Toward a New Approach in Fruit Recognition using Hybrid RGBD Features and Fruit Hierarchy Property

Ema Rachmawati[1], Iping Supriana[2], Masayu Leylia Khodra[3],
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
[1]ema.rachmawati22@students.itb.ac.id, [2]iping@stei.itb.ac.id, [3]masayu@stei.itb.ac.id

***Abstract*— We present hierarchical multi-feature classification (HMC) system for multiclass fruit recognition problem. Our approach to HMC exploits the advantages of combining multimodal features and the fruit hierarchy property. In the construction of hybrid features, we take the advantage of using color feature in the fruit recognition problem and combine it with 3D shape feature of depth channel of RGBD (Red, Green, Blue, Depth) images. Meanwhile, given a set of fruit species and variety, with a preexisting hierarchy among them, we consider the problem of assigning images to one of these fruit variety from the point of view of a hierarchy. We report on computational experiment using this approach. We show that the use of hierarchy structure along with hybrid RGBD features can improve the classification performance.**

Keywords*— *hierarchical multimodal classification; multiclass fruit recognition; hybrid features; fruit hierarchy; color feature; 3D shape feature

I. Introduction

The application of computer vision is very common in agricultural and industrial field, such as fruit harvesting robot [1]–[4], fruit sorting machine, and fruit scanner in supermarket. The fruit scanner is needed to identify the fruit and vegetables in the supermarket at the time of weighing. It is well known that the fruit in the supermarket are not given a barcode like packaged product. Generally, the approach used to identify the fruit or vegetable was using a manual approach and required several previous training sessions for employees. In this case, the fruit recognition system can be applied as a replacement of the employee's duties, therefore the process of weighing and identifying the fruit or vegetables can be done simultaneously and efficiently.

There were four commonly used features in the fruit recognition approach to represent a fruit, namely color, shape, texture and intensity. Some approaches only focus on specific feature like color while others focus on combining two or more features, resulting in several different techniques. However, even using single, two or more features, the problem in fruit recognition might still arise, especially when there are different fruit species which have the same color or shape. Vogl et al. [5] also stated that this condition shows us the requirement for more features in order to increase the fruit recognition performance. Furthermore, there is a consensus among researchers that there is no one single feature that is sufficient for the purposes of object recognition. The reason is, each feature captures a specific object property, which show different things to different object of the class. This is what motivates the exploration for better representation [6].

Meanwhile, we consider the problem of hierarchical multi-classification, which is a prediction system where the target classes are ordered in a hierarchy. It is easily understood that sometimes, the hierarchy itself shows us the resemblance of classes, implicitly. One simplest approach to solve hierarchical multi-classification is using flat classification, which is the simplest method. The flat approach ignores the hierarchical structure of classes. This approach generally adopted the one-versus-rest strategy to elaborate hierarchical classification into multiple binary classification in leaf categories leaf. In contrast to flat classification, a binary classifier is constructed for each category in the hierarchy in the top-down approach. This has the advantage that the total size of the predictive theory become generally smaller, and interdependence among different classes can be involved and made explicit.

In this article, we investigate the importance of integrating object hierarchy in the classification process. In contrast to most fruit recognition system which is based on flat classification, our aim is to introduce an approach for fruit recognition that includes the fruit hierarchy in the recognition process, alongside with hybrid features. Our hybrid feature consists of color feature which is extracted from RGB channel and shape feature which is extracted from Depth channel of RGBD (Red, Green, Blue, Depth) image. This approach is highly desirable since the proposed hybrid feature is very efficient in representing fruit property since human commonly recognize fruit by its color and/or shape. In this case, we use 3D shape feature to represent shape property of multi-view fruit.

