Configurable Solution for Object Recognition and Localization based on Features in YUV Color Space

Xuming Chen, Simon X. Yang
University of Guelph
50 Stone Road East
Guelph, ON N1G 2W1, Canada

ABSTRACT
In this paper, a practicable robot vision solution is presented for multiple kinds of object recognition and localization. The proposed solution generalizes and extends simplicity driven algorithms for recognition and localization of fruits under harvesting environment. In the proposed solution, fruits with shape of quasi round such as oranges, apricots and apples on the tree are segmented from complicated real background with one dimensional V computing in YUV. Further object recognition algorithms are applied to recognize and compute each qualified fruit's localization. For a study of generality, the original solution is extended from fruit recognition to logistics and production industry. One is for recognition and localization of mixed kinds of cookies on conveyor belt in packaging line. Another one is for recognition of parts in warehouse or production place with a synthetic color labeling methods. V-U difference is used to enhance the difference among different kinds of objects. By a certain feature of one-dimension, multiple kinds of objects are distinguished from each other. Based on reconfigurable modeling, the recognition of multiple kinds of objects is realized on a single system.

General Terms
Object Recognition, Image Processing, Algorithms.

Keywords
Robot Vision, Fruit Recognition, Agriculture Application, Production and Logistics

1. INTRODUCTION
Robot vision is an intensely critical portion in robotic systems. A quick (in-time) and right object recognition by efficient and effective algorithms of image processing is the key to realize a real-time system. The requirement of low-cost is common and also necessary to make the application practicable. In [10], a practicable robot vision solution for recognition and localization of ripe tomatoes in green house was proposed with simplicity driven methods. The solution is generalized and extended in this paper for recognition and localization of other fruits on the tree with a quasi-round shape. With the configuration of features in YUV color space and sizes of the objects to be recognized, the proposed algorithm is adaptive to recognize the kinds of detected fruits with corresponding model settings.

The proposed reconfigurable real-time solution realizes multiple applications on a single system and makes its cost as low as possible. The proposed solution is applicable to the certain applications in production and logistics. In [1-8], several researchers studied the solution on object recognition of fruits on the tree. In the field of production and logistics, investigators apply object recognition in robotic operations [16-18]. The paper intends to find the common method to generalize the solution. The proposed solution is the first one to realize multiple applications on a single system within the range of current topic.

The features of U, V color components in YUV color space are unique. V represents redness in an object without affection of illumination and U represents background under redness based fruit harvesting environment. The result of V minus U reflects the better comparison of the color difference in different kinds of certain objects. A certain category of fruit is determined through a simple comparison of featured values. So the feature set is configured to adapt to the corresponding objects at the very beginning. Since algorithms and workflows are almost consistent, the capability of system reconfiguration makes one system to run multiple applications automatically with adaptive configuration.

The following sections represent the proposed solution and its applications in fruit recognition and object recognition in production and logistics. Section 2 studies the recognition algorithms of multiple kinds of objects and the generalized workflow of automatic harvesting in the real world. The first part of section 3 represents recognition algorithms of multiple kinds of fruits on the tree; the second part of section 3 studies the algorithm's applications on recognition and localization of mixed kinds of cookies on conveyor belt in packaging line, and on recognition and localization of parts in warehouse or production place through recognition of defined synthetic color labels. Section 4 states the experiments and results. The last section 5 summarizes key points in this paper.

