

# A Method for Color Classification of Fruits Based on Machine Vision

CHANGYONG LI, QIXIN CAO and FENG GUO

Research Institute of Robotics  
Shanghai Jiao Tong University  
800 Dong Chuan Road Shanghai  
CHINA  
LLCYong@sjtu.edu.cn

*Abstract:* - A dominant color histogram matching method for fruits classification was presented in this paper. In classification of fruits based on machine vision, image was acquired with a color CCD camera that outputted color information in three channels, red, green, and blue. Because traditional RGB color space couldn't meet subjective color sensation of human being, so color image needed to be transformed from RGB to HSV color space which represented human being's subjective color knowledge. However, the conversion result was still three-dimensional information that made determining color grades very difficult. A new color space conversion technique that could be implemented for high-speed real-time processing for color grading was introduced in this paper. The result of this technique was a simple one-dimensional array that represented different color levels. These colors were known as dominant colors of fruits. The technique reduced computation consumption greatly. Color histogram as a statistical feature had visual invariance and high robustness. The dominant color histogram matching method was used for color grading. Grade judgment result was given by calculating and comparing the similarity between the inspected sample histogram and standard template histogram for every grade, fruit sample would be assigned to the grade whose template had the biggest similarity with it. Experiment results show that dominant color histogram matching method has high accuracy in fruits' color classification.

*Key-Words:* - fruit classification, color grading, color classification, histogram matching, dominant color, machine vision, feature extraction

## 1 Instruction

China produces a large number of fruits every year. However fruits quality inspection is still executed manually now, which is characterized as labor intensive, slow, sometimes inconsistent due to operator fatigue and high staff turnover caused by boredom, so machine vision based fruits inspection system has large market potential in China.

Color is an important quality characteristic of fruits. It represents the degree of maturity, sugar content, acidity, taste etc. Some colors are more preferable

than others. For example, in fresh fruit market such as apples and peaches, darker red represents higher quality than the light red colors. Fruits with preferred color generally can be sold for higher price. Color grading based on machine vision technology has become a primary way to maintain consistent quality and increase the fruit value<sup>[1-2]</sup>. Some applications for color grading in Agricultural Products such as fresh market apples<sup>[3-7]</sup>, peaches<sup>[8-10]</sup>, tomatoes<sup>[11-12]</sup>, potatoes<sup>[7]</sup>, peppers<sup>[13]</sup> and cucumber<sup>[14]</sup> have been developed.

The objective of this research is to introduce a new

color grading technique which is simple and easy to implement. In color machine vision applications, image is acquired with a color CCD camera that outputs color information in three channels, red, green, and blue. Because traditional RGB color space can not meet subjective color sensation of human being, so color image need be transformed from RGB to HSV color space which can represent human being's subjective color knowledge. However, the conversion result is still three-dimensional information that makes determining color grades very difficult. A new color space conversion technique that can be implemented for high-speed real-time processing for color grading is introduced in this paper. The result of this technique is a simple one-dimensional array that represents different color levels. These colors are known as dominant colors of fruits. The technique reduces computation consumption greatly.

Color histogram characterizes the distribution frequency of each color in images, as a statistical feature, it is not sensitive to the changes of scale, translation and rotation. So it has a visual invariance and a high robustness. The dominant color histogram matching method is used for color grading. Grade judgment result is given by calculating and comparing the similarity between the inspected sample histogram and standard template histogram for every grade, fruit sample will be assigned to the grade whose template has the biggest similarity with it. Figure 1 shows the classification process with histogram matching method.

Compared with the traditional qualitative classification method, the classification method based on the dominant color histogram matching does not need specific training, and the design of the classification is simplified. The classification system is able to meet the practical requirements.

In this paper, tomatoes were selected as samples to illustrate how the method for color classification of fruits based on machine vision works.

