AIS Linker User Guide

Introduction

This guide provides instructions to use all of the capabilities of the AIS Linker tool, and is intended for all users. Individuals who intend to modify the code of this tool may also wish to review the companion Technical Guide, but it is still recommended to read this guide first.

License

The AIS Linker and all related tools are released under the GNU GPLv3 license. This is a copy-left license, which means that you are free to use, modify, and distribute this software in any way you wish, but if you redistribute this software or any derivative of it, you agree to extend these freedoms to anyone you distribute it to. There is also no guarantee or warranty for this free software or problems that may arise, regardless of whether these are a result of user error or problems with this software.

Details of this license can be found online.

Reference Physical Setup

The original context for this tool is a stationary land-based camera looking over an ocean strait between islands, which took images at regular time intervals. The images that contained vessels would then be sorted out and loaded into the AIS Linker.

This guide is written with this specific setup and these goals in mind. The AIS Linker tool is capable of providing for other research goals and other camera setups with minimal adjustment, but keep this setup in mind when reading this guide.

Installation

This tool requires Python (3.x) to be installed as well as the additional library "PyQt5". On most python installs, this can be installed using pip by entering the command "pip install PyQt5" into a command prompt window. No other installation is necessary.

Ensure all of the files from github are in the AIS_linker folder, then run the file "start.py" using python. You may be able to do this by simply double clicking the file, or you may need to run it from the command line using a command such as "python start.py".

First Start

To preview many of this tool's capabilities, a small set of demo data is packaged with this tool on github. To use this dataset, unzip the "test_data_set.zip" file, then move each of the five folders it contains to the AIS-linker folder before starting the program. Then, start the tool.

Contact and Calibration Databases

When the tool starts, you will be prompted to select first a "contact database" and then a "calibration database". The first time you start this tool, select the defaults provided.

The contact database should be a single camera location and a specific method of analyzing data from that site. Whenever a report is created, it will look at all data in the current contact database.

A calibration database represents exactly one camera location, and accounts for any small changes in the camera's focus, angle, etc. This allows for improved display of the AIS tracks as well as distance calculations in reports where this is important. One calibration database may have multiple contact databases associated with it if the same dataset is analyzed in several different ways, but one contact database will only ever be associated with exactly one calibration database.

Important note: all images associated with one calibration database *must* be the exact same resolution (pixel count in both dimensions). Changing the camera resolution within a single calibration database is likely to cause wildly inaccurate positions and/or tool crashes.

Loading Data

To load a folder of images, select the "Load Images" option in the "File" menu at any time. Navigate to the desired folder in the pop-up window, and press "Select Folder". Images are always loaded by folder. Note that this selection pop-up navigates specifically folders and does not show any other files; don't worry if the selected folder appears to have no contents in this pop-up.

If you are using the demo data, select the "images" folder in this tool's directory. Note that since typically you won't be moving around folders directly in this tool's directory, future images should be placed in a different location.

Only image files, specifically .jpg and .png files, are loaded; any files of other types are ignored. While this means that you can load mixed folders with both images you'd like to view and other file types, this isn't recommended. If this mixed data ever contains images that aren't intended to be loaded, it can cause unpredictable behaviour and/or crashes of this tool.

While any image file can be loaded, many of the functions of the linker depend on the image time which is derived from the filename. The expected format for this is:
   "yyyy-mm-dd_hh-mm-ss"

For example:
   "capture_2018-01-02_07-30-00.jpg"

Note that the program's search for this string is somewhat flexible, and for example doesn't care what the prefix is or where in the filename the datetime string is located. However, for simplicity, it is highly recommended to avoid other numbers in the filename as these may be incorrectly picked up as part of the datetime string.

