Research on Computer Vision-Based Object Detection and Classification

Juan Wu, Bo Peng, Zhenxiang Huang, Jietao Xie

Research on Computer Vision-Based Object Detection and Classification

Juan Wu[1], Bo Peng[1], Zhenxiang Huang[1], Jietao Xie[2],

[1]College of Information and Electrical Engineering, China Agricultural University, 100083, Beijing,China
[2]Suzhou Institute for Advanced Study, University of Science & Technology of China, 215100, Suzhou, China

{juanwu_cau, pengbo_cau, zhenxianghuang, xjtuxjt}@126.com

Abstract. Computer vision techniques become particularly important in agriculture applications due to their fast response, high accuracy and strong adaptability. Two of the most demanding and widely studied applications relate to object detection and classification. The task is challenging due to variations in product quality differences under certain complicate circumstances influenced by nature and human. Research in these fields has resulted in a wealth of processing and analysis methods. In this paper, we explicitly explore current advances in the field of object detecting and categorizing based on computer vision, and a comparison of these methods is given.

Keywords: Computer Vision, Detection, Classification.

1 Introduction

The interest of this topic in agriculture is motivated by the promise of many applications, such as agricultural and food product inspection [1]. Compared with traditional mechanical technologies, computer vision methods enable more efficient detecting and classifying. For example, categorizing fruit for uniformity can benefit from the advances. Based on computer vision methods can rapid grading fruit quality with respect to color [2], shape, crack and defect, and also have sufficient adaptability to variations, while traditional methods cannot.

Impacted by nature or other complicated factors, agriculture products are distinct with each other greatly. Some issues are often unavoidable, for example, lighting conditions vary over time, image blurs or tilts, or object is partly occluded. These are the reason why the topic is challenging, interesting and worth researching. In this paper, a detailed overview of current main advances in the field is provided. Object detection and classification are discussed independently. Moreover, we focus on color and shape these two salient features for detection.

2 Detection

The literature related to this topic is vast, and we attempt to present two main corresponding groups in the following sections: color-based and shape-based approaches. Shape and color features are two key indicators for fruit and vegetables when they are to be inspected or classified. If agriculture products were improperly fertilized or sprayed with insecticide, deformation or abnormal color will be induced.

2.1 Color-Based Approaches

Color information is the most significant visual feature in agriculture products and it has good performance on invariance of size and view point change. Color based approaches usually aim at segmenting out the desired colors of interest region for further processing. The main limitation with these methods is that they are very vulnerable to nature illumination change.

Color Thresholding. Color thresholding is particularly useful and simple for segmenting image with strong contrast between object and background. By setting the color value range and classifying pixels, thresholding can be done and object can be segmented. When a pixel matching, if its color value is among the reference range, the pixel can be called a required object pixel, if not it will be considered as a background pixel. RGB (Red-Green-Blue) is the most intuitionistic color space. Its advantage is fast enough for real-time performance, due to no need for color space transformation.

HIS or Other Color Space Transformation. HIS (Hue-Intensity-Saturation) color space has good consistency with human vision. Significantly eliminating the influence of brightness changes is its most attractive characteristic, by dividing an overall intensity value from saturation and hue two values. But it will take some time and hardware cost to transform the captured image represented by RGB value into HIS value. CIE model takes human visual characteristics and environmental conditions into consideration, and gets more immune to lighting. However, setting parameters is complicated when applied to specific environment. There are other more color spaces like HSV and L*a*b* [3].

Color Indexing. Color indexing is a fast, straightforward and effective detecting method. Its fundamental principle is comparing the colored objects of two images with their color histograms, as color histogram is used to index the object and store into the template database. The potential problem is its computation will grow enormously in complex outdoor scenes.

2.2 Shape-Based Approaches

We divide shaped-based approaches into two categories: boundary-based approaches and region-based approaches [4]. The former encodes the object's boundary, including Chain code, Fourier descriptor [5], Hough transform [6], spline estimation and bending energy. Many statistic features can be used to describe simple shape region, such as area, projection, Euler number, centrifugal rate, rectangular degrees and direction. As for complicate region, we could divide the observation into cells, each of which encodes part of the observation locally.

Object shape detection techniques are usually used for further segmentation. Most of them require big computation and attention about position, viewpoint and partial occlusion.

Fourier Descriptor. Fourier Descriptor is the Fourier transform coefficient of the shape boundary curve, which is a very classical method in transform domain. Its main advantages are based on mature theory of Fourier analysis, and invariant to position, rotation and size under certain circumstances. Also, high accuracy can be acquired with only a few low order coefficients. However, it is very sensitive to noise, and influenced by the curve shape and initial point's position.

Hough Transform. Hough transform is originally used for detecting lines and circles, but gradually its related transforms have been generalized to detect arbitrary shapes. The essence of these approaches is the transformation from image space to the defined parameter space. When the results accumulated in parameter space, it can be used to detect or describe the straight lines or curves in the image. This method is inherently robust, as noise, gap, partial deformation and occlusion can be handled [7], but appear a decreased strength. The potential issue is extensive computation and memory hungry.

Moment. Assuming grayscale image normalized as a probability density of bivariate random variable, the random variable can be described by moment. Moment is a kind of linear characteristic. It can be used to describe local features of binary or gray images. This approach in a manner is invariant to image translation, rotation and scale. Usually, moments include invariant moment, Legendre moment Zernike moment and others. Their shortcoming is difficult to present the object's local information and hard for high order moment to directly associate with object's visual features.

