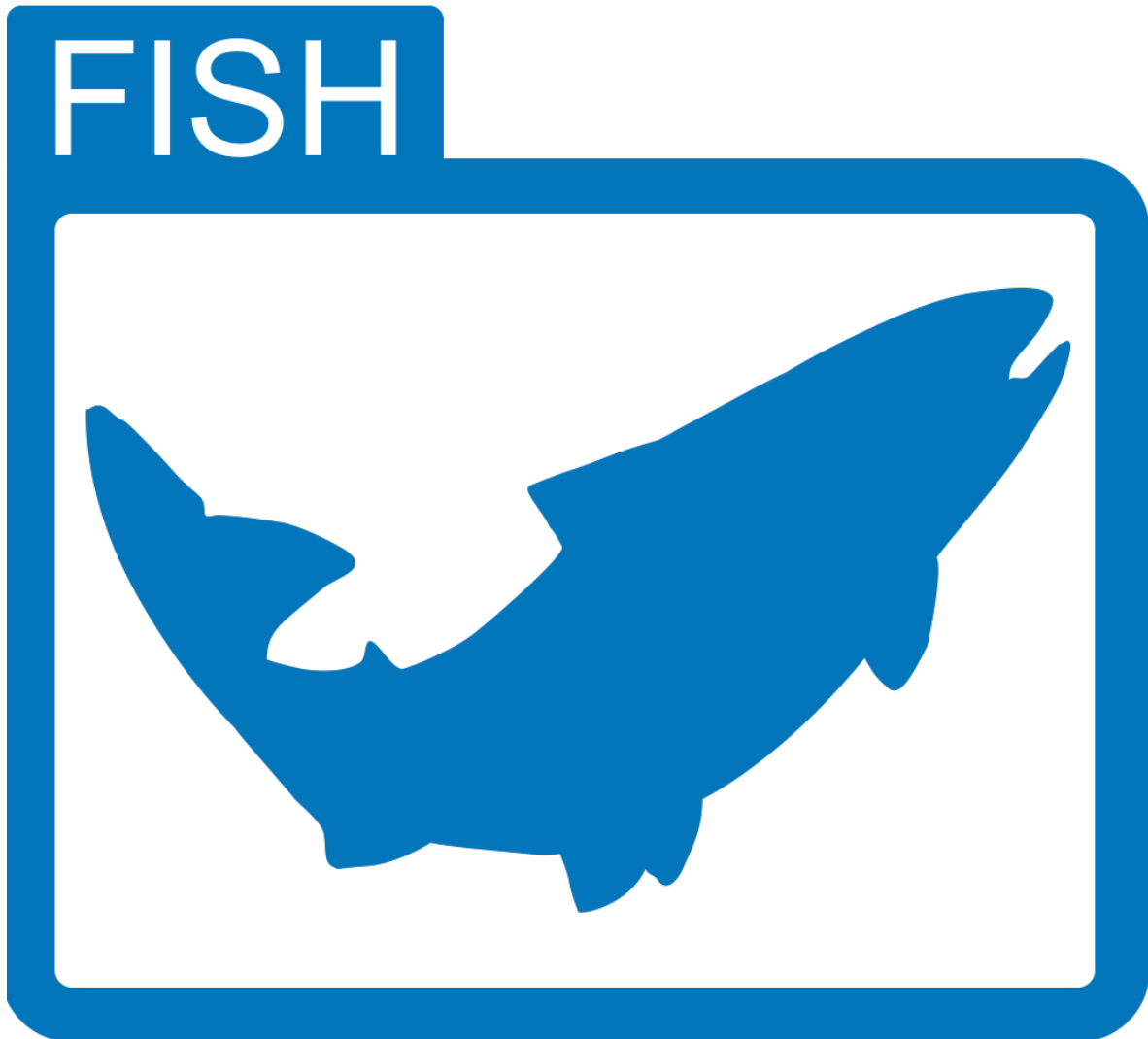

Sonar Annotator

User Manual

Carl T. Pfisterer
Alaska Department of Fish & Game



Introduction

Sonar Annotator is a simple, single purpose application developed to annotate imaging sonar files. Fish Annotator currently only supports DIDSON and ARIS files. The application outputs the frames from the video that have been annotated along with a JSON file with the bounding boxes and annotations themselves and a file containing metadata describing the file. The annotations and frames can be used for training machine learning models used to identify or categorize objects. Sonar Annotator was developed as open source in C/C++ using the Qt frameworks and can be compiled to run on Mac OS, Windows, or Linux.