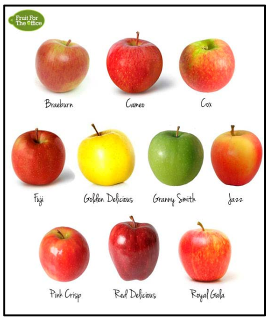

Fig. 1. Shape and color variations of apple

Furthermore, the fruit hierarchy is investigated in the classification process by firstly classifying a fruit into its species class (e.g. apple, banana, pear, etc.) before classifying it into its variety class (e.g. Fuji Apple, Granny Smith Apple, Cavendish Banana, Lisbon Lemon, etc.). It is commonly known that there are many different color and/or shape of a fruit species, as depicted in Fig. 1. Also, a particular fruit species having similarity in color and/or shape with other fruit species, as depicted in Fig. 2. With this approach, by incorporating fruit hierarchy in the classification process, we hope a fruit variety of apple species would not be classified into pear species though it has similarity in color or shape.

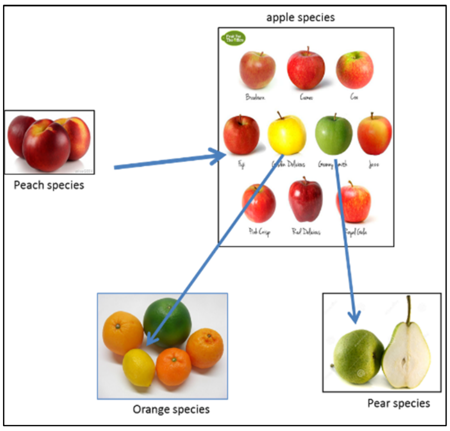

Fig. 2. Peach, orange, and pear that have the same color with apples

The remainder of the article is organized as follows. In Section 2, we briefly explained some works related to our work. Detail explanation of our proposed model can be seen in Section 3. The dataset and experimental evaluation can be seen in Section 4 while conclusion described in Section 5.

II. Related Works

Various researchers have applied many types of image processing and analysis techniques to recognize fruits or vegetables [7]–[11]. Veggie-vision [7], created by Bolle et al., was the first fruit and vegetable recognition, in which uses color (Hue histogram) and texture. Bolle et al. used 48 different produce items (fruit and vegetables) which have very different in shape. Hence, only using color and texture to recognize produce is enough. They did not incorporate shape features, while they suggest the use of shape features in order to improve the classification result. Rocha et al. [8] used a particular approach which combine many features and classifiers to classify 15 fruit categories. For the classification, it uses the Bagging Ensemble of Linear Discriminant Analysis (BLDA) with 17 iterations. Their method shows poor results for some type of fruit and vegetable such as Fuji Apple, though they have achieved good classification accuracy by using top two responses. Zhang and Wu [9] use color, texture, and shape feature to classify 18 fruit categories, using the support vector machine. Zhang et al. [10] also proposed an approach to classify 18 fruit categories based on a fitness-scaling chaotic artificial bee colony. They used color, texture, and shape feature (area, perimeter, Euler number, convex, solidity, minor length, major length, eccentricity). Kuang et al. [11] uses 20 fruit categories and apply PCA to reduce feature dimension. The overall classification performance by using combination of color, texture, shape, HOG, and edge-LBP feature is very good. Though, the recognition rate of using only shape feature is less satisfying, which is 25.7 %.

Meanwhile, many important real-world classification problems might be converted to hierarchical classification problems. In the hierarchical classification problem, the classes to be predicted are organized into a class hierarchy [12]. Further, Silla and Freitas [12] stated that generally the approaches to deal with the classification hierarchy can be grouped into a flat classification approaches, local approaches classifier (top-down), and a global approach classifier (or Big-bang approach). The flat approach is very easy to implement by adopting the existing binary classification algorithm, but this approach assumes that all example should be predicted in a leaf category while many actual applications demand the examples to be predicted also in non-leaf category. In contrast to the flat approach, the top-down approach builds a binary classifier for each category in the hierarchy. The local approaches (top down approaches) consist of a series of local classifiers, which are usually applied in a top–down way. Meanwhile, the global approach learns a single global model for all classes [12].

