J. Fish. Sci. Tech. 4(1), 1~9, 2001

JOURNAL OF FISHERIES SCIENCE AND TECHNOLOGY

Morphological Feature Extraction of Microorganisms Using Image Processing

Hak-Kyeong Kim*, Nam-Su Jeong[1], Sang-Bong Kim[2] and Myung-Suk Lee[3]

Dept. of Mechatronics Eng., Pukyong National Univ., Pusan 608-737, Korea
[1]PROCOM A/S, 652-2, Nambumin-dong, Seo-gu, Pusan 602-020, Korea
[2]Dept. of Mechanical Eng., Pukyong National Univ., Pusan 608-737, Korea
[3]Dept. of Microbiology, Pukyong National Univ., Pusan 608-737, Korea

(Received January 2001, Accepted March 2001)

This paper describes a procedure extracting feature vector of a target cell more precisely in the case of identifying specified cell. The classification of object type is based on feature vector such as area, complexity, centroid, rotation angle, effective diameter, perimeter, width and height of the object. So, the feature vector plays very important role in classifying objects. Because the feature vectors is affected by noises and holes, it is necessary to remove noises contaminated in original image to get feature vector extraction exactly. In this paper, we propose the following method to do to get feature vector extraction exactly. First, by Otsu's optimal threshold selection method and morphological filters such as cleaning, filling and opening filters, we separate objects from background an get rid of isolated particles. After the labeling step by 4-adjacent neighborhood, the labeled image is filtered by the area filter. From this area-filtered image, feature vector such as area, complexity, centroid, rotation angle, effective diameter, the perimeter based on chain code and the width and height based on rotation matrix are extracted. To prove the effectiveness, the proposed method is applied for yeast *Zygosaccharomyces rouxii*. It is also shown that the experimental results from the proposed method is more efficient in measuring feature vectors than from only Otsu's optimal threshold detection method.

Key words: Feature extraction, Otsu's method, Area filter, *Zygosaccharomyces rouxii*

Introduction

Vision system, nowadays, has been widespread technology in several engineering fields since the relevant efforts appeared around 1964 at the J.P.L. (Jet Propulsion Laboratory) and concerned the digital process satellite images coming from the moon (Pitas, 1993). In cell technology, recent methods of digital image analysis include location and enumeration of bacteria in solid foods, *in situ* microscopy and image analysis for on-line monitoring of yeast fermentations (Zalewski and Bildanalyse, 1995), and texture analysis of fungal colonies for subsequent transfer. Notable recent applications include studies on the pulsatile growth of hyphal apices, biochemical differentiation of fungal colonies, and simple structural differentiation of mycelia from submerged fungal cultures (Thomas and Paul, 1996).

All microscopic images have many obstacles on the lens and are directly affected by brightness of light source, staining method for cells, and settings of optical filters, etc. The microscopic image often has several special properties such as non-symmetric contour, a bimodal histogram, and multi-layer objects, etc. Non-symmetric contour is due to light source of the microscope concentrated on the center of the image. Luminance at the center of the background is higher than that of its surrounding. This property affects a gray-level histogram of the image and makes the exact image segmentation difficult. Otsu's optimal threshold detection method is usually used for the

*Corresponding author: hakkyong@dreamwiz.com

image segmentation (Otsu, 1979). But the segmented image by only the Otsu's method includes a lot of noises. So extracting the exact feature vector from the image segmented by only Otsu's method is very difficult.

So this paper shows basic research results to develop vision analysis system which can be applied for bio-process plant and to present a procedure extracting features in order to identify the object cells exactly under background with noise such as obstacles and non-symmetric contour, etc. The classification of object type is based on feature vector such as area, complexity, centroid, rotation angle, effective diameter, perimeter, width and height of the object. So the feature vector plays very important role in classifying objects. Because feature vector is affected by noises and holes, it is necessary to remove noises contaminated in original image to get the feature vector extraction exactly. In this paper, we propose the following method to do so. It presents a procedure which extracts features for vision analysis based on pattern recognition. The procedure works semi-automatically: operator has to find target organisms and adjust focus to get a clean picture manually, and next all following steps work full-automatically. First, several filters are applied to original images in order to clean them. Next, Otsu's optimal threshold selection method and morphological filters such as cleaning, filling and opening filters are employed to separate objects from background and to get rid of isolated particles respectively. The labeling step segments each objects by 4-adjacent neighborhood. The previous labeled image is filtered by area filter. Finally, from this area-filtered image, morphological features such as area, perimeter, complexity, centroid, rotation angle, width, height and effective diameter are extracted by the proposed procedure. To prove the effectiveness, the proposed method is applied for yeast *Zygosaccharomyces rouxii*. It is also shown that the experimental results from the proposed method is more efficient in measuring feature vectors than from only Otsu's optimal threshold detection method.

