An Automated Fish Counting Algorithm in Aquaculture Based on Image Processing

Jiuyi Le[1,a] , Lihong Xu[1,b]

[1]College of Electronics and Information Engineering, Tongji University,

Shanghai, China

[a]lejiuyi91@163.com, [b]xulhk@163.com

Keywords: Fish Counting; Free-swimming; Overlap; Skeleton Extraction

Abstract. A new algorithm based on endpoints of skeleton is presented to efficiently get the number of fish in this paper. Considering the complexity of underwater environment like lack of light, this paper presents an improved adaptive thresholding method to segment the fish image better. In addition, the object of our research is free-swimming fish. The overlapped fish in the image makes the counting result inaccurate often. So after segmentation and morphological processing, this paper adopts image thinning method to extract the skeleton of fish. After that, we get the fish number according to the number of corresponding endpoints in the image. The experimental results show that the method can accurately count the fish population even under high overlapped degree.

Introduction

With the introduction and popularization of the concept of marine ranching, the offshore aquaculture industry has developed rapidly in recent years. The precise counting of the fish stocks provides the basis for the effective management of scientific feeding, sale, transportation, and breeding density control. The traditional counting method uses containers like net bag to sample, which brings lots of disadvantages like low efficiency, limited artificial experience and so on [1]. Since the 80s of the 20th century, a variety of research institutes invented a number of fish counters, which was difficult to apply to the actual production because it can be easily affected by environmental constraints, fish size and other factors. Besides, it is so expensive that few fishermen want to use it in actual production [2].

With the improvement of the computer vision and image processing technology, they are applied to more industries to improve the level of automation of production. At the same time, fish counting method based on image processing is attracting more and more attention due to its high efficiency and accuracy [3]. Some research methods, such as connected area counting, area counting, data fitting [4], curve evolution [5] and neural network [6], did not take the overlapped problem into consideration. So they will not give an accurate number of fish when some fishes get together. In this paper, we propose an algorithm based on endpoints of extracted skeleton, which attenuates the segmented images so that the fish images are refined into some one-pixel wide skeletons. And then the fish population is accurately counted using the number of endpoints.

Experimental Platform

The experimental platform consists of three parts. The whole experiment is carried out in a transparent glass aquarium, the size of the fish tank is $1.2m \times 0.6m \times 0.5m$. The second part is an underwater image acquisition module, which consists of a waterproof camera and an underwater cloud platform. The camera capture the video vertically and we can adjust the bracket to obtain the appropriate focal length. The camera is SHARP 1/4 CCD model, surrounded by a circle of LED compensating underwater light in the dark condition. The third part of the system is the host computer.

The analog video must be converted to digital signal to be saved in the computer. The resolution of all the video in this experience is 640×480. The overall system diagram is shown in Figure 1:
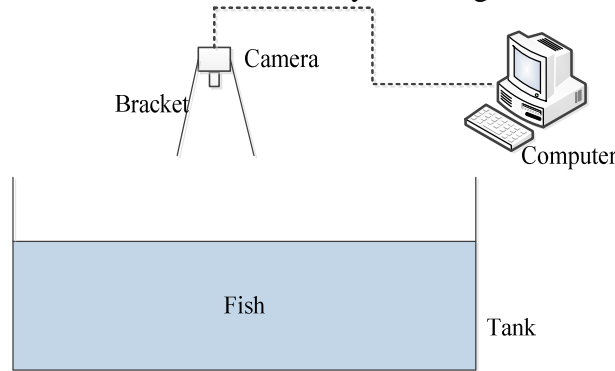

Figure.1 Experimental Platform

Image Preprocessing

Due to the diversity and complexity of the underwater environment, the underwater images usually suffer from severe noise, which reduces the quality of underwater images and affects the accuracy of image analysis. In order to make the underwater image easier to analyze, we need to carry on appropriate pretreatment to the underwater picture [7].

