DETECTION OF FRUI T S ON NATURAL BACKGROUND

Somchai LIMSIRORATANA, Prof. Yoshio IKEDA, and Yoshinari MORIO

Laboratory of Agriculture Process Engineering
Dept. of Bio-production Engineering
Division of Environmental Science and Technology
Graduate School of Agriculture Science
Kyoto University, Kyoto, 606-01 JAPAN
lsomchai@e7sun-1.kais.kyoto-u.ac.jp
ikeda@e7sun-1.kais.kyoto-u.ac.jp
morio@e7sun-1.kais.kyoto-u.ac.jp

ABSTRACT

The objective of this research is to detect the papaya fruits on tree in an orchard. The detection of papaya on natural background is difficult because colors of fruits and background such as leaves are similarly green. We cannot separate it from leaves by color information. Therefore, this research will use shape information instead. First, we detect an interested object by detecting its boundary using edge detection technique. However, the edge detection will detect every objects boundary in the image. Therefore, shape description technique will be used to describe which one is the interested object boundary. The good shape description should be invariant in scaling, rotating, and translating. The successful concept is to use Fourier series, which is called "Fourier Descriptors". Elliptic Fourier Descriptors can completely represent any shape, which is selected to describe the shape of papaya. From the edge detection image, it takes a long time to match every boundary directly. The pre-processing task will reduce non-papaya edge to speed up matching time. The deformable template is used to optimize the matching. Then, clustering the similar shapes by the distance between each centroid, papaya can be completely detected from the background.

Key Word : Edge Detection, Papaya, Elliptic Fourier Descriptors, Deformable Template

INTRODUCTION

Machine visions are necessary to add an intelligent to agricultural machinery. The fruit detection on simple color background can be applied for grading. For harvesting, it requires the detection on natural background. For some kind of fruit, we can easily detect it by color because its color difference from background is distinct. For papaya, which will be used as sample fruit in this research, the detection by color is difficult because it has the similar green color as the backgrounds. Then, we cannot separate it from background by color information.

If we want to detect the shape information, we have to detect object boundary in advance. Many research used the edge detection technique. A good edge detector should be sensitive to localization and step edge, "Canny (1986)". But, the ordinary edge detection technique will detect every objects boundary in the image.

The shape description technique can be used to describe which one is the interested object boundary. The good shape description should be invariant in scaling, rotating, and translating, "Zahn and Roskies (1972)". The successful concept is to use Fourier series, which is called "Fourier Descriptor". It is applied in many areas especially in agriculture, "Tao et al.(1996)", "Ehrlich and Weinberg (1970)", "Segerlind and Weinberg (1973)". Elliptic Fourier Descriptors can completely represent any shape, "Staib and Duncan (1992)", which was selected to describe the shape of papaya.

Using the edge detection image, it takes a long time to match every boundary directly. The pre-processing task will reduce non-papaya edge to speed up matching time. The deformable template is used to optimize the matching.

PROCEDURES

We designed the detecting procedure like Fig. 1.

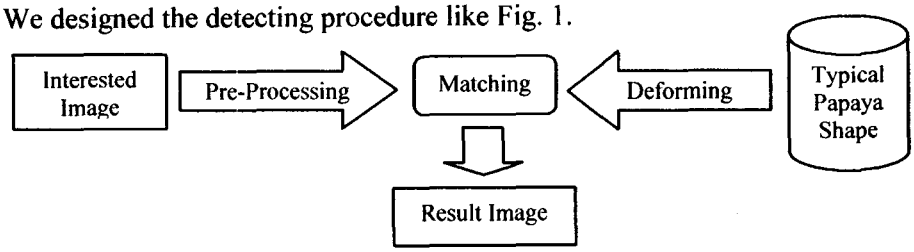


Fig. 1 Detecting Procedure

Pre-Processing

The detection can be performed faster if the image has a few specified pixels of papaya. The aim of pre-processing is to get only specified pixels of papaya edge. Some color ratio and Min/Max filter were used to remove the non-papaya edge from an edge detection image.

Color Ratio

The color information can approximate an object region. For papaya, however, it is very difficult to seperate papaya from its leaves. But, we can detect other objects such as soil and stem by color. Then, we can remove them from an image.