Window Panes and Tabs

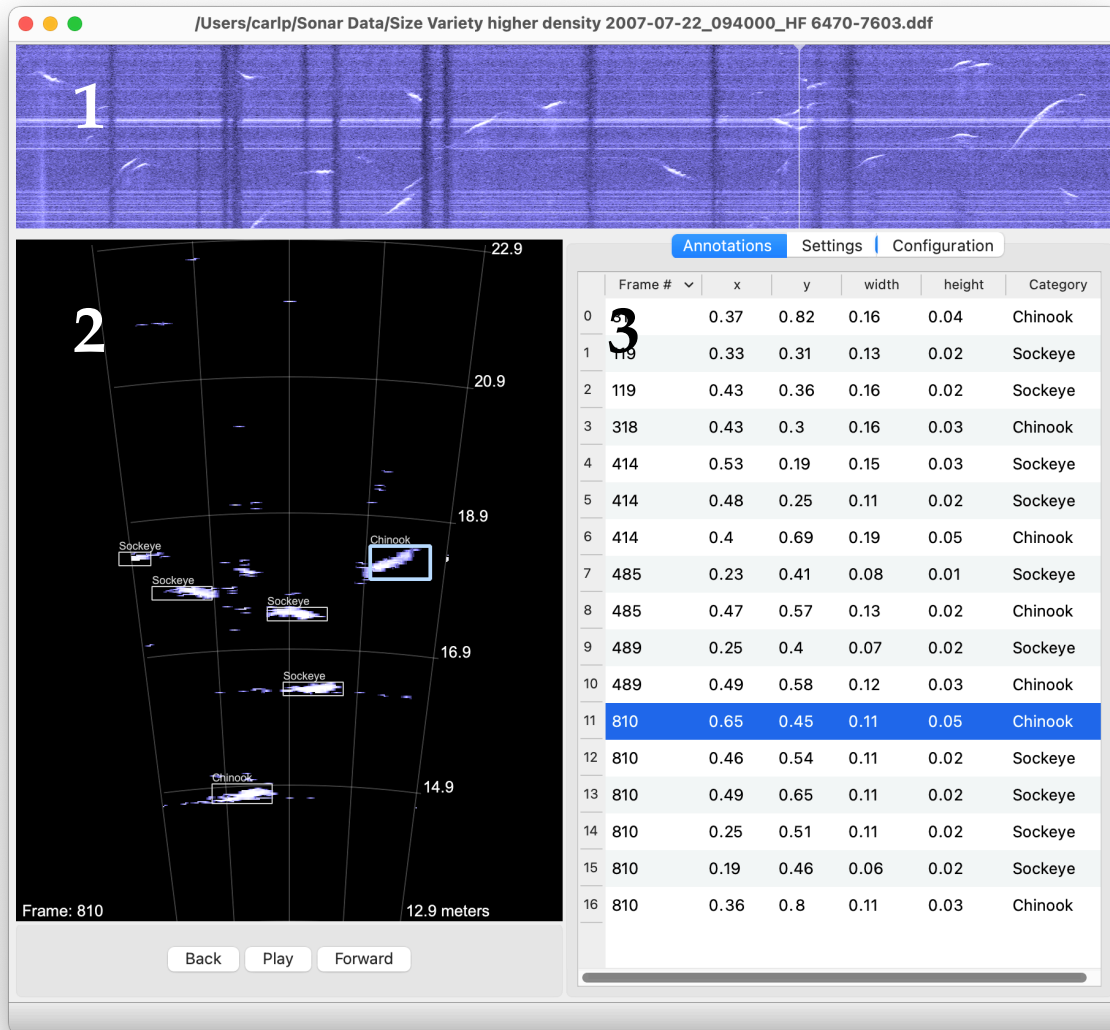


Figure 1—Window showing three panes: 1) echogram with file position indicator, 2) video with annotations, and 3) annotations, settings, and configuration tabs.