3 Classification

Relying on the same features extracted from the object, good classification approaches can get better results. Three commonly used methods exist are described below.

3.1 Template Matching

This method is mainly for searching maximum matching [8]. All required object features to be classified are stored in a database, including considerable variation in performance, viewpoint, and illumination. Each potential object is normalized in size and compared with every template of the same feature. Template matching is fast and easily to modify for new classes. Its potential issue is time-consuming when geometry transform is complex. The issue can be compensated with genetic algorithm due to its powerful searching ability, quick response and less templates for matching.

3.2 Neural Network

This method is currently one of most successful and widely used learning classification algorithms [9]. Its advantage is that it is tolerant to fault and is powerful to classify. Through training the neural networks with specific color, shape pictogram, features of interest region can be recognition. Its commonly models contain Back Propagation Network, Multilayer Perception Network, Hopfield Network, Radial Basis Network, etc.

3.3 Support Vector Machine

Support Vector Machine extended statistic learning theory to classification, employing Structural Risk Minimization principal to improve generalization capability. It is able to solve problems with respect to nonlinear [10], high dimension, local minimum and also invariable with viewpoint.

4 Discussion

Various commonly detecting and classifying techniques based on computer vision have been presented. A comparison of performances of these techniques is given in Table 1.

Color and shape are two essential features for detecting. Color-based approaches are fast and straightforward. The shortcoming is vulnerable with illumination changes, not suitable for weak or reflective conditions. Compared to color-based approaches, shape-based methods face more limitations. When detecting object's boundary, they should be immune to changes, shape imperfect, or partial occlusion. But the issue can be compensated with color-based approaches.

Neural network is very popular for classification. Its performance depends on which what network architecture is applied and how well the network is trained. Support Vector Machine is invariant to rotation, size, slightly tilted and even partial occlusion. Template matching is fast and easily to extend with new classes, but computationally costly when geometry transform is complex. However, if it combined with genetic algorithm, its response rate can be improved. Moreover, the hybrid method wouldn't require abundant templates to be prepared carefully.

Some of these approaches are robust but computation hungry, while others are simple but unable to tackle changes. Research in this field is significant and useful, deserving wider attention.

5 Conclusion

A research of the popular object detection and classification based on computer vision has been present in this paper. The research includes description of: (1) existing object detection methods associated with color or shape, (2) important methods developed to tackle the object classification problem, and (3) comparison of performances of these techniques.

Table 1. Comparison of detection and classification approaches.

Name	Advantage	Limitation
Color Thresholding	Simple, fast.	Sensitive to illumination.
HIS color Transform	Immune to lighting changes.	The transform will cost computation.
Color Index	Straightforward, quick.	Great computation time in complex scenes.
Fourier Descriptor	Invariant to viewpoint.	Accuracy degrades if curve shape or initial point is not desired.
Hough Transform	Allows slightly shape imperfect or occlusion.	Computationally costly and memory hungry.
Moment	Invariant to viewpoint.	Hard to present local information.
Template Matching	Quick and easily to extent for new classes.	Imperfect shape may pose a problem.
Neural Network	Dynamic handle changes, fast matching.	Demand large training data for good network.
Support Vector Machine	Invariant to viewpoint, allow partial occlusion.	Difficult for multi-category classification and massive samples.

References

1. Brosnan, T., Sun, D.W.: Improving Quality Inspection of Food Products by Computer Vision—a Review. Journal of Food Engineering. 61, 3--16 (2004)
2. Lee, D.J., K.Archibald. J, Xiong, G.M.: Rapid Color Grading for Fruit Quality Evaluation Using Direct Color Mapping. IEEE Transactions on Automation Science and Engineering, 8, 292--302(2011)
3. Mendoza, F., Dejmek, P., Aguilera, J. M.: CalibSrated Color Measurements of Agricultural Foods Using Image Analysis. Postharvest Biology and Technology, 41, 285--295 (2006)

4. Moreda, G.P., Muñoz, M.A., Ruiz-Altisent, M., Perdigones, A.: Shape Determination of Horticultural Produce Using Two-Dimensional Computer Vision – a Review. Journal of Food Engineering, 108, 245--261(2012)
5. Guo, C.L., Ma, Q., Zhang, L.M.: Spatio-temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform. In CVPR, pp.1--8, Anchorage, Alaska, USA (2008)
6. Maji, S., Malik, J.: Object Detection using a Max-Margin Hough Transform. In CVPR, pp.1038—1045, Miami Beach, Florida (2009)
7. Eduardo A, M.B., Miguel E, M.R, et al.: Implementation of Hough Transform for Fruit Image Segmentation. Procedia Engineering, 35, 230--239(2012)
8. Ren, F.X., Huang, J.S., Jiang, R.Y., Klette, R.: General Traffic Sign Recognition by Feature Matching. In IVCNZ, pp.409--414(2009)
9. Torres, M., Hervás, C., García, C.: Multinomial Logistic Regression and Product Unit Neural Network Models: Application of a New Hybrid Methodology for Solving a Classification Problem in The Livestock Sector. Expert Systems with Applications, 36, 12225--12235(2009)
10. Li, Z.F., Tang, X.O.: Bayesian Face Recognition Using Support Vector Machine and Face Clustering. In CVPR, vol. 2, pp.374—380, Washington, DC (2004)