For the objects to be studied, we only focus on the contour information of the fish in the image. In order to remove redundant information, the author first converted the underwater color image to grayscale image. In the RGB model, if R = G = B, then the color information represents a grayscale color, where the value of R = G = B is called the grayscale value. So every pixel of the grayscale image only need one byte to store the grayscale value (also called intensity value or brightness value) which range from 0 to 255. The author uses the weighted average method to gray the image. Every pixel get its grayscale value according to the formula: $f(i, j) = 0.30R(i, j) + 0.59G(i, j) + 0.11B(i, j)$. Then the median filter is used to remove the noises. The result after one preprocessing operation is satisfactory. The sample grayscale image is as follows:
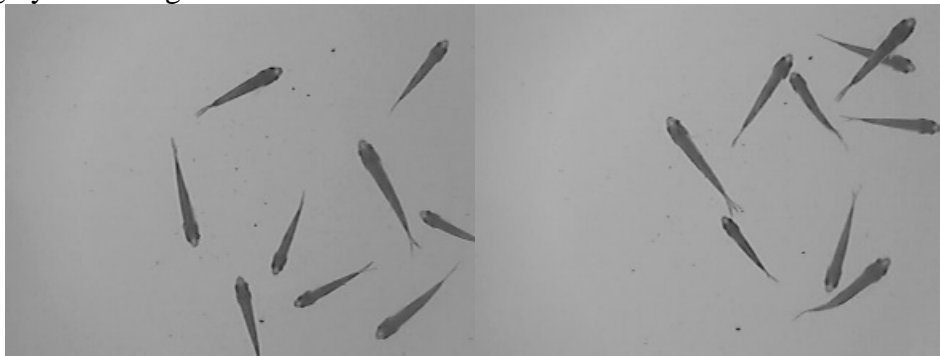

Figure.2 Grayscale image

Image Segmentation

Image segmentation is an important step in the analysis of image data, and in all image segmentation algorithms, threshold segmentation has been widely used for a long time for its extreme simplicity and high practicality. For a better view, all the images appear in this article are images that reverse the black and white pixels of original images.

Otsu Method and Related Derivatives

The maximum between-class variance is an important basis for statistical unsupervised clustering in statistical pattern recognition. The image segmentation method based on maximum between-class variance, which is also called Otsu method, is proposed by famous Japanese scholar NOBUYUKI [8].

Among the global thresholding techniques, the Sahoo et al. study [9] concluded that the Otsu method was one of the better threshold selection methods for general real world images with respect to uniformity and shape measures. This method selects threshold values that maximize the between-class variances of the histogram.

Let the pixels of a given image be represented in L gray levels, and let the probability of the occurrence of a pixel with a grayscale i to be approximated by its frequency. The normalized histogram of the image is:

$$H = \{ p_0, p_1...p_{L-1} \}, \quad \sum_{i=0}^{L-1} p_i = 1 . \tag{1}$$

Now suppose that we dichotomize the pixels into two classes B and O (background and object, or vice versa) by a threshold at level t; B denotes pixels with levels $[1, 2\mathbf{L}, t]$, and O denotes pixels with levels $[t+1\mathbf{L}, L]$. The probability of background and object is $P_B(t)$ and $P_O(t)$, respectively. The average gray level of the background and object is $m_B(t)$ and $m_O(t)$, respectively. So the overall average gray level of the image is:

$$m = \sum_{i=0}^{L-1} i p_i . \tag{2}$$

According to the definition of variance, the total variance of levels of this image is:

$$s^2 = E[(i-m)^2] = \sum_{i=0}^{L-1} p_i (i-m)^2 . \tag{3}$$

Then we can easily find that for a given image, the average gray level m and variance s^2 are both constant. So they are independent of the threshold discussed here.