2. OBJECT RECOGNITION BASED ONE DIMENSION COMPUTING OF COLOR COMPONENTS
Color is the critical feature of separating ripe tomatoes and many kinds of fruits on the tree from the stems, foliages and background in the greenhouse or fields according to the definition of ripeness. In [10]-[13], the authors had a detailed description of color spaces. The RGB color model produces various colors via integrating three base colors - red, green and blue. HSV color space is cylindrical coordinate representations of RGB points that rearrange the geometry of RGB to be more perceptually relevant. Based on CIE XYZ color space coordinates, L*a*b color space represents the luminance with L and colors with *a*b. YUV color model takes human perception into account with coding. The main reasons to select YUV as the color space of our research for real-time application are listed as follows: a) YUV enables masking transmission errors or compression artifacts through reducing bandwidth for chrominance components; b) YUV are the original video color signals following analogue-to-

digital conversion, providing the feasibility of optimization through hardware/software trade-off; c) especially, the V, U signals in YUV color space can be the best representation of pure colors without illumination changes.

A previous research [14] verified that V color component is able to realize object recognition of ripe tomatoes through one-dimension computing. The following is the ROI (Region of Interest) segmentation process for ripe tomatoes: 1) Acquire an image of tomato picking-off area in video from far vision camera; 2) Split the image in YUV space into three sub-spaces Y, U, V; 3) Extract object candidates with a featured threshold value of V in YUV; 4) Execute the object recognition in YUV color space. This section represents extensions and generality of the above methods.

2.1 V Components and Object Segmentation of Fruits on the Tree

Since illumination varies with different environments, the color components V is selected for detection of redness based fruits on the tree without the affecting of illumination. Some of typical picking-off areas of fruits in the orchard are shown in Fig. 1(a), 3(a), 4(a). The k-means classification is used for studying the distinguishing fruit objects from the others and background in the image. Without losing the generality, the oranges in orchid are taken as an example to present the result of the research. From the experimental results with different color spaces and selected color components, R, G and B are used to classify (with K-means) the ripe orange mixed with green leaves and stems, the result of classification is affected by illumination. The HSV classifier gets the same result as RGB classifier, the L*a*b and YUV three-element classifier don't work well for the segmentation of ripe oranges. Without the element L in L*a*b color space, *a*b two-element classifier works for ripe orange segmentation. Without the element, Y in YUV color space, UV two-element classifier is successful for ripe orange segmentation.

The most efficient results come from a single color component V in YUV color space, as shown in Fig. 1(b). In Fig. 2, the histogram is obviously able to be separated from two parts at a certain point, which splits the object with redness and the noise and background without redness. Therefore, ROI segmentation of ripe oranges works with V component in YUV color space. Theoretically, the single V color component classifier is reasonable to realize ROI segmentation of ripe oranges. According to one of the YUV to RGB conversion formula

$$R \quad = 1.164(Y - 16) + 1.596(V - 128)$$

V component of YUV color space is the portion of red color in the true color. V component represents true red color without illumination. By using the V component, the classification calculation is only with one dimension operation. In this way, it will decrease the complexity of design both for hardware and software dramatically. It also makes the object recognition more reliable and easier. By the way, it is possible for us to do certain fine-adjustment with one dimension histogram conveniently for the better performance of V component classifier. With k-means clustering algorithm, the V component classifier separates the biggest ripe orange region as ROI from its background and noises. Through learning and training of samples, the best criteria for classification are available, with which ROI is segmented out from background and other region (see Fig. 1). In Fig. 1, the ripe oranges (class 1) are segmented from green leaves, stems and background (class 2) through a K-means classification of 2 classes, and the part in Fig. 1 (b) is selected as a ROI firstly. In Fig. 3, the ripe apricots (class 1) are segmented from green leaves, stems and background (class 2) through a K-means classification of 2 classes. Since V is a pure color in YUV, a qualified segmentation of ripe apricots is achieved without the affecting of illumination. The case of apples is shown in Fig. 4. The experimental results show us the V component is able to separate oranges, apricots and apples from their stems, foliages and background. Complete results of recognition and localization of oranges are shown in Fig. 1(a-d).

2.2 V-U Difference and Enhanced Object Segmentation

According to the formulas in YUV color space, U and V valued 128 refer to no color. V value of rising from 128 to maximum means more redness; U value of rising from 128 to maximum means more non-redness. Absolute value of V-U difference is a measuring rule to represent the degree of red color versus background color in objects. The results from object segmentation of several kinds of fruit images are listed in the Table 1.

