Imaging Sonar Tracking of Salmon for Size and Tail Beat Frequency

Conference Paper · December 2014

DOI: 10.1109/GlobalSIP.2014.7032297

Matthew Kupilik
University of Alaska Anchorage

Imaging Sonar Tracking of Salmon for Size and Tail Beat Frequency

Matthew Kupilik and Todd Petersen
Department of Electrical Engineering
COE, University of Alaska Anchorage
Anchorage, AK, USA
mkupilik@uaa.alaska.edu, thpetersen@uaa.alaska.edu

***Abstract*—In order to set sustainable harvest limits on anadromous fish, fishery managers require a method for estimating return numbers of individual species. As seasonal migration times of several returning species may overlap, simple fish counts are insufficient. This paper presents an algorithm for tracking fish crossing an insonified area using an extended Gaussian mixture probability hypothesis density filter as well as analyzing both their lengths and the frequency of their tail beats. The combination of the two provides a useful estimate of the return size of a given fish species when combined with seasonal return timing. The algorithm is tested against data collected using a dual frequency imaging sonar on the Kenai River in Alaska and compares favorably to estimates made by manual inspection and sub-sampling.**

***Index Terms*—Sonar Applications, Machine Vision, Acoustic Signal Processing, Tracking Loops**

I. Introduction

Annual migrations of anadromous salmon species are an important resource for many northern hemisphere countries. In addition to being a valuable food source, such migrations act as an ecological biomass delivery system for cycling nutrients from the ocean to terrestrial environments. They are thus an important resource to harvest sustainably. It is the task of fishery managers to determine what a sustainable harvest limit is for each returning salmon species. In order to set commercial and recreational limits they require a fast measurement of the population return size of each species. In addition, multiple species of salmon often migrate simultaneously and share physical characteristics (profile and size). This project discusses the development and application of an acoustic tracker to not only count the number of fish crossing an insonified area but differentiate between two salmon species based on fish length and tail beat frequency using acoustic data from the Kenai River in Alaska.

Previous work using imaging sonar to track and size fish has primarily focused on employing Kalman based trackers [1, 2] with data association to solve the problem of assigning measurements to tracks. The use of tail beat frequency has also been implemented in post processing applications [3]. The effectiveness of these is difficult to compare as they all employ sensors with different resolutions and noise characteristics. This work uses an extended Gaussian mixture probability hypothesis density (GM-PHD) filter [4, 5] to track the fish and image processing techniques to determine the size and tail beat frequency. The GM-PHD filter is chosen as it is better able to account for several aspects of the tracking problem such as false measurements and target birth distributions, which must be handled in an ad-hoc manner using conventional Bayesian trackers [6]. The GM-PHD filter has also been demonstrated to perform well on acoustic data in other applications [7, 8].

The algorithm is functionally separated into three pieces: pre-processing, tracking, and track processing to determine species information. The pre-processing stage covers the elimination of acoustic shadows and determines a set of possible fish centroids in each image frame. The tracker solves the multiple target tracking problem using the GM-PHD filter applied to a set of centroid and orientation measurements. The track processing then takes the resulting sequence of stored intensity maps for each tracked fish and calculates the fish length as well as the tail beat frequency. These two measures were chosen as they can be used to apportion the two species of salmon present in the tracked data, Chinook Salmon (Oncorhynchus tshawytscha) and Sockeye Salmon (Oncorhynchus nerka) [9, 10].

The sensor used to test the algorithm is a Sound Metrics ARIS (adaptive resolution imaging sonar). In the tested configuration, this sonar produces a two dimensional intensity map from acoustic signal strengths generated using 48 beams with 1416 samples per beam over a range of 5.0 meters. The resulting intensity maps are produced with a frame rate of 7-10 Hz. At this resolution each 10 minute segment of data is approximately 360 megabytes in size. Ideally these sensors would be operated continuously over the salmon migration period which can be months in duration. Difficulty in storing and post processing such a volume of information results in current techniques to consist of sub-sampling and manually measuring the fish that are observed [9]. This type of assisted automatic identification has been shown to provide statistically relevant fish count information [11, 12], but remains time consuming and expensive to carry out. An automatic fish tracker that could be run over the entire migration period could provide much faster feedback on estimated species return size, thus increasing the ability of fishery managers to set accurate catch limits.