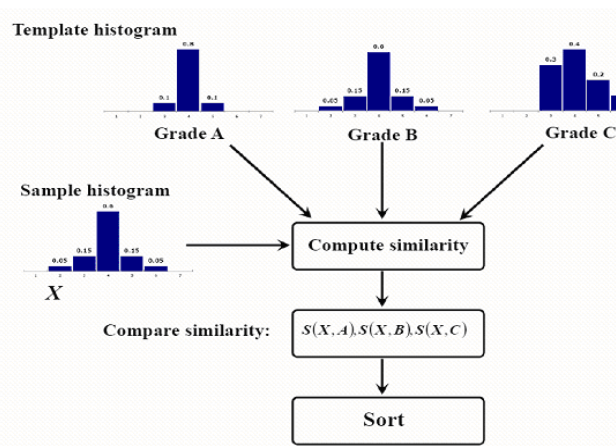


Fig1. Schematic diagram of histogram matching method

## 2 Image pre-processing

### 2.1 Background segmentation

The digital image of fruits (on the production line) grabbed by camera is composed of fruit and background. In order to obtain fruit's feature, the fruit need to be separated from background. It is very important to choose a proper color space for effective image segmentation. The OHTA color space used in this paper was proposed by Ohta [15], who analyzed more than 100 color features which were thus obtained during segmenting eight kinds of color pictures and found a set of orthogonal color features. Compared with other traditional color spaces, the conversion from RGB to OHTA color space is linear and computation is inexpensive. OHTA color space can be effectively applied on color image segmentation. OHTA color space has two different kinds of expression as shown in equation (1). In this paper, Fruit segmentation is achieved by a threshold algorithm with the  $I_2'$  feature (R-B).

$$\begin{cases} I_1 = (R + G + B) / 3 \\ I_2 = (R - B) / 2 \\ I_3 = (2G - R - B) / 4 \end{cases}, \begin{cases} I_2' = (R - B) \\ I_3' = (2G - R - B) / 2 \end{cases} \quad (1)$$

where R、 G and B are RGB color space characters.  $I_1$ 、  $I_2$  and  $I_3$  are OHTA color space characters.  $I_1$  reflects grey character of image ,  $I_2$  and  $I_3$  reflect an

image's color character.

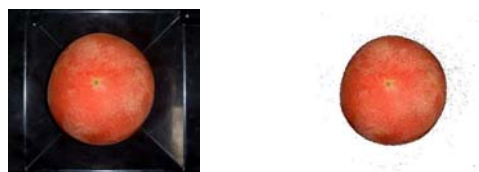
The image taken by camera is transformed from RGB color space to OHTA color space firstly, then threshold based method is selected for image segmentation. Due to the illumination intensity was controlled. So fixed threshold value was selected to do segmentation. The segmentation process is as follows:

$$G_k(i, j) = \begin{cases} G_k(i, j), & \text{if } (T \leq I_2(i, j)) \\ 255 & \text{others} \end{cases} \quad (2)$$

In equation  $i$  and  $j$  are the coordinates of a pixel in an image.  $G_k$  denotes gray value of Pixel,  $k \in \{R, G, B\}$ .  $I_2$  is gray value in OHTA color space.  $T$  is threshold value.

If the  $I_2$  value of a pixel is bigger than the threshold, then the pixel belong to fruit, and it's gray values will not be changed; if the  $I_2$  value of a pixel is smaller than the threshold, then the pixel belong to background and it's gray values will be changed to be 255. In this way, the fruit is separated from the background image, and consumption of computing is significantly reduced because the follow-up algorithm will only work on the area of fruit.

The segmentation result of image based OHTA color space was shown in Fig.2.



(a) Original Image      (b) Segmentation result

Fig.2 image segmentation result

## 2.2 Color Image Filtering

Because of the disturbances from external or internal of the system, there are a certain degree of unavoidable noises in the image obtained from fruit grading lines. These noises will influence the

accuracy of the feature extraction algorithm, so it is necessary to filter out image's noises.

The traditional methods for color image filtering are to filter the noises in each color channel separately, and then compound the results after filtering, export the target image. In this way, there might produce new colors not in the original image so that it distorted the image color. Vector filtering method treats the pixel colors as a three-dimensional vector. It will not create new color which is not in the original image. Therefore, it is generally believed that color image vector filtering method is more reasonable and effective than scalar filtering method<sup>[16]</sup>.

Vector median filtering method<sup>[17]</sup> is an effective non-linear filtering method, have a very good filtering effect on pulse noise, scan noise and so on. It can retain images' details well. This paper, according to the actual situation of fruit classification, improved the vector median filtering method, and filtered out the noises in fruit image. The procedure of filtering noise is listed below:

1) With the R、G、B value of each pixel of the fruit part being obtained from image file, the pixel is taken as the center to create a sliding window whose size is  $(2r+1) * (2r+1)$ . The window radius  $r$  is an integer, bigger than or equal to 1.