Regarding to image analysis and computer vision domain, some researchers have applied the hierarchical classification approach in their solutions [13]–[16]. Barutcuoglu and DeCoro [13] use aggregation of Bayesian Network with KNN (k-nearest neighbor) as the base classifier in 3D shape classification. Their motivation in using hierarchical approach lies in the fact that the class is organized into a hierarchy of the most common shape to the most specific. Dimitrovski et al. [14] use global classifier (GC) approach in the shape classification, while Binder et al. [15] use local classifier per node (LCN). Hernández et al. [16] build a multi-class classifier per each parent node in the hierarchy. Each class in the taxonomy will have a probability in the classification phase, as all local classifiers were applied simultaneously.

According to recent fruit recognition system, in general, the fruit dataset used less representing the multi viewpoint aspect of fruit, where it becomes important in supporting the flexibility of fruit recognition in a supermarket. Shape feature extraction from different shape viewpoints will generate shape features that would be considered differently. In the case of two fruit having same color (yellow apple and lemon), the color feature is no longer be considered, hence determining the shape feature that support multiple viewpoint representation is important. Furthermore, existing research in fruit recognition still relies on the exploitation of 2D aspect of image only, whereas a number of researchers have conducted a variety of approaches to exploit the 3D aspect of image to help improve the classification performance, namely [17]–[19] among others. Therefore, in this article, beside we take the advantages of 2D features combination, we also exploit the 3D feature which is captured using RGBD (red, green, blue, depth) sensor.

III. Hierarchical Multi-Feature Classification

This section presents our approach for building hierarchical multi-feature classification (HMC). The pipeline of our framework consists of several processes as illustrated in Fig. 3.

We first present the multimodal features. Next, we describe the hierarchical classification approach. We use RGBD images captured by RGBD sensor as the input images. The hierarchical property of the fruit is further used in the hierarchical classification.

A. Hybrid RGBD Features

1) Color feature: Color Layout Descriptor (CLD)

CLD is used in this article as color representation, in order to capture the spatial distribution of color in the image [20]. The image is divided into 64 blocks, then the dominant color or the average color is extracted from each of these blocks. Further, each channel of *color space* YcbCr is transformed using 8 x 8 *Discrete Cosine Transform* (DCT) to generate 3 coefficient sets, sized 8 x 8, for each Y, Cb, and Cr. Each set of coefficients will be given a certain weight to produce a feature vector with dimension maximum of 3 x 64 [21].

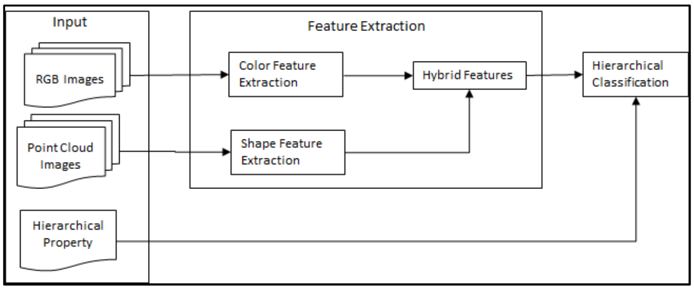

Fig. 3. Proposed framework

2) 3D shape featureViewpoint Feature Histogram (VFH)

Viewpoint features histogram (VFH) [22] is a global shape descriptor, which is extracted from the whole object and have the ability to characterize the global shape of the object with a single vector. This descriptor is working on a set of clusters, with a cluster is defined as a set of 3D points (or, point cloud) which can be represented as an object or scene. VFH provide global orientation of the surface normal of the object. VFH combine the viewpoint direction and Point Feature Histogram (PFH) [23]. PFH considers the geometrical structure of a number of neighboring points to calculate the feature. In particular, PFH generalize the mean curvature around point by means of a multidimensional histogram values. PFH is a histogram that collects pair of pan, tilt and yaw angle between each pair of normal on the surface patch. In detail, for each pair of 3D point $\langle p_i, p_j \rangle$ and estimated surface normal $\langle n_i, n_j \rangle$, normal set of *angular deviations* can be estimated as can be seen in (1), (2), and (3),