General Navigation

When the program is first loaded and the databases are selected, most of the displays will be blank. Load images as specified in the "Loading Data" section to continue.
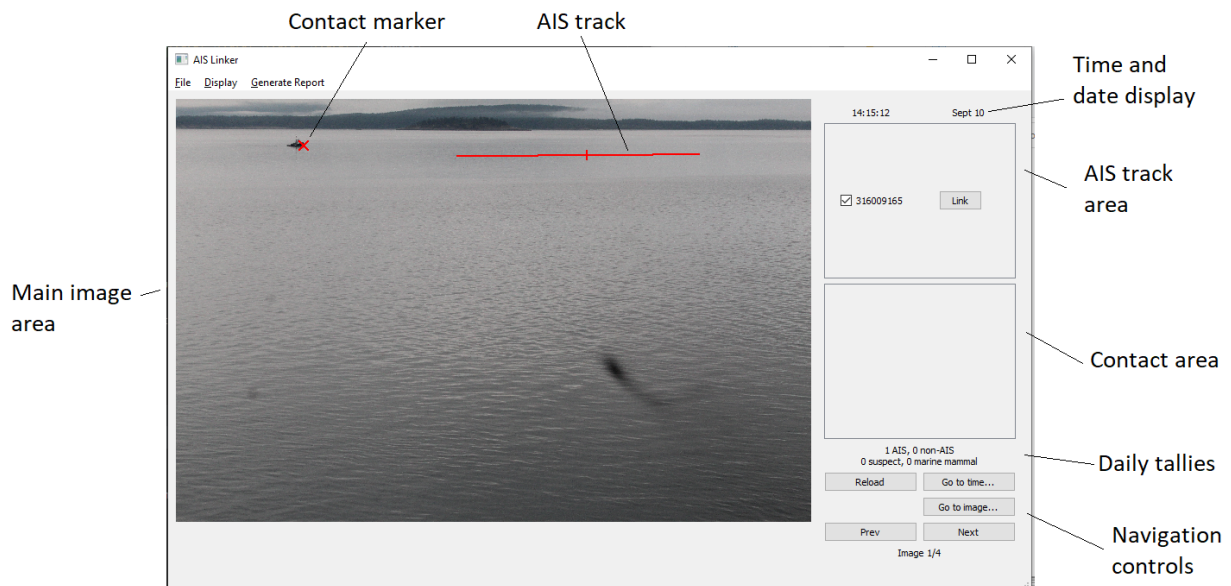

Image 1: Main interface

On the left side of the screen is the main image display, which shows the current image. AIS "tracks" (vessel paths) are shown as lines following the path of the vessel (which are generally

straight, but may vary due to vessel behaviour) with a small vertical tick at the vessel's estimated position at the current image time. Contacts are marked with an "X". Both of these are coloured following the table below:

AIS class A	Red
AIS class B	Magenta
Suspect-AIS Vessel	Pink
Non-AIS Vessel	Yellow
Marine Mammal	Aqua

Table 1: Colour scheme

Right clicking on the image zooms to that location and hides all markings (AIS tracks and contacts). Right click again anywhere on the image to reset the zoom. Left clicking on the image creates a new contact at the specified image. Left clicking on an existing contact marker will open up the dialog for editing or removing this contact.

The right side of the screen contains additional information as well as the navigation controls. In order from the top down are:
1. Time and date display
2. AIS track area
3. Contact area
4. Daily tallies
5. Navigation controls

The time and date display show the date and time of the current image. Note that generally UTC is used for synchronizing the image and AIS times; this isn't necessary, but it is _critical_ that the images and AIS data use the same time zone. If another time such as PST is used, ignoring daylight savings is highly recommended.

The AIS track area displays all AIS tracks that are active within the AoI. Note that some vessels that appear here might not appear as tracks in the main image display area; this can be caused by a vessel approaching the field of view but not quite in it yet, or alternatively by a vessel coming close to the field of view but then turning away.

These tracks each have a checkbox that, when unchecked, prevents the line from being drawn on the image. This can be helpful when trying to parse multiple AIS tracks that transit at the same time. The track then shows the MMSI associated with it, which is generally unique to a certain vessel and can be looked up on 3rd party sites if more information on the vessel is desired. Finally, each track has a "Link" button, which should be pressed when prompted while creating an AIS contact.

The contact area guides the user through adding information to newly created contacts, as well as editing or deleting them. See the "Marking Contacts" section for more information.

The daily tallies show a count of contacts by type in the current day (note: these reset at midnight of whatever time zone the images are using, which may not be the local time zone). This can be useful for a variety of reasons, including:
- Determining if a set of images has already been run through the tool
- Troubleshooting. Notably, if no contacts are showing up in images but the tallies here indicate that they should be, there may be a timing issue between the images and the tool.

The navigation controls provide a number of tools to help with navigating images, as well as a counter that displays how many images and where the user currently is in a folder. The "Reload" button reloads the current image. This can be useful to make sure the current image display is correct after scrolling through images too quickly, as well as to refresh the display after a change such as resizing the window. The "Go to time…" and "Go to image…" buttons provide two ways of jumping to a target image. The "Prev" and "Next" buttons go to the previous and next image in temporal sequence respectively. The left and right arrow keys can be used as a shortcut for these two buttons.

Note that this program loads a small number of images in the background and caches them to improve the performance of viewing images in sequence (both forwards and backwards). Occasionally when scrolling too quickly, this can cause the incorrect image to be shown. Because of this, it is highly recommended to wait for the image to fully load after pressing one of the navigation buttons before pressing another one, such as when pressing the "Next" button a number of times in sequence to go forward a few images.