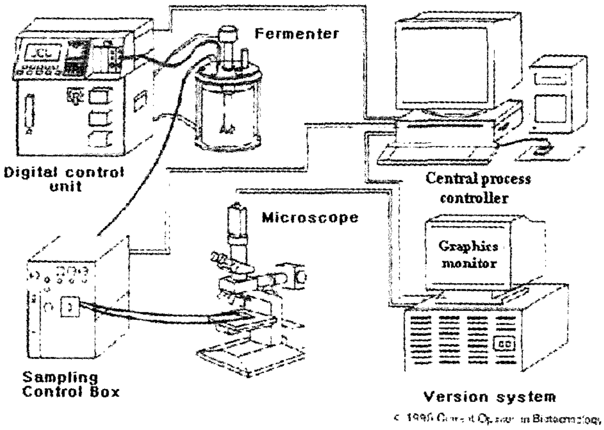


Fig. 1. Typical application of vision system in the cell technology.

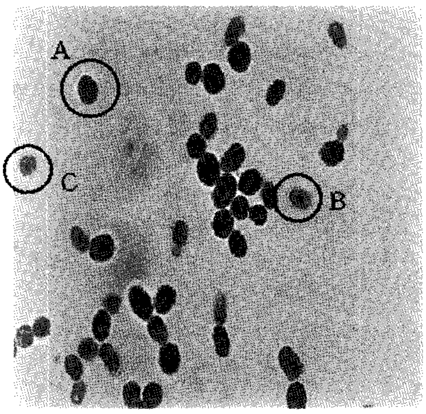


Fig. 2. Microscopic image of *Zygosaccharomyces rouxii*.

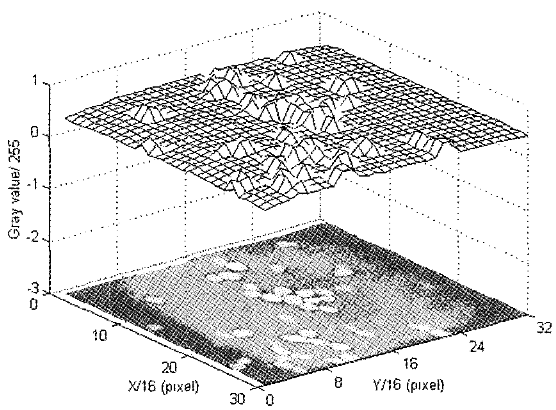


Fig. 3. The spatial representaion of gray value for the image shown in Fig. 2.

Characteristics and Problem Statement of Microscopic Images

Fig. 1 illustrates typical application of vision analysis for fermentation process. All microscopic images have many obstacles on the lens as shown in Fig. 2. They are directly affected by brightness of light source, staining method for cells, and settings of optical filters, etc. Understandings for microscopic image need to adopt the proper methods among a number of image processing techniques.

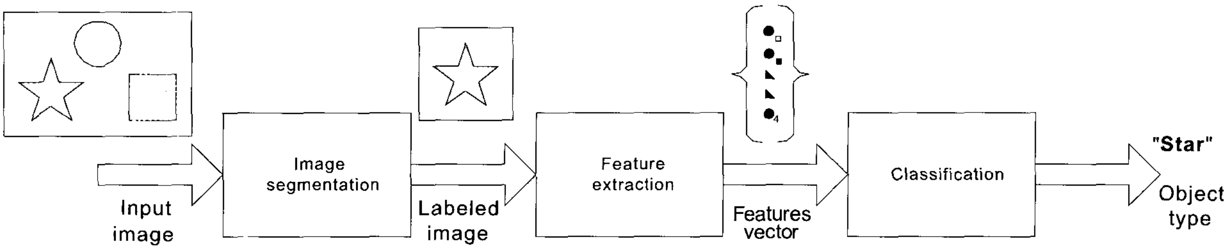


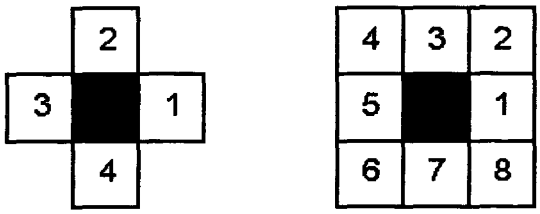
Fig. 4. The three phases of pattern recognition.

