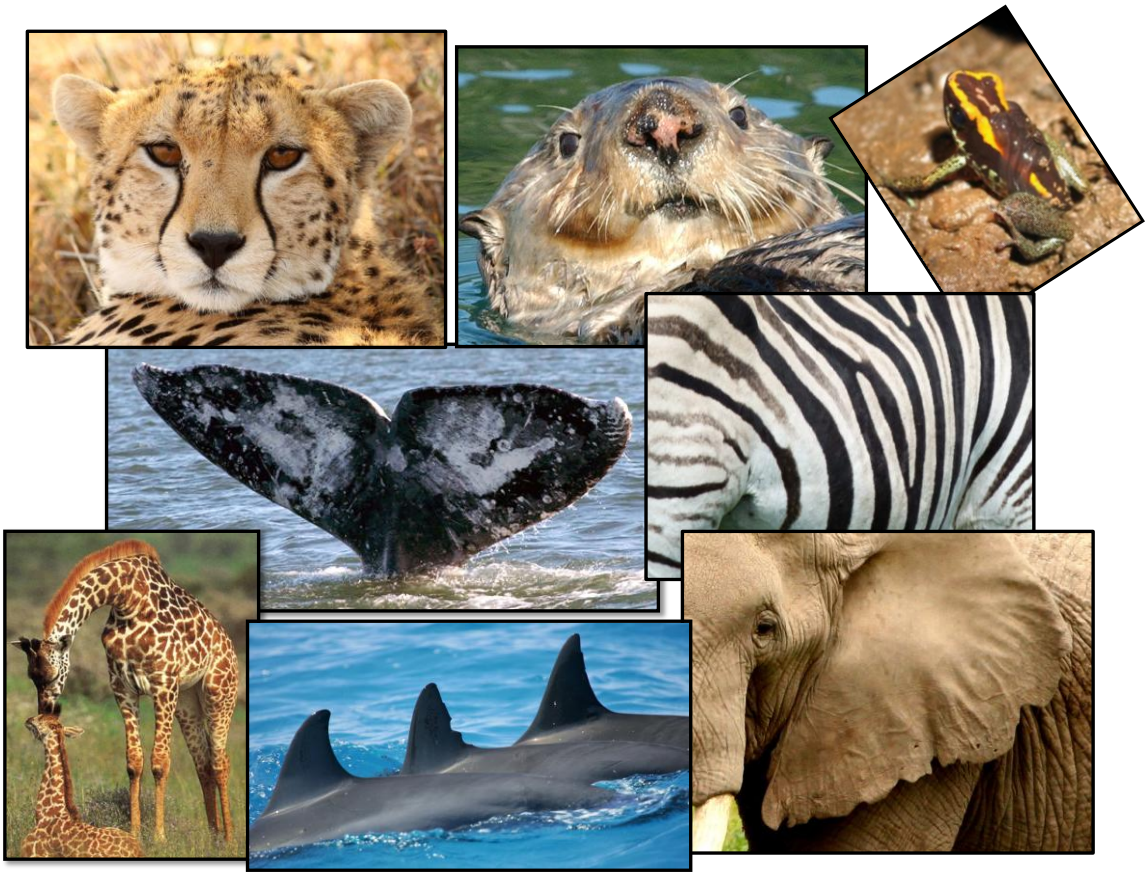
DISCOVERY

PHOTO-IDENTIFICATION DATA-MANAGEMENT SYSTEM FOR INDIVIDUALLY RECOGNIZABLE ANIMALS



ACADEMIC HOME INSTITUTIONS

The Swire Institute of Marine Science, The University of Hong Kong
Mammal Research Institute, University of Pretoria

FUNDING ORGANIZATIONS

National Research Foundation (NRF), South Africa
Research Grants Council (RGC), Hong Kong
Ocean Park Conservation Foundation, Hong Kong (OPCFHK)
The University of Hong Kong

MANUAL COMPILED BY

Simon W.H. Wong, Glenn Gailey, Stephen C.Y. Chan, Olga Sychenko, Carmen K.M. Or, Sze-Wing Yiu, Wei-Lun Chang, Caryn Behrmann and Leszek Karczmarski

Citation details: Gailey G. & Karczmarski L. 2012. Discovery: Photo-identification data-management system for individually recognizable animals.
Web address: http://www.biosch.hku.hk/ecology/staffhp/lk/Discovery

Photo Credits: Caryn Behrmann, Stephen C.Y. Chan, Glenn Gailey, Leszek Karczmarski, Lenin Oviedo, Carmen K.M. Or, Heidi Pearson, Stephanie Plön, Ryan Reisinger, Olga Sychenko, Nana Takanawa, Ricky Tang, David Weller, Simon W.H. Wong, Sze-Wing Yiu and Elaine Yuen.

Table of Contents

List of Figures

List of Tables

Chapter

1

Introduction

Individual photographic identification (photo-ID), where a photograph of an individually recognizable animal constitutes the marking/recapture event, is a powerful and yet non-intrusive technique to study the animals' behavioral and population ecology. It has been applied across species and habitats, gathering a large variety of information, such as population size, population dynamics and trends, group dynamics, social behavior and social structure, geographic range, movement patterns and many others.

This technique involves many hours of labor-intensive processing of photographic material in the lab. Thanks to the recent advances in digital photography, however, there are many avenues to advances in both data collection and subsequent data management and analyses. Digital technology not only allows the photo-ID data to be processed immediately upon the completion of a field day, even in remote locations where processing of traditional photographic material would not have been possible, but it also provides a means for further data management and analyses that could be performed at a more efficient rate by linking different types of data (behavioral, geographic, environmental, etc.) with the photo-ID data.

However, the use of digital photography has also lead to data management issues, as many more high quality digital photographs can be obtained in a short time period at lower cost. For long-term research projects, data management is especially important due to the vast amounts of data accumulated over time, which presents an obvious obstacle to data management.

A number of computer-assisted systems have been developed in recent years to aid individual photographic identification, especially so for marine mammals. Most of them, however, although very useful, have focused on specialized matching routines and lack a more generalized system that could assist researchers at all levels of data collection and management -- a system that would have

not only the ability to assist with matching individuals but also store, manage, and process digital images, provide file naming routines, assist with filtering of raw data and all levels of individual-ID matching, and link sighting information with environmental, geographic, and numerous user-defined parameters. Some systems have built Access forms and queries that are designed for specific projects or species, but lack a dynamic functionality to be applied to other projects without intimate knowledge of the database, form design, and programming language. Clearly, a more generalized dynamic approach was needed to provide a user-friendly platform to assist researchers not only at the matching stage of individual photo-ID data, but also at the multitude of steps of field data collection and the complex data management and analyses that take place after individual matching is completed. The photo-identification data-management system DISCOVERY offers such a computer-assisted approach for photographic ID data and associated field records.

The Discovery data-management system can be used to centralize a database for multiple species and multiple study areas; it is particularly useful for maintaining a single database for research projects collecting data at large geographical scales and between multiple research teams working on different databases. It provides a means to streamline computerized data collection in the field, store photo-identification data along with a suite of associated information, and process and manage all stored data and associated information; it provides a means to link the new system with traditional (film photography-based) datasets to form continuous complete datasets, provides graphical displays of data and basic analytical tools, and facilitates easy integration of all collected and stored data to and from other tools, e.g. currently available applications for further complex population-level analyses. Discovery has been designed so that with a multitude of dynamic functions it can be tailored to suit project-specific requirements and user-specific needs.

Leszek Karczmarski and Glenn Gailey

Hong Kong - Cape Town - Galveston, November 2011

Recommended System Requirements & Installation

Minimum Requirements

Windows operating system (XP/Vista/7/8).

Recommended Software

Microsoft Access, Excel, ESRI ArcMap, EasyGPS (http://www.easygps.com) .

Installation Package and Check for Update

Download the installation package at here.

Use Discovery's check for updates regularly for new versions of the software.

Setup

Open the **Discovery.exe** file  in the installation package and follow the prompts. Once the program has been successfully installed, copy the **Discovery.msi** file to your installation directory (e.g. C:\Program Files(x86)\Discovery\Discovery). Start the application by either going to the shortcut on the desktop or under programs in the start menu.

DISCOVERY Databases

Databases - Discovery creates and uses two types of databases:

1. *System Database (SYSDB.mdb)* - This database is stored in the program directory of Discovery. It will be automatically created if it is absent in the program directory (i.e. upon the first instance the application is started or after the SYSDB.mdb file has been deleted). This database is used to store user-defined parameters, and preferences of Discovery (e.g. changes in "Setup" > "Define" and "Setup" > "Preferences"). *No data* are stored in this database.

2. *Data Database (e.g. MyData.mdb)* - The location and name of this database file is *user-defined*, i.e. a file that users create and name, with the extension .mdb). This database stores all information (e.g. sighting information, images location etc.). Upon initial start of the application, the system will ask the user to create a data database which the user can assign a name. Users can also create / open other data databases by "File" > "Create Database" and "File" > "Open database".

Main Menu

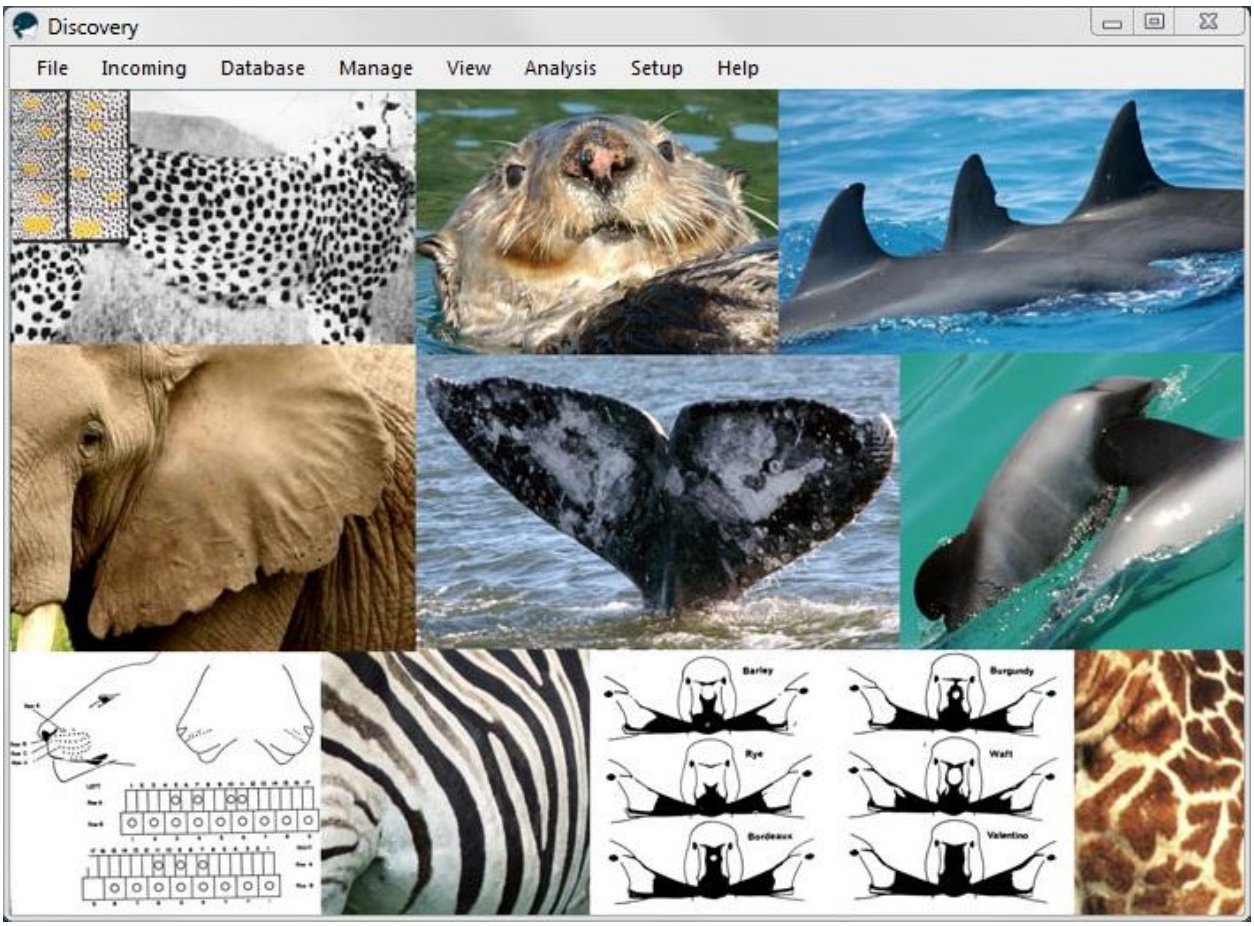


Figure 1. The background image can be user-defined by double clicking on the current image or selecting in the preferences (see Preferences). This is only an aesthetic change, and will not affect any function of the program.

File - Primarily for import/export and creating/opening data databases.

