Sixth International Symposium on Underwater Technology
UT2009, Wuxi, China, April 2009

Automated Acoustic Method for Farmed Fish Counting and Sizing during Its Transfer Using DIDSON

Jun Han[1*], Naoto Honda[2], Akira Asada[1], and Koji Shibata[3]

[1]Institute of Industrial Science, University of Tokyo, Meguro-ku, Tokyo, JAPAN
[2]Natinal Research Institute of Fisheries Engineering, Fisheries Research Agency, Kamisu, Ibaraki, JAPAN
[3]Engineering Department, Toyo Corporation, Yaesu, Chuo-ku, Tokyo, JAPAN

ABSTRACT

Counting and sizing the farmed fish of large-size i.e. tuna, are often performed during the time of its transfer from one net cage to another. DIDSON provides an automated fish counting and sizing tool. However, neither its counter nor its sizer is suitable for measuring the farmed fish during its transfer because the transfer net is waved by currents and the fish image is often broken to several remains. This paper presents a fully automated acoustic method that can count and size farmed fish during its transfer by using DIDSON image. The background is subtracted after being stabilized by a phase only correlation method. Tracing the edges using contour tracing method, the segmentation of fish is obtained. To prevent recounting the same fish, a biologically relevant Kalman filter algorithm was designed and adopted to predict the fish movement. Through analysis of the spatiotemporal trajectory of the track, the automated counting was performed. The broken fish images were searched, and the body length was obtained by summing down the center line segments from the head to the tail of fish. The proposed system was verified to achieve the sizing accuracy within the range of -2.4 to 2.8 cm with the mean -0.4 cm by using the farmed Yellowtail, *Seriola quinqueradiata* (mean total length 83.1 cm).

Keywords

Fish counting, fish sizing, farmed fish, DIDSON

1 INTRODUCTION

A new fully automated software system has been developed for counting and sizing farmed fish during its transfer with a DIDSON (Dual frequency IDentification SONar, Belcher et al 1999). Counting and sizing farmed fish of large-size i.e. tuna and yellowtail during the fish is moved between two net cages are often performed to grasp its number and length. So far the visual count has mainly been taken by a diver for counting. Unfortunately, this kind of counting method is not only labor-intensive, but also its accuracy is low. As for sizing, the visual measurement is also used by the fish being taken out of the farming net, which, however, is tough to the farmed fish. Meanwhile DIDSON can provide those images the quality of which is close to such a level as attained with conventional optics at close ranges even in dark or turbid water, and therefore, has begun to be used for structure inspection, leak and flow detection, underwater surveillance and other underwater applications. The applications for fisheries are also promising such as the observation of fish behaviors. The automatic counting and sizing farmed tuna and yellowtail are the first trial with DIDSON. DIDSON provides an automated fish counting and sizing auxiliary tool. However, neither the counter nor the sizer is adequate for measuring the farmed fish during its transfer, because the background of the net image is difficult to be removed as it is waved by currents and also the fish echo is broken to several remains. This paper presents a fully automated fish tracking system that can count and size multi-fish passing through a transfer net.

2 MATERIAL AND METHOD

2.1 System Overview

Our fish counting and sizing system integrates seven modules (Figure 1) image stabilizer, background subtractor, fish detector, fish tracker, track analyzer, fish counter and fish sizer. The image stabilizer alleviates the shakes of images by using a phase-only correlation technique. The background subtractor estimates and removes the background after it is stabilized. The fish detector detects and extracts candidate fish regions in the input image utilizing region, edge, and shape information. The fish tracker performs prediction and filtering of the fish motion dynamics using a biologically relevant Kalman filter algorithm. The track analyzer compares the outputs of the fish detector and the fish tracker, and takes one of the following actions: creates a new track segment, or updates an existing track, or terminates a track and

* Corresponding author: 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, JAPAN.
Fax: 81-3-5452-6189. Email: hanjun@iis.u-tokyo.ac.jp.