Digital image is a kind of spatial frequency as shown in Fig. 3 which illustrates the relationship between brightness and gray levels of image.

This is useful to get reasonable ways to apply 2-dimensional signal processing technique based on linear theory. The microscopic image often has several special properties such as non-symmetric contour, a bimodal histogram, and multi-layer objects, etc. Non-symmetric contour is due to light source of the microscope concentrated on the center of the image. Luminance at the center of the background is higher than that of its surrounding. This property affects a gray-level histogram of the image. Therefore, to see more clearly the target cells, it is more desirable to get rid of the noise from background. So, the aim.of this paper is to propose an algorithm that can separate the target cells exactly under background with noise.

Morphological feature extraction of microorganisms

Pattern recognition based on statistical decision rule has three stages such as image segmentation, feature extraction and classification in Fig. 4. First, pre-filters improve the quality of the acquired original image so that general noise is eliminated, and objects in the image is clearer than the prior. After preprocessing, image segmentation phase based on Otsu's optimal threshold selection makes binary image from gray-level image. Threshold plays a role in the critical point to separate objects from background. Next, several binary morphological filters such as cleaning, opening and filling filters get rid of isolated particles. And then, the labeling algorithm segments each objects by 4-adjacent neighborhood as shown Fig. 5(a). Finally, morphological feature vectors such as area, perimeter, complexity, centroid, rotation angle, width, height and effective diameter are extracted by the binary image processing.



(a) 4-adjacent (b) 8-adjacent

Fig. 5. Neighborhoods.

Feature vector extraction is to extract a set of the elementary properties to quantify some significant characteristics of the object. Because the classification of object type is based on these feature vectors, feature parameters must be selected as cautious as possible. If feature vectors are decided irrationally, error rate in misclassifying objects is increased.

1. Pre-filters and Image Segmentation

In this paper, morphological filters are applied such as opening, cleaning, etc. Image segmentation is a fundamental technique for image analysis, whose purpose is to identify the regions of image objects and to extract the objects from their background (Morii, 1995). This paper engages *Otsu's variance based thresholding method* (Otsu, 1979). He described three possible discriminant criteria based on ratios of the within-class, between-class and total variance, all of which are equivalent, and thus in a given situation any of the three possible discriminant criteria could be chosen (Hannah et al., 1995). In general case, threshold value is chosen to maximize the between-class variance because it makes the simple calculation.

2. Otsu's Variance Based Thresholding

Assume that an image consists of L gray levels. Total number of pixels in the image is given by $N = n_1 + n_2 + \cdots + n_L$, where n_i means the number of pixels with a given gray level i. The probability density of a given gray level i is thus:

$$p_i = n_i/N \tag{1}$$

where

$$p_i = \frac{n_i}{L},\ p_i \geq 0,\ \sum_{i=1}^{L} p_i = 1 \tag{2}$$

For a single threshold, the optimal thresholding value is the gray value when J, the ratio of the between-class variance to the total variance, is maximized:

$$J = \frac{\sigma^2_B}{\sigma^2_T} \tag{3}$$

where the between-class variance

$$\sigma^2_B = \omega_1(M_1 - M_T)^2 + \omega_2(M_1 - M_T)^2 \tag{4}$$

which can be simplified to

$$\sigma^2_B = \omega_1\omega_2(M_2 - M_1)^2 \tag{5}$$

and the total variance

$$\sigma^2_T = \sum_{i=1}^{L}(i - M_T)^2 p_i \tag{6}$$

where

$$\omega_1 = \sum_{i=1}^{k} p_i,\ \omega_2 = \sum_{i=i+1}^{L} p_i = 1 - \omega_1 \tag{7}$$

$$M_2 = \sum_{i=1}^{k} i\frac{p_i}{\omega_1},\ M_1 = \sum_{i=k+1}^{L} i\frac{p_i}{\omega_2},\ M_T = \sum_{i=1}^{L} i p_i \tag{8}$$

Optimal threshold value t is chosen as k to maximize the ratio J of eq. (3). Since the total variance, Eq. (6) is constant for a given image histogram, it implies maximizing σ^2_B, the between-class variance from Eq. (3). Optimal threshold value t is chosen as k to maximize eq. (5).