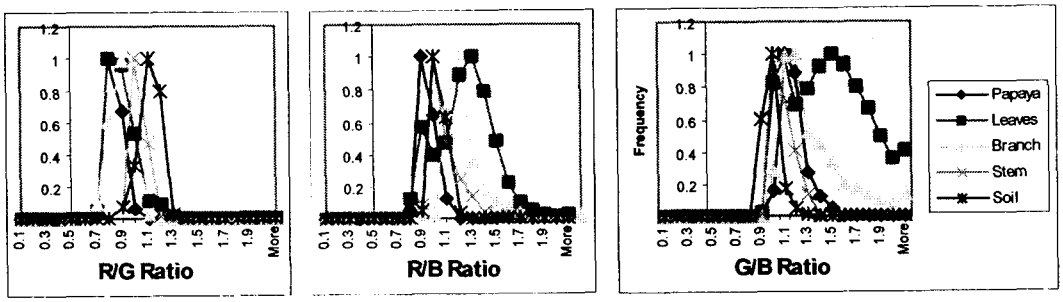


Fig. 2 Histogram of RGB ratio

The histograms in Fig. 2 show that it was difficult to separate a papaya and leaves by color ratio. Color ratio of stem and soil was used to indicate a non-papaya region. The ranges of color ratio for detecting soil and stem are as follows.

Soil Color Ratio Range:

0.667260	< Red/Green	< 1.167260
0.894138	< Red/Blue	< 1.094138
0.835231	< Green/Blue	< 1.035231

Stem Color Ratio Range:

0.883694	< Red/Green	< 1.083694
0.925462	< Red/Blue	< 1.425462
0.942196	< Green/Blue	< 1.442196

Fig. 3 Non-papaya area (IRatio)

$$IRatio(x,y) = \begin{cases} 255 & I_{avg}(x,y)\ in\ Soil\ or\ Stem\ range \\ 0 & others \end{cases} \quad (1)$$

The result is shown in Fig. 3.

Edge Detection

Canny Edge Detection is used to detect the edge by the following equation.

$$\left.\begin{aligned} IEdge(x,y) &= \left|\nabla\left(G(x,y) * IGray(x,y)\right)\right| \\ G(x,y) &= \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \\ IGray(x,y) &= 0.299 \times I_{red}(x,y) + 0.587 \times I_{Green}(x,y) + 0.114 \times I_{Blue}(x,y) \end{aligned}\right\} \quad (2)$$

The results are shown in Fig. 5a and b.

Filtration of non-papaya edge

IEdge is very complex and we cannot know which is the real edge of papaya. But, *IEdge* has low frequency of edge in the papaya area. Therefore, we can use some low pass filter technique to detect this area but we will lose the papaya edge if that low pass filter is un-adjustable. Maximum/Minimum filter is used for this task.

Maximum filter with radius 6 will enlarge pixels. The high frequency of pixels pattern is filtered to be a big area, which is a non-papaya area. We threshold it with 255 and logical OR with *IRatio* image. After that, we delete some small area, which cannot be filtered by this size of filter. The non-papaya areas cover a little bit the inside of papaya. So, the papaya area is enlarged back by Minimum filter with radius 8.

$$IMax = DelSmall\left[Threshold\left(Max(IEdge,6),255\right) \vee IRatio\right] \quad (3)$$

$$IMin = Min(IMax,8) \quad (4)$$

To get the papaya edge, inverse *IMin*, and logical AND with *IMax*, and then logical AND with *IEdge* were computed. Then, we maximized it by Maximum filter radius 1.

$$IPapayaEdge = Max\left(\left(IEdge \wedge \left(IMin_{inv} \wedge IMax\right),1\right)\right) \quad (5)$$

The result is shown in Fig. 5c.

Typical Papaya Shape

We determined the typical shape of papaya (typical template) from 128 images of papaya on white background. From each image, we got the boundary of papaya by Canny edge detection and calculated normalized Elliptic Fourier Descriptors (EFD). Then, the typical shape of papaya was determined by averaged all normalized EFD. The result is depicted in Fig. 6a.

Elliptic Fourier Descriptors (EFD)

Fourier Descriptors concepts assume the object shape (boundary function) as a periodic function. The boundary function has to be a closed curve. We can get Fourier parameters by determination of each harmonics magnitude of Fourier series.