The within-class variance of the background class and the object class is defined as:

$$s_B^2 = \frac{1}{P_B(t)} \sum_{i=0}^{t} (i-m_B(t))^2 * p_i . \tag{4}$$

$$s_O^2 = \frac{1}{P_O(t)} \sum_{i=t+1}^{L-1} (i-m_O(t))^2 * p_i . \tag{5}$$

Let $s_W^2(t)$ represent the within-class variance, which is defined as:

$$s_w^2(\mathrm{t}) = P_B(t) s_B^2 + P_O(t) s_O^2 . \tag{6}$$

Let $s_{OB}^2(t)$ represent the between-class variance, which is defined as:

$$s_{OB}^2(\mathrm{t}) = P_O(t)\left(m_O(t)-m\right)^2 + P_B(t)\left(m_B(t)-u\right)^2 . \tag{7}$$

In fact,

$$s_w^2 + s_{OB}^2 = \sum_{i=1}^{L}\left[(i-m_T)^2 + 2im_T - 2m_T^2\right] p_i = \sum_{i=1}^{L}(i-m_T)^2 p_i = s^2 . \tag{8}$$

Therefore, the total variance of the image is equal to the sum of the within-class variance and between-class variance of the object class and the background class:

$$s^2 = s_W^2 + s_{OB}^2 . \tag{9}$$

Therefore, the threshold selected by Otsu method is a value which maximizes the $s_{OB}^2(t)$. In other words, the formula to select optimal threshold t^* according to Otsu criterion is that:

$$t^* = Arg \max_{0 \le t \le L-1}\left(s_{OB}^2(t)\right) . \tag{10}$$

According to the formula (9), we can easily get the equation that $s_W^2 = s^2 - s_{OB}^2$. So that equation (10) is equivalent to:

$$t^* = Arg \min_{0 \le t \le L-1} \left[P_B(t) s_B^2(t) + P_O(t) s_O^2(t) \right] . \tag{11}$$

Equation (11) shows that maximizing between-class variance and minimizing within-class variance can be equivalent to get the optimal threshold based on OTSU criteria.

However, Hou and other people in the Research Institute of Communications in Singapore have pointed out in document [10] that The Otsu method works well only when the images to be thresholded have clear peaks and valleys—in other words, it works for images whose histograms show clear bimodal or multimodal distributions. In order to overcome the problem that the optimal threshold selected in Otsu always lean to the class with larger variance, a new method based on minimum class variance to select threshold was presented in document [10]. In this document, the minimized objective function is:

$$t^* = Arg \min_{0 \le t \le L-1} \left[s_B^2(t) + s_O^2(t) \right] . \tag{12}$$

It has been proved in document [10] that this method can buffer the tendency of the threshold moving to the class with larger variance.

Wu pointed in the document [11] that Otsu is an optimal approximation based on L2-norm. In this paper, an image segmentation method based on L1-norm is proposed, and its optimal threshold is selected by:

$$t^* = Arg \min \{ S_w(t) \} = Arg \min_{0 < t < L-1} \{ w_0(t) d_0(t) + w_1(t) d_1(t) \} . \tag{13}$$

Where $w_0(t)$ and $w_1(t)$ represent the class probability of the background and object, respectively; $d_0(t)$、 $d_1(t)$ represent absolute difference of the background and object, respectively. They are shown as

$$d_0(t) = \sum_{i=0}^{t} \frac{p_i}{w_0(t)} | i - m_0(t) |, \ d_1(t) = \sum_{i=t+1}^{L-1} \frac{p_i}{w_1(t)} | i - m_1(t) | . \tag{14}$$

An Exponential Improved Algorithm Based on Otsu

Whether the original Otsu method or Hou's and Wu's improved method, all of them have some restricted conditions on gray histogram if they want to get a good result. However, the gray histogram of the underwater image is varied due to various factors such as illumination, water turbidity and so on. According to the improvement research of Hou and Wu, we can find that class probability or histogram gray interval is helpful for Otsu method to segment the grayscale image better. Based on all the above research, this paper presents an improved Otsu algorithm which takes both class probability and histogram gray interval into consideration. The criterion function of threshold selection is:

$$s_w^q(t) = \sum_{i=0}^{t} \frac{p_i | i - m_B(t) |^{q_1}}{\left[P_B(t) \right]^{q_2}} + \sum_{i=t+1}^{L-1} \frac{p_i | i - m_O(t) |^{q_1}}{\left[P_O(t) \right]^{q_2}} . \tag{15}$$