- *Create Database* - Create a new data database.
- *Open Database* - Open an existing data database.
- *Open* - Open either an image file or GPS data file to view.
- *Import* - Import locations of image files, sighting information, database documents and/or another discovery database (see Import).
- *Export* - Export data from the current database into various file formats like MARK and SocProg (see Export).
- *Sync Folders* – Synchronize destination folder by source folder.
- *Update* – Update several optional information into database.
- *Open in Access* - Efficient way to open the current system and data database in Microsoft Access. Note: User must have Access installed on the computer to be able to use this function.

Incoming - The incoming menu contains tools to help the user with renaming images, manipulating EXIF data, filtering, and matching images within a directory. *No data are stored in the data database from these tools.* These tools are designed to help structure files upon return from the field and prepare them for comparison to the primary database of individuals (see Chapter 2 for further details).

> ***Note:*** EXIF data contains details of the image stored in the header of the image file, e.g. camera settings; file detail, geographic location, date/time, etc. With Discovery, user can access and edit EXIF data easily WITHOUT the need of other photo-editing software (see EXIF).

- *Filename* - Allows user to rename images. Using this tool also allows the software to store useful information (such as study area, group ID, individual ID) into each file by use of EXIF.
- *Filtering* - Image manipulation and processing tools.
- *Matching* - Matching within a group (or folder) to reduce the folder to the best quality representatives of each group per survey.
- *Thumb View* - Browser for quick-view of images from incoming folders.

Database - These tools provide a means to enter data into the database by adding individuals or resightings of individuals and other associated sighting information.

- *Match* - Tools to find a potential match to a query image compared to previously identified individuals (see Chapter 4).
- *Sightings* - Tools to record group/survey sighting information (see Chapter 5).
- *Survey Effort* – Record survey effort of each survey (see Chapter 5).
- *Others* – Forms to record user-specified information that may not be appropriate to include in other available datasets (see Chapter 5).

Manage - These tools are designed to assist in managing the data database by viewing and verifying the data entered and reviewing/verifying photo-ID photographs (see Chapter 6).

- *Database* – Explore and edit discovery database.
- *Verify* – Confirm proposed new/matched images of individuals and their new records to the database.
- *Review* – Review the verified catalog of individuals.

View - Tools to use to visualize information contained in the data database (see Chapter 7).

- *Individuals* - Summary of the unique individuals in the database as well as associated image and sighting information.
- *Map* - Geographic plot of location information from individuals, effort, sighting and resightings recorded in the database.

Analysis - Tools that process and/or plot out information derived from the data (see Chapter 8).

Setup - Tools to define user preferences and parameters to be used for the database (see Chapter 2).

Help - Information pertaining to Discovery version and help (this manual) information.

Chapter

2

Parameter Setup

The Discovery Photo-Identification System was designed to be as dynamic as possible to provide researchers a tool to fit their different study types and questions. Due to this dynamic design, the user needs to setup various parameters and settings in order to use the system correctly. The parameters are categorized into four categories: 1. General, 2. Individual ID, 3. Sightings, and 4. Other. Each category is described below.

Note: Users *cannot* type special characters (e.g. =, <, #, @, etc) as a textural description, as the software will not recognize these characters. These descriptions are used as data fields in the Access database where special characters are not allowed and can cause errors later on.

Note: Users should try to avoid using field and parameter names that might cause confusion with program coding (e.g. "Date", "Time" and "Unknown") which may lead to a "Syntax Error". If "Syntax Error" message occurs, attempt to replace the names of relevant fields or parameters (e.g. use "Date_Sight", "Time_Sight" and "Unknown_Behavior" instead).

In the main menu, click "Setup" > "Define"

Figure 2. Setup menu for the four main categories and sub-categories.

General - Important parameters that are used for the entire system (Figure 2).

- *Study Area* - Textual descriptors of the area, catalog, or study area of interest. Users can type the name into the provided box and click "Add"; or select the category to be deleted and click "Delete". Users can add as many study areas as they want, but *a default or current study area of interest must be selected.* After entering the names of the study area, the user can pull down the combo box (labeled "Default") and select the name that the user wants to currently work on. The default study area can be changed at any time via this menu (Figure 3).

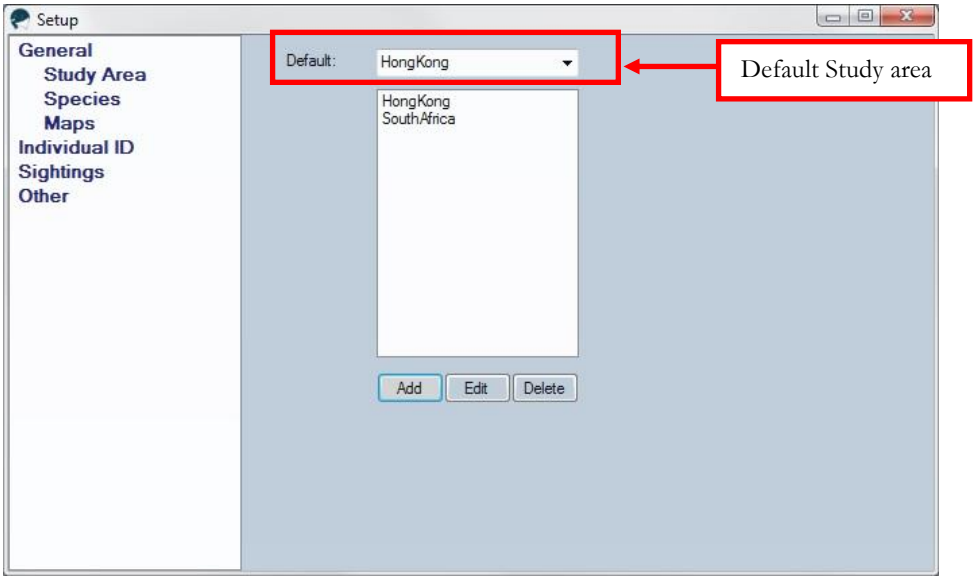


Figure 3. Combo box with the default study area. Other areas that have been added by the user appear in the list box below the default study area name.

- *Species* - Textual descriptors of different species that will be used in the database. Users can type the name into the provided box and click "Add" or select a category that is to be deleted and click "Delete". Users can add as many species as they want, but *a default species must be selected.* After entering several different species names, the user can pull down the combo box and select the name that is to be the default species. Default species can be changed at any time. The interface is similar in design and functionality as Figure 3.

- *Map* - ESRI shapefile maps that can be incorporated to geographically view sighting information. To add a new map, click "Add" and select the shapefile.

Individual ID - Setup parameter related to individual identification (Figure 4).

- *Rename Images* – The Discovery software provides a function to rename (one or a batch of many) image files based on user-defined format (see Renaming in Chapter 3). The new file name of the image files can contain user-defined parameters (e.g. Photographer name, Study Site, Date, etc.) and information extracted from the original filename as created by the camera software. Further textual information can also be added to each filename upon renaming. If the user does not designate a renaming structure, then the following structure will be used as a filename format:

  [ID] + _ + [DATE] + [STUDY_SITE] + [GROUP_ID]

  Filenaming structures are outlined in Defining Directory/Filename Structure below.

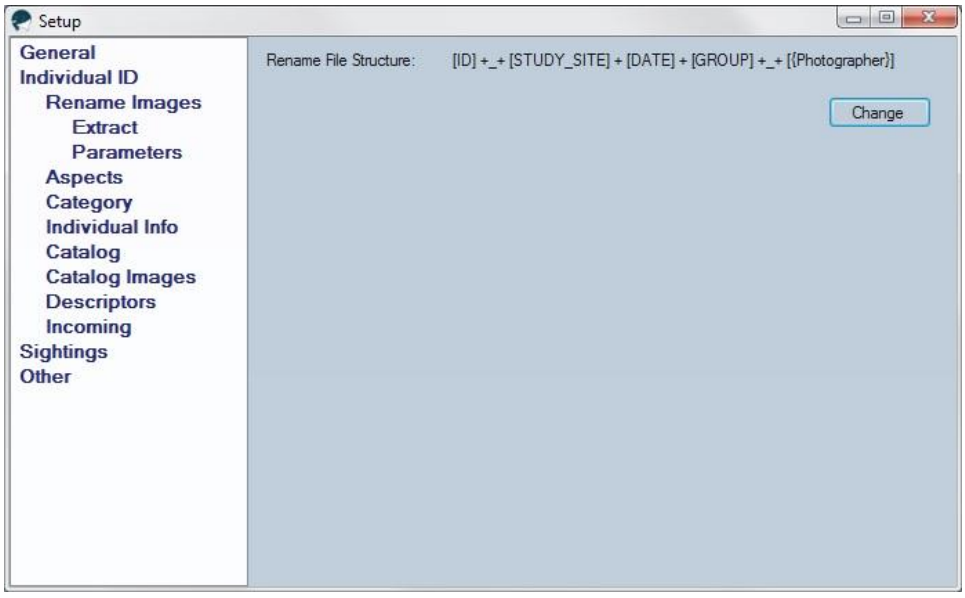


Figure 4. The currently defined filename structure is displayed on top. Users can edit the structure to fit their specific needs.

Example of renaming filenames based on a pre-defined file structure:

Original Image Files:

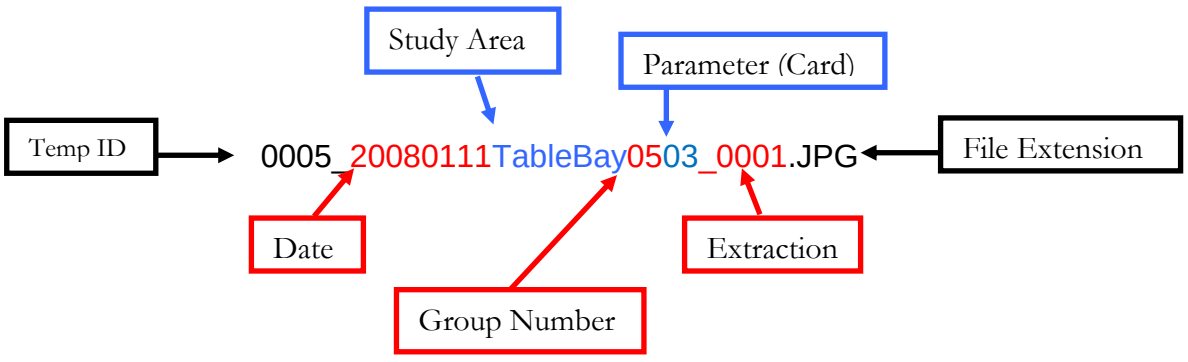
DSC_0001.jpg DSC_0002.jpg DSC_0003.jpg DSC_0004.jpg etc……

Renaming Format

[ID] + _ + [DATE] + [STUDY_SITE] + [GROUP_ID] + [Card] + _ + [Frame]

Parameters	Meaning of the Parameters
[ID]	ID of the individual [A temporary ID (Temp ID) would be assigned automatically in renaming, which would later be replaced by the actual ID when the photo is added to the catalog.]
[DATE]	Date (Automatically extracted from EXIF, if available, of the original image file); users can also define date as needed.
[STUDY_SITE]	Study area (pre-defined by user)
[GROUP_ID]	Group identifier name/number (assigned by user)
[Card]	Card name/number (assigned by user)
[Frame]	Frame number (auto-extracted from original file name in this example)

Renamed Image files



- *Extract* – This part configures how Discovery extracts (all or part of) the original image filename; this is an automatic extraction of information from standard file formats generated by cameras, such as Nikon DSC_0001, where frame number could be extracted automatically. Of course, different manufactures/brands have different default filenaming as well as customized filenaming options.

Quick Procedures:

1. Click on "Extract" to open the extraction menu.
2. In Box A (see Figure 5 below), input number of characters up to the first character to be included (i.e. if we want to extract the "e" in "the", we would start with 2).
3. In Box B, input the number of characters leading up to and including the last character to be included (i.e. if we want to extract "ppy" in "happy" we would input 2 in Box A and 5 in Box B).

Example of Extraction:

Original filename was **DSC_0028_ABCD_9843_96FE.TIF**

Each interval between characters (or space before the first character) represents a position in the name (i.e. space before "**D**" = 0; interval between "**S**" and "**C**" = 1; interval between "**C**" and "**_**" = 2; etc. The first interval (or space) is position 0 (zero).

If the user would like to keep the "**9843**" part of the original filename, in Box A the user would need to type in the lower range of the desired positions (in this case 14) and in Box B it would need to upper range (in this case 18). Therefore, the positions to extract "**9843**" would be 14 to 18.

