

# Counting Method of Heterotrophic Bacteria Based on Image Processing\*

Hong Men, Yujie Wu, Xiaoying Li, Zhen Kou, Shanrang Yang

School of Automation Engineering  
Northeast Dianli University  
Jilin City, China  
menhong\_bme@163.com

**Abstract**—Heterotrophic bacteria is detected and counted by National Standard Method of China at present, i.e. flat dish numberation, which has the disadvantages of subjectivity, big error and low efficiency because of relying on manual counting. A method of rapid determination of Heterotrophic bacteria in industrial cooling water is developed. After converting the original image into a greyscale one, we removed the edge image of culture dish on the basis of the principle that three unco-linear points can determine a circle. Then binarization with the greyscale-weighted threshold method, distance transform and watershed algorithm to segment overlapping colonies are applied, and finally four-connected searching algorithm was used to mark and count the colony. It has been proved that the method greatly improves the accuracy and efficiency of the heterotrophic bacteria counting and the counting results are not affected by shapes or sizes of the colony. It has been revealed that colony counting based on image processing method can be an effective way.

**Keywords**—image processing, colony counting, heterotrophic bacteria, threshold segmentation, Biofouling

## I. INTRODUCTION

Industrial cooling water is a particular environment which is suitable for microorganisms' growing. When the PH value, temperature, illumination of water is suitable, they are able to grow and reproduce rapidly. Once they propagate excessively in the industrial cooling water system, biofouling could form. This causes many problems such as increase of the frictional resistance in tubes and decrease in the heat exchange capacity of heat exchanger. It also leads to under-deposit corrosion, which causes perforation of equipments and influences the regular production. Consequently, it is absolutely necessary to detect microbes in power plant and chemical plant. Heterotrophic bacteria are one main kind of microbes in industrial cooling water. The counting of it is detected by National Standard Method of China-flat dish numberation at present. Because the amount of experimental samples is always large, it has the disadvantages of subjectivity, big error and low efficiency because of relying on manual counting. So it needs to be modified urgently.

With the development of computer vision technology, various kinds of automatic analysis systems of particle images based on image processing and visual analysis are widely used

because they can detect the amount of particles and extract characteristic parameters of particles automatically, rapidly and objectively. For instance, according to the characteristics of colony image, Yingli Zhou, Libo Zeng and their assistants proposed the method of binarization with between class variance (BCV), distance transform and watershed algorithm to segment overlapping colonies, then counting the amount and analyzing the variety, shape and size of the colony [1]. It can replace the traditional manual counting method. Corkidi G and his assistants developed an automatic colony image analyzing system, which can be able to segment, statistic and analyze the overlapping and inhomogeneous colonies [2]. To improve the accuracy of automatic segmentation, You Yu-sai and his assistants studied on erythrocyte image [3]. Zenghua Zhao realized the automatic classification and counting of metarhizium colonies applying the wavelet analysis theory [4]. In the field of industrial production, Qiang Song studied the bar images and realized the automatic counting of bars according to the fact that the surface of bars is round [5].

After capturing the Heterotrophic bacteria images using CCD with 7 million pixels, we removed the edge image of culture dish on the basis of the principle that three unco-linear points can determine a circle firstly. Then we converted it into a binary one with the greyscale-weighted threshold method. Thirdly, overlapping colonies were segmented using the distance transform and watershed algorithm. Finally, the four-connected searching algorithm was used to mark and count the colony. It has been revealed that colony counting based on our method is promising and obtain a robust and fast segmentation of colony images.

## II. IMAGE PROCESSING ALGORITHM

### A. Grey Processing of the Colony Image

The format of our image captured by CCD is real colour JPG, as is shown in Fig. 1. In order to describe the characteristics of each pixel in it, we should use three helves of it: R, G, B. Because of this, it is relatively a time consuming task. While grey level image is a data matrix, the data of which represents the intensity value within a certain range. Therefore, we should convert it into a greyscale one first.

\*This work is supported by CNSF # 60604023 and National Basic Research Program of China (973 Program) #2007CB206904 to to H. MEN

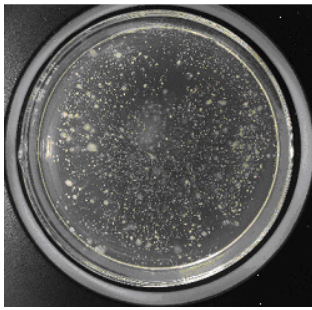


Fig. 1 Primary colony image

### B. Removal of the Edge of Culture Dish

We can see it clearly from Fig. 2 that the edge of culture dish takes up a very big room of the image which badly influences the later counting process. Thus, it should be removed firstly. Because it is a regular circle and the colonies are equably distributed in the circle, we can list the circle equation of the edge of culture dish. This can be achieved by changing the greyscale value of outside the circle into zero. Fig. 3 is the final result of our processing.