Table 1. The average U, V threshold values and V-U difference of fruits on the tree

Fruits	Apple	Apricot	Oranges
U Value	119	115	101
V Value	141	139	151
V-U Diff	21	31	45

Table 2. The average U, V threshold values and V-U difference of cookie on the packaging conveyor belt

	Cookie A	Cookie B	Cookie C
U Value	109	125	122
V Value	143	132	137
V-U Diff	34	7	15

2.3 Distinguishing Multiple Kinds of Objects based One Dimension Computing

The workflow of distinguishing multiple kinds of objects is listed as follows: 1) Acquire an image of object area in video from far vision camera; 2) Split the image in YUV space into three sub-spaces Y, U, V; 3) Extract object candidates with a featured threshold value of V or U in YUV; 4) Calculate the V-U difference of the above extracted objects; 5) Determine the kind of objects and trigger reconfiguration of the corresponding setting for this kind of objects; 6) Do further object recognition and localization.

Certain strategies are necessary to be studied for the different application scenarios as described in the following.

Scenarios for Fruit Recognition

a) Multiple kinds of fruits in one orchard: when there are more than one kinds of fruits in orchard, the proposed method may be used for different kinds of fruits one by one.

b) Multiple kinds of fruits in several orchards: the configuration needs to be done per orchard's situation.

c) Mixed kinds of fruits in packaging line: when two or more kinds of fruits need to be recognized, the proposed method can be repeated the same times as the number of the fruit kinds.

Scenarios of for Object Distinguishing in Production and Logistics

a) Mixed kinds of cookies on conveyor belt in packaging line: a direct recognition can be applied with enhancing the similar method as that used in fruit recognition. The experimental result is shown in Fig. 5.

Scenarios in Warehouse or Production Place

a) Small-size parts on the shelves: assign a label with white or black background on box or object surface or a certain position on the shelf

b) Large-size materials on the floors: assign a label with white or black background at determined position within a certain range of the materials.

c) Colored parts and materials: if the parts are circle shape in the view plane, a direct recognition can be applied with the similar method as that in cookie recognition; if the parts' colors are in a limited range of V-U difference, a special colored label whose V-U difference is apart from the parts' V-U difference can be used; if the parts' colors are distributed with a diversity of V-U difference, a special shape of label with a special curvature can be used to outstand the labels for recognition.

d) Large amount of parts' kinds: when a large amount of parts' kinds are encountered to be recognized, two or more differently sized concentric circles can be used. Theoretically, if 6 colors are used, then totally 6! / (6-2)! = 30 kinds of two concentric circle labels or 6! / (6-3)! = 120 kinds of three concentric circle labels are acquired.

The steps for recognizing two concentric circle labels are described as the following: 1) Recognize the inner circle first based on its color; 2) Change the inner circle to the same color as that of the outer circle; 3) Recognize the outer circle based on its color.

In Fig. 6(a), 7(a) and 8(a), the part HubA is assigned with a bigger sized label, the part HubB is assigned with a medium sized label, and the part HubC is assigned with a small sized label. After background removal as shown in Fig. 6(b), 7(b) and 8(b) separately, the target objects' labels are recognized individually as shown in Fig. 6(c), 7(c) and 8(c). The corresponding target objects' locations are acquired through the recognition of labels instead of direct recognition of the complicated parts.

In Fig. 9(a), 10(a) and 11(a), the part WheelA is assigned with a bright red-colored label, the part WheelB is assigned with a medium red-colored label, and the part WheelC is assigned with a dark red-colored label. All of the three labels are the same sized ones. After background removal as shown in Fig. 9(b), 10(b) and 11(b), the target objects' labels are recognized individually as shown in Fig. 9(c), 10(c) and 11(c). The corresponding target objects' locations are acquired through the recognition of different levels of redness instead of direct recognition of the complicated parts.