978-1-4799-7088-9/14/$31.00

II. Pre-Processing

The inputs to the pre-processing segment consist of the raw ARIS images at a frame rate of 7 Hz. The output of the pre-processing segment is a set of measurements consisting of probable fish centroids, which are then used as the input to the GM-PHD tracker. In order to extract the moving fish and other targets, background subtraction is used to filter any non-changing pixel intensities which are assumed to be rocks or other riverbed features. The algorithm used is the OpenCV implementation of the Gaussian mixture based background segmentation algorithm [13]. The background subtractor is adaptively run over the previous 50 frames and updated every frame.

Acoustic shadows represent a significant problem to the background subtraction. For many of the data segments the sonar insonifies an area starting several meters out from shore. Thus if fish or debris float in between the sonar and the start of the insonified range, they cast an acoustic shadow. In order to compensate for this intensity change, the histogram of intensities for each beam are individually normalized to a uniform brightness. This normalization helps with acoustic shadows that are cast by objects outside the insonified area but does reduce the intensity variation of any fish that are visible in the frame and casting a shadow.

Once the background is removed, the remaining pixels are analyzed for possible fish using connected components and morphological operations. In order to reduce noise the image is opened using a 3×10 pixel structuring element. Any objects with a major axis length of less than 0.2 meters are also eliminated. The centroids of the remaining objects are passed to the tracker as possible fish. Since orientation of the potential fish is available, it is also passed to the tracker and used as a measurement.

III. Tracking

The GM-PHD filter requires several assumptions in order to give a closed form solution to the multiple target tracking problem. Each target must move and generate measurements independent of other targets. The clutter, or measurements generated by non-targets, must be independent of the target measurements. It is also assumed that each measurement is associated with only one target (or targets produce only a single measurement). In this particular application there is no target spawning and birth functions are assumed Poisson.

For this application the state for each target consists of the fish position and velocity in the x and y directions and body orientation θ, $x_k = \begin{bmatrix} x_{pos} & y_{pos} & \dot{x_{pos}} & \dot{y_{pos}} & \theta \end{bmatrix}^T$. The dynamic state transition model is f, and the sensor has a measurement model, g.

$$f_{k|k-1}(x|\xi) = \mathcal{N}(x; F_{k-1}\xi, Q_{k-1}) \tag{1}$$

$$g_k(z|x) = \mathcal{N}(z; H_k x, R_k). \tag{2}$$

Where $\mathcal{N}(\cdot; m, P)$ represents a normal distribution with mean m and covariance P, ξ is a previous state, F is the state transition matrix, Q is the process covariance, H is the observation matrix, and R is the observation noise covariance. Both Q and R were set empirically. Since the state transition model f is nonlinear, the state transition matrix F must be found by calculating the Jacobian:

$$\frac{\partial f}{\partial x_k} = \begin{bmatrix} 1 & 0 & \cos(\theta)dt & 0 & -\dot{x_{pos}}\sin(\theta)dt \\ 0 & 1 & 0 & \sin(\theta)dt & \dot{y_{pos}}\cos(\theta)dt \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where dt is the time between frames, or $\frac{1}{7}$ seconds. The tracker thus weights fish that travel in the direction they are oriented with a constant velocity and orientation. The process covariance and observation noise are:

$$Q = \begin{bmatrix} .0004 & 0 & 0 & 0 & 0 \\ 0 & .0004 & 0 & 0 & 0 \\ 0 & 0 & .0004 & 0 & 0 \\ 0 & 0 & 0 & .0004 & 0 \\ 0 & 0 & 0 & 0 & .0004 \end{bmatrix} \tag{4}$$

$$R = \begin{bmatrix} .003025 & 0 & 0 & 0 & 0 \\ 0 & .003025 & 0 & 0 & 0 \\ 0 & 0 & 2.42 & 0 & 0 \\ 0 & 0 & 0 & 2.42 & 0 \\ 0 & 0 & 0 & 0 & .1 \end{bmatrix}. \tag{5}$$

The observation matrix H is chosen such that position in both x and y, as well as orientation are measured directly. At each time step k the PHD (first moment approximation) is a weighted sum of normal distributions

$$v_{k|k}(x) = \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} \mathcal{N}(x; m_{k|k}^{(j)}, P_k^{(j)}) \tag{6}$$

where $m_{k|k}^{(j)}$ is mean of component j, and $P_k^{(j)}$ is the covariance of component j.