Where $s_w^q(t)$ is the absolute within-class difference with exponential parameter. And t is the best threshold when it minimizes $s_w^q(t)$. q_1 and q_2 are two adjustable parameters, which are used as an index of within-class absolute difference and class probability. As can be seen from the formula, the threshold selection criteria in this paper fully consider the class probability and gray level difference, so that it can adapt to different situations for image segmentation. When q_1 and q_2 take some special value, the method presented in this paper is equivalent to the Otsu algorithm, Hou's method and Wu's improved method respectively. They are proved as followed:

① When $q_1=2$ and $q_2=0$, formula (15) is transformed into the following cases:

$$s_w^q(t) = \sum_{i=0}^{t} \frac{p_i \, |i - m_B(t)|^2}{[P_B(t)]^0} + \sum_{i=t+1}^{L-1} \frac{p_i \, |i - m_O(t)|^2}{[P_O(t)]^0}$$

$$= P_B(t) \sum_{i=0}^{t} \frac{p_i [i - m_B(t)]^2}{P_B(t)} + P_O(t) \sum_{i=t+1}^{L-1} \frac{p_i [i - m_O(t)]^2}{P_O(t)} = P_B(t) s_B^2 + P_O(t) s_O^2 = s_w^2(t). \qquad (16)$$

Compared with formula (6), we know that the method of this paper is equivalent to the original Otsu method at this moment.

② When $q_1 = 1$ and $q_2 = 0$, we can also prove that this method is equivalent to the minimum within-class absolute difference method proposed by Wu. The process of derivation is similar to the derivation of the above formula (16), which will not be repeated here.

③ When $q_1 = 2$、 $q_2 = 1$, this method is equivalent to the minimum variance method proposed by Hou.

From the above proof, we find that the original Otsu algorithm, Hou's minimum variance method and Wu's minimum within-class absolute difference method are only one of those specific form of method presented in this paper [12]. These three methods have a solid theoretical foundation of mathematical statistics, which also shows the rationality of the method presented in this paper. This method is more adaptive when you adjust the parameters. For the grayscale image in Figure 2, the result using Otsu and method presented in this paper is show as (b) and (c) in figures 3 respectively.

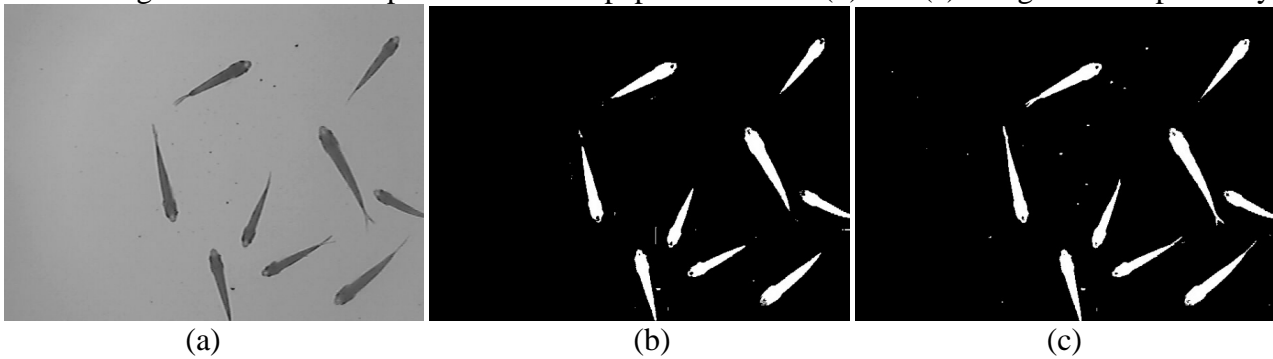(a)   (b)   (c)