4. The user then clicks "ADD" and specifies a textural description (such as 'frame number') of the extraction.

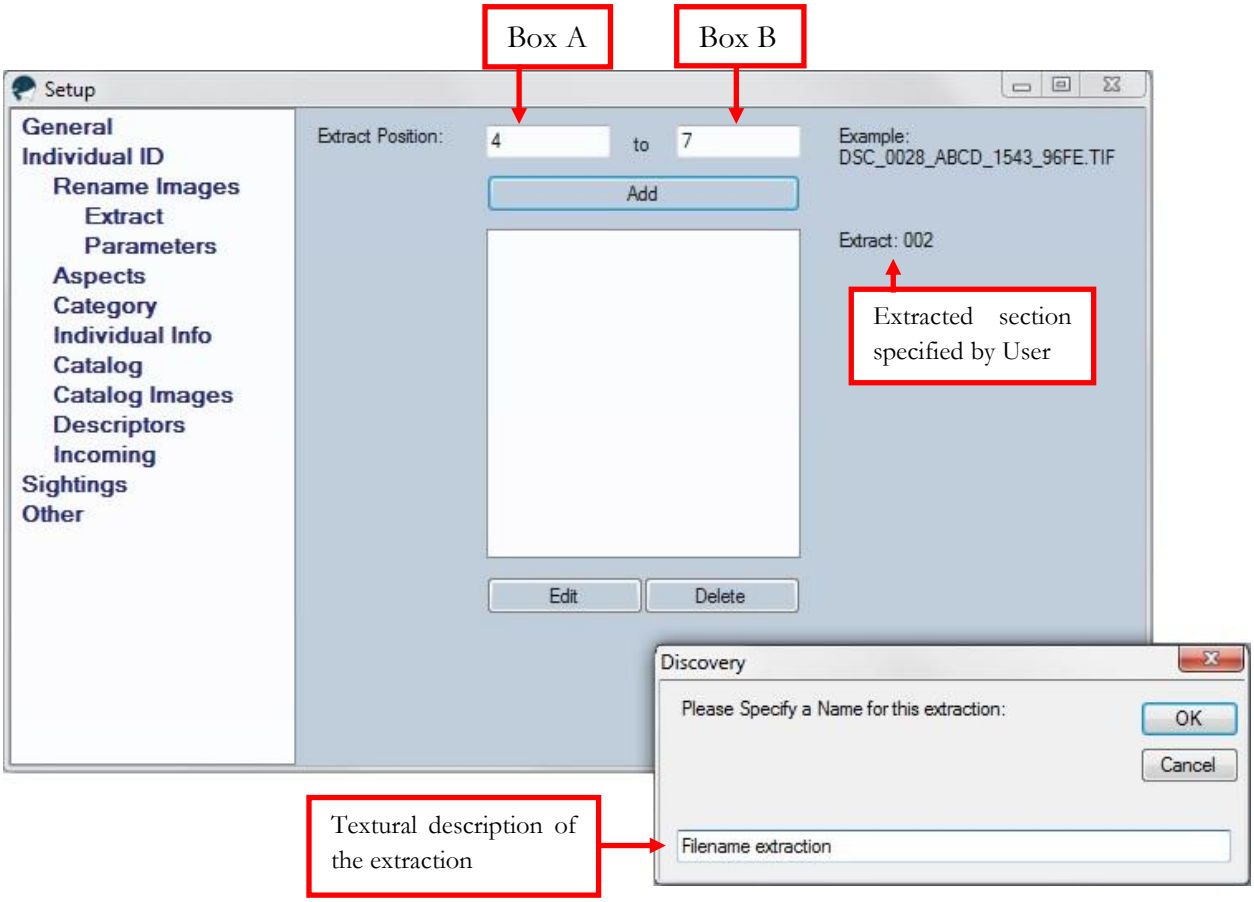


Figure 5. Window illustrating setup for automatic extraction from original filenames.

- *Parameters* - User-specified parameters to include as part of the new filename; for example, the card number or initials of the photographer could be incorporated into the filename upon renaming the images.

Quick Procedures:

1. Click on "Parameters" to open the parameters menu.
2. Click "ADD".
3. Type in the name of the parameter (i.e. Card) and click "OK".
4. The parameter will be listed in the box.
5. Repeat until all parameters are included.

Note: In the dynamic renaming structure, parameters are designed with a "[{" Name "}]" designation. Therefore, for the parameter card, it would be designated as "[{Card}]" in the renaming structure.

- *Aspects* - Categorical description of the different aspects that may be used for photo-identification; for example, left and right sides of the dorsal fin for dolphins; or left, right, and flukes for whales. Users may add as many Aspects as needed, but only one aspect can be assigned to an image (i.e. an image can be of either the left or right side of the dorsal fin, not left and right side of the fin at the same time).

Quick Procedures:

1. Click on "Aspects" to open the aspects menu.
2. Click "Add".
3. Type in the name of the aspect (e.g. Left) and click "OK".
4. The aspect will be listed in the menu.
5. Repeat until all desired aspects have been included.

- *Category* - Textual classification of the primary identification features. For dorsal fins, these could be "top notch", "1 notch", "2 notches", etc. (notches referring to obvious indentations or protrusions from the fin's edge). Other identification features may include scarring, permanent/long lasting pigmentation, etc.

Quick Procedures:

1. Click on "Category" to open the category menu.
2. Click "Add".
3. Type in the name of the category and click "OK".
4. The category will be listed in the menu.
5. To add a subcategory, right click on the category. Select "Add subcategory". Type in the name of the subcategory and click "OK".

Note: User cannot use special character (e.g. =, <, #, @, etc) as a textural description. The software will not recognize these characters as they can cause errors later on.

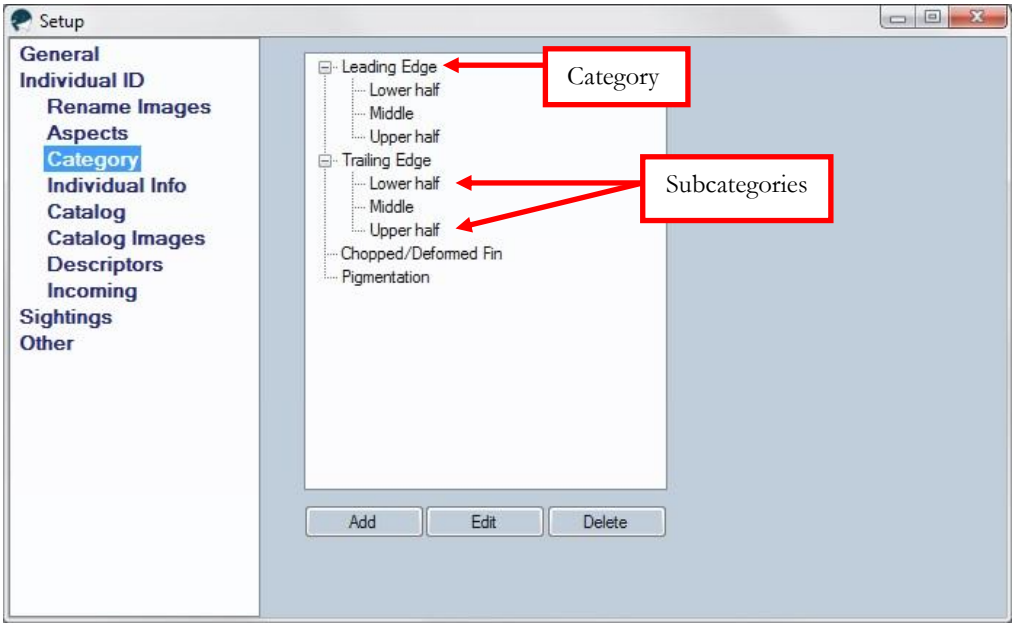


Figure 6. Setting up categories that will be used for classifying primary identification features for filtering images for potential matches in the database.

- *Individual Info* - Information pertaining to the individual ID and not to a specific image. Examples include sex, age, name, etc. Procedures are similar to that of ASPECTS.

- *Catalog* - The main directory where images of individuals are stored for the photo-identification catalog.

Click on "Catalog" to open the catalog Menu.

Click on "Change**"** and select a folder location of the catalog.

Note: When the individuals are identified (i.e. matched to an individual in the catalog) the images are *moved* automatically from the location of the current folder to a separate folder created in the directory defined by the user. The folder created will be named based on the user's defined catalog structure. *Hence, users are advised to create a separate new folder to store the newly formed catalog.*

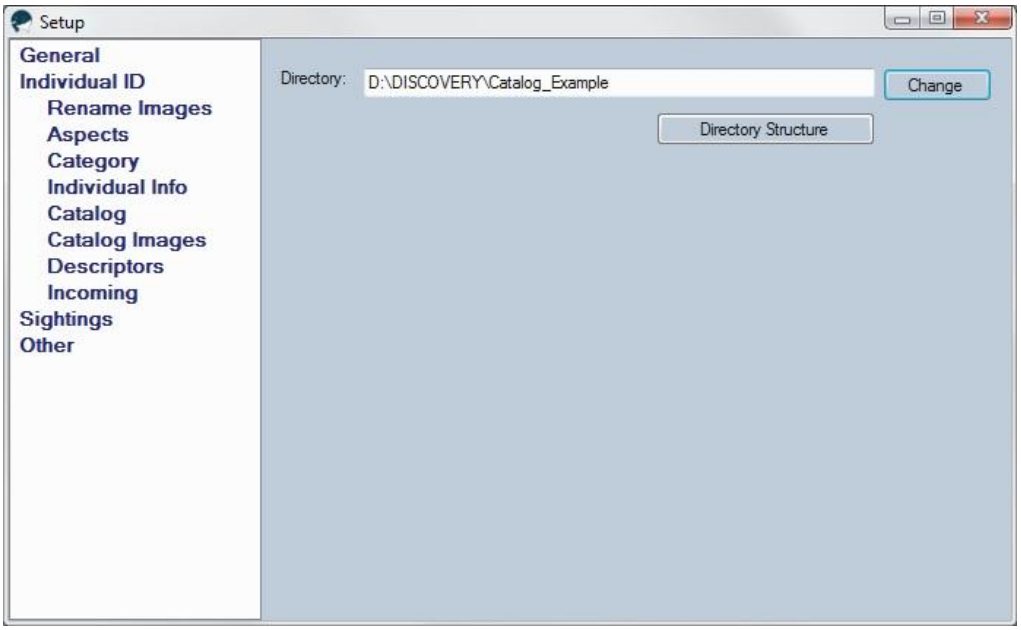


Figure 7. Screen capture for defining the location of the directory of database images.

Catalog Structure:

Users can define the directory structure of the catalog as they wish.

The default structure of the catalog is as follows:

User-Defined Catalog Root Directory\[SPECIES]\[ID]

where all matched individuals are moved into a subfolder of their individual ID that is located within a directory of the defined species. All species folders are put in the catalog directory folder (selected by user).

Note: "Catalog Root Directory" is the folder user has defined for the catalog.

Example: If the user defines their catalog directory to be in the "C:\Photo-ID" directory with a new individual's ID of 001 ([ID]) and species "Spinner" ([SPECIES]), then the new individual would be added to "C:\Photo-ID\Spinner\001" folder.

Users can change the structure of the catalog by clicking "Directory Structure".

For further information, see Defining Directory/Filename Structure.

- *Descriptors* - Additional descriptions of identifiable features that may not be part of the primary classification scheme, such as scars, cookie cutter shark markings, killer whale tooth rakes, etc. Procedures that apply are similar to that of ASPECTS.

- *Incoming* - This is the area where image data are processed and prepared for matching to the catalog of individuals. To setup incoming information, double click "Incoming" in Figure 8.

    - Directory - This is where the images are stored for initial processing prior to being entered into the "master" or catalog database (i.e. matching). The user defines its location and this area is where all the images from temporary images are stored.

      To change, double click "Incoming", click "Directory", click "Change" and select the folders that contains the desired directory to process your incoming images.

      The directory creation structure format is used in the incoming directory when applying renaming operations. The structure is based on user-defined parameters or a default structure (Incoming Root Directory\[DATE]\[GROUP_ID]).

      For further information, see Defining Directory/Filename Structure.

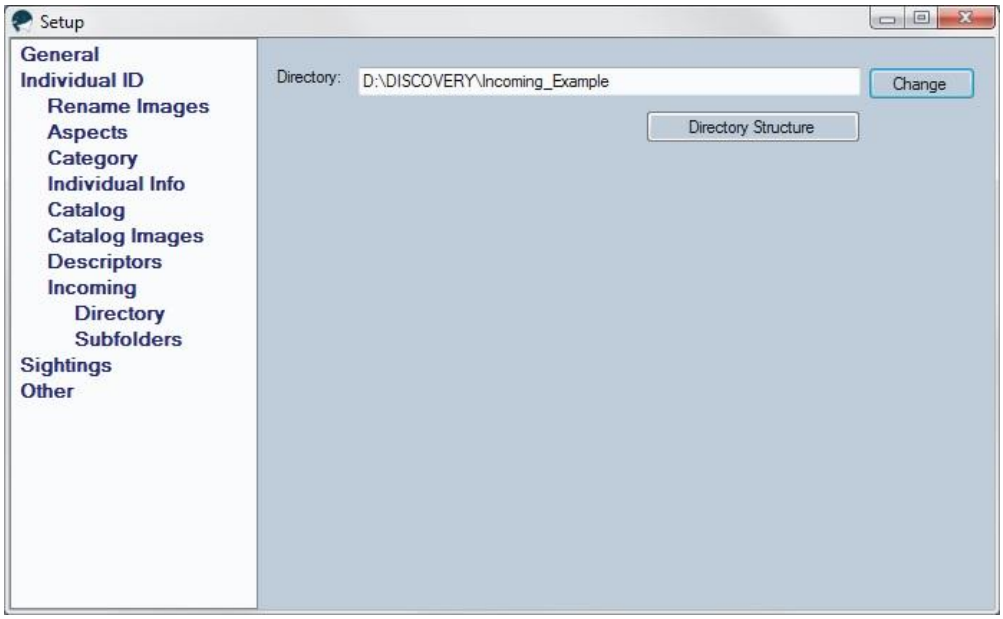


Figure 8. Defining incoming directory.

