when using ImageSet3 is superior to when using the other image sets. The mean accuracies when using ImageSet3 generated by all classifiers are 92–96% whereas using the other image sets achieved only 78–84% mean accuracies.

The classification results produced by GP of using different image sets were compared using McNemar's test and the Friedman test. Figure 5 shows Z values analysed by McNemar's test: all Z values when using ImageSet1 and ImageSet2 are lower than Z_{crit} (1.96), meaning that performance differences are insignificant. In contrast, all Z values from ImageSet3 compared with using the other image sets indicate that differences in performance are significant, and those Z values are in the range 2–6, above Z_{crit}.

In addition, the Friedman test was employed to analyse performance differences of using different image sets based on GP classification results. They generated $\chi^2_r(2) = 32.5$ (significantly more than the critical value of 5.99) and $P < 0.0001$. This result indicates that there is a significant deference between at least two conditions. CD_F was calculated to be 15.1; if any of the differences between the sums of ranks is equal to or greater than CD_F, it means that a comparison is significant. With respect to the three pairwise comparisons, the difference scores between the sums of ranks of ImageSet1 vs. ImageSet3 was 35.0, ImageSet2 vs. ImageSet3 25.0 and ImageSet1 vs. ImageSet2 10.0; the two former pairs are greater than CD_F whereas the latter pair is lower. These results indicate that the performance differences from ImageSet3 compared with the other image sets are significant. Statistical analysis by McNemar's test and the Friedman test for this experiment demonstrates that the performance of using the image set generated by the proposed technique was significantly superior to those of employing the other image sets for apple defect grading.

B. Guava defect discrimination

The colour of guava defects is noticeably different from the colour of good-quality skin. Some 48 good guavas and 106 defective guavas were examined; a defective guava is shown in Figure 3. 30% of the samples were randomly selected for training and the remainder used for testing. Table I presents the classification results produced by kNN, BPNN SVM and GP on the different input image sets, as in the previous experiment. SVM was the leader of the group with ImageSet1, achieving 86.1% mean accuracy, while GP took the lead for both ImageSet2 and ImageSet3 with mean accuracies of 92.5% and 95.1% respectively. The kNN classifiers achieved the lowest mean accuracies for all input sets. However, overall the performances of all classifiers do not differ greatly. Z values from McNemar's test based on the GP results are shown in Figure 6. The Z values of ImageSet1 vs. ImageSet2 exceed the critical value 11 times, with values in the range 2.1–3.7. Performance differences of ImageSet1 vs. ImageSet3 are significant 17 times, and the Z values are 2.7–4.4. Conversely, almost Z values produced by ImageSet2 vs. ImageSet3 are lower than Z_{crit}.

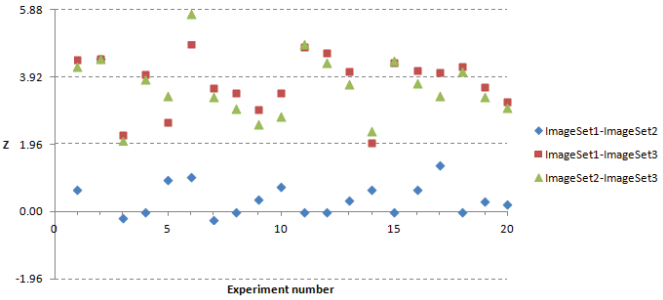


Fig. 5: Z scores of the experiment on apple grading

TABLE I: Classification accuracy results

	kNN		NN		SVM		GP	
Apple grading	Ave.	sd	Ave.	sd	Ave.	sd	Ave.	sd
ImageSet1	78.0	3.4	79.8	5.3	82.3	4.0	80.6	3.0
ImageSet2	78.3	3.3	82.4	4.6	83.4	3.3	82.0	1.7
ImageSet3	92.8	2.0	95.7	1.4	95.2	2.1	95.2	1.8
Guava grading								
ImageSet1	80.2	5.0	85.1	3.5	86.1	3.1	81.6	4.1
ImageSet2	85.9	4.1	91.2	3.7	92.1	3.6	92.5	1.8
ImageSet3	88.5	3.5	91.1	4.1	92.1	3.8	95.1	2.0