2) Firstly, whether the center pixel of the window is a noise pixel or not should be distinguished by the equation (3). If it is not a noise pixel, then the algorithm return to step1 and continue to deal with the next pixel; If it is a noise pixel, the other pixels of the window also need to be judged whether it is a noise pixels or not by the equation (3). Then algorithm implement step 3.

$$f(i, j) = \begin{cases} 1 & \text{if } (|R(i, j) - \mu_R| \geq \frac{3}{4} \Delta_R) \\ 1 & \text{if } (|G(i, j) - \mu_G| \geq \frac{3}{4} \Delta_G) \\ 1 & \text{if } (|B(i, j) - \mu_B| \geq \frac{3}{4} \Delta_B) \\ 0 & \text{others} \end{cases} \quad (3)$$

Where  $\mu_i$  is the average value of gray level of all pixels in a sliding window and calculated with

equation (4).

$$\mu_i = \frac{\sum_{j=1}^N c_{ij}}{N} \quad (4)$$

$\Delta_i$  is the difference between the maximum gray value and minimum value of all pixels in a sliding window and calculated with equation (5).

$$\Delta_i = \max(c_{ij}) - \min(c_{ij}) \quad (5)$$

In which  $i \in \{R, G, B\}$ .  $j$  is the total number of pixels in window.  $C_{ij}$  denotes each color value in R、G、B. For a color pixel, if  $f(i, j)=1$ , the pixel is a pixel noise; otherwise the pixel is not a pixel noise.

3) Calculate the RGB component's weighted average value of all the pixel in the sliding window to make a weighted average vector  $V_A = (\bar{r}, \bar{g}, \bar{b})$ , in which  $\bar{r}, \bar{g}, \bar{b}$  values are defined as follows:

$$\begin{cases} \bar{r} = \frac{1}{N_w} \sum_{i=1}^M w_i r_i \\ \bar{g} = \frac{1}{N_w} \sum_{i=1}^M w_i g_i \\ \bar{b} = \frac{1}{N_w} \sum_{i=1}^M w_i b_i \end{cases} \quad (6)$$

In the equation:  $w_i$  is Weight value, the weight value

of noise pixels  $w_n = 0$ ,  $N_w = \sum_{i=1}^M w_i$ ;  $M$  is the

number of pixels in the window.

4) For all color vectors of the non-noise pixel, their distances to the weighted average vector  $V_A$  are calculated. The color vector with the smallest distance is used to replace all noise pixel color vector in the window. The color vector of center pixel in window ( $V_C$ ) is calculated with equation (7).The algorithm will return to step 1 till all pixels of fruit part are filtered out.

$$V_c = \arg \min \|V_i - V_A\| \quad i=1,2,\dots,N \quad (7)$$

Filtering result is shown in Figure 3:



(a) image with noise (b) image after filtering

Fig.3 noise filtering result

### 3 Dominant color feature extraction

#### 3.1 Color space selection

RGB color space is used widely. The digital images grabbed by camera are generally based on the expression of RGB color space. The RGB color space is the same with the color space showed on computer and the computation is simple. However, traditional RGB color space can not meet subjective color sensation of human being. The human eyes discriminate objects through the difference of brightness, hue and color saturation. They can not feel the proportion of the primary colors directly. Therefore non-linear color space (HSV) which can represent human being's subjective color knowledge is often selected.

HSV color space is consistent with subjective color sensation of human being<sup>[18]</sup>. It is calculated and described with Hue, Saturation and Value which is related to visual characteristics. Hue(H)is determined by the preponderant reflection wavelength from objects, different wavelengths have a different color perception, such as red、green、blue, and so on. Hue is the most important attribute of colors. It determines the fundamental characteristics of color nature. Saturation (S) refers to the color shade, namely the percentage of white light in the solid color. The less white light, the greater the color saturation is. Value (V) is the brightness level of the color, measured by percentage usually, from 0% which means the darkest black, to 100% for the brightest white. The color's chrominance character is defined by Hue H and color saturation S together. Figure 4 is the conversion diagram from RGB color space to HSV color space. H represents hue ( $H \in [0^{\circ}$ ,

360°), S is saturation (S ∈ [0,1]), V on behalf of value (V ∈ [0,1]).