- Subfolders - Directory folders that are created within a given directory in order to reduce the number of images needed to be entered into the database. The user can create subfolders which will be useful to segregate images into different categories. For instance, during the "Filtering" process (see Filtering in Chapter 3), the user may want to create subfolders with different quality criteria i.e. 60-69%, 70-79%, 80% and above for quality or categorical names, such as blanks, no-info, etc. The names of the folders are user-defined and can be named according to any criteria system. Another usage of subfolders can be for duplicate images of the same individuals, which can be put into a subfolder "Extra". Blanks or other photographs containing no information for photo-identification purposes can be placed in a subfolder named "No Info".

  To add/modify/delete specified subfolders, double click "Incoming", click "Subfolders", click **"**Add", "Edit" or "Delete.

  Note: After the subfolders are added as described above, there are no folders created yet. Subfolders will be automatically created when incoming images are sorted in filtering and the user designates placement of an image into the subfolder. This usually occurs in the filtering process, when the user clicks on the button of an appropriate subfolder and the software will either: *(a)* generate the folder (if it has not already been created earlier) or *(b)* transfer the images into that folder (if the subfolder already exists) within the current directory. Once these folders are created, when sorting incoming images and placing the images into these different folders, the images are automatically transferred to the designated Subfolder.

Sightings - Sightings refers to information about the group encounter (see Chapter 4)

- *Group Info* - User-defined descriptions pertaining to details about the group, such as number of Adults, number of Calves, etc. There are several data types that can be assigned to each parameter (see Data Types below).

  To add, type the name of the new parameter in box of "Parameter Name", then choose "Data Type", and click "Add".

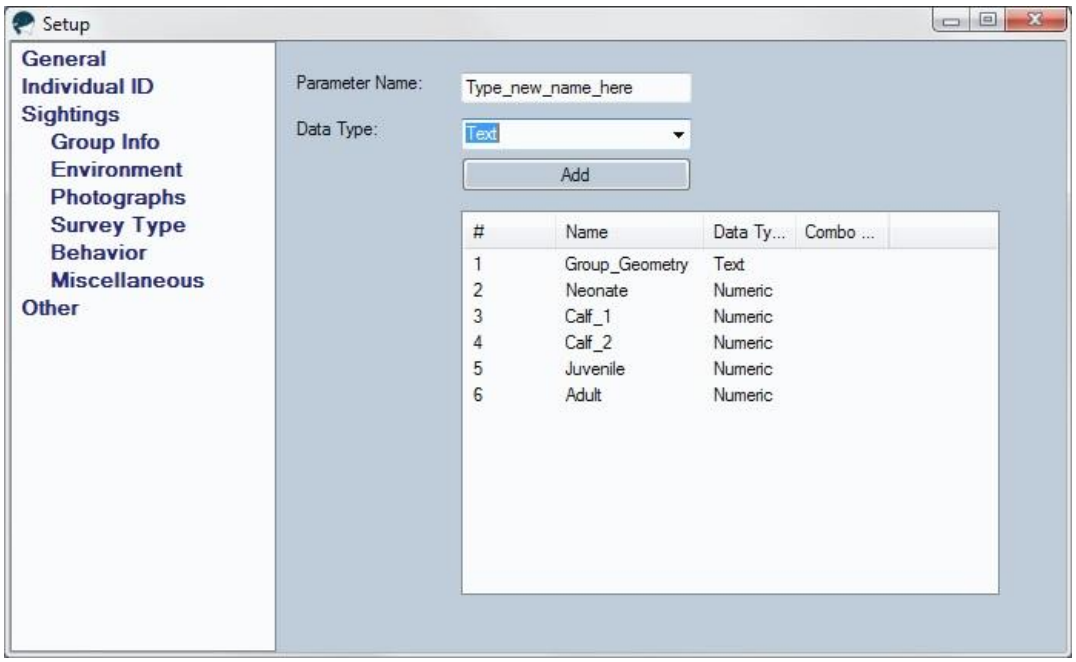


Figure 9. Adding new data types for sightings.

- *Environment* - Environmental parameters recorded at the time of sighting. Several data types can be assigned to each parameter (see Data Types below).

  To add, type the name of the new parameter in box of "Parameter Name", then choose "Data Type", and click "Add".

- *Photographs* - Information pertaining to the photographs taken during the group encounter.
    - Recorder - The name of the person recording the data.
    - Media - The type of device used to record the data (i.e. video or still camera, or type of cameras used).

- *Survey Type* - The categorical description of the type of survey being conducted for the group encounter (i.e. Photo-ID, Line-Transect, etc.).

- *Behavior* - Textural description of the behavioral state of the group encounter.

- *Miscellaneous* - A dynamic setup that allows users to create additional tables/fields that may not be included above. For example, a field "Genetics" could be created, with parameters such as "Biopsy Taken" with defined data type of Yes/No and "Reaction" with defined data type of combo with "Strong", "Moderate", "Weak", and "None" as categories of the combo.

  To add a new field, fill in a new "Type" or choose a "Type" in the drop down menu, then type the name of the new parameter in box of "Parameter Name", then choose "Data Type", and click "Add".

Note: To add categories of a combo, right click on the parameter and click "Category".

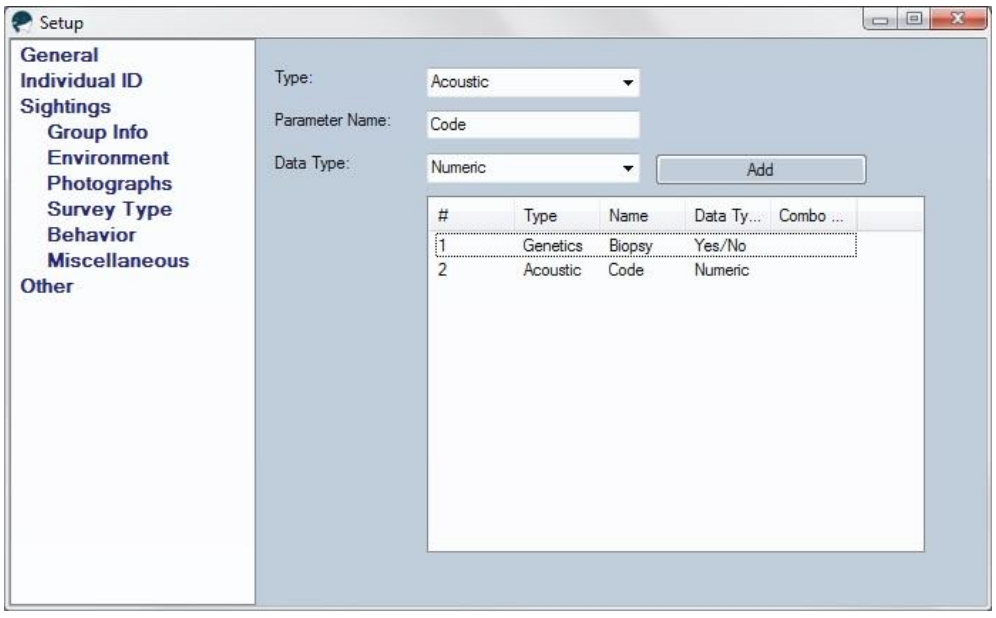


Figure 10. Adding miscellaneous sightings parameters (user-defined).

Other - A dynamic setup that allows users to create additional tables and fields which may not be provided or appropriate to include with the photo-IDand/or sighting tables. For example, stranding events can be recorded by creating type "Stranding" with parameters "StrandingDate" with defined data type of Date/Time, "ID" with defined data type of Text, and "StrandingCodes" with defined data type of Combo with "Code1", "Code2", "Code3" and "Code4" as categories of the combo.

Fill in a new "Type" to create a new table or choose existing "Type" in the drop down menu, type the name of the new parameter in box of "Parameter Name", then choose "Data Type", and click "Add".

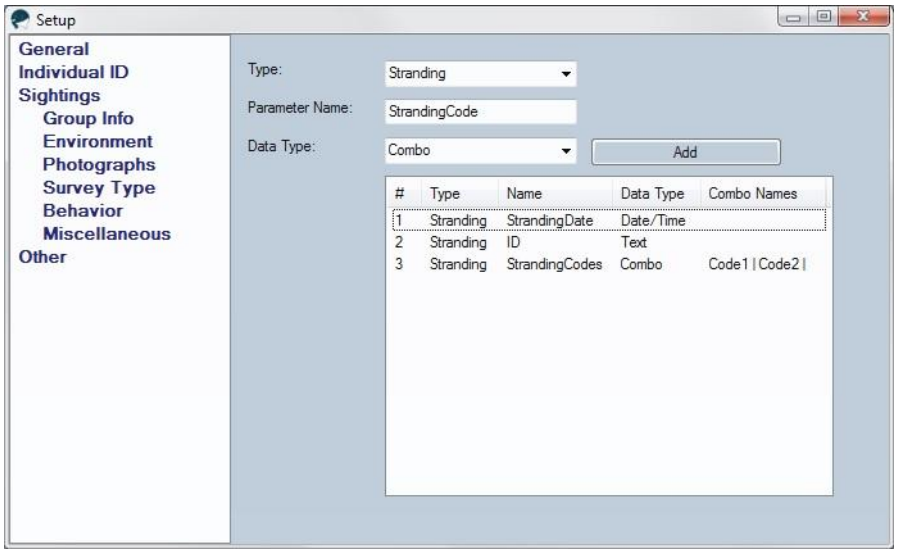


Figure 11. A representation illustrating adding user-specified "Other" parameters to the database.

Note: When all fields of a particular "Type" in "Miscellaneous" or "Other" are deleted (for example, in Figure 12, all fields of the miscellaneous type "Genetics" are deleted), users will be ask to delete the table from the system database. This will not delete any records or tables from the data database but rather remove the parameters from the system information.

Figure 12. Message of deleting table from system database

Defining Directory/Filename Structure

- The directory/filename structure routines are designed to allow users to define their parameters to maintain filename formats and directories for images in the incoming and catalog folders.

  The directory structure can be found in "Setup" > "Individual ID" > "Catalog". Click on the button "Directory Structure" to open.

  The filename structure can be found in "Setup" > "Individual ID" > "Rename Image". Click on the button "Change" to open.

- **Variables**
  - For the catalog structure, five variables are offered:
    1. ID - individual identification designation
    2. STUDY_SITE - the defined name of the study area
    3. GROUP - the designated group identification name/number
    4. DATE - the defined date of the image
    5. SPECIES - the current species defined for the image.
  - For the incoming and filename, additional non-standard variables are offered. These additional variables consist of the user-defined descriptions specified in setup under Rename Images > Parameters and Rename Images > Extract (see Extract).
  - The user can add a variable by selecting the number of variables desired and clicking the "Add" button. Right click on the directory allows users to delete previously defined variables.
  - Standard variables are enclosed with brackets (i.e. [SPECIES]). Incoming parameters are defined with brackets and braces (i.e. [{PHOTOGRAPHER}]). Information extracted from the original filename is defined by brackets and a vertical bar (i.e. [|FrameNumber|]).
  - In addition to the various parameters offered, the user can also add standard text information. Textual information should be added without parentheses.

- To include multiple parameters in the directory/filename, press and hold Ctrl button while selecting the parameters in the user-defined sequence (e.g. to create directory in Figure 14, hold Ctrl button and click "Group", "{Card}" and "|Test|" in the exact sequence).
- Some Examples:
  - [STUDY_SITE] + _ + [SPECIES] + _ + [ID]
  - [ID] + _ + [DATE] + [STUDY_SITE] + [{PHOTOGRAPHER}] + [GROUP_ID] + [|FrameNumber|]
  - [ID] + MyStudyName + Spinner + [DATE] + [GROUP]

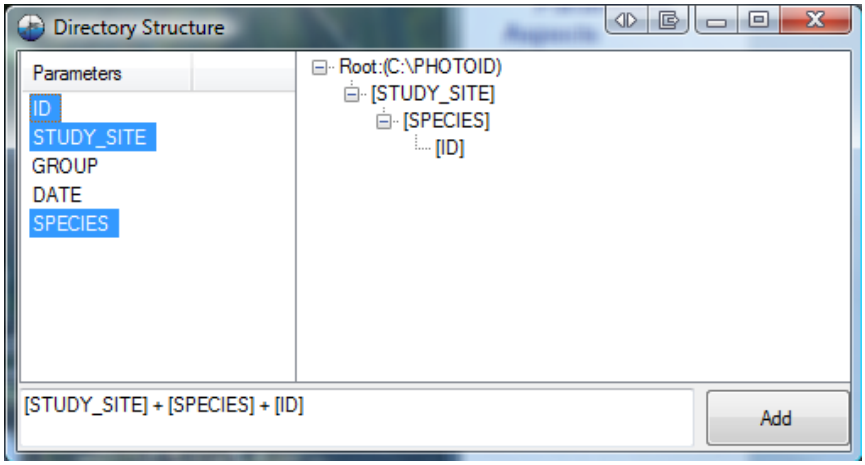


Figure 13. Example of defining directory structure for the catalog directory.