The basic Fourier series is:

$$\left.\begin{aligned} f(t) &= a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt) \\ a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(t)dt \qquad a_n &= \frac{1}{\pi}\int_0^{2\pi} f(t)\cos nt\, dt \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\, dt \end{aligned}\right\} \tag{6}$$

The basic model of Fourier Descriptors is $R(\theta)$ where R is a radius from centroid to boundary at angle θ. This model has some problem when the shape re-entrance, which gives many R's at one θ, "Fong et al.(1979)".

Another model is $\phi(l)$ and normalize by letting $t = l(2\pi/L)$, "Zahn and Roskies (1972)". l is arc length and $0 \le l \le L$. $\phi(l)$ is a net amount of angular bend between starting point and point l. This model contains discontinuities. Therefore, a_n and b_n decrease rather slowly, "Persoon and Fu (1977)" and the reverse transform give non-closed curve, "Jones (1983)", "Strackee and Nagelkerke (1983)".

Elliptic Fourier Descriptors or $(x(s), y(s))$ model can completely represent any shape. x and y are co-ordinates of boundary. s is arc length, $0 \le s \le S$. And $t = s(2\pi/S)$.

$$\left.\begin{aligned} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} &= \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos kt \\ \sin kt \end{bmatrix} \\ a_0 = \frac{1}{2\pi}\int_0^{2\pi} x(t)dt \qquad a_k &= \frac{1}{\pi}\int_0^{2\pi} x(t)\cos kt\, dt \qquad b_k = \frac{1}{\pi}\int_0^{2\pi} x(t)\sin kt\, dt \\ c_0 = \frac{1}{2\pi}\int_0^{2\pi} y(t)dt \qquad c_k &= \frac{1}{\pi}\int_0^{2\pi} y(t)\cos kt\, dt \qquad d_k = \frac{1}{\pi}\int_0^{2\pi} y(t)\sin kt\, dt \end{aligned}\right\} \tag{7}$$

Normalized Elliptic Fourier Descriptors

From an image, the boundary of object can be any size, any rotation, and place at any co-ordinate. Before use, we should normalize the EFD in four parameters, that is starting point, rotation, phase-shift, and scale.

- Starting Point

Starting point of EFD is a point at $t = 0$. The centroid is the zeroth harmonic, a_0 and c_0. It is changed to (0,0).

$$x_{start} = \sum_{k=0}^{n} a_k \qquad y_{start} = \sum_{k=0}^{n} c_k \tag{8}$$

- Rotation and Phase-Shift

The shape is paralleled to X-axis by rotating the first harmonic. rx is the major axis. If not, rotate 90 degree and change the starting phase.

$$rx = \sqrt{a_1^2 + b_1^2} \quad ry = \sqrt{c_1^2 + d_1^2} \quad \text{starting phase} = \pi - \arctan\left(\frac{d_1}{c_1}\right)$$

(9)

- Scaling

To normalize the scaling, all parameters are divided with the magnitude of first harmonic on X-axis, rx.

$$as_k = \frac{a_k}{rx} \qquad bs_k = \frac{b_k}{rx} \qquad cs_k \frac{c_k}{rx} \qquad ds_k = \frac{d_k}{rx} \tag{10}$$

Deformation

After normalizing a shape from image, a simple way to detect an object is to compare with the typical template. This simple model can be processed when the interested image is not complex, which we can get a continuous and closed-curve boundary. But, natural background image is complex and we cannot get the complete boundary. Therefore, normalized shape of the interested object cannot be determined.

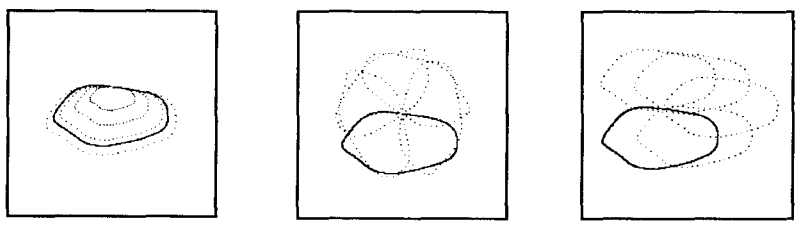
Another way is to determine any possible transformation of the typical template, which give the maximum likelihood. We call it "Deformable Template". This concept is flexible for any equation and can be applied widely in many areas, for example in medical image, "Chakraborty et al.(1996)", "Strackee and Nagelkerke (1983)".