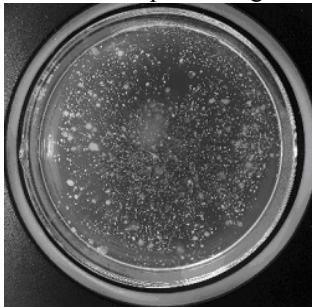


Fig. 2 Greyscale colony image

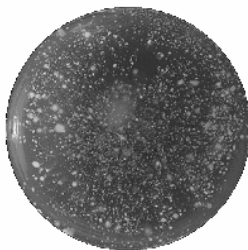


Fig. 3 Colony image after removal of the edge of culture dish

### C. Binarization of Colony Images

It can be seen from Fig. 3 that the colonies are equably distributed in every corner of the culture dish and the value of greyscale is even and coherent; the contrast between colonies and background is very high. According to this characteristic, we can consider the method of threshold segmentation to process the images by binarization, by which the colonies can be initially separated from background. In order to show it conveniently, we process the binary images by reverse colour, what is to say, setting the colony greyscale into 0 and the background greyscale into 1. Here we will talk two methods of processing images by binarization: the iteration threshold method and the greyscale-weighted threshold method [6].

#### 1) Iteration threshold method

It is supposed that the minimum and maximal intensity value are  $Z_{\min}$  and  $Z_{\max}$  respectively. Firstly, we suppose an original threshold

$$T_0 = (Z_{\min} + Z_{\max}) / 2 \quad (1)$$

Then segment the image into target and background using it. At the same time, it evaluates the average greyscale value of the two portions and makes it be the new chromatic value. We can attain the final chromatic value 92 by programming. The result attained by using this method is shown in Fig. 4. We can see from it that the value is somewhat low which makes some background points recognized as target ones, especially in the left portion of the picture. Then there is becoming a big block of black area which goes against the proceeding of the later counting.

#### 2) Greyscale-weighted threshold method

The greyscale-weighted threshold method is basically similar to the first one. Both of them obtain a threshold through finite iteration. The difference between them is the original threshold value

$$T_0 = \omega Z_{\min} + (1 - \omega) Z_{\max} \quad (2)$$

and new threshold value

$$T_{K+1} = \omega Z_0 + (1 - \omega) Z_B \quad (3)$$

So the latter algorithm can determine the segmenting threshold by choosing different threshold weight coefficients  $\omega$  according to specific case of the image.

In contrast to the method of evaluating optimal chromatic value by iteration, the greyscale-weighted threshold method's real-time performance is better. As the weight coefficient  $\omega$  changes, the greyscale-weighted threshold method can be suitable for both the segmentation of dark and bright objects. The value of bright object's weight coefficient  $\omega$  is taken beyond 0.5; while the value of dark object is taken below 0.5; the value of moderates brightness image's weight coefficient  $\omega$  is taken by 0.5. So we choose the greyscale-weighted threshold method to segment the colony image in this article. It can be seen from Fig. 3 that the contrast between colony and background is much higher, so  $\omega$  is taken beyond 0.5. After many experimental studies, we discovered that when  $\omega$  is taken by 0.65 and the threshold is taken by 129, the effect of segmentation is promising. The final result is shown in Fig. 5. Compared with Fig. 4, the segmentation effect of Fig. 5 is much better. However, one part of background in left part is wrongly divided into object, whose reason is caused by the problem of light source. We can ignore it as noise with the method of statistic and amendment method in the later counting process.

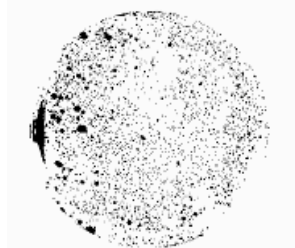


Fig. 5 Segmentation result using the greyscale-weighted threshold method

#### D. Segmentation of Overlapping Colonies

As can be seen from the binarization colony image (Fig.5), there are many colonies which overlap together apparently. In order to ensure the accuracy of the later counting work, we should segment the overlapping ones. We applied the method of the distance transform and watershed algorithm.

Distance transform is aiming at binary image [7]. In two-dimensional space, a binary image can be thought to include two kinds of pixels: object and background pixels. The value of object pixel is 1, while the background one is 0. The result of distance transform is not another binary image, but a greyscale one is distance image. The value of each pixel in the image is the distance between it and the nearest nonzero-valued one.