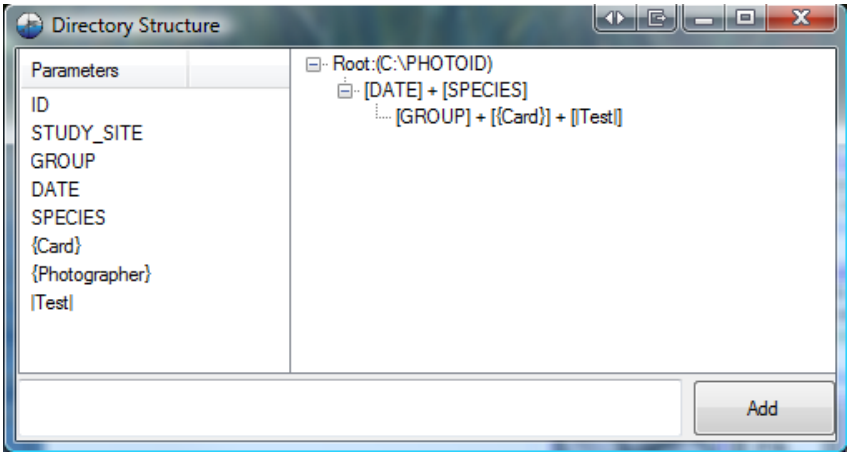


Figure 14. Example of defining directory structure for the incoming directory.

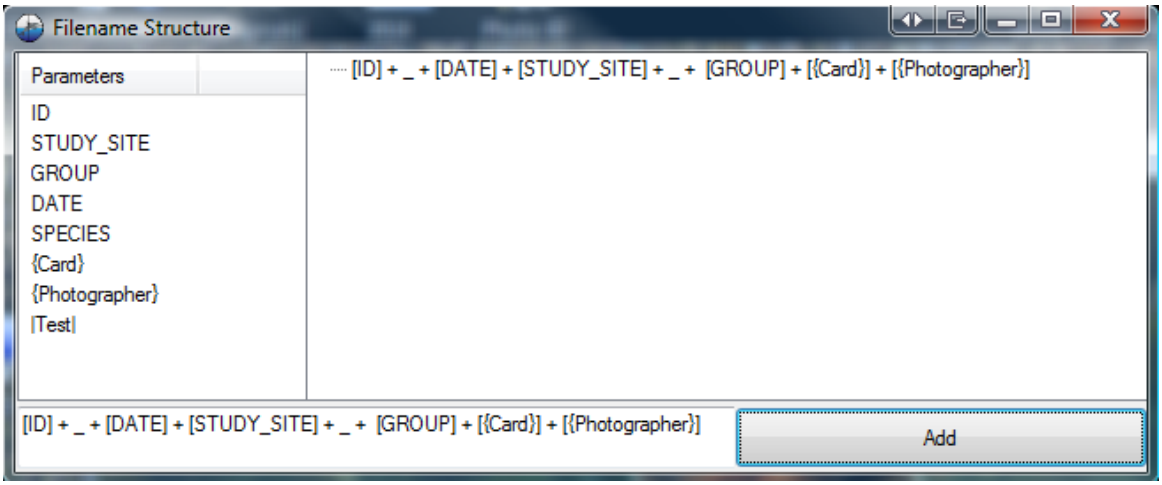


Figure 15. Example of defining filename format for renaming images.

Data Types

Table 1 illustrates different data types that the user can define for their parameters. These consist of textual, numeric, boolean (yes/no, true/false), date/time and textual combo parameters.

Table 1. Data types available in Discovery.

Type	Description
Numeric	Only numerical values are used for the definition of the parameter. Paramaters such as # adults, sea state, salinity, etc. can be defined as numerical types. A parameter defined as numerical will ONLY allow numerical values to be recorded upon data entry.
Text	Textual description of the parameter, such as wind direction.
True/False	The parameters contain only true or false responses.
Yes/No	Parameters that have only a yes/no definition.
Date/Time	Parameters that only have date and/or time values.
Combo	Categorical definitions of a parameter. For example, a parameter for habitat with various habitat types as standard categories. Classification of combo will result with a display of a pull-down box of the categories further defined by the user to allow for quick entry of data.

Preferences

In Preferences users can customize the interface to a suitable working style.

Click on setup and then preferences to open the preferences menu.

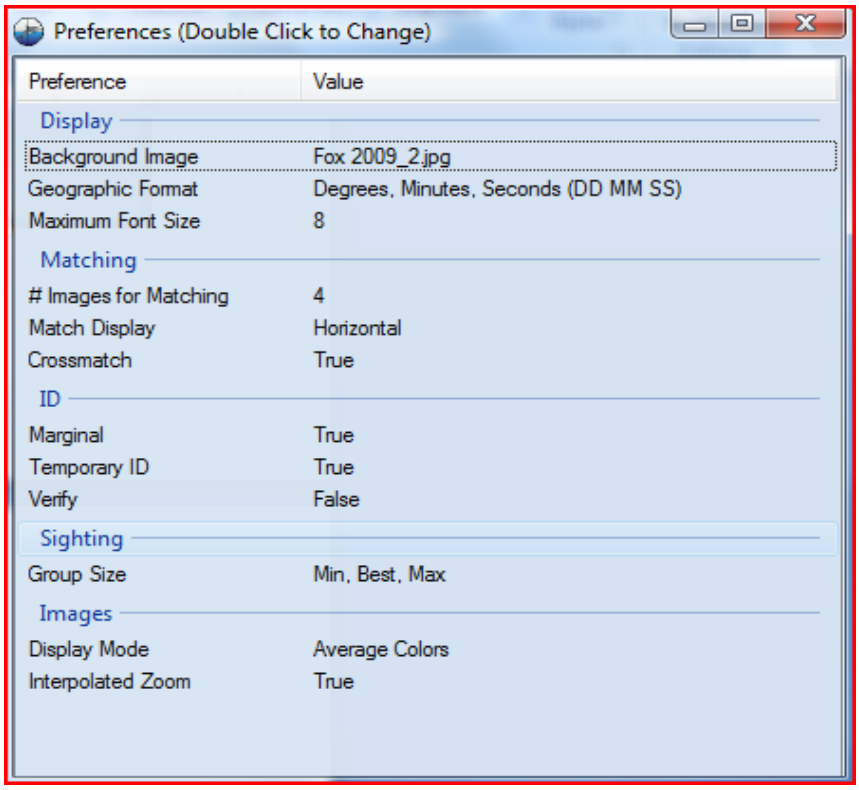


Figure 16. An illustration of the preferences screen. Preference items that can be changed are located in the left column and the user-defined choices in the right column.

Preferences allow the user to select various display and options associated with the Discovery software system. To change an item, double click on the desired preference in the left column to bring up the available options. Once changed, the new option associated with the item will be shown in the column to the right. Each preference is described below:

Display

- *Background Image* - Users can select an image to display in the main menu as the background image. This image is used for aesthetic purposes only and has no effect on the program itself.
- *Geographic Format* - Users can specify the format in which the coordinates of geographic positions are to be displayed:

Table 2. Format types for geographic (Latitude/Longitude) positions.

Format	Syntax	Example
Degrees, Minutes, Seconds	DD MM SS	52° 42' 22"
Decimal Minutes	DD MM.SSS	52° 42.444'
Decimal Degrees	DD.MMMMM	52.33333°

To shift to other format, double click "Geographic Format"

- *Maximum Font Size* - Maximum font size that the program will use during resizing to display on the screen. Smaller fonts allow more information to be displayed while larger fonts may be necessary if working on a small screen. Be aware that if the font size becomes too large, information may be lost from the trailing end of the text when the form resizes.

Matching

- *Number of Images for Matching* - The default number of cataloged individuals that are displayed for matching purposes (see Matching in Chapter 4).
- *Match Display* - Cataloged images can be displayed in either horizontal or vertical modes. The Horizontal mode illustrates the cataloged individuals/images on the right hand side of the screen with the query individual on the left hand side of the screen. The Vertical display mode shows the query image on top of the screen with the database of cataloged individuals that were previously identified on the bottom portion of the screen (see Matching in Chapter 4 for further details).

- *CrossMatch* - Multiple catalogs may have different identification designations. The crossmatch routine allows the database to link one catalog to another by assigning individual IDs from one catalog to the designated IDs of another catalog. Choose true if you would like to use this option or false if you prefer not to use crossmatch.

ID

- *Marginal* - Include the option to specify that the quality of the image or quality of the match is of marginal value and therefore should be separated from the main database of information.
- *Temporary ID* - Classification of a new individual with uncertainty. These identifications may need to be revisited at a later stage when more information is available to confirm the identification of the individual, or to be verified by another person (e.g. to have a consensus with an ID-dataset manager).
- *Verify* - A processing technique that allows the data manager to verify matches of new individuals that are being entered into the database. This is useful to ensure accurate matches. Ideally, the data manager should be the most experienced person with matching and will verify/check matches or new individuals that are added to the catalog of individuals.

Sightings

- *Group Size* - Estimated group size, either as one estimate or by minimum-best-maximum group size estimation approach.

Images

- *Display Mode* - Images are reduced in size to fit into the visual displays. This reduction can be done in various ways. Several options are provided for the user to choose their preferred display:

    - EliminatePixels – the display color for a region matches one of the pixels in the region and the remaining pixels are ignored. This is a very fast technique that gives satisfactory results with color photographs, but may eliminate necessary detail with some monochrome images.

    - FavorBlack – the display color for a region will match the pixel with the lowest intensity (brightness) in the region. This technique is primarily designed for black on white monochrome images.

    - FavorWhite – the display color for a region will match the pixel with the highest intensity (brightness) in the region. This technique is primarily designed for white on black monochrome images.

    - AverageColors – the display color for a region is the average color for the pixels in the region. This technique is slow compared to *EliminatePixels* but results in the best color reproduction.

- *Interpolated Zoom* - results in each display pixel being calculated from a linear interpolation of the surrounding pixels in the source image. This technique results in a refinement of the detail seen in many magnified images, but is relatively slow. Setting Interpolated Zoom to False results in a much faster display of a zoomed image, but blockiness or pixilation will occur with high magnification.

Once you have chosen your preferences, click the close button, and your changes are automatically stored.

Chapter

3

Incoming Processing

When researchers return from the field, they typically download their images to a specific folder which in Discovery is termed the "Incoming" folder. Some researchers may create subfolders in this directory with the date of the survey, and subsequent subfolders containing the card number or group number. Discovery offers several tools that can help researchers in processing and filtering images, and can reduce the number of images that are needed for matching to an existing catalog. These tools are Renaming, Filtering, and Incoming Matching. They are designed to assist researchers in processing their images BEFORE they begin matching them to a database. Each of these tools works independently from the database and from each other.

Renaming

Renaming routines are designed to provide meaningful filenames that contain the individual ID, date of survey, group ID, study area, and any other user-specified parameters (see Individual ID in Parameter Setup). To access the renaming tool, select Incoming from the Main Menu and then open the Filename menu. Then, select the images/frames of a specific group. The application will automatically extract the date from the EXIF data if the image contains such information. The user then specifies the group ID and other user-defined parameters. The default start number of temporary ID is "0001", which can be changed to user-defined numbers by double-clicking the word "TmpID". Auto-extraction routines can be setup to extract information that is standard in a filename. For example, Nikon cameras filenames are usually DSC_0001, with the number specifying the frame number taken. Users can automatically extract such information from the filename by setting the option in the setup (see Individual ID in Parameter Setup). Once the frames for a specific group are selected, click the Rename Images button; filenames are then renamed with a temporary ID, date, study area, group number, and any other user specified parameters and auto-extraction options.

Example: Frames 1-42 were taken on Card 02 of group 01 on 9 Aug 2009 for our study area "WGW". We setup the program to also include the Card # and auto-extract the frame number from standardized Nikon filename. We select these frames and specify the group ID and card # and click rename images. Those files are then renamed as something like "0001_20090809WGW0102_0005.tif". The "0001_" is the temporary ID, 20090809 is the date, "WGW" is the study area, "01" is the group, "02" is the user specified card #, and "_0005" is the frame number that was automatically extracted from the standardized Nikon filename format from the user-specified parameters.

Once images are renamed with Discovery, the study area, individual ID, and group number are embedded into the EXIF part of the image. EXIF functions to manipulate the date, time, and GPS position of the images are also included. Currently, GPS exchange format (GPX) can be processed and GPS track data, waypoints, or route information can be used to find the GPS position that is closest to the image date/time for processing. If the user uses these routines, the closest GPS position can be embedded in the EXIF data. *Note that changes done to the original photos in this step are irreversible; although they can be changed again, once completed they cannot be undone (!)* so keep a separate backup of all the original images and folders

Note: EXIF data can be changed by either using this renaming function again or using the function in "File"> "Update"> "Exif Based on DB".

When images are renamed, they are moved to a new folder located in the same sub-directory as the previous folder. The new folder is labeled based on the user-defined incoming data structure. The default incoming data structure is Date + Group ID (i.e. Incoming folder\[DATE]\[GROUP ID]). See Incoming Directory Structure in Chapter 1 to change the data structure.