Marking Contacts

A contact is a detected and marked vessel. Generally, each vessel should be marked only once, but there are specific research cases where you might want to mark a vessel more than once (for example: if you are counting the number of vessels staying in an area over time, you might want to mark each vessel each minute that it remains in the camera).

To mark a vessel, left click on the vessel in the image and a new contact will be created in the contact area. This will guide you through adding basic information to the new contact. Press the "Cancel" button at any time to delete.

When marking the vessel, you may want to first right-click to zoom in on it to place the marker more accurately. If any of the positional information is important for the results (latitude/longitude, distance), it is very important to put the marker precisely at the waterline of

the vessel. The horizontal position of the vessel is largely a matter of convention, but the front of the vessel is an easy convention to use.

The four types of contacts are:
- AIS: vessels that are actively broadcasting AIS, in which case one must specify the AIS track using the "Link" button next to that vessel's MMSI.
- Suspect AIS: vessels that do not appear to be broadcasting, but that seem like they definitely should. This is useful for example when there's a hole in the AIS data, but a tanker vessel that clearly should be broadcasting passes through the image.
- Non-AIS: any vessel that is not actively broadcasting AIS data.
- Marine Mammal: for non-vessel contacts, to mark items of interest that should not appear in the other categories.

When creating a vessel, the prompts past the initial AIS classification add "tags" to the contact, which are aggregated or otherwise used in the various reports. These are not mandatory; you may sometimes wish to scroll to the bottom of the prompt to find the "confirm" button to create a contact without filling in the extra information

There is also a text entry box, which allows users to enter their own tags. Note that these are stored in a comma-delimited format, which means that tags cannot include commas in them. However, if you wish to add multiple custom tags, you can do so by entering them here separated by commas. For example, entering "transport,idling at anchor" would add both tags. Note the lack of space after the comma.

When you left-click on an existing contact, information on it is displayed in the contact area instead of the usual prompts for creating a new contact. This is primarily an informational and debugging display: while you can edit the contact here, it is highly recommended that you instead delete the contact and create a new one with the new information.

Menu Options

The file menu provides the following options:
- Load Images: loads a new folder of images into the tool. See the "Loading Data" section in this document for more details.
- Save contacts: saves all of the added (and edited or removed) contacts to the current contact database. Typically this is not necessary because the program automatically saves on a day change, when creating a report, and when exiting the program. However, a tool or computer crash will *not* save any work-in-progress, which means that sometimes users may wish to save data manually
- Change contacts database: This allows the changing of the current contact and calibration databases without restarting the program. See the "Contact and Calibration Databases" section for more details.

- Calibrate mode…: changes the tool over to the calibration mode, which is used to precisely calibrate the distance measurements over time. See the "Calibration Mode" section for more details.

The "display" menu provides a couple of ways to adjust the current how AIS lines are drawn:
- Limit AIS display distance: this caps the distance at which AIS lines are drawn to 2km. This is sometimes useful in busy waters or when only near vessels are being considered
- Display time: this changes the length of time on either side of a vessel that the track is drawn for. For example, when set to one minute, the vessel track will be drawn from one minute before the image's time until one minute after the image's time. This is sometimes useful when dealing with timing issues.

The "Generate Report" menu allows users to create different types of reports. See the "Creating Reports" section below for more details on the different types of reports.

Creating Reports

Several types of reports are available, each created for a different research goal. See the Technical Guide for details on how to add new report types if needed.

All reports are formatted as .csv files, and the file will be created in the tool's main directory. The three types of reports are:
- Daily reports: These reports sum each category by day, with each line corresponding to a single day. Basic categories are always summed in the same order, then each unique tag is also added as a separate category. Use the header in these reports to determine which columns correspond to what. These files also include sums of the entire report period at the end.
- Minute reports: These reports are a specialized report type for determining how many of which types of vessels are within the camera's field of view at any given time (accurate to the closest minute). To generate proper results, these reports require a database in which each vessel is marked each minute (as opposed to marking each vessel exactly once per transit, which is assumed in the other report types). Unique tags are ignored. Near shore and far shore are defined as <1km and >3km respectively.
- Vessel reports: These reports print one line per contact (which is expected to represent one line per vessel), in which it includes positional data and any unique tags on that vessel. These reports generally require a user to do some sort of analysis afterwards, but contain the most information about each contact. Also note that while positional data is included, these can be subject to quite a bit of error depending on distance from the camera and the camera's calibration, and should be handled with care if used.

The process for creating each type of report is the same:
1. Ensure you are using the contact database that you want to generate the report for.