The first step for the filter is a prediction step where the probability of target survival is combined with the state transition matrix and the process noise covariance. Fish survival probability is experimentally set at 0.9. The survival probability represents the chance that a fish will not leave the insonified area. Once existing targets have been predicted possible new targets are assigned weights according to a birth function. The birth function has a large impact on how measurements are assigned, either as new targets, or as continuing measurements from existing targets. For our application it is expected that fish will appear on the edges of the insonified region with very high probability. In addition the flow direction of the water is known. Since these salmon are spawning, the targets are most likely swimming in the upstream direction. For these reasons the birth function was chosen as one of the following Gaussians:

$$\mathcal{N}(\begin{bmatrix} 1.4623 & 15.8115 \end{bmatrix}, \begin{bmatrix} 5 & 30 \\ 30 & 200 \end{bmatrix}) \tag{7}$$

$$\mathcal{N}(\begin{bmatrix} -1.4623 & 15.8115 \end{bmatrix}, \begin{bmatrix} 5 & -30 \\ -30 & 200 \end{bmatrix}) \tag{8}$$

depending on if the sonar is located on the left or right side of the river. The choice of this birth function heavily weights the tracker to extract targets that appear at the image edge; fish that enter the insonified area from any of the other edges are often assumed to be clutter.

IV. Track Processing

The output of the tracker is a set of fish tracks consisting of fish location, velocity, and orientation at each frame it was visible. Using this estimated location a smaller image sequence of the fish at each frame is cropped and saved until the respective track is processed. These tracks are the input to the track processing segment, the output of which are the two measures chosen as useful for differentiating the salmon species (length and tail beat frequency).

A. Fish Length

Calculating the actual length of the fish is a two step process. First the fish must be segmented from the background, then it must be determined if the segmented blob shows the full length of the fish. The tracker provides an estimate of the centroid of the fish and the background subtraction algorithm provides an estimate of the fish contour. The motion mask produced from the background subtraction algorithm does not provide an accurate enough shape to extract the actual fish length. This is due to the fact that the tail is often detached from the fish body as its intensity is often much lower. However, using the motion mask area to dynamically set a threshold results in an acceptable segmentation. For situations where the motion mask is lost to noise but the tracker still estimates a fish is present, a threshold is calculated using the average intensity of a 5×5 pixel box around the estimated fish centroid.

Once the fish is segmented the blob's skeleton is found and the location of all skeleton end points are calculated. To calculate the length of the blob the maximum distance between any two skeleton end points is found, as shown in Fig 1.

Fig. 1. Segmented fish (white) and measured length segment (black) on raw data background.

This process is repeated for every frame, building a vector of lengths for the fish blob. The frame length vector does not provide an easily calculated fish total length. The pixel length of any individual frame varies with the fish position relative to the sonar as well as with distance from the center and background noise. In manual implementations the user selects a frame where the tail is easily recognized and the full fish profile is present. Although the algorithm could be extended to include some shape recognition, the median of the length vector is taken as the estimate. The fish are then binned into two categories, greater than 0.8m and less than 0.8m, often the dividing line between Chinook and Sockeye salmon [10].

B. Tail Beat Frequency

The tail beat frequencies for upstream migrating Chinook salmon range between 1.0 and 2.0 beats per second. The tail beat frequencies for upstream migrating Sockeye salmon range between 2.0 and 3.5 beats per second. The range for both species is also independent of the individual fish length [10]. Previous methods of calculating the tail beat frequency involve calculation of the intensity image sequence for each tracked fish overlaid in time, known as an echogram of the fish track. The maximum value for each beam is used for each pixel. The result is shown in Fig 2. If the fish shows any vertical

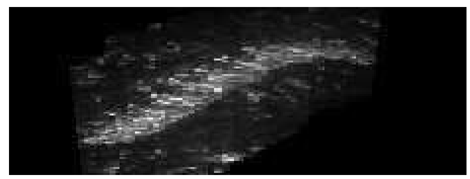


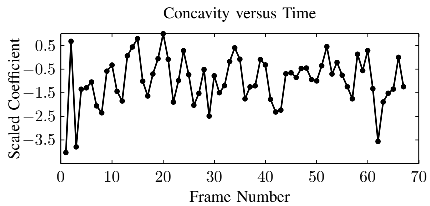
Fig. 2. Fish echogram with countable protrusions indicating tail beats.

movement, the tail extensions become quite visible as extrusions from the echogram, and can be automatically counted. The path of the fish is segmented from the background and the image skeleton found. The segment with maximum length is taken as the main path of the fish body, then the number of branches off the main path on both sides are counted. The side with the most extrusions is then divided by the length of time the fish is tracked.