The deformable template was created by inverting the typical template at each possible scaling (Fig. 4a), rotating (Fig. 4b), phase-shift (Fig. 4c), scale of the first harmonic, or any others.

$$T(scale) = scale \times T \tag{11}$$

$$\left.\begin{aligned} T(rotate) &= \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} + \sum_{k=1}^{n} \begin{bmatrix} ar_k & br_k \\ cr_k & dr_k \end{bmatrix} \begin{bmatrix} \cos kt \\ \sin kt \end{bmatrix} \\ ar_k &= a_k \cos(rotate) - c_k \sin(rotate) \qquad br_k = b_k \cos(rotate) - d_k \sin(rotate) \\ cr_k &= a_k \sin(rotate) + c_k \cos(rotate) \qquad dr_k = b_k \sin(rotate) + d_k \cos(rotate) \end{aligned}\right\} \tag{12}$$

$$\left.\begin{aligned} T(phase) &= \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} + \sum_{k=1}^{n} \begin{bmatrix} ap_k & bp_k \\ cp_k & dp_k \end{bmatrix} \begin{bmatrix} \cos kt \\ \sin kt \end{bmatrix} \\ ap_k &= a_k \cos(k\,phase) - b_k \sin(k\,phase) \qquad bp_k = a_k \sin(k\,phase) + b_k \cos(k\,phase) \\ cp_k &= c_k \cos(k\,phase) - d_k \sin(k\,phase) \qquad dp_k = c_k \sin(k\,phase) + d_k \cos(k\,phase) \end{aligned}\right\} \tag{13}$$



a b c

Fig. 4 a:Scaling, b:Rotating, and c:Phase-Shift

Deforming Optimization

If we can create many possible transforms, we will get a perfect template, which can be used in any case. But, it is not possible because we have to keep all possible point, which require more memory.

To decrease memory requirement, only necessary points should be kept, which involve the unique point, and do not include points around starting point. Then, we link these points to their related transformation. Only possible angle, phase-shift, scale, and scale of the first harmonic were calculated. Do not rotate it every degree and size every scale. The result is shown in Fig. 6b.

Matching

At each possible transformation, we have to calculate a total energy (likelihood) by summing the energy at each pixel of that transform. Then, the maximum value indicates the best fitting as shown in Eq. 14.

This processing takes a long time because at each possible transform parameter p, we have to sum the energy at every point on the transformed boundary. Each point is determined by transforming the typical template and inverse. Some point will be recalculated because it overlaps to others possible parameters.

$$M(I,T_p) = \max_p \left[\ln \Pr(T_p) + \ln \Pr(I|T_p)\right]$$
$$= \sum_{i=1}^{N}\left[\ln\left(\frac{1}{\sigma_i\sqrt{2\pi}}\right) - \frac{(p_i - m_i)^2}{2\sigma_i^2}\right] + \frac{k}{\sigma_n^2}\int_0^S I(x(p,s),y(p,s))ds \quad (14)$$

I = *Image* T_p = *Template of parameter* p

N = *Number of parameters* m_i = *mean of* p_i

σ_i^2 = *Variance of* p_i s = *arc length along the curve*

k = *Magnitude of* T_p *at any point,constant*

Matching Optimization

To decrease the processing time, it is not necessary to calculate the first term because we use the same set of parameters. It will be a constant value. The k, and σ are constant too. Then, the optimized matching equation will be only the sum of the value of image at each point on the possible transformation edge boundary.

$$M(I,T_p) = \int_0^S I(x(p,s),y(p,s))ds \quad (15)$$

The papaya will be found at the scale, rotation, phase and other transforms, which give the maximum of M. Using the deformable template can reduce the processing time because it reduces the number of point, which has to calculate energy (likelihood).

RESULTS

Pre-processing

From the gray level image (Fig. 5a), we detected the edge by Canny edge detection as show in Fig.5b. Using color ratio from Eq. 1 and the Minimum/Maximum filter in Eq.'s 3 and 4, we got the papaya edge as Fig. 5c. This pre-processing takes about 15

seconds.