2. Select the type of report you would like to create from the "Generate Report" dropdown menu.
3. Select the date range over which you would like to create the report.
4. Change the name of the new report if desired.

If you are using the sample data set, a small amount of data is included for September, 2019. Include this month in the date range to create sample reports.

Calibration Mode

Getting exact positions of the contacts is difficult due to the sharp angle of the camera: at a camera height of 20 meters and distances of several kilometers, even a single pixel in high resolution images can be hundreds of meters. The method of calculating distances used in the program follows the Earth-as-a-sphere model, which is accurate to within a few percent of the Earth's actual spheroid shape, but which does not account for any atmospheric conditions or rising/falling tides. Precisely calibrating for small differences is important whenever the distances are used in the results.

This tool is capable of accounting for lens field of view adjustments (such as camera zoom) and lens tilt (a non-flat horizon line, which is common to some extent even if the horizon looks flat, especially as camera resolutions increase).

For locations where the distances aren't important in the results, they are still used in drawing the AIS tracks, but pixel-perfect calibration isn't needed. For these locations, it is recommended to calibrate the first image to start, then two images whenever the camera is adjusted: one just before, and one just after the adjustment. This is important because the program uses linear interpolation between calibration images, meaning that if only an image after the adjustment is calibrated, the program's calibration will slowly drift over potentially a long period of time.

To enter calibration mode, select the "Calibration mode" option from the File menu. Note that many of the tool's functions are dramatically altered and others are not available when in calibration mode. To exit calibration mode, select the "Exit calibration mode" from the File menu at any time.

Using the same image as in Image 1: Main interface, note the following differences when calibration mode is entered:

1. AIS tracks and contact markers are not shown.
2. These markers show the marked reference points when an image is already calibrated. Right-clicking on the screen still zooms in, but unlike in normal operation, the calibration markers are still displayed when the image is zoomed in.
3. The date and counts are still displayed

4. The contact area is replaced by a new calibration area
5. The pixel position of the two marked points is displayed here.
6. Either point can be reset individually if the user isn't happy with their position.
7. The functionality of the navigation buttons are altered. Reload still reloads the display as normal
8. The "Prev" and "Next" buttons now skip to the previous or next image that either (a) is already calibrated, or (b) the program recommends calibrating.
9. The "Calibrate next" image functions like the "Next" button, but skips any images that are already calibrated and instead jumps straight to the next image that the program recommends calibrating.
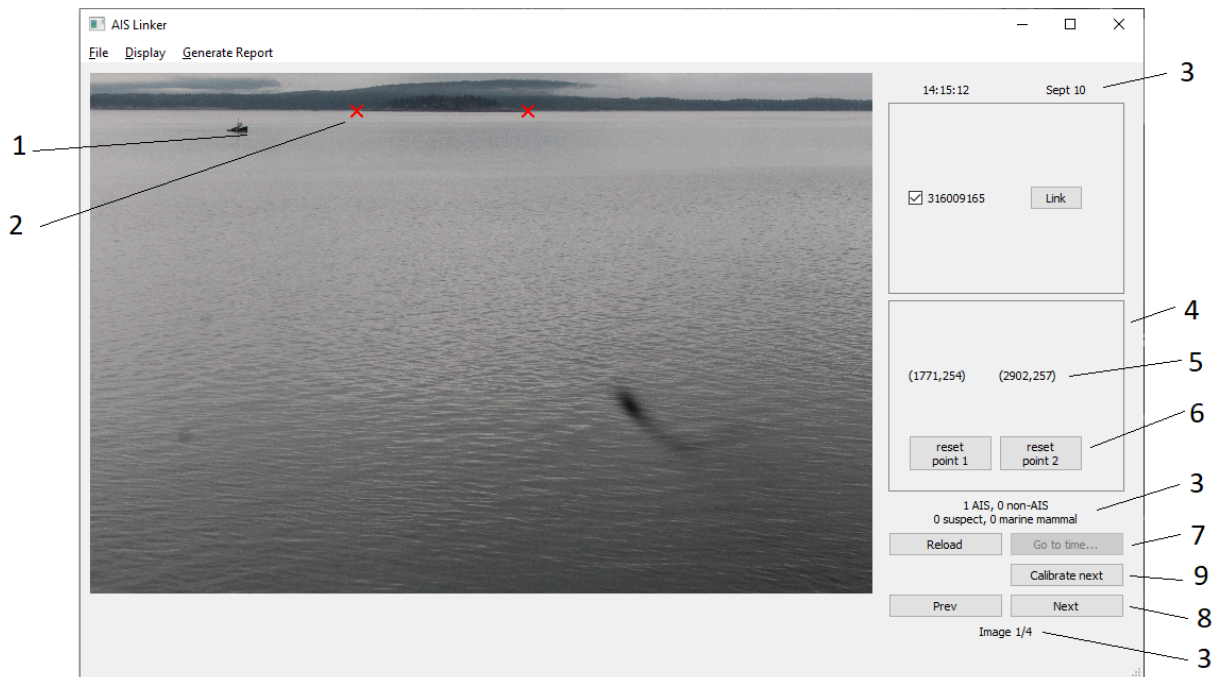