2.4 General Model and Configuration for Multiple Kinds of Objects

2.4.1 General Model and Configuration for Multiple Kinds of Fruits

The general model is generated by general feature sets of all kinds of fruits to be recognized. The certain evidences of recognition are criteria to meet in the recognition tasks. Some of evidences of recognition can be relaxed if other evidences are enough to distinguish the target object from other objects in the candidate pool. The model is built as follows: a) Value of V-U difference for distinguishing the kinds of fruits; b) Value of V component for object segmentation with redness; c) Value of U component for enhanced object segmentation; d) Curvature for object size recognition.

2.4.2 General Model and Configuration for Parts and Materials

The general model is generated by feature sets of all kinds of synthetic labels to be recognized, the labels are attached at a certain position on parts, shelves or floors. The evidences for recognition are the criteria to distinguish target object from other objects in the candidate pool. The following model is built as follows: a) Value of V-U difference in the biggest circle of object label distinguishing corresponding to further object recognition options; b) Value of V component for object label segmentation; c) Value of U component for object label segmentation; d) Area for object label size recognition.

3. RECONFIGURABLE SYSTEM SOLUTION FOR RECOGNITION OF MULTIPLE KINDS OF OBJECTS

3.1 Recognition of Multiple Kinds of Fruits

The proposed solution is adaptive and reconfigurable for different kinds of fruits with corresponding settings as first step of recognition tasks.

3.1.1 Adaptive Configuration

V-U difference distinguishes the different kinds of fruits from each others. The adaptive configuration switches the settings based on V-U difference values of the segmented object images. According to the V-U difference, the system loads the corresponding model settings. Table 1 and 2 list the model settings for certain fruits.

3.1.2 Feature-matching for Object Recognition

The proposed solution is for recognition of fruits with quasi-round shape such as apples, oranges, apricots etc. V single component can be used to segment out the fruits completely without foliages, stems and background left. Through detecting the arc edges with a certain curvature, circle objects are found where they are. Each center point can also be calculated through three points on the selected arc edge that is part of the perimeter of a fruit object. This is a simplicity driven recognition and localization solution for fruit on the tree in the real world that is extended from our solution for ripe tomatoes in greenhouse [14].

3.2 Reconfigurable Solution for Parts Recognition in Production and Logistics

The object recognition is widely applied in automated robotic system in the field of production and logistics [16-18]. In this paper, the proposed method can be extended into the applications in production and logistics. Parts may be very small or very big and have simple or very complicated shapes. Robotic systems have different operation processes to handle the corresponding parts. A labeling method is proposed to identify the different kinds of parts in warehouse to realize the reconfiguration of robotic system.

3.2.1 Synthetic Colored Labels

The following synthetic colored labels are applied for the proposed solution:

Designated Shapes

a) The shape of labels can be one round pie (A) with a certain radius R;

b) Two differently sized concentric round pies, one's radius is R (A) and another one's radius is 2/3*R (B), and one pie overlays another one. Or,

c) Three differently sized concentric round pies, the big one's radius is R (A), the median one's radius is 2/3*R (B) and the small one's is 1/3*R (C)

Designated Colors

Three round pies (A), (B) and (C) are painted with different levels of colors for different kinds of parts. If eight levels of colors are designated, they are numbered as 1, 2…, and 8 individually.

Encoding of the Labels

a) One round pie is used as a label: the number of the pie colors are the codes for labels;

b) Two or three round pies are used as a label: the combination of the numbers indicating the pies' colors is encoded as recognition IDs. For example, the codes of a two-pie labels with four levels of colors are 12 13 14 23 24 34 21 31 41 32 42 43;

c) The labels are used for a large amount of parts' kinds: three round pies with different values of radius are encoded with changing the ratio of the radius of the three round pies. Theoretically this proposed label solution is capable to meet the recognition requirements of a large amount of parts' kinds.