Figure.3 Segmentation result

The figures above show that the original Otsu method is not good enough to process the detail of the images. The Otsu method mistaken a lot of fish tail pixel points as the background points. Nevertheless, the improved method presented in this paper has a better performance. The binary image has some noises because of the impurities in the water. In addition, there are also some black pixels in the fish area and some fish area is split. So we conduct a closed operation (corrosion after expansion) and then a median filter on the binary image [13]. And finally, a very ideal binary image is shown in figure 4.
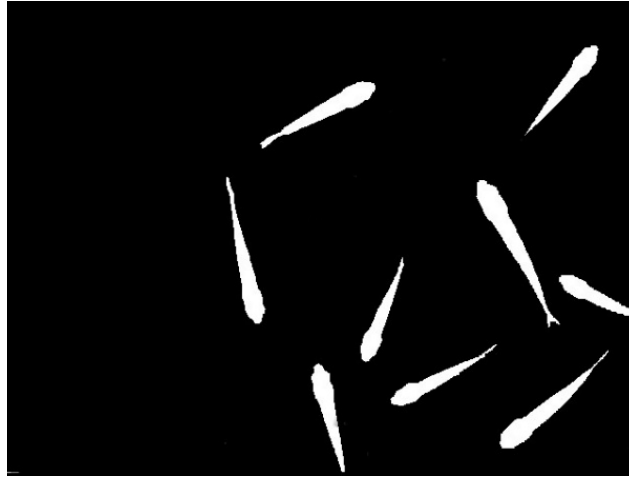
Figure.4 Binary image after been filtered

Skeleton Extraction

This paper is aimed to count the fish precisely. So we are going to thin the fish image in this part to make fish counting easier, which is also called skeleton extraction. The so-called skeleton, can be understood as the axis of the image, such as rectangle's skeleton is the center line of the long direction, circle's skeleton is its center point [14].

Figure 5 is the object pixel p_1 and its 8 neighborhood. Every pixel is either 0 or 1. Our method for extracting the skeleton of a picture consists of removing all the contour points of the picture except those points that belong to the skeleton. In order to preserve the connectivity of the skeleton, we divide each iteration into two subiterations [15].
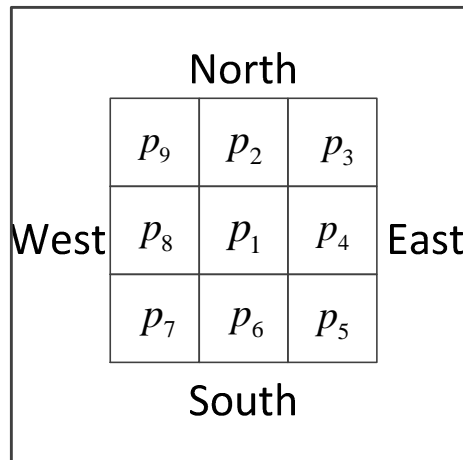

Figure.5 8-neighborhood

In the first subiteration, the contour point p_1 is deleted from the digital pattern if it satisfies the following conditions:

(a) $2 \le B(p_1) \le 6$, (b) $A(p_1) = 1$, (c) $p_2 * p_4 * p_6 = 0$, (d) $p_4 * p_6 * p_8 = 0$.

Where $A(p_1)$ is the number of $0 \rightarrow 1$ patterns in the order set $p_2, p_3...p_8, p_9$ that are the eight neighbors of p_1. $B(p_1)$ is the number of pixels which equals to 1.