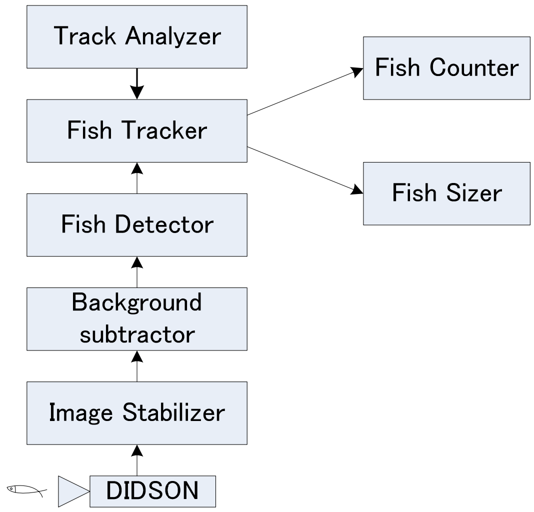


Figure 1 System block diagram for counting and sizing fish during its transfer using DIDSON.

predicts its state for next frame. For new track segments, initial motion states are initialized. The fish counter oversees the entire tracking history and establishes the complete fish trajectories and counts the disappearing fish. The fish sizer first calculates minimum bounding box, then obtains center line of candidate fish polygon by connecting the midpoint of equally spaced section lines along the thinner side of the bounding box. Extrapolating the center line, we search and find the broken remains. Finally the body length is obtained by summing down the center line segments from the fish head to the fish tail.

2.2 Image Stabilizer

In order to count and size fish, it is necessary to extract and remove the background. If the background is static, it can be performed by subtraction as usual. In case of fish transfer, however, as the background net is waved by current, the calculation of movement between the sequential frames is needed for image stabilization.

To calculate movement between two sequential images, the pattern matching method is widely used, in which the normalized correlation is calculated. Here we adopt a robust and high position accuracy POC (Phase-Only Correlation) method which is based on phase information only (Zitova´ & Flusser 2003).

Assume $f(n_1, n_2)$ and $g(n_1, n_2)$ are two images, and $F(k_1, k_2)$ and $G(k_1, k_2)$ are discrete Fourier transforms of $f(n_1, n_2)$ and $g(n_1, n_2)$ respectively, then POC $r_{fg}(n_1, n_2)$ can be defined as follow.

$$r_{fg}(n_1, n_2) = IDFT\left[\frac{F(k_1,k_2)G^*(k_1,k_2)}{\left|F(k_1,k_2)G^*(k_1,k_2)\right|}\right] \qquad (1)$$

where IDFT is an abbreviation of inverse discrete Fourier transform, the asterisk * means complex conjugation. The phase limitation correlation has two features. One is that the autocorrelation is a delta function, and the other is that the peak position of the correlation will shift same value as the displacement between two images. The amount of the gap between two images can be measured by searching the peak position of the POC. The calculation flow is shown Figure 2.

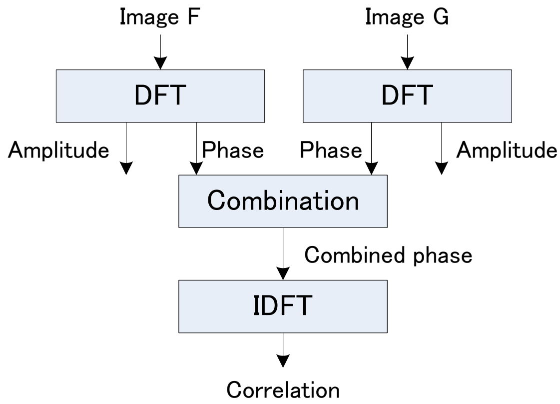


Figure 2 Phase only correction processing flowchart.