3.2.2 Adaptive Configuration

V-U difference distinguishes the different kinds of synthetic circle labels from each other for corresponding models for the recognition and further robotic operations. The adaptive configuration switches the settings based on V-U difference values of the segmented, synthetic and colored label images. According to the V-U difference, the system loads the corresponding model settings. Table 3 lists the model settings for certain Parts.

3.2.3 Feature-Matching for Object Recognition

According to different scenarios, the proposed solution is described as follows: a) The parts are of round shape and painted with different colors directly. The same solution can be used as that for fruit recognition. The circle objects' edges are recognized by detecting the arcs with certain curvatures. The coordinate values of center points are calculated through every three points on the selected arc edges that are portions of the perimeters of the parts objects. b) The parts are not of round shape, in this scenario, certain synthetic circle labels are created as recognition objects for disguising the target objects from other objects. Each of the different parts needs further processing for robotic operation; this is beyond the topic of this paper.

4. CONCLUSIONS

Fruits such as oranges, apricots and apples can be segmented from complicated and uncontrolled harvesting environments through a single V color component in YUV color space and then be recognized and localized. The proposed solution can be extended to the applications in logistics and production industry. V-U difference enhanced the solution's effectiveness. Multiple kinds of objects can be distinguished from each other by a certain feature of one-dimension. The recognition of multiple kinds of objects can be realized on a single system based on reconfigurable modeling. The proposed reconfigurable real-time solution can realize multiple applications on a single system. The current application designs are in the field of fruit recognition and in production and logistics, and extendable into more fields or industries. The considerations for the solution design are simplicity driven, low cost and practicability.

Table 3. The average V threshold values and V-U difference of the parts in warehouse

	Hub A	Hub B	Hub C	Wheel A	Wheel B	Wheel C
U value	128	128	127	127	127	128
V Value	129	129	129	131	130	129
V-U Diff	1.3	1.43	1.71	3.78	2.32	1.10

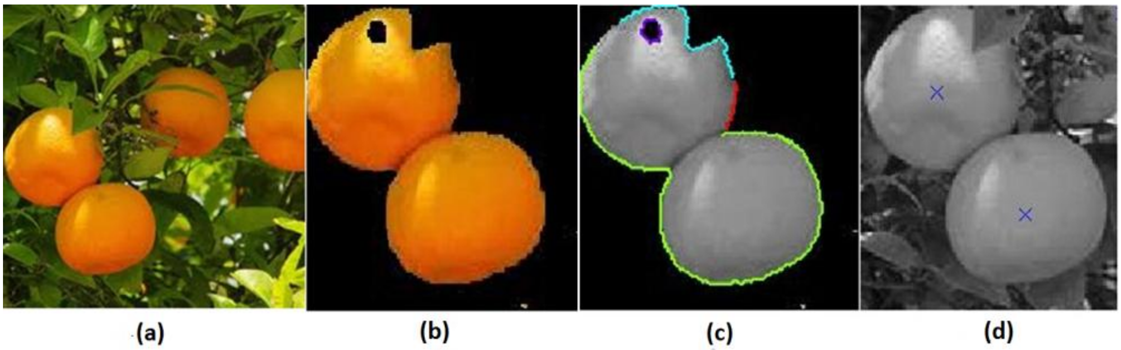

Figure 1 (a) an original image (b) the result of segmentation with V (c) extraction of featured edges (d) the result of localization

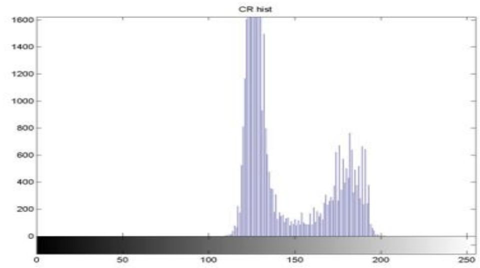

Figure 2 Histogram of V for the orange image in fig. 1(a)