Echogram

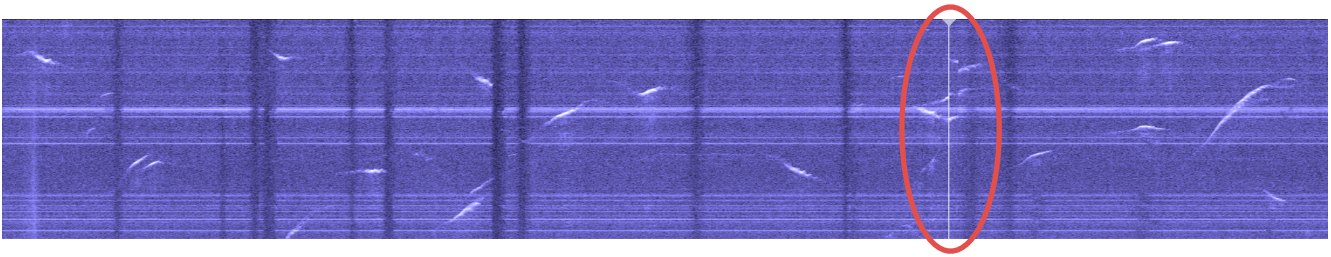


Figure 2—Echogram showing file position indicator.

The echogram pane at the top of the window shows the entire file compressed into a single timeline. Clicking anywhere in the echogram will jump to that location in the file. The file location itself is shown by the file position indicator line. There is a small triangle at the top of the file position indicator that allows the line to be moved by clicking in the triangle and dragging.

Video

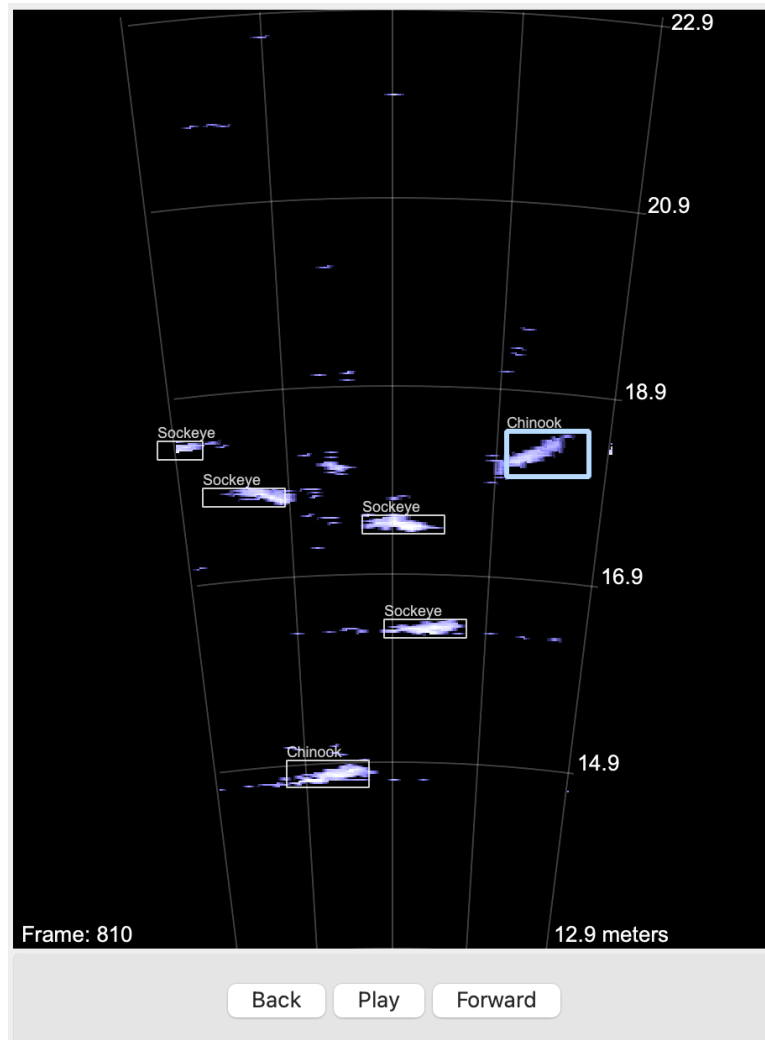


Figure 3—Video pane showing the frame number and annotations. Currently selected annotation is highlighted using a thicker border.

The video pane shows the video and images can be annotated by left-clicking in the video and drawing a rectangle. To enter the annotation, hit the number of the category (1–5) on the keyboard, or hit the enter/return key to enter as category 1. The play button will play the video, the back and forward buttons will step backward or forward one frame respectively.