To decrease the influence of discontinuity of the image when discrete Fourier transform is performed, the input image $f(n_1, n_2)$ and $g(n_1, n_2)$ were multiplied a two dimension Hanning window function shown as fellow.

$$w(n_1, n_2) = \frac{1+\cos\frac{\pi n_1}{M_1}}{2}\,\frac{1+\cos\frac{\pi n_2}{M_2}}{2} \qquad (2)$$

2.3 Background Subtractor

Background subtraction is a widely used approach to detect moving objects in optical videos from a static camera (Zhang et al 2003 and Wren et al 1997). Several methods for performing background subtraction have been proposed. All of these methods try to effectively estimate the background model from the temporal sequences of the frames. Piccardi 2004 has conducted a good survey about background subtraction. The stabilized DIDSON background image remains almost still, hence the background subtraction can be used. On account of the future use in real time, a robust adaptive background subtraction method based on the running Gaussian average has been implemented. In the following, we describe the method briefly.

This method detects targets by computing the difference between the current frame and a background image for each pixel. A thresholding operation is performed to classify each pixel as background if

$$\left|I_t(x,y) - B_t(x,y)\right| < n \qquad (3)$$

where $I_t(x, y)$ is a vector representing the intensity value of the pixel at image position (x, y). $B_t(x, y)$ is the background (mean intensity) of the pixel, n is the noise threshold. The operation in (3) is performed for all image pixels (x, y).

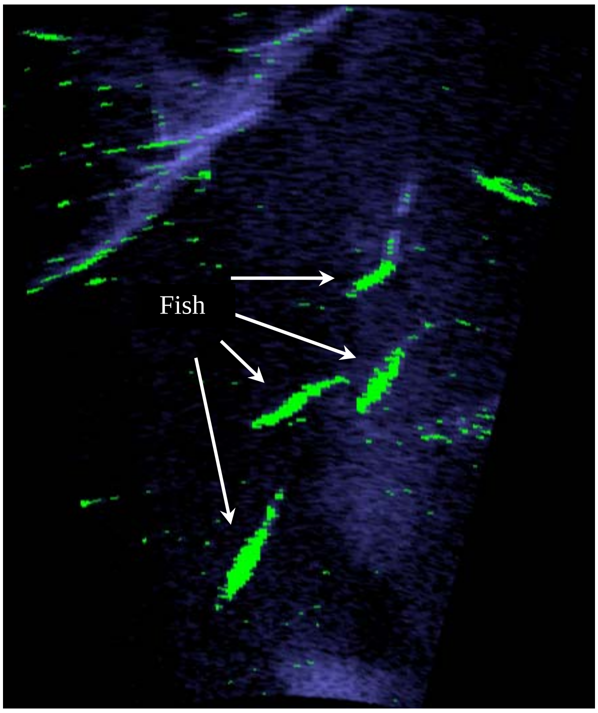


Figure 3 An example image shows the extracted foreground image in green overlapped on the original image.

To take into account slow condition changes which are necessary to ensure longtime tracking, the background image is subsequently updated by

$$B_{t+1}(x,y) = aI_t(x,y) + (1-a)B_t(x,y) \qquad (4)$$

where α is a parameter used to determine the weights in the background calculation.

2.4 Fish Detector

After the background subtraction, a black and white binary image can be obtained. Assume that only the white pixels are to be considered, and the others are treated as a background. Figure 3 shows an example image in which the foreground is painted with green and overlapped on the original image. To eliminate the isolated pixels, dilation and erosion (Ritter G.X. & Wilson J.N. 2001) with a 3x3 mask are performed. Tracing the edges using contour tracing method, we achieve the segmentation of objects from the background. However, in the objects there are some smaller regions having only several pixels, which are obviously different from fish. These objects are treated as noise and removed. After above processes, the remains are considered as targets. The centroid (x_c, y_c) of a contour is calculated by using moments.

$$(x_c, y_c) = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}}\right) \qquad (5)$$

where, the moments are defined as

$$m_{p,q} = \sum_x \sum_y x^p y^q I(x,y) \qquad (6)$$

The centroid is taken as the position of the target for later processing. The other information about the candidate fish such as the polygon area, the bounding rectangle and minimal bounding rectangle are also calculated. We only take the relatively large target whose area is between the given thresholds as candidate fish for tracking and the others as fish remains.