$$\alpha = v \cdot n_j \tag{1}$$

$$\emptyset = u \cdot \frac{(p_j - p_i)}{d} \tag{2}$$

$$\theta = \arctan(w \cdot n_j, u \cdot n_j) \tag{3}$$

with u, v, w represent coordinate system of Darboux frame at p_i. Further, PFH on a *patch* from point $P = \{p_i\}$ with $i = \{1, \ldots, n\}$ capture all sets $\langle \alpha, \phi, \theta \rangle$ between all pair $\langle p_i, p_j \rangle$ of P, and save the result on a histogram. In the VFH, statistic of relative angle between surface normal at each point to surface normal at object centroid is used as additional viewpoint component which is calculated by collecting histogram of the angle at which the direction of each viewpoint to the surface normal.

3) Hybrid Features

We adopt the early fusion approach to combine several image descriptors [24], [25], by concatenate the feature vectors into a single vector. An image I is represented as $I = I(f)$, where f is a set of low level feature, namely color (f_{cld}) and shape (f_{VFH}). Basically, each feature can be modeled by several representations, in this research we use color layout descriptor (CLD) for representing color feature and viewpoint feature histogram (VFH) for representing 3D shape feature, as seen in (4). Each representation is itself a vector with multiple components.

$$f = [f_{cld}, f_{VFH}] \tag{4}$$

B. Hierarchical Multi-feature Classification

In hierarchical classification, we were given: (a) Set of class C which is organized into IS-A hierarchy called class hierarchy. (b) Set of fruit images. (c) Class proximity $\beta(C_i, C_j)$, which represents error value of classification error from class C_i to class C_j. The purpose is to find some rules, called classifier, which determine the class of fruit image with low error value. Quantitatively, the selection of $\beta(C_i, C_j)$, basically a measure of the proximity of the two members in the hierarchy, is very dependent on the application. Commonly, the shortest distance from C_i to C_j is used.

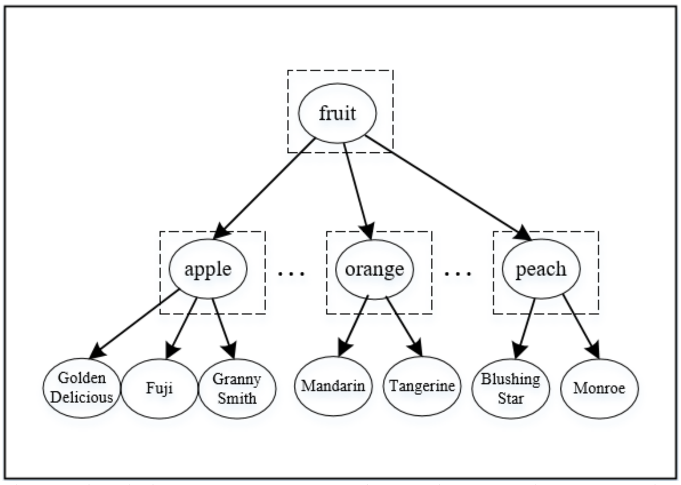

Fig. 4. The circles represent parent nodes. Dash squares in parent nodes represent multiclass classifiers, to predict their child classes

In our approach, we adopt the top down hierarchical classification approach, especially local classifier per parent node (LCP) approach [12]. A simple illustration of LCP can be seen in Fig. 4. In the LCP, for each parent node in the fruit hierarchy, a multiclass classifier is trained to distinguish between its child nodes. For example, suppose that the first level classifier assigns the test image to the class apple. The second level classifier, which was only trained with the children of the class apple, in this case Golden Delicious, Fuji, and Granny

Smith, will further perform its class assignment of the test image. This approach avoids the inconsistent predictions and gives value to the natural constraint of fruit membership.

C. Performance Measure

The accuracy is used in measuring performance of local classifier approach, as can be seen in (5), with TP = the number of true positives (correctly predicted positive examples), FP = the number of false positives (positive predictions that are incorrect), FN = the number of false negatives (positive examples that are incorrectly predicted negative), and TN = the number of correctly predicted negative examples.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (5)$$

Fig. 5. Dataset used in the experiment: 7 fruit species, 32 fruit variety.