Watershed transformation is a classical method of image segmentation in mathematical morphology [8]. In geography, a watershed is the ridge that divides areas drained by different river systems. A catchment basin is the geographical area draining into a river or reservoir. The watershed transformation applies these ideas to greyscale image processing in a way that can be used to solve a variety of image segmentation problems. Understanding the watershed transform requires that we think of a greyscale image as a topological surface, where the values are interpreted as heights. If we imagine rain falling on this surface, it is clear that water would collect in the two areas labelled as catchment basins. Rain falling exactly on the labelled watershed ridge line would be likely to collect in either of the two catchment basins. The watershed transform finds the catchment basins and ridge lines in a grey-scale image.

Using the method mentioned above, we can segment the overlapping colonies. The final image is as shown in Fig. 6. Because the size of original image is big and the length of our article is limited, the length and width of displaying image is only 20% of the original one. Because of this, the effect after segmentation is not evident. Fig. 7 is just a part of Fig. 6. As can be seen from it, overlapping colonies are separated reasonably and effectively except a few colonies which overlap with each other seriously.

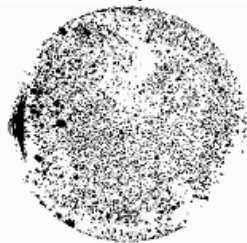


Fig. 6 Segmentation result of overlapping colonies

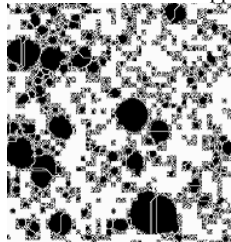


Fig. 7 The partial enlargement of Fig. 6

#### E. Heterotrophic Bacteria Counting

1) *Connected region labelling*: In order to obtain the number of the colonies, we should perform label processing to the image. Label processing is such a method that assigns the whole pixels which is contained in the same area the same label, while different labels to different area ones. Therefore, the number of connected areas can be obtained by counting the different labels. In this paper, we apply the four-connected searching algorithm to perform our label processing [9]. Its steps are summarized as follows:

a) *Initializing the variables*: label  $K$  is 0; list  $L$  is empty; the mark of current point  $(x, y)$ ,  $k(x, y)$  is set to be 0.

b) *Searching the image*: in the order of left to right, up to down, until finding the last pixel. If the current pixel  $(x, y)$  is black and the intensity value  $k(x, y)$  is 0, we plus 1 to the value of  $K$ , assign a new  $K$  to  $k(x, y)$ , save its adjacent pixels in four directions: up, right, down and left into the list  $L$ . Then go to step *c*), otherwise (when the pixel is white or it is black but  $k(x, y)$  is not equal to 0) go to step *b*) and scanning the next pixel continually.

c) *Picking out a pixel from list  $L$  and neglect it*: If it is black and  $k(x, y)$  is equal to 0, we assign  $K$  to its label and add the connected pixels of it in four directions to  $L$ . If not, we neglect it and return to step3 until the list is empty. Then continue to go to step 2.

d) *Finally, output the number of connected areas  $K$* .

2) *Colony counting*: If the colony image after segmentation doesn't include pseudo-objects like impurity and noise, the number of colonies in the image is the maximum number of labels actually. But in fact, there might be more or less pollution in previous colony inoculation operation, so besides colonies, a few subtle impurities and large area of pollution are also included in the segmented image. In order to count colonies accurately, we adopt the method of statistic and amendment to decrease the impact of disadvantageous factors as much as possible. In the process, we introduced three critical parameters of the connected areas: area, perimeter, shape factor.

a) *Connected region area  $A$* : The area of connected region can be simply defined as the number of pixels surrounded by the boundary of labelled parts. It could be obtained by scanning the labelled image only one time. For example, the area  $i$  connected domain equal to the number of all pixels whose label  $k(x, y)$  is  $i$  in the image.

b) *The perimeter of connected area  $C$* : The perimeter of connected area can be calculated from the number of pixel points surrounded by the boundaries. As to the labelled image using four-connected searching algorithm, the boundary points can be defined as follows: If the point itself is included in the labelled area and is adjacent to the pixels which represent background, it is boundary point. Concretely, if the label of current point  $(x, y)$ :  $k(x, y)$  is equal to  $i$  and during its four neighbourhoods, there is at least one point whose labelled value is 0 (representing background), it is one of boundary

points of  $i$ th connected region. The number of boundary points which belong to the same area is namely the perimeter of it.

c) *Shape factor  $R$* : The ratio of the area of connected region ( $A$ ) and the square of its perimeter ( $C^2$ ), that is,

$$R = 4\pi A / C^2, \quad (4)$$

Where, the upper limit value of  $R$  is 1. It is able to describe the degree to which the shape of the area closes a circle. The more it closes a circle, the greater  $R$  is.