3. Labeling and Morphological Filters

In general, two labeling algorithms are used to identify objects into binary image by 4-adjacent or 8-adjacent neighborhood. This paper has applied labeling algorithm called 'frame propagation' using queue to store positions of labeled pixels. $g(x,y)$ is a binary image that we want to do labeling. (s,t) is a position coordinate of a pixel removed from the queue list of $g(x,y)$. (u,v) is a four-neighbor coordinate of (s,t). The queue has two major member functions: function insert (u,v) appends a pixel (u, v) to the end of the queue list and (s,t)←remove() excludes a pixel (s,t) from the begin of the queue list. The labeling pseudo-code and simulation result are shown in List 1.

List 1. A pseudo-code of the frame propagation labeling

```
L←−1        (Initialize label.)
scan the image from left to right and top to
 bottom for all it(x,y)
 if g(x,y)≥0 then insert (x,y)
 while list is not empty do
  (s,t)←remove()
  for each 4-neighbor (u,v) of (s,t) do
   if (u,v) is unlabeled and g(u,v)=g(x,y) then
    insert (u,v)
   end if
  end for
 end while
 L←L÷1      (Get new label.)
end scan
```

4. Feature Vector Extraction

In Fig. 6, feature vector is determined to identify object as follows:

$$F := [A\ L\ e\ I_c\ J_c\ \theta\ W_x\ W_y\ d_e]^T \tag{9}$$

where A is area of each labeled object, L is the length of perimeter, e means complexity, the parameters $(I_c,\ J_c,\ \theta)$ are centroid of the object and primary angle of its axis, $(W_x,\ W_y)$ are the width and the height. The effective diameter (d_e) is a diameter of the object that is considered as circle.

The object area A is calculated by summing pixels in the image as follows:

$$A = \sum_{i=0}^{M}\sum_{j=0}^{N} f(i,j) \tag{10}$$

where $f(i,j)$ is a binary image that has 1 inside object and 0 inside background and $M \times N$ is a size of an image.

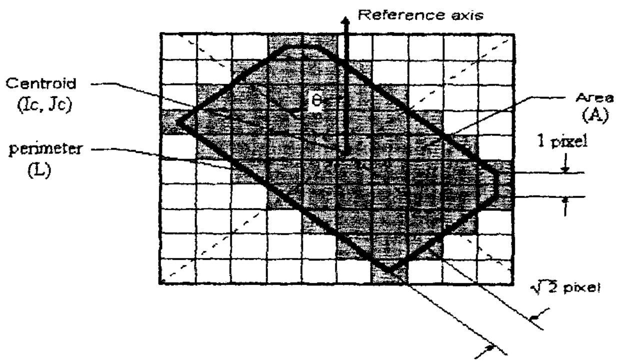


Fig. 6. Concepts of moment and geometric parameters.

The perimeter L is obtained by counting boundary pixels, where the distance between 8-adjacent pixels is 1 or $\sqrt{2}$ pixel unit. Using chain code with 8 directions shown as Fig. 5(b), L is calculated as follows:

$$L = m + \sqrt{2}\, n \tag{11}$$

where m is the number of odd number and n is the number of even number in chain code.

Simply counting boundary pixels can not make reasonable result. Therefore, the directions of the boundary must be kept to the tracks of the path. Complexity of the object is taken as following:

$$e: = \frac{L^2}{A} \tag{12}$$

Morphological features of shapes include the moment parameters as follows:

$$M(p,q): = \sum_{(i,j)} i^p j^q f(i,j) \tag{13}$$

There are several parameters to be calculated by combinations of Eq. (13). However, in this paper, the centroid (I_c, J_c), central moments μ_{ij} taken the center of gravity as the origin and the rotation angle (θ) are only measured as follows:

$$I_c = \frac{M(1,0)}{M(0,0)} \tag{14}$$

$$J_c = \frac{M(0,1)}{M(0,0)} \tag{15}$$

$$\mu_{ij} = \sum_{i=0}^{M} \sum_{j=0}^{N} (i - I_c)^i (j - J_c)^j f(i,j) \tag{16}$$

$$\tan 2\theta = \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \tag{17}$$