For the Friedman test based on the GP results, $\chi^2_r(2) = 35.63$, $P < 0.0001$ and $CD_F = 15.1$. The differences between the sums of ranks of ImageSet1 vs. ImageSet2, ImageSet1 vs. ImageSet3 and ImageSet2 vs. ImageSet3 are 22.5, 37.5 and 15.0 respectively. Pairwise comparisons indicate that there are significant differences between ImageSet1 vs. ImageSet2 and ImageSet1 vs. ImageSet3 because their differences of the sums of ranks exceed CD_F. This experiment shows that using ImageSet1 achieved the least accuracy, and its performance significantly differed from using ImageSet2 and ImageSest3, while the performance differences from ImageSet2 and ImageSet3 tend to be insignificant for this task.

VI. Conclusions

This paper presented techniques to estimate the quality of fruit using colour and texture properties extracted from top view images alone, multiple view images and multiple view images without duplicated object regions. Royal gala apples and guavas were employed for this task. Four machine learning approaches, GP, kNN, BPNN and SVM, were used for classification. McNemar's test and the Friedman test were

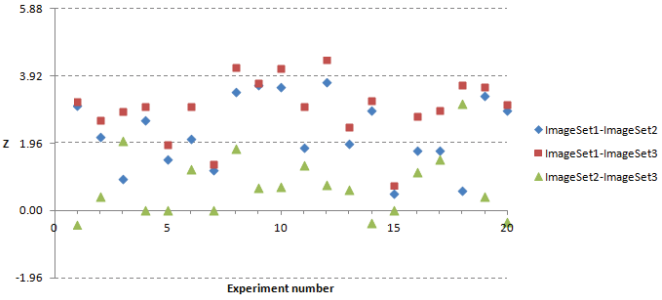


Fig. 6: Z scores of the experiment on guava grading

applied to assess performance differences in classification using extracted features from the three image sets.

For the experiment on apple quality discrimination, it was found that the performance of using the features extracted from the multiple view images without duplicated object areas achieved the highest accuracy for all learning algorithms, and its performance significantly differed from those of using the other image sets. As more spot blush regions were probably found in multiple images, the good skin samples would be misclassified for this apple type. Thus, the proposed technique to extract features from multiple views without object region duplication was appropriate to apply to apple grading.

The experimental results of guava inspection also show that using multiple view images was significantly superior to employing only single-view images because defective areas might not be found in one viewpoint. The performance of employing multiple view images and multiple views without object area duplication did not differ significantly. This case is different from the previous task as the colour of guava defects was obviously different from that of good skin. However, features extracted from multiple images were more appropriate to employ for guava defect inspection than those extracted from only one viewpoint.

VII. Acknowledgements

This research is supported by a grant from the Faculty of Science and Social Sciences, Burapha University Sakaeo Campus, Thailand.

References

[1] D. Wu and D.-W. Sun, "Colour measurements by computer vision for food quality control–a review," *Trends in Food Science & Technology*, vol. 29, no. 1, pp. 5–20, 2013.

[2] P. Jackman and D.-W. Sun, "Recent advances in image processing using image texture features for food quality assessment," *Trends in Food Science & Technology*, vol. 29, no. 1, pp. 35–43, 2013.

[3] Z. Xiaobo, Z. Jiewen, and L. Yanxiao, "Apple color grading based on organization feature parameters," *Pattern Recognition Letters*, vol. 28, no. 15, pp. 2046–2053, 2007.

[4] Q. Xul, X. Zou, and J. Zhao, "On-line detection of defects on fruit by machinevision systems based on three-color-cameras systems," in *Computer and Computing Technologies in Agriculture II, Volume 3*, pp. 2231–2238, Springer, 2009.

[5] G. Feng and C. Qixin, "Study on color image processing based intelligent fruit sorting system," in *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on*, vol. 6, pp. 4802–4805, IEEE, 2004.

[6] P. S. Rao, A. Gopal, R. Revathy, and K. Meenakshi, "Color analysis of fruits using machine vision system for automatic sorting and grading," *J. Instru. Soc. India*, vol. 34, no. 4, pp. 284–291, 2004.