Figure 3 (a) an original image of apricot (b) the result of segmentation with V

Figure 4 (a) an original image of apple (b) the result of segmentation with V

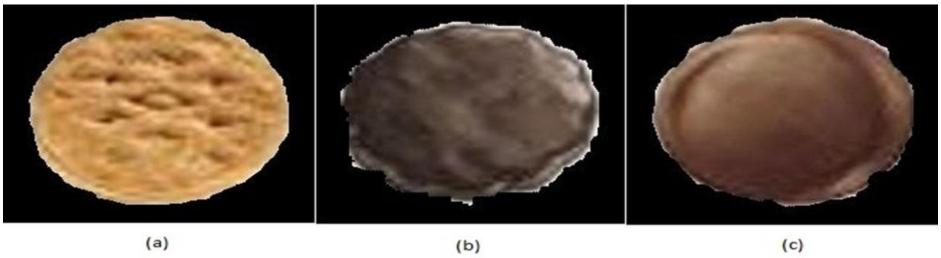

Figure 5 Cookies are recognized by V-U (a) cookie 1(b) cookie 2(c) cookie 3

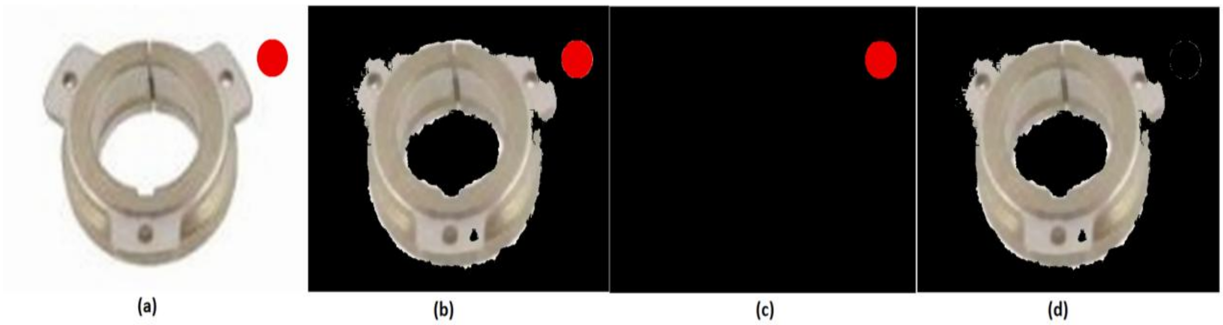

Figure 6 (a) original HubA + big size label (b) segmented HubA + label (c) segmented label only (d) the part segmented

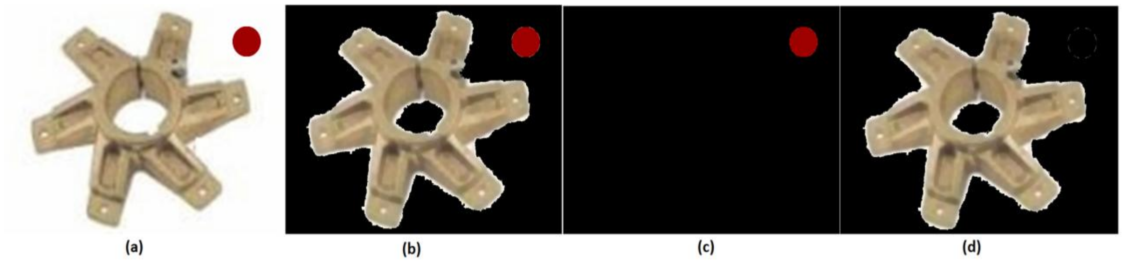

Figure 7 (a) original HubB + medium size label (b) segmented HubB + label (c) segmented label only (d) the part segmented