2.5 Fish Tracker

In order to prevent recounting the fish, tracking the moving fish is essential and very important in the image processing procedures. From this view point, we have tried to detect the candidate fish. Then, there arises a problem as to how to identify the same candidate fish in a consecutive image sequence. Image object matching techniques are commonly used to solve this kind of problem. However, in this case, the candidate fish images are not so stable because of their movements, and, therefore, the object matching method is not adequate. We perform prediction and filtering of the fish motion dynamics using a biologically relevant Kalman filter algorithm (Kalman 1960, Soreson 1970 & Brookner 1998).

Kalman filter is based on a motion model that describes the target dynamics, and a measurement model that relates states to measurements. The motion model and the measurement model are defined as

Motion model: $s_k = Fs_{k-1} + w_{k-1}$

Measurement model: $z_k = Hs_k + v_k$

Here s_k is the state vector of a fish in frame k, which consists of the centroid position, velocity of the fish. The corresponding measurement vector $z_k = (x_{k,} y_k)^T$ contains the measured centroid position. Note that the up-right character "*T*" denotes vector or matrix transposition. *F* is the state transition matrix of model, and *H* is the measurement matrix that relates states to measurements. w_{k-1} and v_k are the process and measurement noise vectors, which are uncorrelated zero-mean Gaussian processes with covariances *Q* and *R* respectively. The filtering calculation consists of two stages: prediction and correction. The prediction stage predicts the state $\hat{s}_{k|k-1}$ at time *k* based on the state history up to time *k-1*; the correction stage generates a refined estimate $\hat{s}_k$ by incorporating the newly obtained measurement z_k.

Prediction: To compute the state prediction $\hat{s}_{k|k-1}$ and covariance $P_{k|k-1}$ are calculated as follow.

$$\hat{s}_{k|k-1} = F\hat{s}_{k-1} \qquad (7)$$

$$P_{k|k-1} = FP_{k-1}F^T + Q \qquad (8)$$

Correction: Given the predicted states, covariance, and measurement, the Kalman filter is used to obtain the updated state $\hat{s}_k$ and covariance P_k.

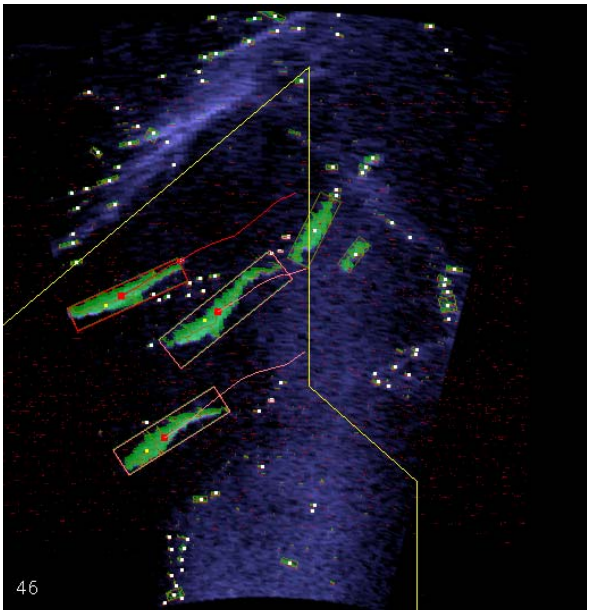


Figure 4 An example of tracking fish trajectories. The yellow lines show the region of tracking.

$$K_k = P_{k|k-1} H^T (H P_{k|k-1} H^T + R)^{-1} \quad (9)$$

$$\hat{s}_k = \hat{s}_{k|k-1} + K_k (z_k - H \hat{s}_{k|k-1}) \quad (10)$$

$$P_k = (I - K_k H) P_{k|k-1} \quad (11)$$

where K_k is the Kalman gain.