The width(W_x) and height(W_y) are determined by measuring the distances between each first detected pixels from left, right, top, and bottom side after rotating the object region for the rotation angle θ obtained from Eq. (17). The width and height is measured after the clock-wise rotation step using rotation transformation matrix T as follows:

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{18}$$

$$W_x = \text{rightmost pixel} - \text{leftmost pixel}; \tag{19}$$

$$W_y = \text{bottommost pixel} - \text{topmost pixel}; \tag{20}$$

The effective diameter (d_c) can be defined as follows:

$$\frac{\pi}{4} d_c^2 = A \tag{21}$$

Fig. 7 shows us that the actual programmed procedure includes area filters to cancel next extraction procedure for very small object in order to obtain higher calculation efficiency than the method using

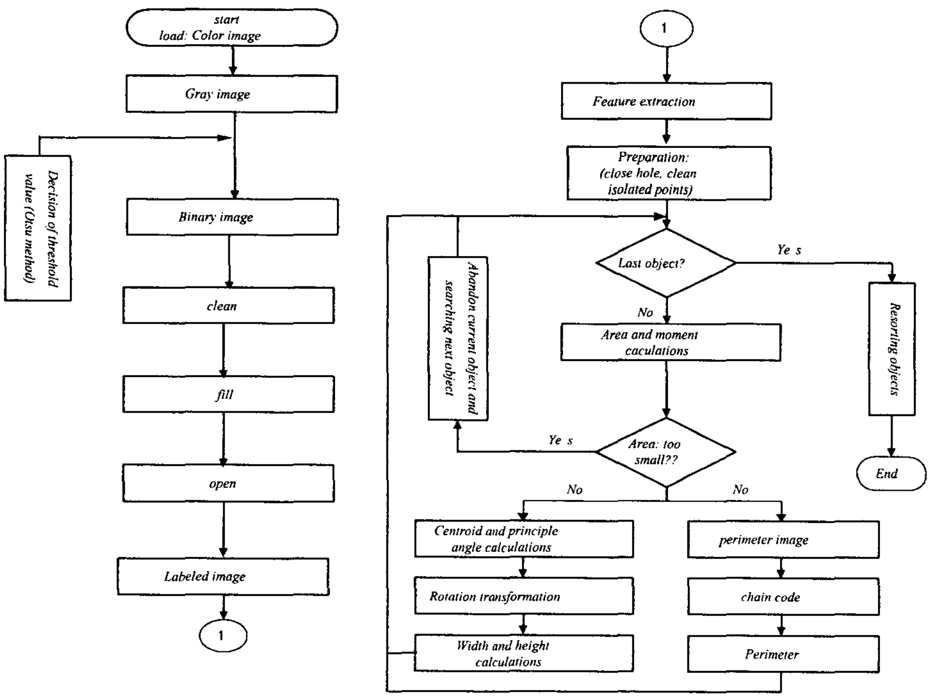


Fig. 7. Flowchart for extractions of moment and geometric parameters.

only Otsu's method. The preparation step in the flowchart also contains several filters based on binary morphological procedure such as filling hole and cleaning isolated point.

Experiment results

Table 1 shows the specifications of devices used in this experimental image processing system.

Fig. 8 shows the theshold criterion(J) and the optimal threshold value obtained by Otsu's method. The optimal threshold for the image was determined as $t=85$ at $J_{max}=0.681167$. The result graph is in the following Fig. 9.

Fig. 10 shows the image processing results obtained by the proposed method. In Fig. 10, (a) shows the original image as shown in Fig. 2, (b) is a binary image obtained by thresholding value, $t=85$, (c) is an image obtained after removing the isolated object and filling holes in the binary image(b) and (d) shows an image obtained by filtering area less than 50 pixels in the filtered image(c).

Table 1. The specifications of devices used in the image processing system

Items	Models	Spec.
Microscope	BX40, Olympus	Lens sets: ×1000
CCD-Camera	IK-642K, Toshiba	1/2 inch CCD cell 470H×350V or more
Frame grabber	Oculus TCX/MX Coreco	Color image grabber 4MB image RAM
Host computer	GT53H, LG Electronics	Pentium Windows 98 32MB RAM