IV. Experiment

In this experiment, we demonstrate the feasibility of our proposed approach for the multi-class fruit classification problem. We compare the recognition performance of fruit recognition using flat and hierarchical classification. Moreover, we also show the recognition performance of local and hierarchical classification approach.

Fig. 6. Point cloud images

The dataset we use in this experiment consists of 32 fruits, categorized into 7 fruit species (apple, banana, lemon, lime, orange, peach, and pear), as depicted in Fig. 5. This fruit is arranged into fruit hierarchy, as can be seen in Fig. 7. Meanwhile, samples of point cloud images are depicted in Fig. 6. In total, we use 21284 RGB images and 21284 point cloud images of fruit.

We build a 1-nearest neighbor classifier over hybrid features, with 10-fold cross validation, and for the performance measure, we use accuracy as can be seen in (5).

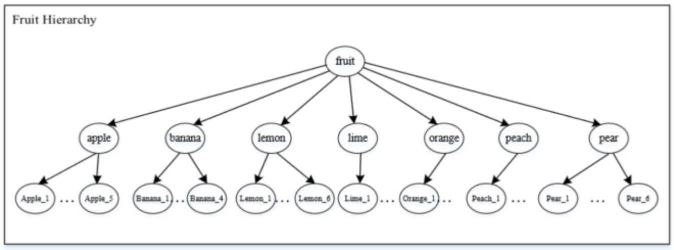

Fig. 7. Fruit hierarchy used in the experiment; 1 node at root level (level 0); 7 nodes at level 1, which is fruit species level (apple, banana, lemon, lime, orange, peach, pear); 32 nodes at level 2, which is fruit variety level.

A. Hybrid Descriptor

For the color feature, we use the CLD implementation as standardized by MPEG7 [26] as the color representation which is extracted from RGB images. Particularly, we use Y-coefficient = 10, Cb-coefficient = 3, and Cr-coefficient = 3.

Meanwhile, we use VFH implementation of Point Cloud Library (PCL) [27] for the shape feature descriptor which is extracted from point cloud images. In our experiment, we use 45 binning subdivisions for each of the three extended FPFH (Fast Point Feature Histogram) values [28], plus another 45 binning subdivisions for the distances between each point and the centroid and 128 binning subdivisions for the viewpoint component, which results in a 308-byte array of float values. For a given point cloud, only a single VFH descriptor will be estimated. In the normal estimation, the searching method used is kd-tree [29], with radius of search is 0.03.

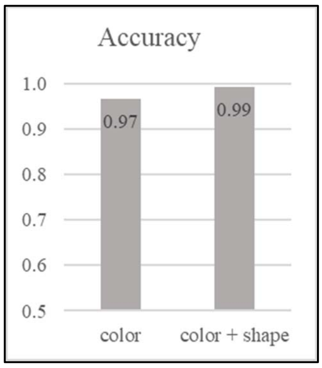

Fig. 8. Accuracy of hybrid descriptor in LCP approach compared to color descriptor

B. Result and Analysis

In the HMC approach, basically we conduct two classification steps based on the tree structure of our dataset. A total of 8 classifiers is used to train the training data to predict the fruit species and fruit variety. One classifier is used to train the training data to predict 7 fruit species, namely apple, banana, lemon, lime, orange, peach, and pear. Further, other 7 classifiers are used in the second step, to predict the fruit variety on each fruit species node. In particular, 1 classifier is used to classify 5 variety of apple, 1 classifier is used to classify 4 variety of banana, 1 classifier is used to classify 6 variety of lemon, 1 classifier is used to classify 4 variety of lime, 1 classifier is used to classify 4 variety of orange, 1 classifier is used to classify 3 variety of peach, and 1 classifier is used to classify 6 variety of pear. Generally, in LCP approach, the hybrid descriptor is able to improve the accuracy of 2.59 %. By using the color descriptor (CLD) only, the accuracy reached 96.70 %, whereas the accuracy reached 99.29 % if we use color and shape descriptor, as depicted in Fig. 8.