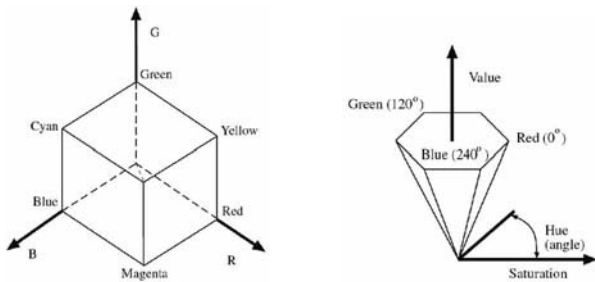


Fig. 4 color space conversion from RGB to HSV  
The image can be transformed from RGB color space to HSV color space based on equation (8):

$$h = 60 * \begin{cases} (5 + b'), (r = m) \cap (g = n) \\ (1 - g'), (r = m) \cap (g \neq n) \\ (1 + r'), (g = m) \cap (b = n) \\ (3 - b'), (g = m) \cap (b \neq n) \\ (3 + g'), (b = m) \cap (r = n) \\ (5 - r'), \text{others} \end{cases} \quad \begin{matrix} s = \frac{m - n}{m} \\ v = m / 255 \end{matrix} \quad (8)$$

In which: m=max(r,g,b), n=min(r,g,b)

$$r' = \frac{m - r}{m - n}, g' = \frac{m - g}{m - n}, b' = \frac{m - b}{m - n}$$

r, g, b, represent the color value of each pixel(r, g, b ∈ [0...255]).

### 3.2 Color space quantization

There are large numbers of colors whether in the RGB and in HSV color space. People’s ability to distinguish colors is limited. In color space, when the difference between two kinds of colors is less than a certain value, human eyes can not distinguish the difference between them, and regard them as the same color. So color space need to be quantized, namely to use fewer colors to present the image’s color space, and then built the color map from the original color space to the selected dominant color collection which was utilized to represent color information of fruits<sup>[19]</sup>.

Three components (H, S, V) in HSV color space are quantized unevenly according to the color sensation of human being, as the equation (9) shows below:

$$H = \begin{cases} 00 & \text{if } h \in (345, 10] \\ 01 & \text{if } h \in (10, 20] \\ 02 & \text{if } h \in (20, 30] \\ 03 & \text{if } h \in (30, 50] \\ 04 & \text{if } h \in (50, 70] \\ 05 & \text{if } h \in (70, 105] \\ 06 & \text{if } h \in (105, 140] \\ 07 & \text{if } h \in (140, 165] \\ 08 & \text{if } h \in (165, 180] \\ 09 & \text{if } h \in (180, 215] \\ 10 & \text{if } h \in (215, 255] \\ 11 & \text{if } h \in (255, 275] \\ 12 & \text{if } h \in (275, 290] \\ 13 & \text{if } h \in (290, 315] \\ 14 & \text{if } h \in (315, 330] \\ 15 & \text{if } h \in (330, 345] \end{cases} \quad \begin{matrix} S = \begin{cases} 0 & \text{if } s \in (0, 0.3] \\ 1 & \text{if } s \in (0.3, 0.7] \\ 2 & \text{if } s \in (0.7, 1] \end{cases} \\ V = \begin{cases} 0 & \text{if } v \in (0, 0.3] \\ 1 & \text{if } v \in (0.3, 0.7] \\ 2 & \text{if } v \in (0.7, 1] \end{cases} \end{matrix} \quad (9)$$

According to the equation (10), three-dimensional HSV color vectors can be synthesized to be one-dimensional vector:

$$I = 9H + 3S + V \quad (10)$$

A new color collection I is come into being. It is of range [0,144]. Since the illumination intensity of the fruit classification is controlled and it is uniform illumination constant, we can disregard the effects of brightness on color classification, therefore the equation 10 can be:

$$I' = 3H + S \quad (11)$$

The number of color in color collection is reduced to be [0, 48].The color map from HSV to I<sub>i</sub>' is shown in table1:

Table1 Color map table from HSV to I<sub>i</sub>'

I <sub>i</sub> '	Quantization space of H and S
0	$h \in (345, 10] \text{ and } s \in (0.0, 0.3]$
1	$h \in (345, 10] \text{ and } s \in (0.3, 0.7]$
2	$h \in (345, 10] \text{ and } s \in (0.7, 1.0]$
.....	.....
46	$h \in (330, 345] \text{ and } s \in (0, 0.3]$
47	$h \in (330, 345] \text{ and } s \in (0.3, 0.7]$
48	$h \in (330, 345] \text{ and } s \in (0.7, 1.0]$

### 3.3 Dominant Color Clustering Algorithm

In the applications of color grading, because the minor difference between colors doesn't effectively reflect the level of fruit and the standard for classification is different for different type of fruits, the interval of quantization often need to be expended by combining adjacent color spaces, and the same time several dominant colors as color grading characteristics of the fruits are extracted. We have adopted a dominant color clustering algorithm to complete this process. The process is as below:

1) Get m sample images of fruit from each grade fruit;

2) Based on equation (9) and equation (11), transfer the fruit surface color from HSV color space into the corresponding new color collection  $I'$  and calculate the number of pixel in each color, then normalize the results using equation (12);

$$N_{si} = \frac{N_{I'_i}}{N_F} \quad N_{si} \in [0,1] \quad (12)$$

In which  $N_{si}$  is the result of normalization, s is serial number of sample, i is serial number of color in new color collection  $I'$ ;  $N_{I'_i}$  is the number of pixel in each color belonging to  $I'$ ;  $N_F$  is the total number of pixel belonging to fruit part .

3) According to equation (13), merge normalization results of the same color in all sample images.

$$N_{pi} = \frac{\sum_{s=1}^m \sum_{i=0}^n N_{si}}{m} \quad (13)$$

4) Set threshold  $T_p$  ; The colors whose pixel value  $N_{pi}$  is smaller than threshold  $T_p$  are filtered out from color collection  $I'$  . A new color collection(  $I''$   $\{ C_q \} q \in [0,n]$ ) of remaining colors is made, and the map relationship between  $I'$  and  $I''$  is established. For example, if the remaining colors are 0,3,6,9,13,

then they will correspond to new color, 0,1,2,3,4 in  $I''$  respectively.

5) Set the distance threshold  $T_d$ ;  $C_q$  denotes fruit color in  $I''$  ( $q \in [0, n]$ );  $Z_p$  is clustering center( $p \in [0, n]$ ); Initially, set  $p = q = 0$ ,  $Z_p = C_{q+1}$ , the distance  $D_{p,q+1}$  between  $Z_p$  and  $C_{q+1}$  is calculated according to the equation (14);

$$D_{p,q+1} = \frac{1}{2} \sqrt{(x_{1p} - x_{1q+1})^2 + (x_{2p} - x_{2q+1})^2 + (x_{3p} - x_{3q+1})^2} \quad (14)$$

In which  $x_1 = s \cos(h), x_2 = s \sin(h), x_3 = v = 0$

6) If  $D_{p,q+1} \leq T_d$ , merge color quantization space of each of H and S channels corresponding to  $Z_p$  and  $C_{q+1}$  ( $Z_p \cup C_{q+1}$ ) and make  $q = q + 1$ , then calculate the distance  $D_{p,q+1}$  between  $Z_p$  and  $C_{q+1}$ , repeat step 6; otherwise if  $D_{p,q+1} > T_d$ , make  $p = p + 1$ ,  $q = q + 1$ .  $Z_p$  becomes the new clustering center, and  $Z_p = C_q$ , then calculate the distance  $D_{p,q+1}$  between  $Z_p$  and  $C_{q+1}$ , repeat step 6; if  $q = n - 2$ , then go to step 7;

7) Calculate the distance  $D_{0,n}$  between  $C_0$  and  $C_n$ ; if  $D_{0,n} \leq T_d$ , then merge color quantization space of each of H and S channels corresponding to  $Z_0$  and  $C_n$  ( $Z_0 \cup C_n$ ).

8) Finally a new group color  $Z_i$  ( $i \in [0,p]$ ) come into being, they are the dominant grading color of fruit.

And a color map table from HSV to  $Z_i$  is created.

The color map relation is listed in table 2:

Table2 Color map table from HSV to  $Z_i$

$Z_i$	Quantization space of H and S
0	$h \in (h_{0b}, h_{0t}]$ and $s \in (s_{0b}, s_{0t}]$
1	$h \in (h_{1b}, h_{1t}]$ and $s \in (s_{1b}, s_{1t}]$
.....	.....
i	$h \in (h_{ib}, h_{it}]$ and $s \in (s_{ib}, s_{it}]$

The process of getting a group of dominant color characteristics of fruits which are ready to be classified is as below:

Firstly, the image of fruit can be converted into HSV color space from RGB color space by equation (8); then the dominant colors  $Z_i$  are abstracted through color map table as table 2 show. These dominant colors represent main color feature of a fruit.

### 3.4 Expression of color histogram

#### 3.3.1 Dominant color histogram

Color is a kind of visual feature of object's surface. Each object has its own unique color feature, and the same type of things has similar or same color features usually. So color features can be used to distinguish between different objects. Color is the most important features of image, and color histogram characterizes the distribution frequency of each color in images, so color histograms can describe the overall color features of the image. Since histogram is a statistical feature, it is not sensitive to the changes of image scale, translation and rotation, so it has visual invariance and high robustness, which is good at accurate extraction of fruit color characteristics and grading.