As the fish passing a transfer net gate has almost constant velocity, the state vector s_k, the system matrices *F* and *H* can be defined as follow.

$$s_k = (x_k, \dot{x}_k, y_k, \dot{y}_k)^T \quad (12)$$

$$F = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (14)$$

where T_s is the time between the measurement, in this case it is the frame interval. With the system matrices defined, the noise covariances *Q*, *R*, and the initial error covariance matrix P_0 can be estimated from training sequences using the expectation-maximization algorithm (Bishop 2008).

In order to do tracking undisturbedly, a ROT (region of tracking) can be defined by users with mouse in the tracking window on the computer screen.

2.6 Track Analyzer and Fish Counter

The track analyzer compares the outputs of the fish detector and the fish tracker, and takes one of the following actions: creates a new track segment, or updates an existing track, or terminates a track. For continuing track segments, the track analyzer updates the fish motion state, and predicts its state for next frame. For new track segments, initial motion states are initialized. The fish counter oversees the entire tracking history and establishes the complete fish trajectories and counts the disappearing fish as shown in Figure 4.

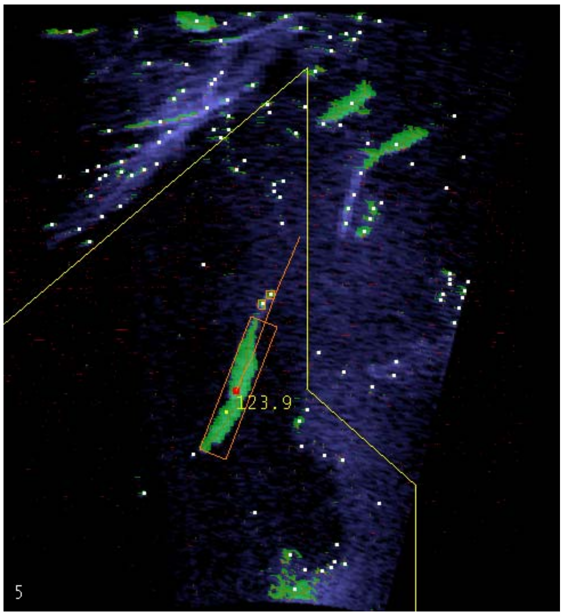
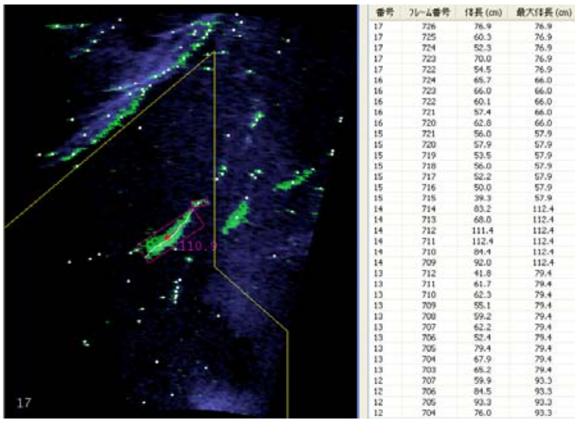


Figure 5 The broken fish body echo.



番号	フレーム番号	体長 (cm)	最大体長 (cm)
17	726	76.9	76.9
17	725	60.3	76.9
17	724	52.3	76.9
17	723	70.0	76.9
17	722	54.5	76.9
16	724	65.7	66.0
16	723	66.0	66.0
16	722	60.1	66.0
16	721	57.4	66.0
16	720	62.8	66.0
15	721	56.0	57.9
15	720	57.9	57.9
15	719	53.5	57.9
15	718	56.0	57.9
15	717	52.2	57.9
15	716	50.0	57.9
15	715	39.3	57.9
14	714	83.2	112.4
14	713	68.0	112.4
14	712	111.4	112.4
14	711	112.4	112.4
14	710	84.4	112.4
14	709	92.0	112.4
13	712	41.8	79.4
13	711	61.7	79.4
13	710	62.3	79.4
13	709	55.1	79.4
13	708	59.2	79.4
13	707	62.2	79.4
13	706	52.4	79.4
13	705	79.4	79.4
13	704	67.9	79.4
13	703	65.2	79.4
12	707	59.9	93.3
12	706	84.5	93.3
12	705	93.3	93.3
12	704	76.0	93.3

Figure 6 Showing tracking status, listing the frame number, the body length and the maximum length of the passed fish.