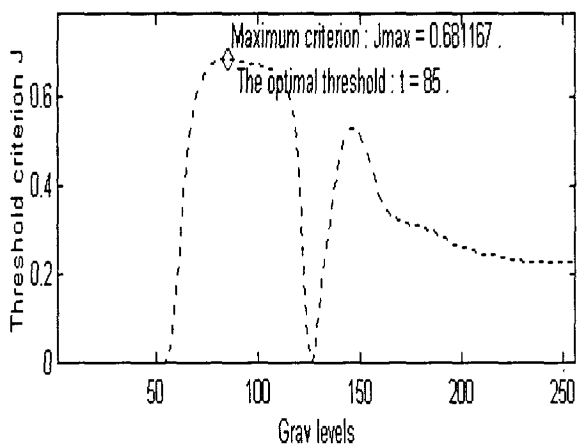


Fig. 8. Otsu's optimal threshold selection for Fig. 10(a).

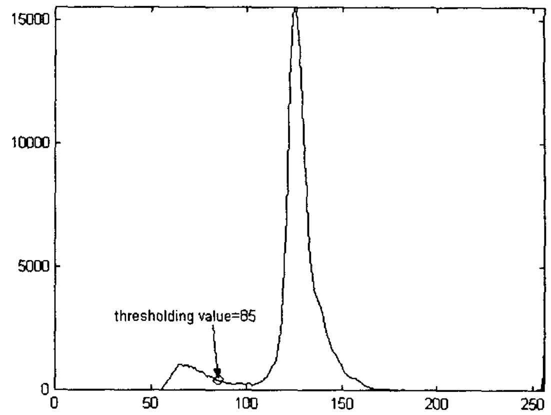


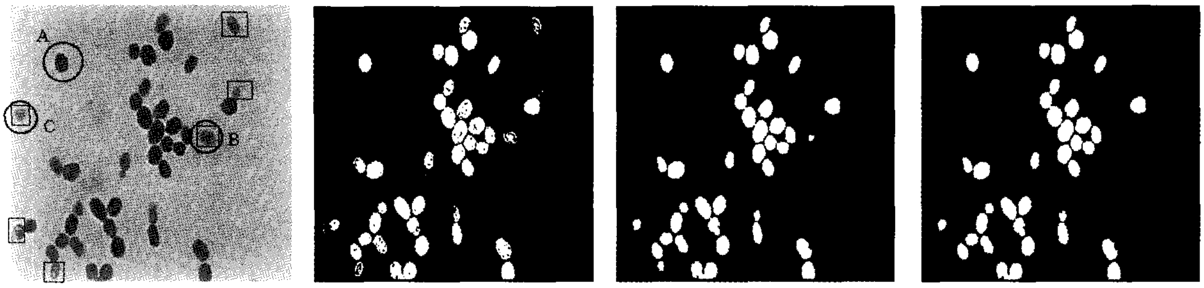
Fig. 9. Histogram and thresholding value for image Fig. 10(a).

Fig. 11 describes the labeled images for the image processing results, Fig. 10. In Fig. 11, (a) is an labeled image obtained by only Otsu's method as shown in Fig. 10(b), (b) is an labeled image obtained after cleaning, filling and opening filters as shown in Fig. 10(c) and (c) are labeled images obtained after filtering area less than 50 pixels as shown in Fig. 10(d). The counted numbers on the labeled images in Fig. 11 represent the number of labeled microbial cells. We can find that some cells are not detected by only Otsu's method in Fig. 11(a), especially, such as non-focused cells like marked cell 'C' in Fig. 10 (a). Fig. 11(b) shows that many cells are eliminated from Fig. 11(a) by canceling isolated points and filling holes by morphological filters such as cleaning, filling, and opening filters. Fig. 11(c) shows that marked cells (□) in Fig. 11(b) less than the area of 50pixels are erased by area filter. After labeling objects as shown in Fig. 11(c), feature extraction vectors such as area, perimeter, complexity, centroid, rotation angle and effective diameter are done.

Fig. 12 shows the area distributions of Fig. 11(b) and Fig. 11(c). In Fig. 11(c), The areas marked by □ as shown in Fig. 12 are removed.