If the size digital image I is  $M \times N$ , pixel gray level is  $m$ ,  $m \in [0, L-1]$ , then the color histogram of image is given by below equation(15):

$$h'(m) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \begin{cases} 1, & \text{if } G(i, j) = m \\ 0, & \text{if } G(i, j) \neq m \end{cases} \quad m \in [0, L-1] \quad (15)$$

Where  $G(i, j)$  is the gray level of pixel  $(i, j)$

In practical application, the color histogram is usually normalized by equation (16).

$$h(m) = \frac{h'(m)}{M * N} \quad m \in [0, L - 1] \quad (16)$$

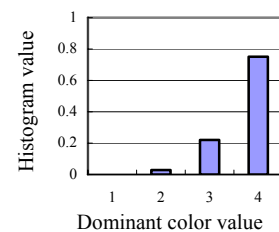
In the equation  $h(m)$  is the probability of pixel with gray level  $m$  in the image.

Tomatoes have a wider range of colors from dark green to dark red. In general, tomato color is a very good indication for maturity stage. Color grading is used to determine the maturity stage. Here tomatoes were selected to illustrate how these algorithms work. According to above algorithms, a dominant color map was obtained as table 3 shown. Then dominant colors of tomatoes were extracted through the table.

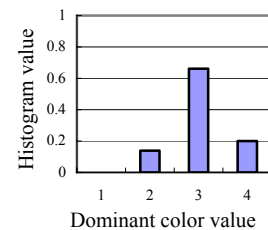
Table3 color map from HSV to dominant color

$Z_i$	Quantization space of H and S
1	$h \in (345, 10]$ and $s \in (0.3, 0.7]$
2	$h \in (10, 30]$ and $s \in (0.3, 1.0]$
3	$h \in (30, 50]$ and $s \in (0.3, 0.7]$
4	$h \in (50, 70]$ and $s \in (0.7, 1.0]$

Dominant color histogram of tomatoes with different ripeness is shown in Fig5.



(a) 10%-20%



(b) 30%-40%

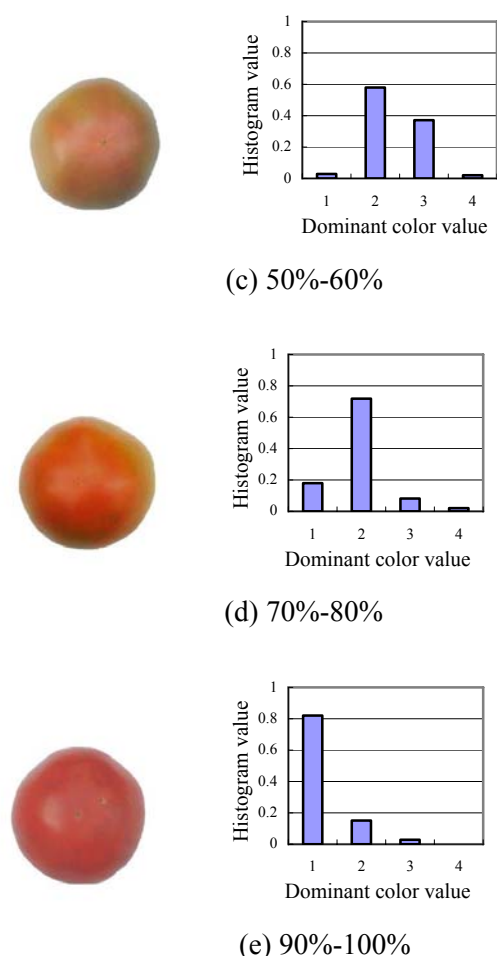


Fig.5 Extracted Dominant color histogram of tomatoes with different ripeness

### 3.3.2 Histogram similarity measure method

Histogram similarity measure is a function defined on pairs of histograms indicating the degree of resemblance of related fruit features. In order to reduce computation time, histogram intersection method is selected<sup>[20]</sup>.

The similarity  $S(i, j)$  between two histograms

$h_i, h_j$  can be calculated using equation (17):

$$S(i, j) = \sum_{k=1}^m \min(h_i(k), h_j(k)) \quad (17)$$

in which  $m$  is the number of histogram bins;

$S(i, j)$  is the similarity between two histograms

$h_i, h_j$ . It is ranged from 0 to 1.