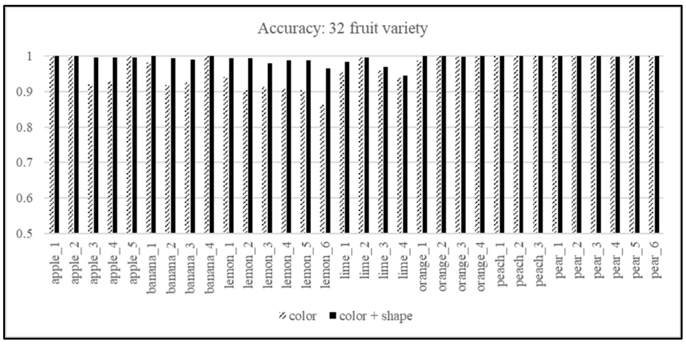

Fig. 9. Accuracy of 32 fruit variety by using color descriptor only and HMC

Further, in Fig. 9 we show the accuracy of each fruit variety, by using color and hybrid descriptor. We can see that the accuracy of some fruit variety is increased by using hybrid descriptor. This improvement is occurred in fruits that have similar color, such as apple_3 and apple_4, banana_2 and banana_3, and also all variety of lemon. In details, the accuracy improvement of some fruit variety is depicted in TABLE 1.

Although the color feature plays an important role in the fruit recognition system, but on the condition that there are many different fruits with similar color, applying color feature only to recognize test image is not enough. It can be found, for example, in apple_3 which has a very similar color to apple_4, or in banana_2 which has a very similar color to banana_3. In this case, the addition of shape feature is very important to recognize objects with colors that are very similar.

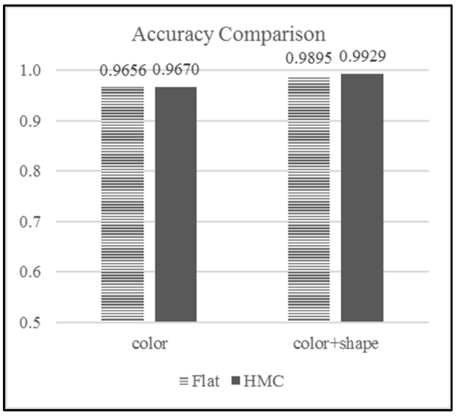

Fig. 10. Accuracy comparison of flat classification and HMC approach

Further, we compare our proposed HMC approach with flat classification. In the flat classification, we also make a hybrid of color and shape descriptor. We use CLD and VFH for color and shape descriptor, respectively. In the flat classification approach, 32 fruit variety is assumed to not have any relationship with other fruit variety. Nearest neighbor classifier is used in the classification process, with 10-fold cross validation. In Fig. 10 we show the accuracy comparison of flat classification and our HMC approach. We get 0.34 % accuracy improvement by using HMC approach. In particular, the fruit variety which have an increasing accuracy is lemon_1, lemon_2, lemon_3, lemon_4, lemon_5, lemon_6, lime_1, lime_2, and lime_3, as can be seen in Fig. 11. As we can see, the hybrid descriptor alone is efficient in recognizing fruit variety. But by improving it into HMC approach, we can still get the improvement in system accuracy.

TABLE I. ACCURACY IMPROVEMENT

Fruit variety	Accuracy improvement (%)
apple_3	7.75
apple_4	6.95
banana_1	1.79
banana_2	7.565
banana_3	6.297
lemon_1	5.23
lemon_2	9.02
lemon_3	6.68
lemon_4	7.97
lemon_5	8.41
lemon_6	10.19
lime_1	2.86
lime_3	0.83
lime_4	0.47
orange_1	1.13
orange_4	0.14