In the second subiteration, only conditions (c) and (d) are changed as follows:
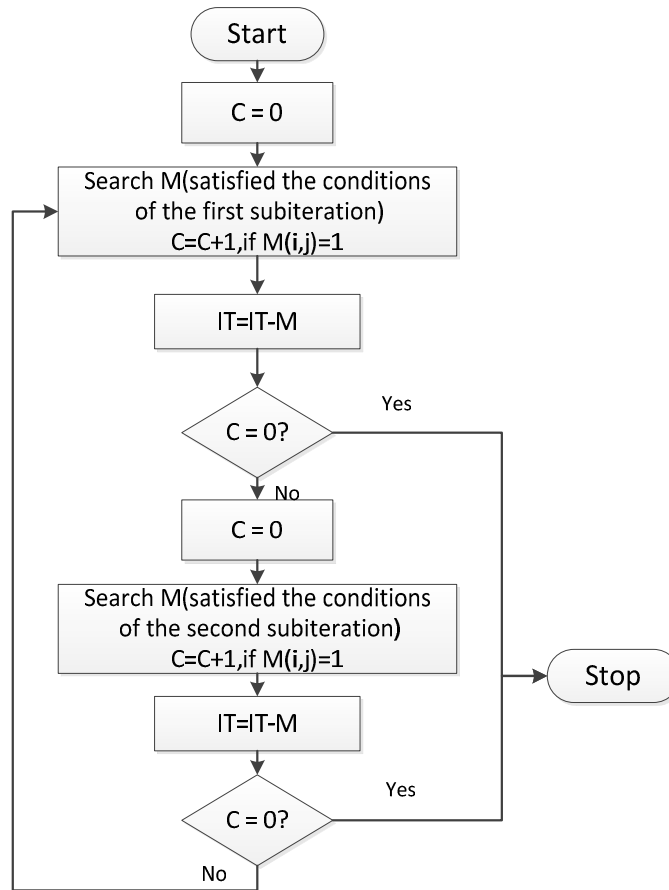
(c') $p_2 * p_4 * p_8 = 0$, (d') $p_2 * p_6 * p_8 = 0$.

Figure.6 Flowchart of the thinning algorithm

The solution to the set of equation (c) and (d) are $p_4 = 0$ or $p_6 = 0$ or ($p_2 = 0 \&\& p_8 = 0$). So the point p_1, which has been removed, might be an east or south boundary point or a north-west corner point. Similarly, it can be proved that the point p_1 deleted in the second subiteration might be a north-west boundary point or a south- east corner point. By condition (a), the endpoint of a skeleton line are preserved. Also condition (b) prevent the deletion of those points that lie between the endpoints of a skeleton line.

A flowchart of the thinning algorithm is shown in figure 6. The original image is stored in matrix IT and a counter C is set to 0. The result of the processed picture is stored in matrix M. The iteration stops until no points can be deleted. The image of fish skeleton is shown in figure 7.
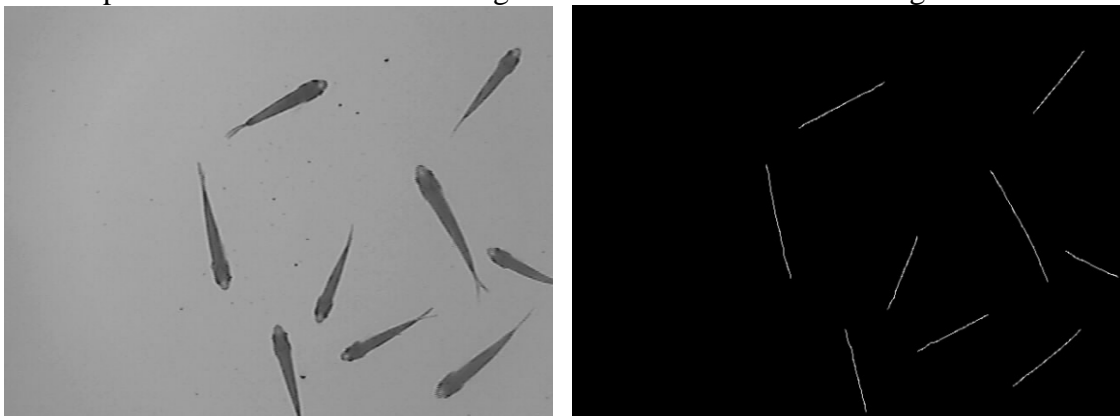

Figure.7 Skeleton image

Fish Counting

When the refinement process is completed, the fish population can then be counted based on the number of skeletal endpoints. And then we process the connected area one by one. The skeleton shape got by the former algorithm is 8-connected. So if there are only one pixel from the eight neighborhood whose value is 1, then we treat this pixel as the endpoint of the skeleton.