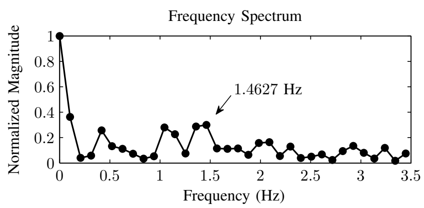
The echogram method is highly dependent on accurately segmenting the fish path from the background. To increase the robustness of the measurement while maintaining an entirely automatic processing, changes in body concavity are also found. Body concavity is much easier to calculate in the presence of noise, and can be found at each frame, whereas the echogram requires the entire track to be post processed. Changes in body concavity are observed to change at the same frequency as tail beats calculated using the echograms for selected fish, which is also verified using acceleration measurements taken from tagged fish in [14].

To measure the concavity of a target fish, a second order polynomial is fit to the same image skeleton segment that was used to determine length. The sign of the leading coefficient then indicates if the body is concave upwards or downwards. The leading coefficient value is tracked for every frame and changes indicate changes in body concavity. To estimate the tail beat frequency from changes in concavity, a Fast Fourier Transform (FFT) is applied to the value of the leading coefficient with a Tukey window applied ($\alpha = 0.5$). The maximum peak in the frequency plot between 1 Hz and 3.5 Hz is identified and returned as the estimated tail beat frequency.

Example data from a fish track is plotted showing the change in concavity as measured by the leading coefficient versus frame number (Fig 3(a)) and the processed FFT of that data with dominant peak highlighted (Fig 3(b)).



(a) Scaled leading coefficient as measure of concavity



(b) Frequency of leading coefficient as measure of tail beat frequency

Fig. 3. Tail beat frequency found from body concavity frequency.

V. Results

In order to test the tracker, the algorithm is run on sonar data collected from the Kenai River in Alaska using a Sound Metrics ARIS sonar device. The results of the tracker are then compared to the manual fish counts obtained by the Alaska Department of Fish and Game [15]. The data used for testing consists of 5 different 10 minute segments taken from the 2013 salmon migration up the Kenai River in Alaska. The data is broken into 10 minute segments chosen to reflect a variety of fish densities. These segments were also chosen as manual fish count data provided by the Alaska Department of Fish and Game was available to verify algorithm performance.

Over the fifty minutes of tracked data the manual count results in 76 fish crossing the insonified area. The extended GM-PHD tracker estimated 64 fish passed in the same time frame. The histogram of fish sizes for the manual and automatic tracker are shown in Fig 4. The automatic tracker has a tendency to undersize fish, as often the tail is not visible. Errors for the tracking segment include fish that were not tracked for a sufficient length of time to be considered a viable target (10 frames or 1.0 meters). This constraint was imposed to prevent double counting of fish that enter the frame and leave the frame (thus being tracked multiple times) as well as fish that enter the frame and remain stationary.

Total fish plotted as points in length and tail beat frequency space is shown in Fig 5. Not all fish that are successfully tracked produce a useable tail beat frequency measurement. This discrepancy is due to fish that are tracked for an insufficient time to produce a full tail beat period. For the sockeye salmon the frame rate of 7 Hz is also very close to the Nyquist rate, so any missed measurements introduce error. For

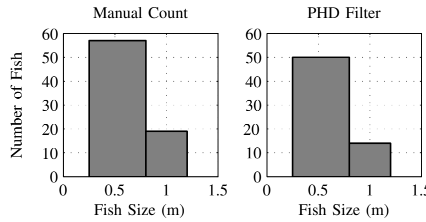


Fig. 4. Manually counted fish and automatically tracked fish, total count broken into possible Chinook (length > 80 cm) and smaller.

those that a measurement can be produced, there is a strong correlation between the larger fish (assumed to be Chinook) and a slower tail beat frequency (1 to 2 Hz). The larger fish are also most likely to produce a useable echogram as they are resolved much better and generally spend more time being tracked. The normalized correlation coefficient between the size and tail beat frequency is found as -0.4121. The tail beat frequency measurement would be a useful addition to a species classifier or an additional distribution to include in a mixture model approach in determining species.