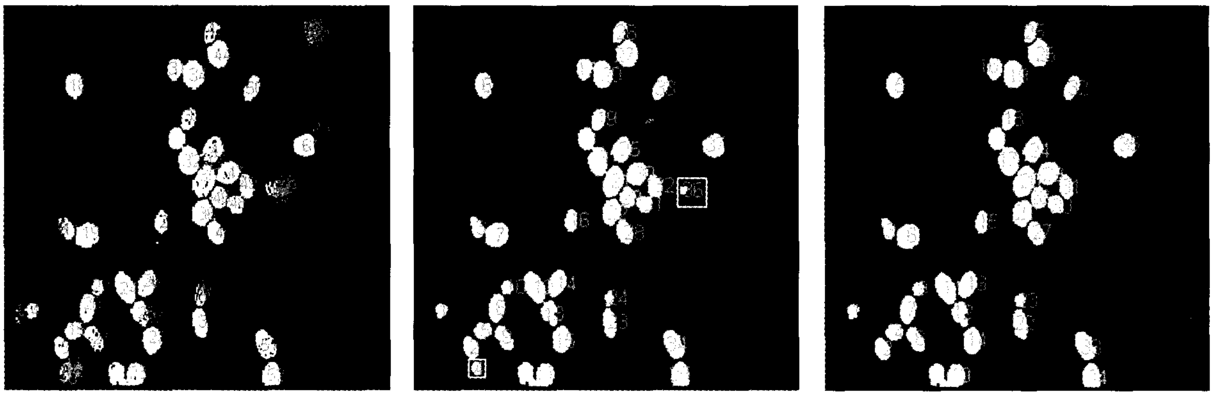
Fig. 13 shows sampled images from Fig. 12(a). Fig. 14 describes the same images as Fig. 13 sampled from Fig. 12(c). They include two ellipsoidal objects((a) and (b)) and two overlapped images((c) and (d)) in Fig. 13 and Fig. 14.

Fig. 15 shows perimeter image and rotation process. In Fig. 15, (a) is labeled image, (b) is peri-



(a) Original image (b) Binary image (c) Filtered image (d) Final image

Fig. 10. Image processing results obtained by the proposed method.



(a) Only Otsu's method (b) Opening, cleaning and filling hole (c) Proposed method

Fig. 11. Labeled images.

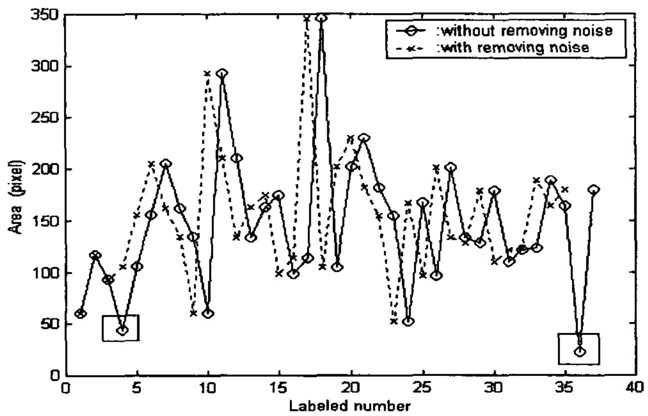
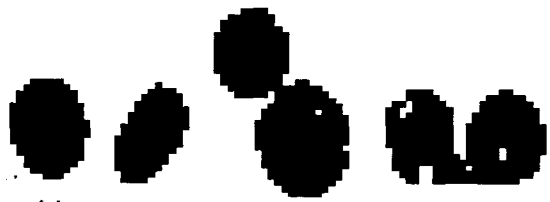
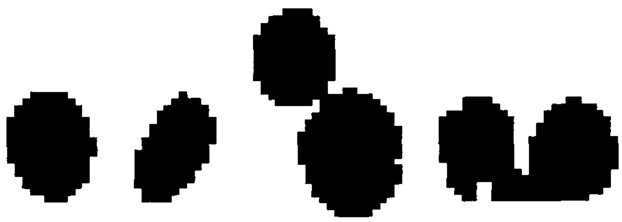


Fig. 12. Area plot after labeling.



(a) (b) (c) (d)

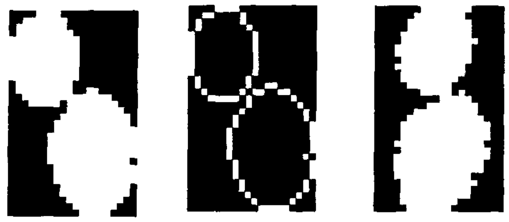
Fig. 13. Example images by only Otsu's method.