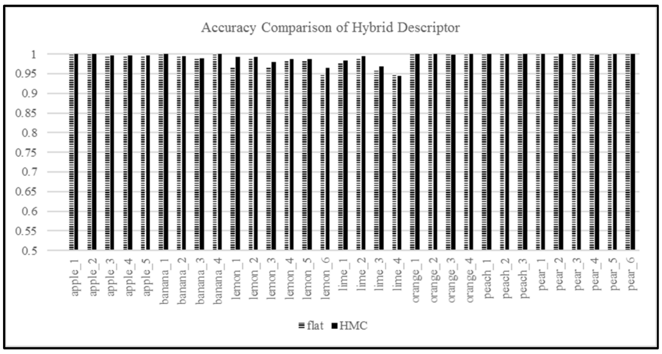

Fig. 11. Accuracy comparison of hybrid descriptor in flat classification and HMC approach

V. CONCLUSION

This paper presented an approach for the hybrid RGBD features and fruit hierarchy in the fruit recognition system, named as hierarchical multi-feature classification (HMC) approach. The approach adopts the local classifier per parent node approach applied to hybrid descriptor of RGBD images. Particularly, we combined the advantage of color descriptor and 3D shape descriptor into the hybrid descriptor. From the evaluation of the conducted experiment on fruit dataset, given the positive result, we can conclude that this approach should definitely be considered in the fruit recognition tasks.

ACKNOWLEDGMENT

We would like to thank the Ministry of Research Technology and Higher Education of the Republic of Indonesia (Kementerian Riset Teknologi dan Pendidikan Tinggi Republik Indonesia) for the financial support in conducting this research.

REFERENCES

[1] D. M. Bulanon, T. F. Burks, and V. Alchanatis, “Fruit visibility analysis for robotic citrus harvesting,” *Trans. ASABE*, vol. 52, no. 1, pp. 277–283, 2009.

[2] P. Li, S. Lee, and H.-Y. Hsu, “Review on fruit harvesting method for potential use of automatic fruit harvesting systems,” *Procedia Eng.*, vol. 23, no. 0, pp. 351–366, 2011.

[3] K. Kapach, E. Barnea, R. Mairon, Y. Edan, O. Ben-Shahar, and K. K. E. B. and R. M. Y. E. O. Ben-Shahar, “Computer Vision for Fruit Harvesting Robots: State of the Art and Challenges Ahead,” *Int. J. Comput. Vis. Robot.*, vol. 3, no. 1/2, pp. 4–34, Apr. 2012.

[4] W. Ji, D. Zhao, F. Cheng, B. Xu, Y. Zhang, and J. Wang, “Automatic recognition vision system guided for apple harvesting robot,” *Comput. Electr. Eng.*, vol. 38, no. 5, pp. 1186–1195, 2012.

[5] M. Vogl, J. Y. Kim, and S. D. Kim, “A fruit recognition method via image conversion optimized through evolution strategy,” in *Proceedings - 17th IEEE International Conference on Computational Science and Engineering, CSE 2014, Jointly with 13th IEEE International Conference on Ubiquitous Computing and Communications, IUCC 2014, 13th International Symposium on Pervasive Systems,* 2014, pp. 1497–1502.

[6] A. Toshev, “Shape Representations for Object Recognition,” University of Pennsylvania, 2011.

[7] R. M. Bolle, J. H. Connell, N. Haas, R. Mohan, and G. Taubin, “Veggie Vision: A produce recognition system,” in *IEEE Workshop on Applications of Computer Vision*, 1996, pp. 244–251.

[8] A. Rocha, D. C. Hauagge, J. Wainer, and S. Goldenstein, “Automatic Fruit and Vegetable Classification from Images,” *Comput. Electron. Agric.*, vol. 70, no. 1, pp. 96–104, 2010.

[9] Y. Zhang and L. Wu, “Classification of fruits using computer vision and a multiclass support vector machine.,” *Sensors (Basel).*, vol. 12, no. 9, pp. 12489–505, Jan. 2012.

[10] Y. Zhang, S. Wang, G. Ji, and P. Phillips, “Fruit classification using computer vision and feedforward neural network,” *J. Food Eng.*, vol. 143, pp. 167–177, 2014.