Fig. 5 a: Gray level image(IGray), b: Edge detection image(IEdge), and c: Papaya edge(IPapayaEdge)

Deformation and Matching

Fig. 6a is the typical papaya shape. It was determined from 128 images. The deformable template was created by scale from 60 to 90 at every 2 steps, rotate from 75 to 105 degrees at every 3 degrees, phase-shift from 0 to 360 degrees at every 10 degrees and scale of the first harmonic 10%. And, 200 points were generated for each transform. (Fig. 6b). This preparation takes about 30 minutes. Fig. 6c is the result of matching, which takes about 33 seconds.

Fig. 7a and 7b are another example and result.

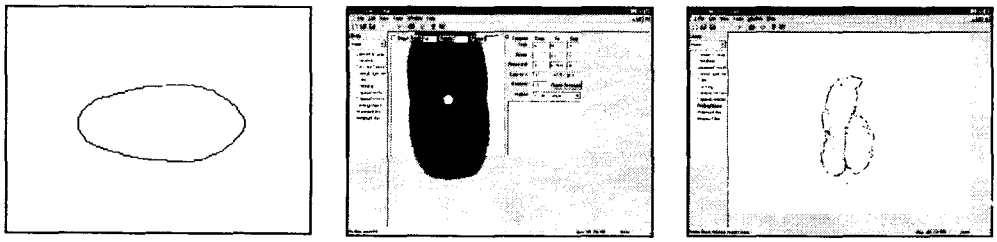


Fig. 6 a: Typical Papaya Shape(Typical Template), b: Deformable Template, and c: Result of matching image in Fig. 5a

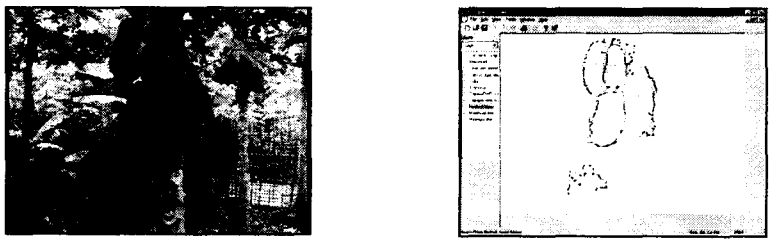


Fig. 7 a: Another example gray level image, and b: Result of matching image in a

CONCLUSIONS

This technique has to prepare a deformable template, which calculate only once. Then, we can use it as a shape database. The pre-processing task should be changed if it is applied to other fruit. Accuracy of this technique depends on the accuracy of pre-processing task and deformable template.

REFERENCES

1. Canny, John, 1986, A Computational Approach to Edge Detection, IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI-8(6): 679-698.
2. Chakraborty, Amit, Lawrence H. Staib and James S. Duncan, 1996, Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information, IEEE Transactions on Medical Imaging, 15(6): 859-870.
3. Ehrlich, Robert and Bernhard Weinberg, 1970, An Exact Method for Characterization of Grain Shape, Journal of Sedimentary Petrology, 40(1): 205-212.
4. Fong, S. T., J. K. Beddow and A. F. Vetter, 1979, A Refined Method of Particle Shape Representation, Powder Technology, 22: 17-22.
5. Jones, S., 1983, The Problem of Closure in the Zahn-Roskies Method of Shape Description, Powder Technology, 34: 93-94.
6. Persoon, Eric and King Sun Fu, 1977, Shape Discrimination Using Fourier Descriptors, IEEE Transaction on Systems, Man, and Cybernetics, SMC-7(3): 170-179.
7. Segerlind, Larry J. and Bernhard Weinberg, 1973, Grain Kernel Identification by Profile Analysis, Transaction of the ASAE, pp. 324-327.
8. Staib, Lawrence H., and James S. Duncan, 1992, Boundary Finding with Parametrically Deformable Models, IEEE Transaction on Pattern Analysis and Machine Intelligence, 14(11): 1061-1075.
9. Strackee, Jan and Nico J. D. Nagelkerke, 1983, On Closing the Fourier Descriptor Presentation, IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI-5(6): 660-661.
10. Tao, Y., C.T. Morrow, P.H. Heinemann and H.J. Sommer III, 1995, Fourier-Based Separation Technique for Shape Grading of Potatoes Using Machine Vision, Transactions of the ASAE, 38(3): 949-957.
11. Zahn, Charles T. and Ralph Z. Roskies, 1972, Fourier Descriptors for Plane Closed Curves, IEEE Transaction on Computers, C-21(3): 269-281.