Annotation Tab

Annotations

Settings

Configuration


	Frame # 	x	y	width	height	Category
0	31	0.37	0.82	0.16	0.04	Chinook
1	119	0.33	0.31	0.13	0.02	Sockeye
2	119	0.43	0.36	0.16	0.02	Sockeye
3	318	0.43	0.3	0.16	0.03	Chinook
4	414	0.53	0.19	0.15	0.03	Sockeye
5	414	0.48	0.25	0.11	0.02	Sockeye
6	414	0.4	0.69	0.19	0.05	Chinook
7	485	0.23	0.41	0.08	0.01	Sockeye
8	485	0.47	0.57	0.13	0.02	Chinook
9	489	0.25	0.4	0.07	0.02	Sockeye
10	489	0.49	0.58	0.12	0.03	Chinook
11	810	0.65	0.45	0.11	0.05	Chinook
12	810	0.46	0.54	0.11	0.02	Sockeye
13	810	0.49	0.65	0.11	0.02	Sockeye
14	810	0.25	0.51	0.11	0.02	Sockeye
15	810	0.19	0.46	0.06	0.02	Sockeye
16	810	0.36	0.8	0.11	0.03	Chinook

Figure 4—Annotations tab showing currently selected annotation.

The annotations tab shows a table for each annotation. Clicking a row will select that annotation, jump to the frame the annotation is from, and highlight that annotation. A selected annotation can be removed by hitting the delete key. The table can be sorted by clicking on the header of any of the columns. Note that the x, y, width, and height values are proportional to the width and height of the video image itself and are referenced with (0,0) starting in the upper left. The values in this table are the same as exported in the CSV file, but different from the JSON file. The JSON file has actual pixel locations with (0,0) referenced from the lower left left of the image.

Settings Tab

The Settings Tab interface includes the following controls:

- Annotations** | **Settings** | Configuration
- Threshold: 55 (with up/down arrows)
- Intensity: 100 (with up/down arrows)
- Color Scale: Blue (with a dropdown arrow)
- Background**
 - ☒ Remove (with a value of 10.0 and up/down arrows)
 - ☐ Subtract
 - ☐ No Removal
- Across Frame Filters**
 - ☒ Average
 - ☐ Median
 - ☐ None
- Time varied gain**
 - Spreading: 40.00 (with up/down arrows)
 - Alpha: 0.10 (with up/down arrows)
- Annotation Categories**

Chinook	Category 1
Sockeye	Category 2
Chum	Category 3
Fish	Category 4
Fish	Category 5

Figure 5—Settings tab

The settings tab contains fields and controls that affect the image and is where the annotation categories are defined. The current settings are saved upon quitting the application and loaded when it is started. The settings and how they affect the image are as follows:

Threshold and Intensity:

These are similar to the threshold and intensity in the SMC DIDSON software. The threshold defines the lowest value visible in the video, anything less than the threshold is set to black. The intensity controls the maximum value, anything higher is set to the lightest values of the color scale. All the colors are scaled between these values. Note that these values are not exactly equivalent to the SMC values in the file so they will be different than the values used in the manufacture's software.

Color Menu:

Drop down menu to select the color scale for the image. In most cases, this is probably best left as Black & White for the purposes of machine learning models. There is no actual

color values associated with the raw data and it likely wouldn't be desirable if machine learning training algorithms keyed on color as opposed to amplitude, size, and shape.

Background:

The values in the Background box define how the background is removed. No Removal doesn't attempt to remove static objects at all. Remove compares each pixel value to the same location in a reference background image and if it is within plus or minus the value of the Remove Threshold value, it sets the displayed value to zero (or black). The amplitude of the remaining pixels is not affected at all. Subtract, as the name suggests, subtracts the value of the pixels in the reference image from the values in the current image. This does not completely remove values and the amplitude of the remaining pixels is affected as well which might require adjusting the threshold and intensity.

Across Frame Filters:

Selects whether the pixels of the current frame should be the average of the current, previous, and next frame, whether it should be the medial value, or whether it should only be the current frame. Using the average or median removes quite a bit of noise.

Time Varied Gain:

Values to use for correcting spherical spreading and attenuation losses. Typical value for Spreading is 40 or $40 \log_{10} R$ where R is range. There may be some question about whether a $40 \log_{10} R$ correction is appropriate in a narrow river but that is a good starting point. Alpha is the attenuation coefficient, this can be adjusted to normalize signal strength across the range.

Annotation Categories:

These are user defined categories. To annotate with a particular value, press the number on the keyboard corresponding to the Category number.