2.7 Fish Sizer

The fish sizer first calculates minimum bounding box, then extracts center line of candidate fish polygon by connecting the midpoint of equally spaced section lines along the thinner side of the minimum bounding box. The center line can be taken as the length of the fish if the tracked candidate fish is a whole fish. As shown in Figure 5, however, the fish echo often breaks up into several pieces. To get the whole fish length, we search and find the broken remains by extrapolating the center line. Finally the body length is obtained by summing down the center line segments from the fish head to the fish tail. Figure 6 shows the example tracking status, which lists

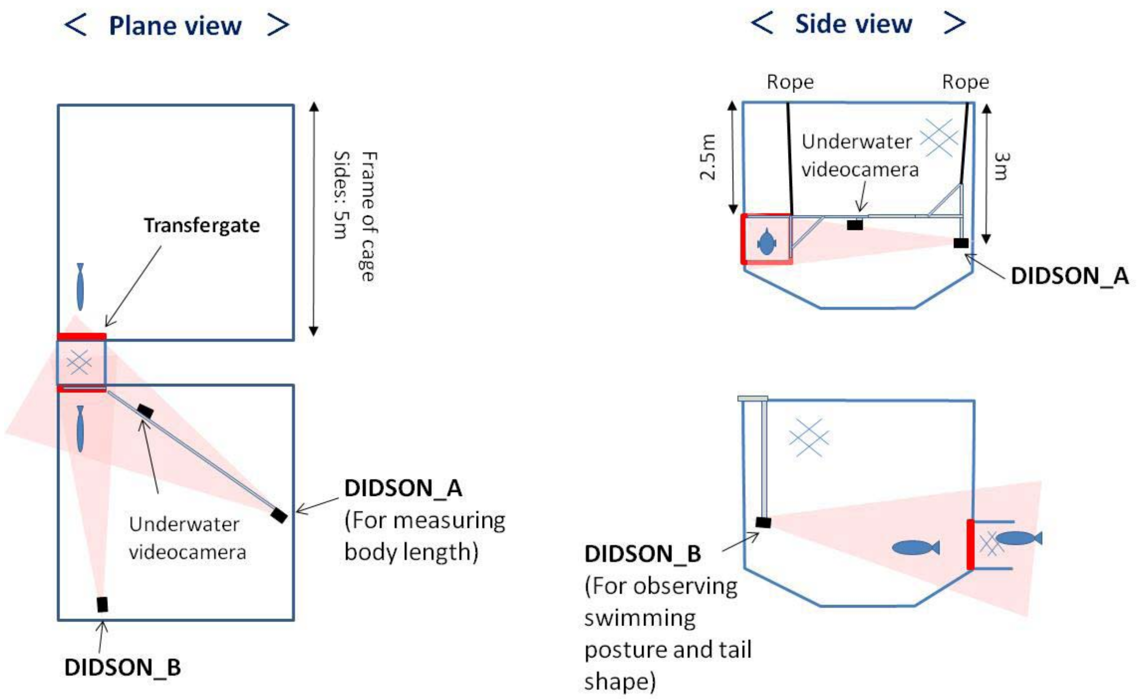


Figure 7 Depicting the layout of DIDSONs used in the experiment.

the frame number, the body length and the maximum length of the passed fish.

2.8 Verification Experiment

The verification experiment was conducted at the marine facility in Goto station, National Center for Stock Enhancement, Fisheries Research Agency. The two square net pontoons (5m x 5m) were used, which are connected by two square transfer gate (1 m x 1 m) at a depth of approximately 3 m. The layout of the pontoons is depicted in Figure 7.