After all steps, the number of fish in the connected area is determined by the number of endpoints:

If there are 2 endpoints in the connected area, and then we believe the number of fish is 1.

If there are 3 endpoints in the connected area, and then we believe the number of fish is 2.

If there are 4 endpoints in the connected area, there may be 2 or 3 fish. And then we believe the number of fish is 2.5.

If there are n (n > 4) endpoints in the connected area, and then we believe the number of fish is n/2+1.

Finally, the number of fish in all connected areas is summed up to get the total number of fish [16].

Conclusion

In order to solve the problem of fish counting, this paper introduces the computer vision technology to improve the automation level. In view of the characteristics of underwater images, a more general adaptive thresholding method is proposed, and the segmentation of fish image is more accurate. Secondly, this paper adopts skeleton extraction method to solve the overlapped-fish problem cleverly. Three typical cases are shown as figure 8. From (a) to (c), more and more fish gather together which makes our result less accurate. We get hundreds of image from the videos, and experiments prove that the average counting error is less than 6%, which is much better than some traditional method like connected area method and so on.

But in the actual production, the underwater environment is much complex. For example, low visibility and lack of light and other issues are the urgent problem to be solved. When the fish density is too large, lots of them attach together. At that time, the result will not be precise as before. In the future work, we plan to conduct more experiments with the simulation of underwater environment.
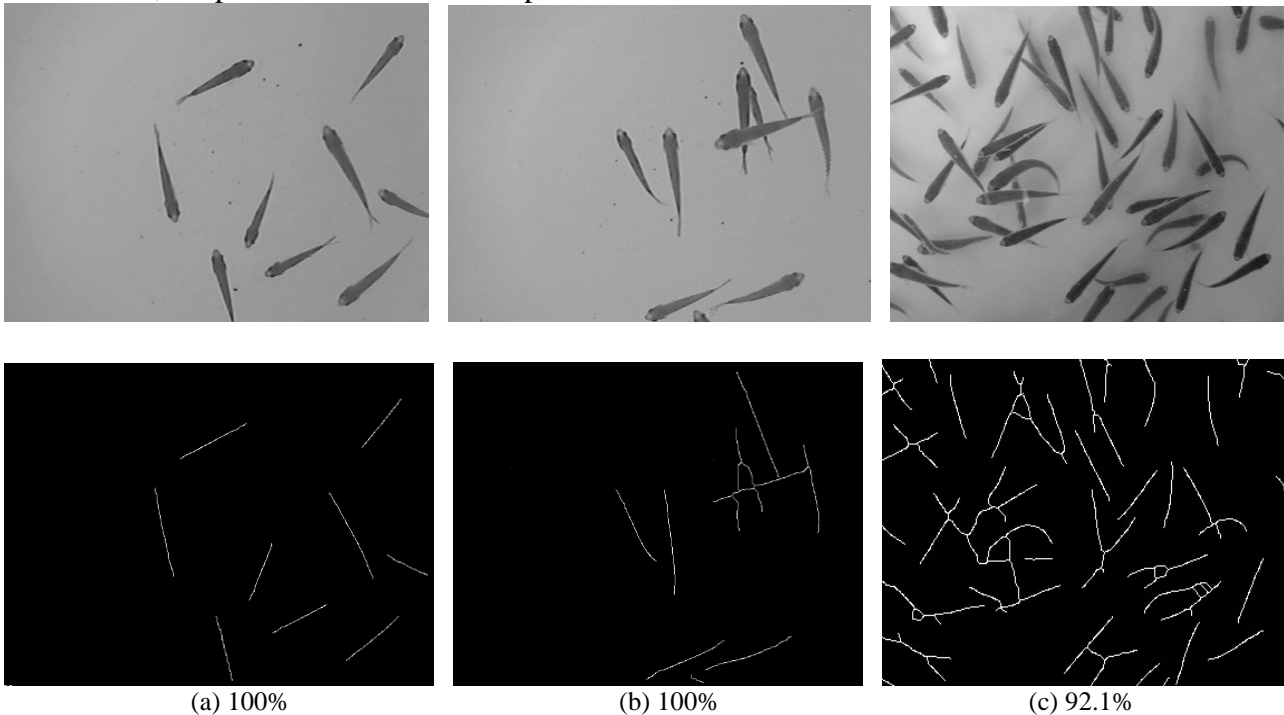