Image 2: Calibration mode

It is recommended to calibrate a set of images before marking contacts on them due to how this improves the accuracy of the drawn AIS lines. However, distance information in the reports is always generated from the calibration of the tool when the report is created. This means that calibrating for a set of images even after marking contacts in them is useful for reports where positional accuracy is important.

Calibration Mode Workflow

When calibrating a set of images, the following workflow is recommended:

1. Open the tool and load the calibration database that you wish to save the current calibration data to.
2. Load the set of images you wish to calibrate using the file menu option
3. Enter calibration mode using the file menu option
4. Starting at the first image:
    a. Zoom in on reference point 1 using the right mouse button, then mark it using the left mouse button.
    b. Zoom in on reference point 2, then mark it using the left mouse button.
    c. Click the "Next" button to jump to the next image.
5. Continue the above process with each image that the tool jumps to.

If the reference points are ever obscured (for example, by a large transiting vessel), press the "Next" button until both reference points are clearly visible. It is very important to never "guess" where a reference point approximately is - these marked reference points should be as accurate as possible.

Creating and Deleting Databases

New databases (both calibration and contact) can be created from their corresponding selection menus. Existing databases that are no longer in use can be archived: this doesn't delete the data, but instead hides them from the display. While generally not recommended, a database can be fully deleted by removing the single folder that contains a database's name from this tool's source directory.

Databases can also be edited from these menus. Names and descriptions can be freely updated. In general, calibration information (latitudes and longitudes) should only be edited if there is a mistake in the original data; if new calibration points are chosen, create an entirely new calibration database with the new information.

Adding AIS data

Due to a wide variety of source AIS formats, this tool expects AIS data to be preprocessed to a specific format. See the "Processed_ais" folder in the sample data for an example of this format.

The AIS data must be organized into daily files named "condensed_ais.txt", with one file corresponding to exactly one day. These files must be organized from a root folder into yearly, monthly, and then daily folders that are each named with only the number of the year/month/day. For example, AIS data for February 1-3, 2020 might look like:

Root_folder_name

```
 -> 2020
      -> 2
            ->1
                  ->condensed_ais.txt
            ->2
                  ->condensed_ais.txt
            ->3
                  ->condensed_ais.txt
```

Within any given ais file, each AIS point corresponds to one line. Only a limited amount of information from the original AIS message is preserved, and must be in the exact comma-delimited format:

date, MMSI, AIS class A/B, latitude, longitude

For example, a few lines of AIS data might look like:

2019-09-11_00-00-00,369035000,A,48.64402,-122.869226667
2019-09-11_00-00-00,366896490,A,48.7705766667,-122.847611667
2019-09-11_00-00-01,338200899,B,48.6520766667,-123.154066667

Note the formatting of the datetimes. Also note that the timezone *must* match the timezone of the images: UTC is recommended for simplicity, but the important part is that they do match.

These files also must have each line corresponding to any specific MMSI in temporal order, but the ordering of lines with distinct MMSI values does not matter. The two easiest ways to have the data sorted in this way are to either have the entire file in temporal order, or to have the file sorted by MMSI with each MMSI being in temporal order.

No additional filtering is required (removing stationary vessels, removing duplicate points, etc), but some might be helpful in specific use cases.

Unlike other databases, AIS folders can be stored either in this tool's root folder, or in a separate location. Keeping these folders in this tool's root folder allows the use of relative file paths for this folder, which can make it easier to copy or move between deployments, but there are use cases where an external AIS folder is the preferred option.

Running Remotely and in Parallel

If this tool is stored in a central location that can be accessed by multiple computers, the tool supports multiple people using it simultaneously, generally without concern. This allows multiple people to use the same data source and even contribute to the same contact database. However, specifically multiple people loading images from the same day in the same contact database *will* cause data loss (one person's work will overwrite the other's), and must be

avoided in this setup. Editing and archiving/recovering databases should also never be done when another user is using those specific databases.

Running the program remotely like this, including accessing the database files it stores internally, was found to have minimal performance impacts in practice. However, loading large image files is the tool's biggest performance limitation, meaning that these images should be stored locally whenever possible.