Two DIDSON cameras were used, one (DIDSON A) for measuring fish body length and the other (DIDSON B) for observing fish swimming posture and tail shape. As shown in Figure 8, DIDSON A was installed to the pontoon by using an iron pole to prevent the transfer gate from deviating out of the observing range, and echo image from shaking due to relative movement between the gate and the acoustic camera by waves or swells.

Figure 8 Showing DIDSON A installed to the pontoon by using an iron pole.

Fixed at the same depth as that of the gate, DIDSON A was headed in horizontal direction to monitor the fish passing through the transfer gate from the side. The system was operated in high-frequency (1.8MHz) mode with a window length of 2.55 m, starting 3.40 m from the transducer and extending to 5.95 m. The range setting was to cover the transfer gate so as to ensure that fish were detected during they were passing through it. All the image data was recorded to a hard disk with a rate of 10 frames per second.

Farmed yellowtail, *Seriola quinqueradiata* with total length 75.0 to 90.0 cm and folk length 69.0 to 84.0 cm, whose body lengths were manually measured in advance, were repeatedly moved from one net cage to the other, and were observed by DIDSON when passing through the transfer gate net.

3 RESULTS AND DISCUSSIONS

The difference of total length measured by proposed method was -2.4 to 2.8 cm with mean errors -0.4 cm and standard deviation 2.3 cm, compared with the manually measured one.

The DIDSON image is based on echo strengths and slant distances from the camera head to targets, so installing the camera head properly and adjusting the sonar parameter are very important to get fine image data for counting and sizing fish. The sound velocity directly affects the size measuring, thus the fish sizes must be corrected in accordance with the measured sound velocity.

At this time, the image data taken by the horizontal DIDSON A is only used. Now the development of 3D measuring method for fish by combination of two DIDSON data is in progressing, the measuring accuracy can be expected to be improved.

4 CONCLUTION

We have developed an automated acoustic method for counting and sizing farmed fish during its transfer using DIDSON. The background is subtracted after being stabilized by a POC method. Tracing the edges using contour tracing method, the segmentation of fish is obtained. To prevent recounting the same fish, a biologically relevant Kalman filter algorithm was designed and adopted to predict the fish movement. By analyzing the track of fish, the automated counting was performed. The broken fish images were found, and the body length was obtained by summing down the center line segments from the fish head to the fish tail.

The proposed system was verified to achieve the sizing accuracy within the range of -2.4 to 2.8 cm with the mean -0.4 cm by using the farmed Yellowtail, *Seriola quinqueradiata* with the mean total length of 83.1 cm.

REFERENCES

Belcher E. O., Dinh H. Q., Lynn D. C., Laughlin T. J. (1999). 'Beamforming and imaging with acoustic lenses in small, high-frequency sonars'. OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century, pp1495-1499. Seattle, WA, USA.

Zitova´ B., Flusser J. (2003). 'Image registration methods: a survey'. Image Vision Computing. 21:977-1000.

Zhang C., Chen S. C, Shyu M. L., Peeta S. (2003). 'Adaptive background learning for vehicle detection and spatio-temporal track'. ICICS-PCM. pp15-18, Singapore.

Wren C., Azarbayejani A., Darrell T., Pentland A. (1997). 'Pfinder: Real-time tracking of the human body'. IEEE Trans. on Pattern Analysis and Machine Intelligence, **19**(7):780-785.

Piccardi M. (2004). 'Background subtraction techniques: a review'. IEEE Conference on Systems, pp.3099-3104.

Ritter G. X. & Wilson J. N. (ed.) (2001). Handbook of computer vision algorithms in image algebra. CRC, New York.

Kalman R. E. (1960). 'A new approach to linear filtering and prediction problems'. Trans. ASME J. Basic Eng. **82**(Series D):35-45.

Soreson H. W. (1970). 'Least-squares estimation from Gauss to Kalman'. IEEE Spectrum 7:63-68.

Brookner E. (1998). Tracking and Kalman filtering made easy. John Wiley & Sons, New York.

Bishop C. M. (2008). Pattern recognition and machine learning. Springer-Verlag, Tokyo (in Japanese)