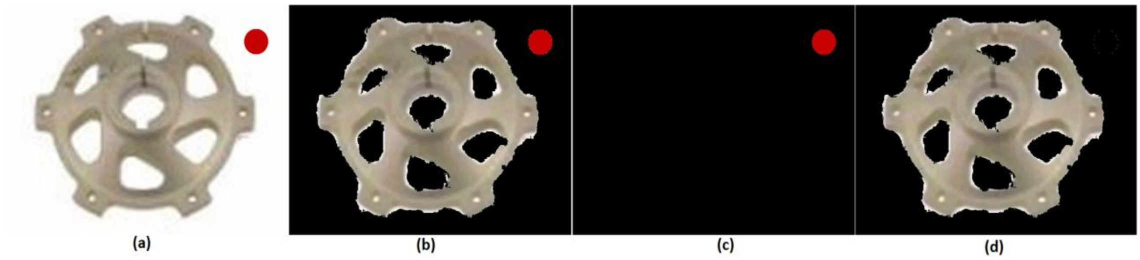

Figure 8 (a) original HubC with small size label (b) segmented HubC + label (c) segmented label only (d) the part segmented

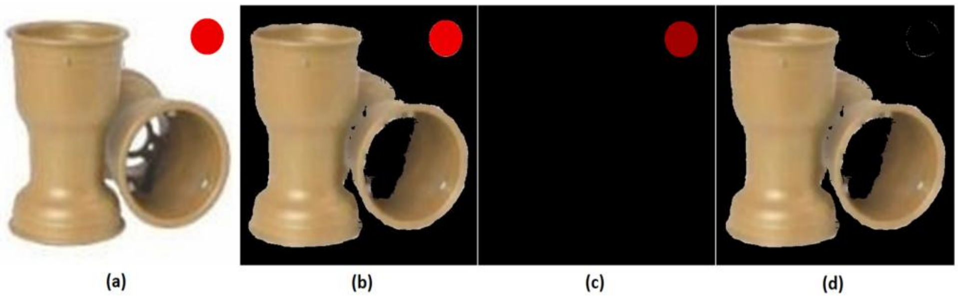

Figure 9 (a) original WheelA + high CR label (b) segmented WheelA + label (c) segmented label only (d) the part segmented

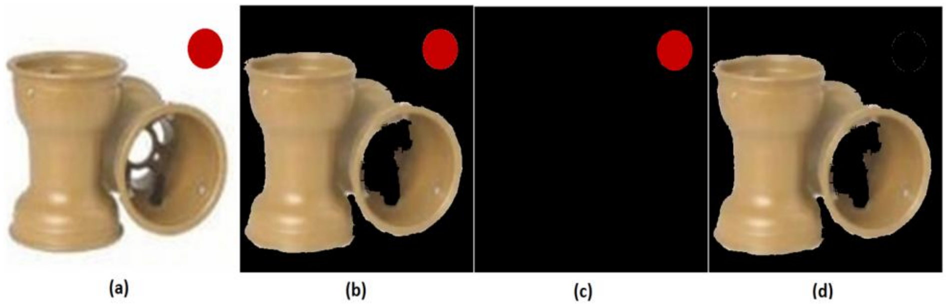

Figure 10 (a) original WheelB + medium CR label (b) segmented WheelB + label (c) segmented label only (d) the part segmented