(a) (b) (c) (d)

Fig. 14. Example objects by the proposed method.

meter image and (c) is rotated image. With chain code of Fig. 15(b), the perimeter of labeled image (a) is measured by Eq. (11). First, we make perimeter image (Fig. 15(b)) from the labeled image (Fig. 15(a)) to produce the chain code. We find the



(a) Labeled image (b) Perimeter image (c) Rotated image

Fig. 15. Perimeter image and rotation process.

first pixel of the perimeter image and start the first pixel in the counter-clockwise direction. The direction numbers (Fig. 5(b)) along the perimeter as shown in Fig. 15(b) become chain code as following:

chain code of Fig. 14(a)=66677777777777878811812
12232333333343344455555

chain code of Fig. 14(b)=6667676767777787111222
32233233343454

chain code of Fig. 14(c)=6567677777787881111287
6676777787787888111122232334233334344454 5
654323333334434555

chain code of Fig. 14(d)=7656777777877788133218
7711111111111121233233334334454565676767 7
7776443333333444555

From the rotated images (Fig. 13(c) and Fig. 14(c)) that the labeled images (Fig. 13(a), Fig. 14(a)) are rotated by rotation angle θ in the clockwise direction, widths and heights of the labeled images are measured by Eq. (19) and Eq. (20). The feature vectors of the selected example objects are obtained from Eq. (10)~(21) as the following Table 2 by Otsu's method and Table 3 by the proposed method respectively. Feature vectors of objects having no hole or few holes such as (a), (b) and (c), by only Otsu's method and by the proposed method are same, but the feature vectors of objects having a lot of holes such as (d) by only Otsu's method and the proposed method is different. Because feature vectors is affected by noises and holes, it is necessary to remove noises contaminated in original image to get feature vector extraction exactly.

Conclusion

This paper shows basic research results to develop vision analysis system which can be applied for bio-process plant and to present a procedure extracting features in order to identify the object cells exactly under background with noise such as obstacles and non-symmetric contour, etc. To prove the effectiveness, the proposed method is applied for yeast, *Zygosaccharomyces rouxii*. The segmented image by only Otsu's method includes holes and noises, but the segmented image by the proposed method is clearer than that by only Otsu's method. So we can see that extracting the feature vector such as area, perimeter, complexity, centroid, rotation angle, width, height and effective diameter obtained by proposed method is more exact and more efficient to express geometric information of the cell than that obtained by only Otsu's method.

Table 2. Feature vectors of selected example objects by only Otsu's method

Object	A	L	e	(I_c, J_c, θ)	(W_x, W_y)	d_e
Unit	pixel	pixel	·	(pixel, pixel, °)	(pixel, pixel)	pixel
a	204	57.4264	16.1656	(7, 10, 4.4818)	(13, 19)	16.1165
b	124	44.2843	15.8153	(6, 9, −22.7614)	(15, 11)	12.5651
c	347	100.8823	29.3292	(11, 17, 24.6943)	(15, 32)	21.0194
d	290	106.0244	38.7627	(13, 9, −1.7231)	(24, 16)	19.2156

Table 3. Feature vectors of selected example objects by only proposed method

Object	A	L	e	(I_c, J_c, θ)	(W_x, W_y)	d_e
Unit	pixel	pixel	·	(pixel, pixel, °)	(pixel, pixel)	pixel
a	204	57.4264	16.1656	(7, 10, 4.4818)	(13, 19)	16.1165
b	124	44.2843	15.8153	(6, 9, −22.7614)	(15, 11)	12.5651
c	348	100.8823	29.2449	(11, 17, 24.7252)	(15, 32)	21.0496
d	297	106.0244	37.8491	(12, 9, −3.0738)	(24, 16)	19.4461

Acknowledgements

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Research Center for Ocean Industrial Development (RCOID) at Pukyong National University.

References

Hannah, I., D. Patel and R. Davies. 1995. The use of variance and entropic thresholding for image segmentation. Pattern Recognition, 28, 1135~1143.

Morii, F. 1995. An image thresholding method using a minimum weighted squared-distortion criterion. Pattern Recognition, 28, 1063~1071.

Otsu, N. 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern., 9, 62~66.

Pitas, I. 1993. Digital Image Processing Algorithms. Prentice Hall, N.Y., 254~281.

Thomas, C.R. and O.C. Paul. 1996. Application of image analysis in cell technology. Curr. Opin. Biotech., 7, 35~45.

Zalewski, K. and E.E. Bildanalyse. 1995. System zur on-line Charakterisierung von Hefesuspensione. Master of Diploma. TU-Berlin, VDI-Verlag GmbH.