Configuration Tab

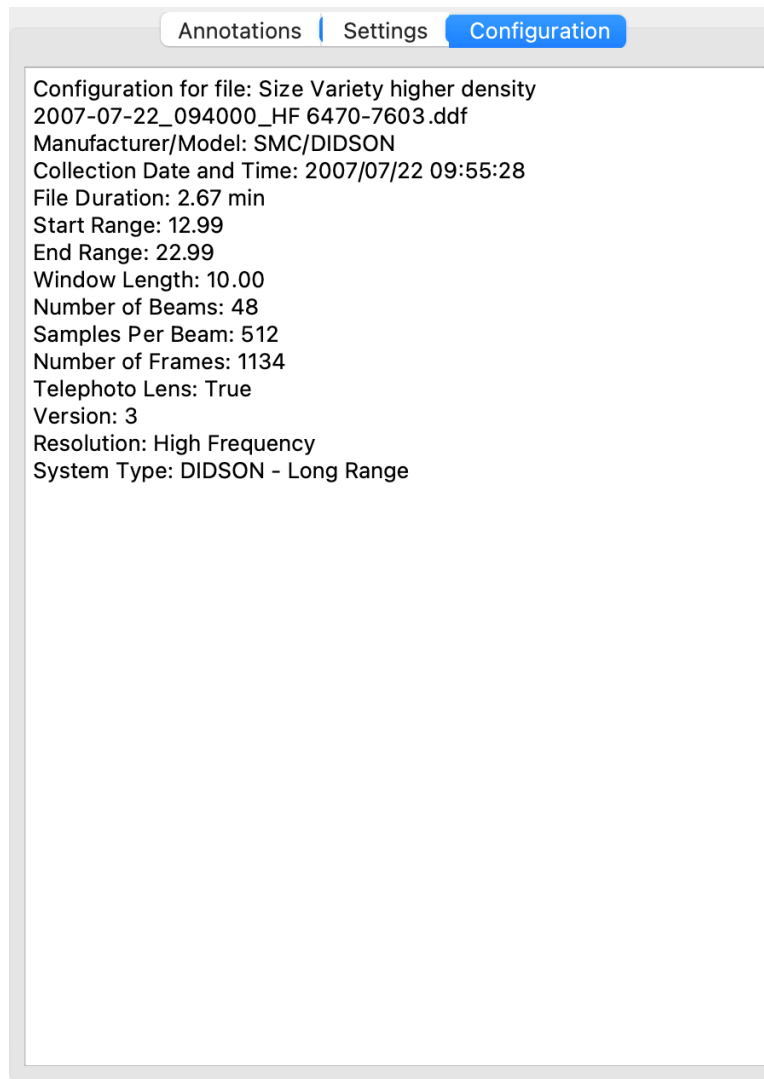


Figure 6—Configuration tab

The configuration tab shows the metadata for the file. When a file is opened for the first time, or when a previous .cfg file does not exist, it defaults to settings from the file header. This is just a regular text field so all values can be edited and additional information can be entered. For example, it may be helpful to include information on where the data was collected and species present. When the user selects File->Export Annotations, this is saved to a text file in the same directory as the original file but with a .cfg extension appended to the name. When a file is opened, the application looks for a .cfg file in the same directory with the same name and if it exists, the contents are read.

Annotating Files

Before beginning a project, it is recommended that a separate folder be created for each file that will be annotated. This is not a requirement, but it will likely make organization easier. To begin, open a file to be annotated. Adjust the settings as desired, edit the categories, and edit the text in the Configuration tab if necessary. To annotate, simply draw rectangles around objects by left-clicking and dragging. Press the number on the keyboard corresponding the appropriate category and the annotation will be entered into the table in the Annotations tab.

Annotations can be reviewed and edited by selecting the annotation in the annotations table. The video will jump to the appropriate frame and the annotation will be highlighted in a different color and with a thicker line. The annotation can be removed by hitting the delete key.

Once the desired annotations have been completed, select File->Export Annotations. The application will create a separate JPEG image for each frame with an annotation. The name of the image will be same as the original file but with an underscore, followed by the frame number and the .jpg extension. A .csv file is created so that the annotations can be loaded by the application if the original file is reloaded. Applications for training machine learning models often take a JSON file which is also exported. Finally, the metadata associated with the file is exported as a flat ASCII file with a .cfg extension. When a sonar file is loaded into Sonar Annotator, the application looks for corresponding files with a .cfg and / or .csv extension. If it finds these files, they are automatically loaded.

Notes