## 4 Experiments and Discussion

### 4.1 Materials

Tomato samples were selected randomly from farm. They were classified into three grades according to color ratio, which is the ratio between surface showing red color and the whole surface: The sample, whose more than 90 percent of the surface, in the aggregate, shows red color was A, more than 70 percent but not more than 90 percent of the surface, in the aggregate, shows red color was B, more than 50 percent but not more than 70 percent of the surface, in the aggregate, shows red color was C, more than 20 percent but not more than 50 percent of the surface, in the aggregate, shows red color was D, and not more than 20 percent of the surface, in the aggregate, shows a definite change in color from green to red was E.

### 4.2 Intelligent Fruit Sorting System

The histogram matching method was tested on intelligent fruit sorting system developed by Research Institute of Robotics Shanghai Jiao Tong University. It could carry out quality inspection of three varieties of fruit: apple, tomato and sweet pepper, and could complete fruit loading, conveyance and sorting automatically. The system consisted of three sections: fruit conveyance and sorting mechanism, image acquisition and control system. The diagram of fruit sorting system is shown in fig6. Because the research objective is to test algorithm performance, tomato samples were put on the sorting line manually and only the top image of sample was acquired.

The image acquisition and control system includes cameras, frame grabber, lights, computer and so on. The CCD color camera having 6mm/F1.2 lens was installed above the inspected fruit to capture fruit image. A ring-shaped fluorescent tube was installed in an illumination chamber whose inner surface was



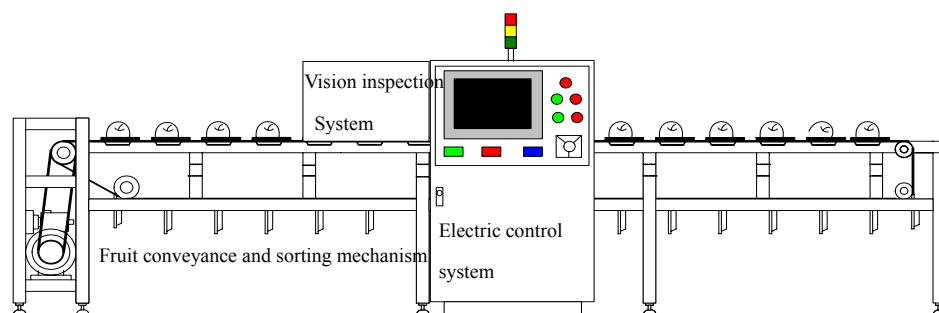


Fig 6 Diagram of Intelligent Fruit Sorting System

painted matt white to provide diffused lighting and to eliminate shades.

CCD camera was assigned to frame grabber that was used for image acquisition task. An optical fiber sensor was utilized to provide trigger signal for the imaging system. A compatible computer was utilized as controller PC, An in-house Visual C++ for software application was used to execute image analysis task. The system software was run under Windows system.

### 4.3 Color classification experiment

Experiment was carried out to test the classification performance of histogram matching method for tomato color classification using dominant color histogram. 250 tomato samples (50 A grade, 50 B grade, 50 C grade, 50 D grade and 50 E grade) were selected from each grade for tomato color classification experiment. The color classification experiment result is listed in Table 4.

Table 4 Color classification experiment result

Sample grade	The number of samples classified into each grad					Correct rate(%)
	A	B	C	D	E	
A	50	0	0	0	0	100
B	0	48	2	0	0	96
C	0	2	48	1	0	94
D	0	0	2	47	0	96
E	0	0	0	0	50	100

Experiment results indicate that dominant color histogram matching based method can achieve accuracy of 100% for A grade tomatoes, accuracy of 96% for B grade tomatoes, accuracy of 94% for C grade tomatoes, accuracy of 96% for D grade tomatoes and accuracy of 100% for E grade tomatoes. Experiment result shows dominant color histogram matching method for color classification has high accuracy and meets the farmer's requirement.

## 5 Conclusions

We have presented a dominant color histogram matching method for fruit classification in this paper. Image segmentation algorithm is used to extract fruit's image from background in OHTA color space. And we have improved the vector median filtering algorithm in order to greatly reduce the impact of noise in the color image. Images grabbed by camera are transformed from RGB color space to HSV color space and three components in HSV color space (H, S, V) are quantized unevenly according to the color sensation of human being. Then the dominant color histogram feature was picked up via color clustering method in HSV color space. Experiment result showed histogram matching method has high accuracy and is suitable for real-time application.