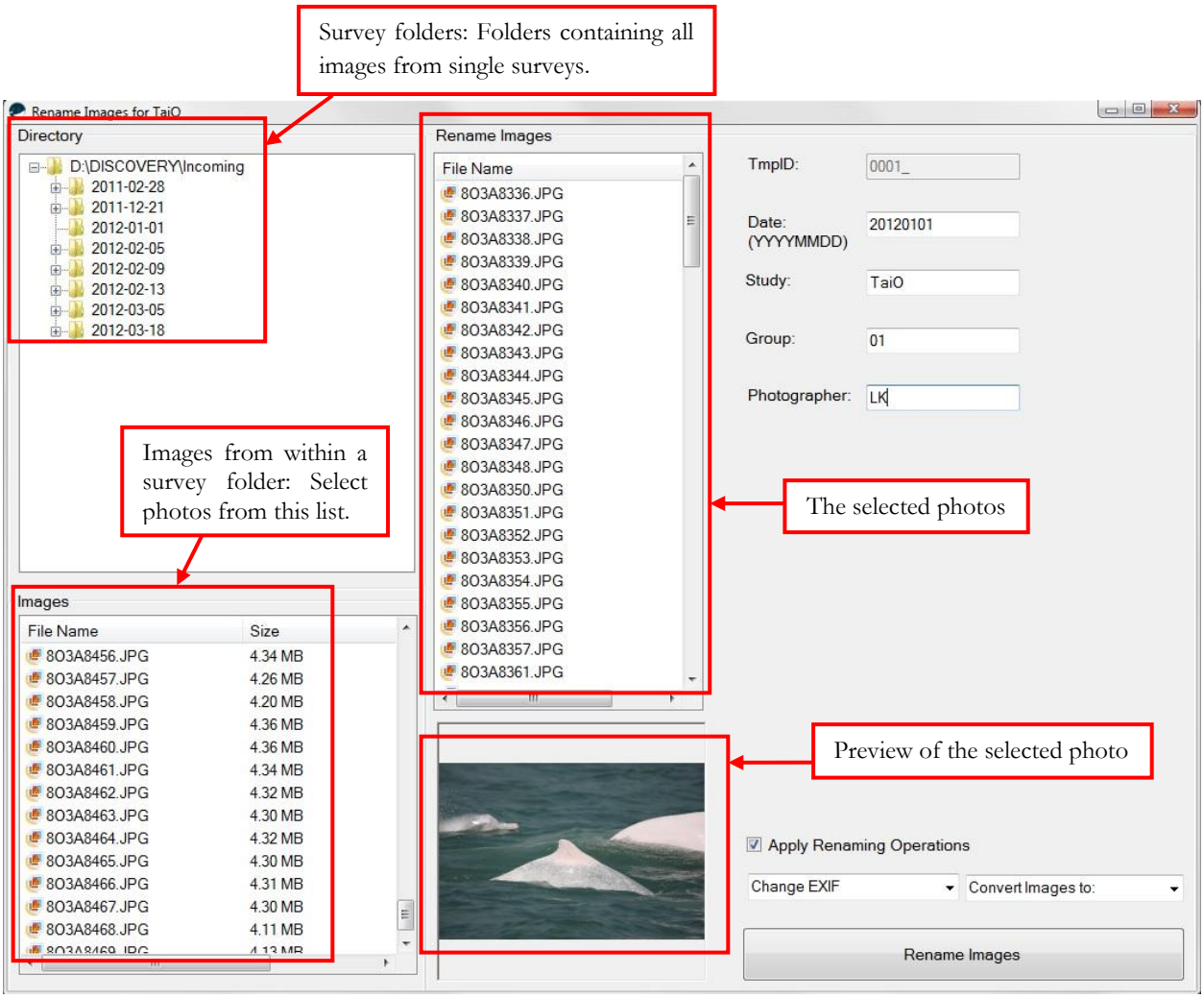


Figure 17. Illustration of renaming images with Discovery.

Quick Procedures:

1. In the Main Menu, click on "Incoming" and then "Filename" to open the menu to rename the images.
2. In top left box, click on the '+' sign to view subfolders and choose the folder that contain images you want to rename.
3. Select one or more images in bottom left box for renaming.
4. Selected images will be shown in the middle box. Click on the image name for preview.
5. Fill in the information of parameters on the right hand side (Date is extracted automatically from EXIF, however users can modified the date, if needed; temporary ID, which by default starts from "0001", can also be changed by double-clicking the word "TmpID").
6. Select "change EXIF" or "convert image to…" on the bottom right, if needed to batch modifies EXIF data or convert images to different formats.
7. Tick the box 'Apply renaming operations' if auto-extraction is to be completed.
8. Press "Rename Images" and wait for the green progress bar to finish.
9. In this process, the software will rename the images, add new information in the EXIF files of the images, and move the image based on the user-defined incoming directory structure.

Note: This function also keeps part of the original filenames according to the setting by the users in "Setup"> "Define"> "Individual ID" > "Rename Images" > "Extract"

Note: Parameters on the right can be set in "Setup"> "Define"> "Individual ID" > "Rename Images" > "Parameters"

Note: After images are renamed, a new folder is created and labeled with the date of the survey. It will be located in the same sub-directory as the previous folder. All renamed images will be automatically transferred into this new folder.

Filtering

A filtering application was developed for the user to manipulate and further process images to reduce the time needed for matching. The user usually works in a subdirectory in the Incoming folder (i.e. Incoming folder\20090809\02"). The images usually contain all the photographs related to a particular group of individuals on a particular day.

The filtering application is designed to allow users to manipulate images with contrast, brightness, sharpening routines, and cropping images to a specific size that eliminates other non-useful information (i.e. surrounding environment). There are other functions such as auto-contrast, auto-level, mirror, rotation, etc. that can be used to further manipulate images. Users can also transfer images to user-specified subfolders (see in Parameter Setup in Chapter 2) which can help further minimizing the processing time at a later stage. For example, "water shots" or "blank" images can be placed in user specified folders called "No Info" and "Blanks", respectively. Other subfolders can be created if the user wants to separate images with no identifiable information but useful for other purposes (such as calculation of the proportion of identifiable individuals); for example subfolder "Clean Fins" for images of dolphin dorsal fins that do not contain identifiable features.

Upon startup of the filtering application, the system will open up the *Incoming Directory* that was specified in the Setup (see Incoming Directory in Parameter Setup in Chapter 2). The user can select various folders or images that are displayed. Each folder has a number in parentheses that indicates the number of images in that directory (see Image Formats in Chapter 10 to understand which formats Discovery supports). The image manipulation tools are as follows:

Adjust

A. *AutoContrast* - Optimizes the contrast of an image by analyzing intensity coverage in the analysis area of an image and then readjusting it so that the intensity distribution range is as wide as possible. The analysis area can be the entire image or the current selected area. The operation adjusts the brightest pixels to be white, and the darkest pixels to be black, with the intensity of the intermediate pixels raised or lowered accordingly.

B. *Autolevel* - Balances the colors in an image by analyzing the intensity coverage of each color component (red, green, and blue) in the analysis area of an image and then readjusting the overall intensity so that each color's distribution range in that region is the same. The analysis area is of the entire image or the current selection rectangle. The operation adjusts the brightest pixels to be white, and the darkest pixels to be black.

C. *Mirror* - Performs mirroring operations (flipping the image).

D. *Rotate* - Rotates the image by either 90 degrees clockwise, 90 degrees counter clockwise, or by a user-defined angle. Angles can be positive (clockwise) or negative (counter clockwise), but must be between 0 and 360.

E. *Contrast* - Performs contrast adjustment on the current image. Values range from –100 to 100 reflecting the percent change to be made to the contrast of the image.

F. *Brightness* - Alters the intensity of the image. Values range from –100 to 100 reflecting the percent change to be made to the brightness of each pixel.

G. *Sharpen* - Performs sharpening operations on the image using convolution kernel technique. The values range from 1 to 100, which define the amount of sharpness increase to be applied.

Mouse Operation Tools

A. *Pan* - The primary mouse button is used for panning the current image. Note that the middle mouse button is always used for panning.

B. *Area Select* - The primary mouse button is used for drawing a selection rectangle on the image.

C. *Zooming* - The primary mouse button is used for zooming. The middle mouse scroll button can be used to zoom as well.

D. *Crop* - The crop application removes all image information outside of the boundaries of the current selection rectangle. For example, if the selection rectangle is positioned at pixel coordinates "0, 0, 15, 15", after crop the image will be 16 pixels wide and 16 pixels high.

Note: Users can right click on the toolbar for other options such as undo, redo, sign (watermark the image), copy image, etc.

Subfolders

User-specified subfolders (see Parameter Setup in Chapter 2) are created on the toolbar at the top of the image. If you want to move an image from its current directory to a subfolder within that directory named as the button (or subfolder selected by the user) you can just select the subfolder button located along the top bar of the window. By using this subfolder method, files can be deleted from the source folder and copied to the target subfolder, so that there are no duplicates, which makes it easier to track what images have been processed.

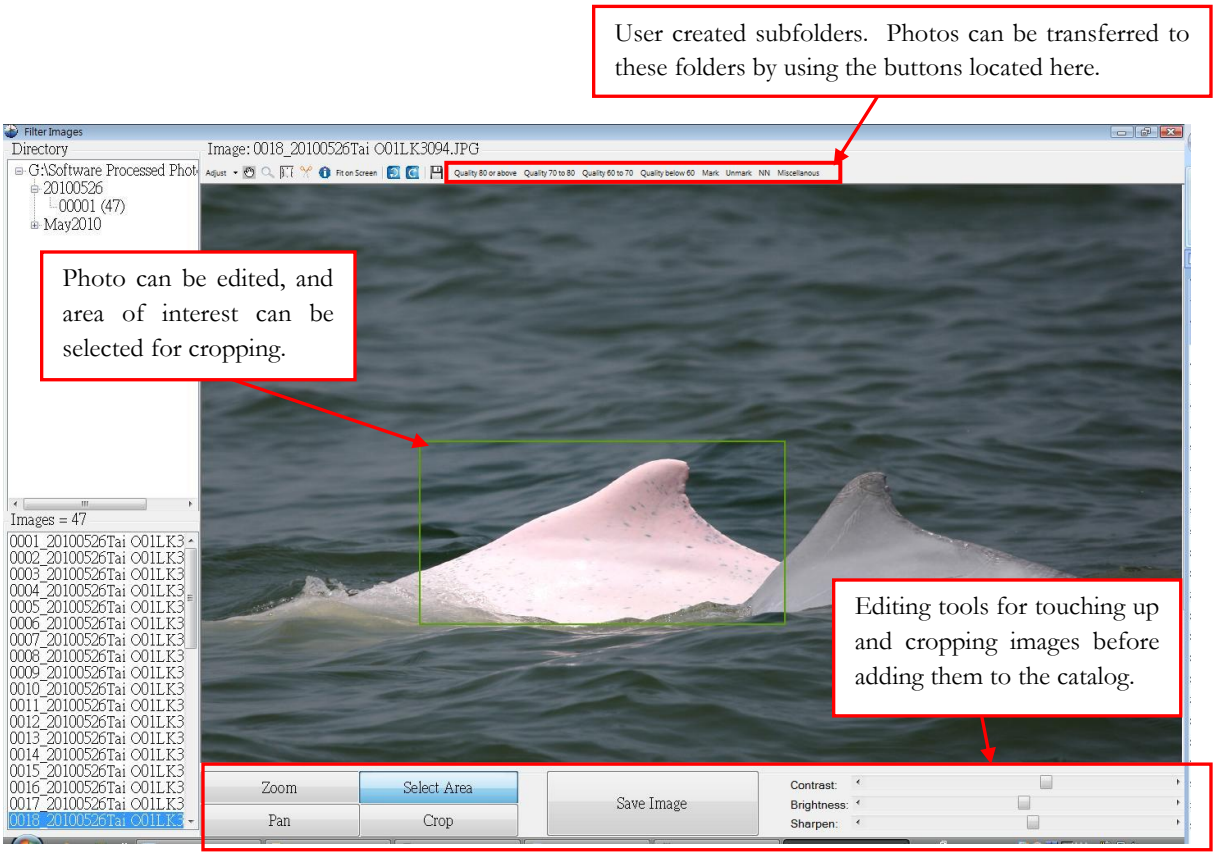


Figure 18. Filtering image form with a photograph of an Indo-Pacific humpback dolphin mother-calf pair.

Deleting/Renaming Images

Images can be deleted or renamed in the image list. To delete, the user simply selects the image and press the delete key on the keyboard. To rename, double click on the image and change the filename (similar to the Windows Explorer process).

Drawing Tools

Drawing tools are available on the top bar of the filtering page above the image.

Quick Procedures:

1. In the Main Menu, click on "Incoming" and then "Filter" to open the filtering menu.
2. Select appropriate folder in top left box.
3. Select an image for process in the bottom left box.
4. Use the image editing tools at the top left above the image (for quick tools); bottom left below the image (for manipulating the image); and bottom right below the image (for fine adjustments).
5. Save the adjusted image by clicking the "Save Image" button at the middle bottom below the image. To replace the original image, save with the same name and click "Yes" when prompted. To save as a new image (or save as a duplicate), give the adjusted image a new name.
6. Select an appropriate button with a folder name (at the top above the image) to transfer the image. Click on the button and the image will be transferred into the selected folder.