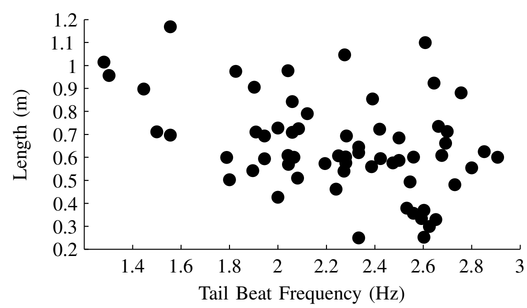


Fig. 5. Calculated tail beat frequency versus algorithm length estimate.

VI. Conclusion

The automatic tracker is capable of estimating numbers of fish crossing an insonified area and estimating their length and tail beat frequency in near real time. Although the tracker underestimated the number of fish, it is capable of running continuously throughout the migration season. Thus allowing fishery managers to more intelligently choose sub samples of data to manually verify. The measurement of size and tail beat frequency also gives fishery managers a near real time estimate of the Chinook salmon return size, which would aid in setting catch sizes.

Acknowledgement

The authors would like to thank Debby Burwen and the Alaska Department of Fish and Game for generously providing the sonar data. We would also like to thank Bill Hanot at Sound Metrics for helping with the Matlab code to interface with the ARIS data.

REFERENCES

[1] Nils Olav Handegard and Kresimir Williams. Automated tracking of fish in trawls using the DIDSON (Dual frequency IDentification SONar). *ICES Journal of Marine Science*, pages 636–644, 2008.

[2] Jun Han, Naoto Honda, Akira Asada, and Koji Shibata. Automated acoustic method for counting and sizing farmed fish during transfer using didson. *Fisheries Science*, 75(6):1359–1367, 2009.

[3] Myoung-Hee Kang. Semiautomated analysis of data from an imaging sonar for fish counting, sizing, and tracking in a post-processing application. *Fisheries and Aquatic Sciences*, 14(3):218–225, September 2011.

[4] B.-N. Vo and W.-K. Ma. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, November 2006.

[5] Kusha Panta, DE Clark, and BN Vo. Data association and track management for the Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Aerospace and Electronic Systems*, 45(3):1003–1016, 2009.

[6] Ronald P. S. Mahler. *Statistical Multisource-Multitarget Information Fusion*, pages 1–888. Artech House, Inc., Norwood, MA, USA, 2007.

[7] Ramona Georgescu and Peter Willett. The GM-CPHD tracker applied to real and realistic multistatic sonar data sets. *Oceanic Engineering, IEEE Journal of*, 37(2):220–235, 2012.

[8] Daniel E. Clark and Judith Bell. Bayesian multiple target tracking in forward scan sonar images using the PHD filter. *IEE Proceedings-Radar, Sonar and Navigation*, pages 327–334, 2005.

[9] Debby L. Burwen, Steven J. Fleischman, and James D. Miller. Evaluation of a dual-frequency imaging sonar for detecting and estimating the size of migrating salmon. Technical Report Fishery Data Series No. 07-44, Alaska Department of Fish and Game, 2007.

[10] Anna-Maria Mueller, Deborah L. Burwen, Kevin M. Boswell, and Tim Mulligan. Tail-Beat Patterns in Dual-Frequency Identification Sonar Echograms and their Potential Use for Species Identification and Bioenergetics Studies. *Transactions of the American Fisheries Society*, 139(3):900–910, May 2010.

[11] Debby L. Burwen, Steven J. Fleischman, and James D. Miller. Accuracy and Precision of Salmon Length Estimates Taken from DIDSON Sonar Images. *Transactions of the American Fisheries Society*, 139(5):1306–1314, September 2010.

[12] Ian C. Petreman, Nicholas E. Jones, and Scott W. Milne. Observer bias and subsampling efficiencies for estimating the number of migrating fish in rivers using Dual-frequency IDentification SONar (DIDSON). *Fisheries Research*, 155:160–167, July 2014.

[13] Zoran Zivkovic and Ferdinand van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780, May 2006.

[14] H Tanaka, Y Takagi, and Y Naito. Swimming speeds and buoyancy compensation of migrating adult chum salmon Oncorhynchus keta revealed by speed/depth/acceleration data logger. *Journal of Experimental Biology*, 204(Pt 22):3895–904, November 2001.

[15] James D. Miller, Debby L. Burwen, and Steven J. Fleischman. Estimates of Chinook Salmon Passage in the Kenai River Using Split-Beam and Dual-Frequency Identification Sonars, 2010. Technical Report Fishery Data Series No. 13-58, Alaska Department of Fish and Game, 2013.