Generally, the noisy pixels are much smaller than colonies, so we can separate them from colonies by pre-assigning an area threshold:  $A_0$ . If the area of one region  $A$  is smaller than  $A_0$ , we consider it as an impurity and neglect it. For noise and pollution with big area, we can see from figure 6 that there are noise pixels in its left part. The shape factor of it  $R$  is quite small. In order to know whether it is pollution and noise or not, we can choose a shape factor threshold  $R_0$ . If the shape factor of one domain  $R$  is smaller than  $R_0$ , we consider it as pollution and neglect it.

### III. RESULTS AND DISCUSSIONS

We did many experiments using the algorithm that has been discussed above. For 18 Heterotrophic bacteria samples whose sample dilution is from  $10^0$ - $10^{-5}$ , we did human counting and automatic counting respectively. The results are shown in Table I. As can be seen from it, compared with human counting, the maximum relative error of automatic counting is 2.9%; the minimum relative error is 0% and the average one is 1.7%. The main reason with error is that some colony areas after segmentation is within the range of the area threshold we set and few colonies which overlap with each other were not segmented effectively. The error can be eliminated by man-machine interaction. Since the relative error of automatic counting result is smaller than the allowable error and during the manual counting of actual big samples, the relative error of which is greater than the one of automatic counting, the automatic counting result using our algorithm is acceptable.

TABLE I. COMPARISONS BETWEEN HUMAN COUNTING AND AUTOMATIC COUNTING RESULTS OF HETEROTROPHIC BACTERIA SAMPLES

Sample dilution	Human counting	Automatic counting	Relative error %
$10^0$	3495	3596	2.9
	2650	2586	2.4
	3056	3117	2.0
$10^{-1}$	1498	1529	2.1
	1237	1209	2.3
	983	998	1.5
$10^{-2}$	332	326	1.8
	379	387	2.0
	289	293	1.4
$10^{-3}$	154	156	1.3

	130	128	1.4
	179	182	1.7
$10^{-4}$	107	108	0.9
	96	95	1.0
	123	120	2.4
$10^{-5}$	59	58	1.7
	46	46	0
	63	62	1.6

### IV. CONCLUSIONS

In allusion to the present Heterotrophic bacteria counting method having the disadvantages of subjectivity, big error and low efficiency, a simple and practical algorithm for automatic colony counting is proposed. After grey processing of the acquired colony image, we removed the edge image of culture dish, converted it into a binary one with the greyscale-weighted threshold method and then segmented overlapping colonies using the distance transform and watershed algorithm. In order to statistic the number of colonies, we use the four-connected searching algorithm to mark the colony area firstly. Then we adopted the method of statistic and amendment to decrease the impact of disadvantageous factors as much as possible. It has been revealed that colony counting based on image processing method is feasible. Not only it has the advantages of fast, simple and accurate, but also the counting result does not be affected by using inoculation, variety, shape and size of the colony. Furthermore, the method is also widely applicable for counting other kind of microbial and it greatly improves the accuracy and efficiency.

### REFERENCES

- [1] Y. Zhou, L. Zeng, and J. Liu, et al, "A method for automatic colony counting based on image processing and its realization," *Journal of Data Acquisition & Processing*, vol. 18, no. 4, pp. 460-464, December 2003.
- [2] G. Corkidi, R. Diaz-uribe, J L. Folch-mallol, et al, "An image analysis method that allows detection of confluent microbial colonies of various sizes for automated counting," *Applied and Environmental Microbial*, vol. 64, no. 4, pp. 1400-1412, 1998.
- [3] Y. You, H. Yu, "A separating algorithm for overlapping red blood cell images," *Journal of Image and Graphics*, vol. 10, no. 6, pp. 736-740, 2005.
- [4] Z. Zhao, Y. Shu, J. Wang, "Research for industrial microbial germ automatic recognition based on wavelet transformation," *Journal of Tianjin University (Science and Technology)*, vol. 10, no. 6, pp. 736-740, 2005.
- [5] Q. Song, K. Xu, J. Xu, et al, "Automatic counting technique for steel bars based on image processing," *Iron and Steel*, vol. 39, no. 5, pp. 34-37, 2004.
- [6] F H Y. Chan, F K. Lam, "Adaptive threshold by variational method," *IEEE Transactions on Digital Object Identifier*, vol. 7, no. 3, pp. 468-473, 1998.
- [7] C. Olivier, "Locally adaptable mathematical morphology using distance transformations," *Pattern Recognition*, vol. 39, no. 3, pp. 405-416, 2006.
- [8] H. Sun, J. Yang, M. Ren, "A Fast Watershed Algorithm Based on Chain Code and Its Application in Image Segmentation," *Pattern Recognition Letters*, vol. 26, no. 9, pp. 1266-1274, 2005.
- [9] W. Lou, T. Yu, F. Gu, "A method of automatic protozoan counting based on image processing technology and its realization," *Research and Exploration in laboratory*, vol. 26, no. 4, pp. 24-28, 2007.