Note: You can add/delete/edit the folder buttons in "Setup"> "Define"> "Individual ID" > "Incoming" > "Subfolders".

Incoming Matching

Once users have filtered the images, they should have a folder containing all the identifiable individuals in a group. An incoming matching application was developed to assist researchers with finding the highest quality image of each individual in a group in order to process this identification to the existing catalog.

Two dynamic multi-image windows can be setup containing the "Query" and "Potential Match" images within a group. Users flip through the images and can transfer images of lower quality or less desirable images into a user-specified subfolder. Users can move the image by right-clicking on the image in the List of Images and selecting the subfolder to which they would like to move the image. Multiple query and/or Images can be selected by holding down the CTRL key and selecting all the images the user wants to view.

The logic behind this processing approach is to end up with the highest quality image of a particular individual that was identified in a group in order to match this image to the existing catalog. This is an optional process and users are not bound by any processing approach as this system attempts to be as dynamic as possible to accommodate alternative processing strategies.

Quick procedures:

1. Select a folder at the top left box.

2. Select an image of a certain individual in the "Query Images" box.

3. Select one or more images of what appear to be the same individual for comparison in the "Potential match" box at bottom left.

4. Right click on the images with lower quality and in the right-click menu either

   a. Choose an appropriate folder (setup in Incoming Subfolders; see Subfolders in Parameter Setup) where the selected image(s) are to be transferred.

   b. Choose "Change TempID" and then input a new ID to change the temporary ID in both filename and EXIF, so as to differentiate images of the same individuals from the others in the same folder.

   c. Choose "TempID folder">"New" to create a new temporary ID folder and move the selected image(s) to the folder. Subsequent temporary ID (starting from TempID0001, TempID0002……) will be assigned automatically. Choose the existing temporary ID folders if it is a match.

   d. Choose "Move To">"Create" and input folder name to create new folders (not necessarily a temporary ID folder) and move the selected image(s) to the folder. Choose existing folders if it is a match

5. Repeat steps 3 and 4 until finish.

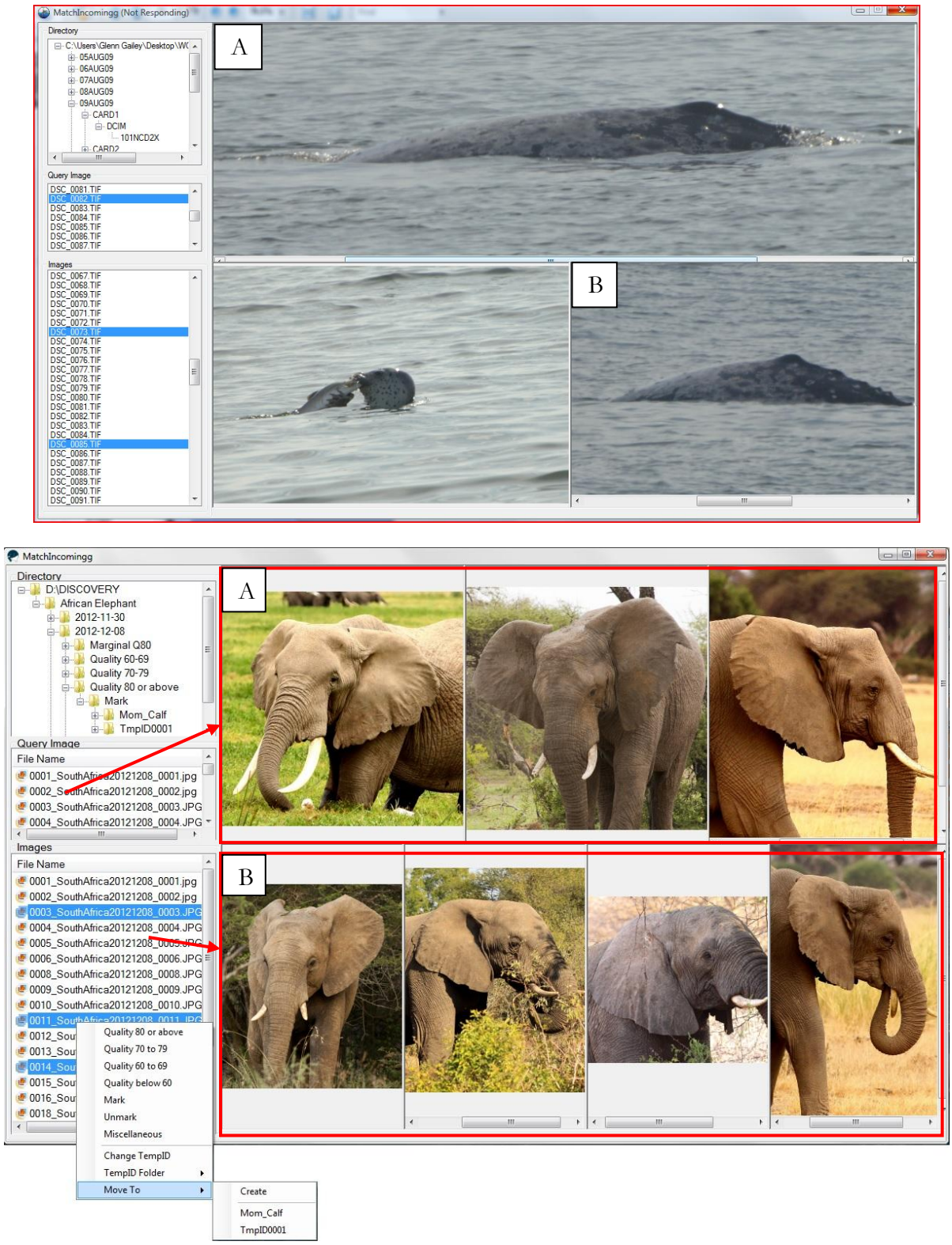


Figure 19. Two illustrations of incoming matching procedures for: 1. left body side of a gray whale and 2. African elephant ears. Box A represents the query image while Box B represents images selected for matching.

Thumb View

Thumb View is a tool to quickly flip through your images in the Incoming directory. Multiple images can be selected by holding down the Ctrl key and select all the images you want to view. You can right click on the thumb image to delete the image if you wish.

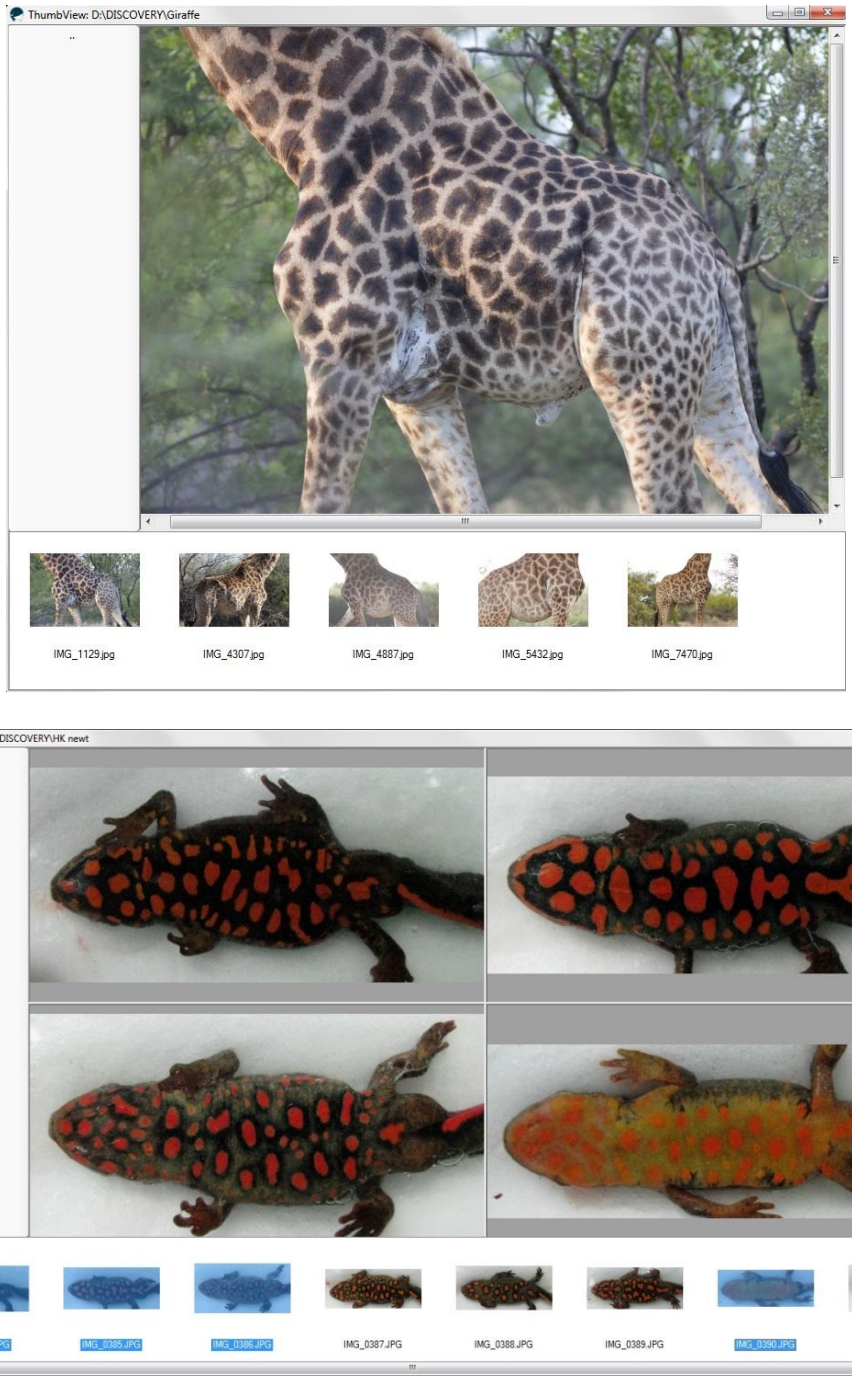


Figure 20. Two thumb views illustrating identifiable features in giraffe and Hong Kong newt images.

Chapter

4

Matching

Once the incoming images are processed (see Chapter 3), the individuals are ready to be matched to the existing catalog of verified individuals (see more in Verifying Individuals in Chapter 6). Users could choose to include the non-verified individuals by choosing "False" for "Verify" in Preference (see Verify in Preferences). The user loads the query image in the matching application and compares this image to either the "Type Specimen" or "All images" in the database. The term "Type Specimen" refers to the most representative image of an individual. The most representative image typically is the highest quality most recent image of that individual. "Type Specimens" can be defined in View Individuals (see View Individuals in Chapter 7) by clicking on the individual and selecting the image and selecting that image as a type specimen. *Only one type specimen may be defined per aspect (i.e. left/right/fluke) of an individual.*

Match Displays

Match displays can be either Horizontal or Vertical, dependent on the user selected preference (see Chapter 1).

1. *Horizontal* display shows the query image on the left hand of the screen and the catalog images (tiled) on the right-hand screen.
2. *Vertical* display shows the query image on the top and the catalog images on the bottom.

For small cetaceans, such as dolphins, the horizontal display may be more functional for viewing dorsal fins. For larger animals, such as for instance gray whales where the entire side of the body is used for identification purposes, the vertical display maybe preferable.

The user can specify the number of cataloged images that will be initially displayed at one time. The default number of images displayed can be specified in the "Preferences" (see Preferences

in Chapter 2). The number of displayed images can be modified at any time within the matching programs.

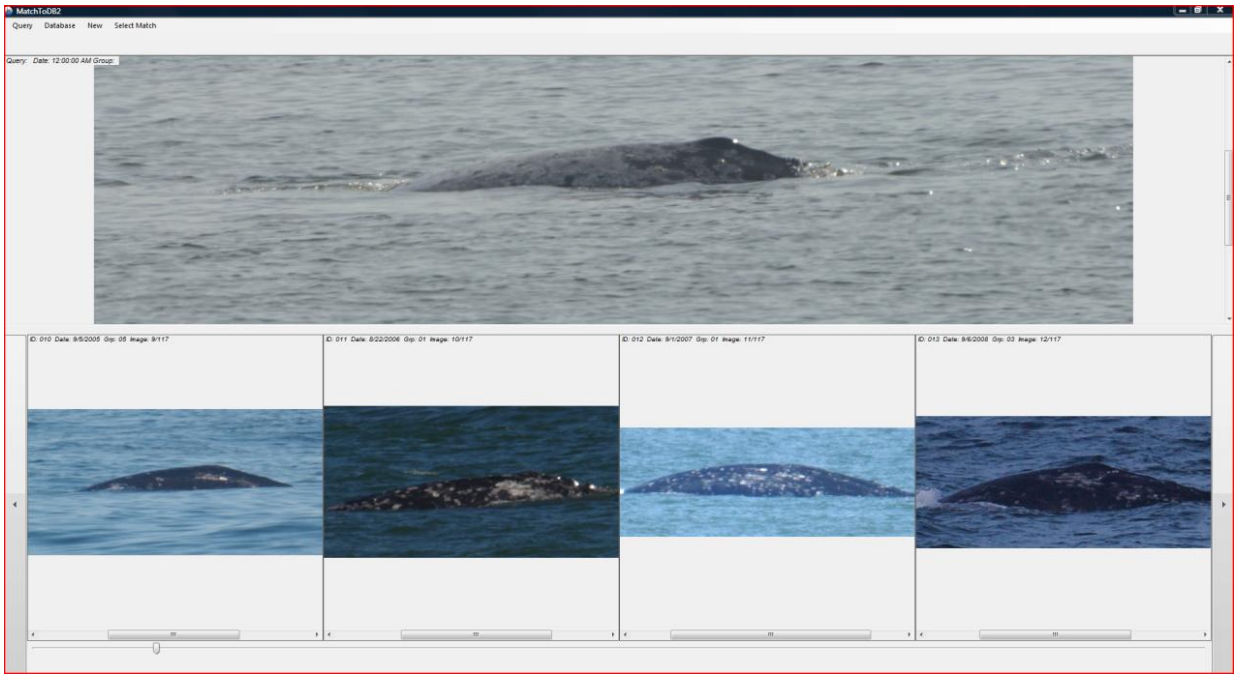


Figure 21. Vertical matching of gray whales with a query image at the top and four database images at the bottom of the screen.