(a) 100%   (b) 100%   (c) 92.1%

Figure.8 Accuracy of counting with different degrees of overlap

Acknowledgements

References

[1] Assis J, Claro B, Ramos A, et al. Performing fish counts with a wide-angle camera, a promising approach reducing divers' limitations[J]. Journal of Experimental Marine Biology & Ecology, 2013, 445:93-98.

[2] Yi Jin-geng，Huang Gui-lin．Design of Optical-Electric Computer Controlled System for Multi-channel Fry Counter ［J］. Transactions of the Chinese Society of Agricultural Engineering, 1997, 13（4）：171-173.

[3] CHATAIN B, DEBAS L, BOURDILLON A．A photographic larval fish counting technique：comparison with other methods, statistical appraisal of the procedure and practical use ［J］. Aquaculture, 1995, 141（1-2）：83-96.

[4] Cong rong Zhu. A method for fish fry automatic counting based on machine vision [J]. Fishery Modernization, Volume 36, Issue 2, 2009.

[5] WANG Shuo, FAN Liangzhong, LIU Ying. The research of turbot fry counting method based on computer vision [J]. Fishery Modernization, 2015, 42(1):16-19.

[6] Zheng X, Zhang Y, Zheng X, et al. A Fish Population Counting Method Using Fuzzy Artificial Neural Network[C]. IEEE International Conference on Progress in Informatics and Computing. 2010:225-228.

[7] Jiang H F, Xu Y. Research Advances on Security Problems of Underwater Sensor Networks [J]. Advanced Materials Research, 2011, 317-319:1002-1006.

[8] OTSU N. A threshold selection method from gray level histograms [J]. IEEE Trans on SMC, 1979, 9(1): 62-69.

[9] Sahoo, P.K., Soltani, S., Wong, A.K., Chan, Y.C., 1988. A survey of thresholding techniques. Computer Vision, Graphics and Image Processing 41, 233–260.

[10] Hou Z, Hu Q, Nowinski W. L. On minimum variance thresholding [J]. Pattern Recognition Letters, 2006, 27(14): 1732-1743.

[11] WU Yi-quan, PAN Zhe. The Image Thresholding Algorithm Based on Minimum Within-Cluster Absolute Difference and Maximum Difference [J]. Journal of Signal Processing. 2008(06): 943-946.

[12] Yang Shuhong. Study on the Adaptive and Fast Algorithm of Gray Scale Image Thresholding[D]. Chongqing University, 2014.

[13] Fu X, Zeng D, Huang Y, et al. A fusion-based enhancing method for weakly illuminated images [J]. Signal Processing, 2016, 129:82-96.

[14] Liang Yu, Guangwei Wang. A Fast Thinning Algorithm for Binary Image [J]. Science & Technology Vision, 2012(14):111-112.

[15] Zhang T Y, Suen C Y. A fast parallel algorithm for thinning digital patterns [J]. Communications of the Acm, 1984, 27(3):236-239.

[16] Fan S, Liu J, Yang Y. Research and Realization of Image Recognition Technology in Fry Counting [J]. Fisheries Science, 2008, 27(4):210-212.