[11] H. Kuang, L. L. Hang Chan, C. Liu, and H. Yan, “Fruit classification based on weighted score-level feature fusion,” *J. Electron. Imaging*, vol. 25, no. 1, p. 13009, 2016.

[12] C. N. Silla and A. A. Freitas, “A survey of hierarchical classification across different application domains,” *Data Min. Knowl. Discov.*, vol. 22, no. 1–2, pp. 31–72, 2011.

[13] Z. Barutcuoglu and C. DeCoro, “Hierarchical Shape Classification Using Bayesian Aggregation,” in *IEEE International Conference on Shape Modeling and Applications 2006 (SMI’06)*, 2006, vol. 2006, pp. 44–44.

[14] I. Dimitrovski, D. Kocev, S. Loskovska, and S. Džeroski, “Hierarchical annotation of medical images,” *Pattern Recognit.*, vol. 44, no. 10–11, pp. 2436–2449, 2011.

[15] A. Binder, M. Kawanabe, and U. Brefeld, “Efficient classification of images with taxonomies,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 5996 LNCS, no. PART 3, pp. 351–362.

[16] J. Hernandez, L. E. Sucar, and E. F. Morales, “Multidimensional hierarchical classification,” *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7671–7677, 2014.

[17] K. Lai, L. Bo, X. Ren, and D. Fox, “A large-scale hierarchical multi-view RGB-D object dataset,” in *ICRA*, 2011, pp. 1817–1824.

[18] M. Blum, J. T. Springenberg, J. Wulfing, and M. Riedmiller, “A learned feature descriptor for object recognition in RGB-D data,” in *2012 IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 1298–1303.

[19] S. Tang, X. Wang, X. Lv, T. Han, J. Keller, Z. He, M. Skubic, and S. Lao, “Histogram of Oriented Normal Vectors for Object Recognition with a Depth Sensor,” in *Computer Vision – ACCV 2012*, vol. 7725, K. Lee, Y. Matsushita, J. Rehg, and Z. Hu, Eds. Springer Berlin Heidelberg, 2013, pp. 525–538.

[20] B. S. Manjunath, J.-R. Ohm, V. V Vasudevan, and A. Yamada, “Color and texture descriptors,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703–715, Jun. 2001.

[21] P. A. S. Kimura, J. M. B. Cavalcanti, P. C. Saraiva, R. S. Torres, and M. A. Gonçalves, “Evaluating Retrieval Effectiveness of Descriptors for Searching in Large Image Databases,” *J. Inf. Data Manag.*, vol. 2, no. 3, pp. 305–320, 2011.

[22] R. B. Rusu, G. Bradski, R. Thibaux, and J. Hsu, “Fast 3D recognition and pose using the Viewpoint Feature Histogram,” in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 2155–2162.

[23] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, “Learning informative point classes for the acquisition of object model maps,” in *2008 10th International Conference on Control, Automation, Robotics and Vision, ICARCV 2008*, 2008, pp. 643–650.

[24] L. Piras and G. Giacinto, “Information fusion in content based image retrieval: A comprehensive overview,” *Inf. Fusion*, vol. 37, pp. 50–60, 2017.

[25] J. Yu, Z. Qin, T. Wan, and X. Zhang, “Feature integration analysis of bag-of-features model for image retrieval,” *Neurocomputing*, vol. 120, pp. 355–364, 2013.

[26] T. Sikora, “The MPEG-7 visual standard for content description-an overview,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 696–702, 2001.

[27] R. B. Rusu and S. Cousins, “3D is here: Point Cloud Library (PCL),” in *2011 IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 1–4.

[28] R. B. Rusu, N. Blodow, and M. Beetz, “Fast Point Feature Histograms (FPFH) for 3D registration,” *IEEE Int. Conf. Robot. Autom.*, pp. 3212–3217, 2009.

[29] J. L. Bentley, “Multidimensional Binary Search Trees Used for Associative Searching,” 1975.