**Acknowledgements** We would like to express our gratitude to The Ministry of Shanghai Prospering Agriculture through Science and Technology Major Project for giving the collaborative research funds. For the help extended by the engineers in Shanghai

Agricultural Machinery Research Institute, we are also grateful and we appreciate them a lot.

### References

- [1] D. J. Lee and R. Anbalagan, High-speed Automated Color Sorting Vision System, *SPIE Optical Engineering Midwest'95*, Chicago, IL, April 1995.
- [2] Min Zhang, L.I. Ludas, M. Morgan et al, Applications of color machine vision in the agricultural and food industries, *Precision agriculture and biological quality*, Boston, Massachusetts, USA, 3-4 November, 1998, Proceedings of SPIE, 3543, pp. 208-219, 1998.
- [3] Z. Varghese, C. T. Morrow, P. H. Heinemann, H. J. Sommer III, Y. Tao, R. M. Crassweller, Automated inspection of Golden Delicious apples using color computer vision, *Paper American Society of Agricultural Engineers*, No. 91-7002, 1991, 16pp
- [4] Yang Xiukun Chen Xiaoguang Ma Chenglin et al Study on Automated Color Inspection of Apples Using Genetic Neural Network *Transactions of the CSAE*, vol.12,no.2,1997,pp.17-176
- [5] He Dongjian, Yang Qing, Xue Shaoping, et al, Computer vision for color sorting of fresh fruits, *Transactions of the CSAE*, vol.14,no.3,1998,pp.202-205.
- [6] Y. C. Hung, K. Morita, R. Shewfelt, A. Resurreccion, S. Prussia, Color evaluation of apples, *Transactions of the ASAE*, No. 93-6541, 1993 ,pp.15.
- [7] Y. Tao, P. H. Heinemann, Z. Varghese, C. T. Morrow, H. J. Sommer, Machine vision for color inspection of potatoes and apples, *Transactions of the ASAE*, vol.38,no.5, 1995,pp.1555-1561.
- [8] B. K. Miller, M. J. Delwiche, Automatic grading of fresh peaches using color computer vision, *Acta-Horticulturae*, no.254, 1989,pp. 161-166.
- [9] B. K. Miller, M. J. Delwiche, *A color vision system for peach grading*, Transactions of the ASAE, vol.32, no. 4, 1989,pp. 1484-1490.
- [10] N .Singh, M . J. Delwiche, R. S Johnson, J. Thompson, Peach maturity grading with color computer vision, *Paper American Society of Agricultural Engineers*, No. 92-3029,1992, pp. 23
- [11] K. H. Choi, G. H. Lee, Y. J. Han, J. M. Bunn, Tomato maturity evaluation using color image analysis, *Transactions of the ASAE*, vol.38, no.1,1995, pp. 171-176.
- [12] S. Laykin,V. Alchanatis, E. Fallik, Y.Edan. 2002. Image-Processing algorithms for tomato classification. *Transactions of the ASAE* vol.45,no.3, 2002, pp. 851-858.
- [13] S. A. Shearer, F. A. Payne, Color and defect sorting of bell peppers using machine vision, *Transactions of the ASAE*,vol.33, no.6, 1990, pp. 2045-2050.
- [14] W. C. Lin, J. W. Hall, A. Klieber, Video imaging for quantifying cucumber fruit color, *HortTechnology*, vol.3:,no.4, 1993, pp. 436-439.
- [15] Yu-Ichi,Ohta,TakeoKanade. Color information for region segmentation. *Computer Graphics and Image Processing* vol.13, no.4, 1980 , pp. 222-241.
- [16] P .E.rahaniyas, I .Pitas, A .N. Venetsanopoulous. Color image processing. In: CT Leondoes, *Control and Dynamic Systems* . Academic Press, 1994.45-90.
- [17] J. Astola, P .Haavisto, Y .Neuvo. Vector median filters. *Processing of the IEEE*, vol78. no.4.1990,pp.678-689.
- [18] A.R .Smith. Color gamut transformation pairs . *Computer Graphics*, vol.12, no.3 ,1978, pp.12-19.
- [19] P .S. Heckbert. Color image quantization for frame buffer display. *Computer Graphics*, vol.16, no.2 ,1982, pp.297-307.
- [20] Richard O.Duda, Peter E. Hart, David G. Stork. *Pattern Classification (Second Edition)*. John Willey & Sons, USA, 2000.