[7] I. Kavdır and D. Guyer, "Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features," *Biosystems engineering*, vol. 89, no. 3, pp. 331–344, 2004.

[8] V. Leemans and M.-F. Destain, "A real-time grading method of apples based on features extracted from defects," *Journal of Food Engineering*, vol. 61, no. 1, pp. 83–89, 2004.

[9] Q. Li, M. Wang, and W. Gu, "Computer vision based system for apple surface defect detection," *Computers and electronics in agriculture*, vol. 36, no. 2, pp. 215–223, 2002.

[10] J. C. Noordam, G. W. Otten, T. J. Timmermans, and B. H. van Zwol, "High-speed potato grading and quality inspection based on a color vision system," in *Electronic imaging*, pp. 206–217, International Society for Optics and Photonics, 2000.

[11] M. Khojastehnazhand, M. Omid, and A. Tabatabaeefar, "Development of a lemon sorting system based on color and size," *African Journal of Plant Science*, vol. 4, no. 4, pp. 122–127, 2010.

[12] D. G. Kim, T. F. Burks, J. Qin, and D. M. Bulanon, "Classification of grapefruit peel diseases using color texture feature analysis," *International Journal of Agricultural and Biological Engineering*, vol. 2, no. 3, pp. 41–50, 2009.

[13] E. Boyer, "Object models from contour sequences," in *Computer Vision: ECCV'96*, pp. 109–118, Springer, 1996.

[14] Y. Kuzu and V. Rodehorst, "Volumetric modeling using shape from silhouette," *Proc. the 4th Turkish-German Joint Geodetic Days*, pp. 469–476, 2001.

[15] S. M. Seitz and C. R. Dyer, "Photorealistic scene reconstruction by voxel coloring," in *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, (Washington, DC, USA), pp. 1067–, IEEE Computer Society, 1997.

[16] S. Seitz and C. Dyer, "Photorealistic scene reconstruction by voxel coloring," *International Journal of Computer Vision*, vol. 35, no. 2, pp. 151–173, 1999.

[17] K. N. Kutulakos and S. M. Seitz, "What do photographs tell us about 3D shape?," tech. rep., Technical Report TR692, Computer Science Dept., U. Rochester, 1998.

[18] K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," *Int. J. Comput. Vision*, vol. 38, pp. 199–218, July 2000.

[19] W. B. Culbertson, T. Malzbender, and G. Slabaugh, "Generalized voxel coloring," in *Vision Algorithms: Theory and Practice*, pp. 100–115, Springer, 2000.

[20] C. Leung, B. Appleton, and C. Sun, "Embedded voxel colouring," in *Digital Image Computing-Techniques and Applications*, pp. 623–632, 2003.

[21] C. Leung, B. Appleton, M. Buckley, and C. Sun, "Embedded voxel colouring with adaptive threshold selection using globally minimal surfaces," *Int. J. Comput. Vision*, vol. 99, pp. 215–231, Sept. 2012.

[22] A. O. Balan, "Voxel carving and coloring-constructing a 3D model of an object from 2D images." http://cs.brown.edu/~alb/papers/balan03voxel.pdf. Accessed, Sep. 2012.

[23] K. M. Cheung, S. Baker, and T. Kanade, "Shape-from-silhouette across time part i: Theory and algorithms," *International Journal of Computer Vision*, vol. 62, pp. 221 – 247, May 2005.

[24] R. Kehl, M. Bray, and L. Van Gool, "Full body tracking from multiple views using stochastic sampling," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, pp. 129–136, IEEE, 2005.

[25] T. Chalidabhongse, P. Yimyam, and P. Sirisomboon, "2d/3d vision-based mango's feature extraction and sorting," in *Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference on*, pp. 1–6, 2006.

[26] P. Yimyam and A. Clark, "Agricultural produce grading by computer vision using genetic programming," in *Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on*, pp. 458–463, 2012.

[27] P. Yimyam, *Agricultural Produce Grading by Computer Vision Based on Genetic Programming*. PhD thesis, University of Essex, 2015.