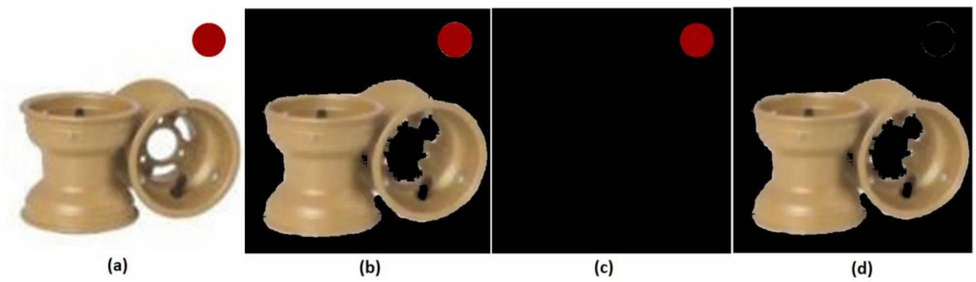

Figure 11 (a) original WheelB + low CR label (b) segmented WheelC + label (c) segmented label only (d) the part segmented

5. REFERENCES

[1] Schertz, C.E., Brown, G.K.: Basic considerations in mechanizing citrus harvest. Transactions of the ASAE. 11, 343–348(1968)

[2] Monta, M., Kondo, N., Fujiura, T.: Fruit harvesting robots in japan. Advances in Space Research, 18,181–184 (1996)

[3] Sakaue O., Hayashi, S.: Basic considerations in mechanizing citrus harvest. Annual Inter-national Meeting, ASAE. 963–967 (1996)

[4] Kondo, N., Taniwaki,S., Tanihara K., Yata, K., Monta. M., Kurita, M., Tsutumi, M.: An End-Effector and Manipulator Control for Tomato Cluster Harvesting Robot. ASABE. 073114 (2007)

[5] Kondo, N., ENDO, S.: Studies on Fruit Recognizing Visual Sensor. Journal of the Japanese Society of Agricultural Machinery, 49(5), 476-482 (1987)

[6] Kondo, N., Shibano Y.: Studies on Image Recognition of Fruit Vegetables (Part 1). Journal of the Society of Agricultural Machinery Japan, 53(5), 52-58 (1991)

[7] Kondo, N., Shibano, Y., Mohri, K., and Monta, M.: Masic Studies on Robot to Work in Vineyard (Part 2). Journal of the Japanese Society of Agricultural Machinery, 56(1), 45-53 (1994)

[8] Yin, Hongpeng, Yi, Chai, Yang, Simon X., Mittal, Gauri S.: Ripe Tomato Recognition and Localization for a Tomato Harvesting Robotic System. International Conference of Soft Computing and Pattern Recognition, 557-562 (2009)

[9] Chen, X., Yang, S. X.: High Efficient ROI Segmentation with Single Color Classifier. The 11th International Conference on Control, Automation, Robotics and Vision: ICARCV (2010)

[10] T. Horiuchi: Grayscale Image Segmentation Using Color Space. IEICE T. Inf. Syst., Volume E89-D 3, 1231-1237(2006)

[11] M. D. Fairchild: Color Appearance Models. Addison-Wesley, Reading, Mass. (1998)

[12] Keith Jack: Video demystified: A handbook for the digital engineer. Brooktree Corporation (1993)

[13] M. Tkalcic and J. Tasic: Colour spaces - perceptual, historical and applicational background, 304-308(2003)

[14] Chen, Xuming, Yang, Simon X.: A practical solution for ripe tomato recognition and localisation. Journal of Real-Time Image Processing: Volume 8, Issue 1, 35-51 (2013). [DOI>10.1007/ s 11554-011-0222-9]

[15] N. Kondo, S. Taniwaki, K. Tanihara, K. Yata, M. Monta, M. Kurita, M. Tsutumi: An End-Effector and Manipulator Control for Tomato Cluster Harvesting Robot. ASABE. Paper no. 073114 (2007)

[16] James A.Tompkins and Jerry D. Smith: The warehouse management handbook,Tompkins Press(1998)

[17] Rene de Koster, Tho Le-Duc, and Kees Jan Rood-bergen: Design and control of warehouse order picking: A literature review. European Journal of Operational Research, 182(2):481-501(2007)

[18] Oytun Akman and Pieter Jonker: Object recognition and localisation for item picking, In Automation in Warehouse Development, 153-162(2012)