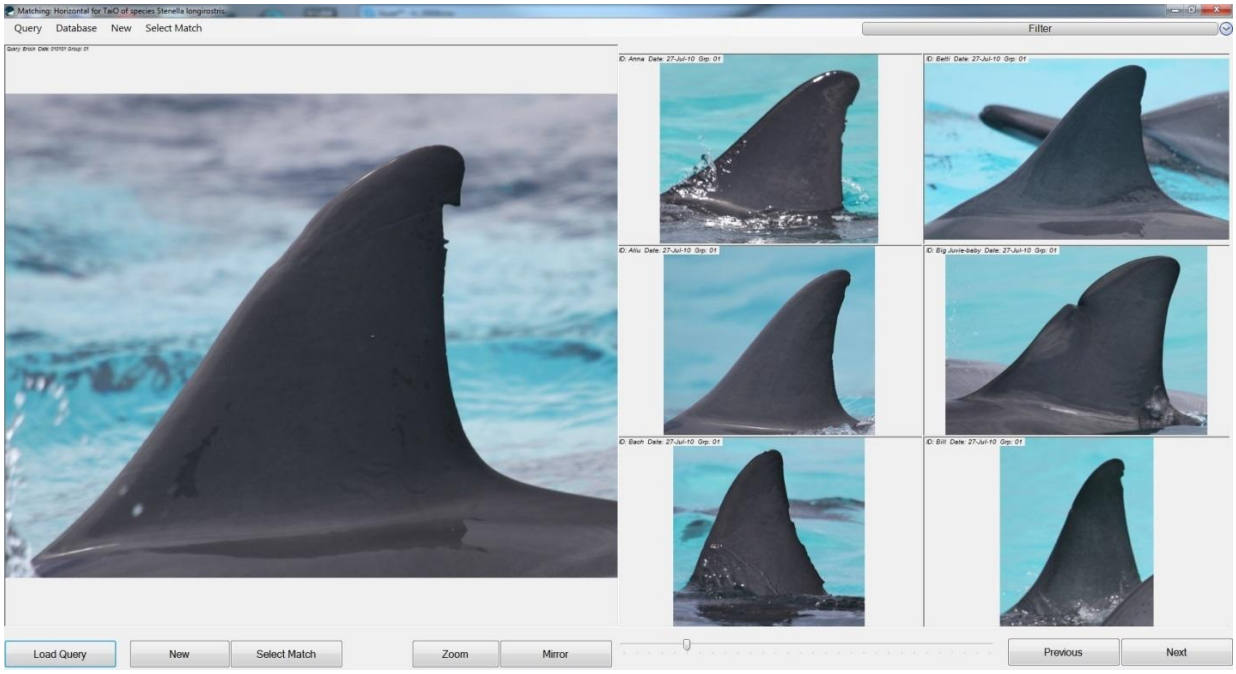


Figure 22. Horizontal matching of Hawaiian spinner dolphins with a query image on the left and database images on the right.

Filtering Routines

As the number of individuals increases in a database, the time to find a potential match to a query image correspondingly increases. There are a number of classification and processing schemes that can be employed to minimize the number of individuals to find a match; however, to find a new individual that is not currently in the catalog, users will need to evaluate all potential matches first before classifying an image as a new individual. Discovery offers several filtering tools to assist users to find a potential match. These are:

1. *Aspects* - Filter the type specimens or images by the user-defined aspect categorizations. For example, if "Fluke" was an aspect, the user can filter the database to show only the fluke aspects of individuals.

2. *Categories* - This is intended to be used as the primary classification scheme. For example, for small cetaceans, categories may be defined as "1 notch", "2 notches", "3 notches", etc. Subcategories can also be defined within a category for further filtering.

3. *Species* - If there is a need to inspect images of multiple species in classification, then this option can be selected.

4. *Study Area* - A user can select which study area(s) will be used to find a potential match.

5. *Descriptors* - These are secondary identifiable features that may be used to describe an individual; for example a scar, tooth rakes, pigmentation, etc.

6. *Type Specimen or All Images* - This allows the filtering process or list of potential matches to either show just the type specimen or all images of all individuals in the database.

7. *Empty Values* - If it is uncertain what is to be done with parameters that are not defined, these empty values can either be included or excluded from the list of potential matches. Below is a list of values that are considered as well as their default parameters.

    A. *Exclude Study Area* - If a study area is not defined for an image, then that image is by default not included in the list of potential matches. Users can select this option to include empty values if desired.

B. *Exclude Species* - If species is not defined for an image, then that image is by default not included in the list of potential matches. Users can select this option to include empty values if desired.

C. *Include Aspect* - If aspect is not specified for an image, then that image is included by default in the list of potential matches (regardless of the filtering method).

D. *Include Categories* - If a category is not specified for an image, then that image is included by default in the list of potential matches regardless of the filtering method.

8. *Distinctiveness & Image Quality* - Filter the type specimens or images by the level of distinctiveness and quality of the images. Users need to specify the operator, which must be either "<" (smaller than), ">" (larger than), "<>" (not equal to), "=" (equal to), ">=" (larger than or equal to), or "<=" (smaller than and equal to), before inputting the value of distinctiveness or image quality.

Note: For filtering, the relationship among parameters is and/or relationship, with the exception of descriptors which is only an *and* relationship. This means that if, for example, more than one sub-category and several descriptors are selected for a filtering procedure (e.g. one category, the Trailing Edge of a dorsal fin; two sub-categories, Upper Half and Middle; and three other descriptors), individuals selected by Discovery will be those that are listed in the database under at least one or more of the indicated sub-categories. For descriptors, however, all selected descriptors will have to be present to be included and displayed for matching purposes. Therefore, the relationship for descriptors filtering is *and* only, which is why in the earlier example the individuals selected for matching purposes must have all the three descriptors *and* at least one of the indicated sub-categories.

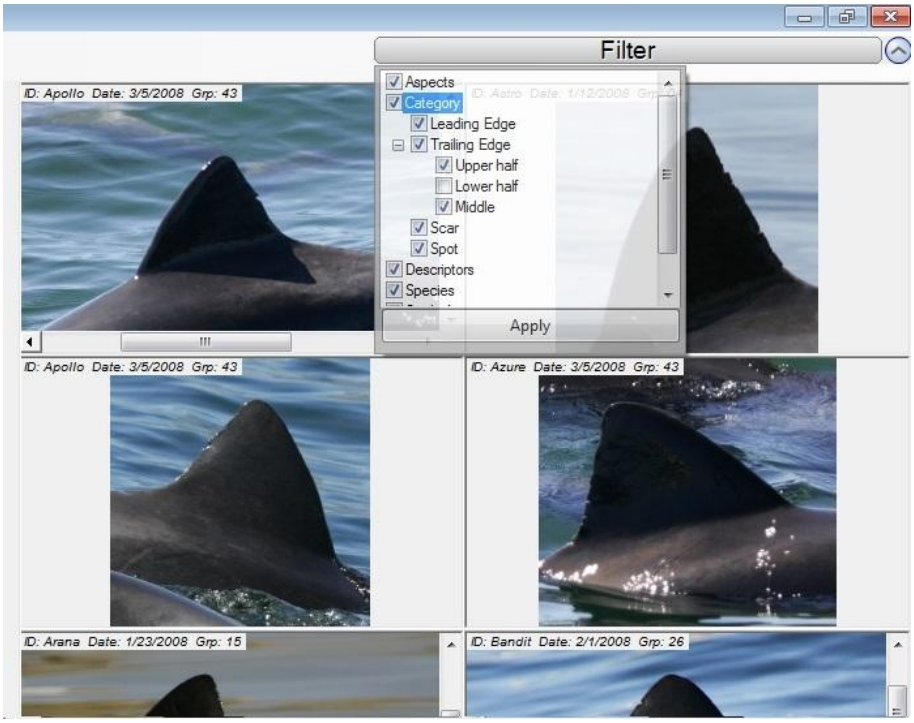


Figure 23. Filtering of a Benguela dolphin (Heaviside's dolphin) catalog. User can apply their filtering criteria by using the "Filter" button on top right hand corner of the window. Sub-categories within a category can be displayed by double-clicking on the category and they too can be used to further specify the criteria and narrow the search options.

Finding Individuals

There are two functions to find a specific individual or individual(s) with certain features that may be in the list of individuals in the database or filtered database. These two options are:

1. *Find Individual* - Find the individual in the list by its identification ID.

2. *Find By Feature* - Find individual(s) with a specific feature. For example, find all Females or all Individuals with calves (adult-calf pairs). Once a feature is specified and there are more than one individual in the database with that feature (if more than one, Discovery automatically goes to that individual), then the Find by Feature option will show all individuals with that feature. Another example is Find by Name. Note that in Discovery an individual-ID can be alphabetic, numeric or alpha-numeric. We suggest users to use numeric or alpha-numeric for individual-ID; but users can create a feature called "Name". The user can then search the database or filtered database by using either individual names (using Find by Feature) or individual IDs (using Find Individual).

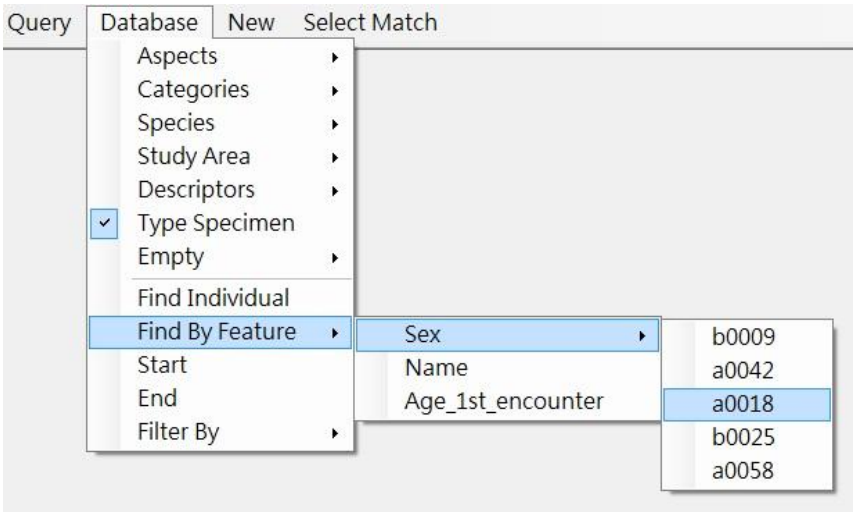


Figure 24. Illustration of finding individuals by a feature (e.g. "Sex").

Note: The function "Find" will find individual(s) with specific features from database, and display the associated catalog image(s) or list the ID(s) as shown above; while other individual(s) in catalog will also be displayed, though appearing in later sequence. On the other hand, the function "Filter" will find individual(s) with specific features and display ONLY the image(s) of the selected individuals.

Note: Various users have various preferences in naming the individually recognizable animals in their catalog; some use names (e.g. Bent-Fin, Mary, etc.) while others use numbers. In Discovery, for the purpose of data management, we recommend the use of numbers for individual-ID (so the ID is either numeric, e.g. 001, 002; or alpha-numeric, e.g. LK001, LK002) as it simplifies various data management procedures and export of data into other programs. However, users can still use names, which many may prefer as they are much easier to memorize and associate with a specific individual (e.g. Bent-Fin, Chopped-Fin, White-Spot, etc.), but in Discovery these could be entered as an individual feature (feature called "Name") and, subsequently, these features (just as individual-IDs) can be used for filtering the database and searching for specific individuals. With this approach, both numbers and names can be associated with an individual's ID.

New/Match Individual

Once a match or new individual has been determined, the user will need to add the image to the database. To classify the image as a new individual, simply press the New Individual button. To classify the image as a Match to another individual in the database, click on the Select Match button and then double click on the image to which you matched the query image. If only one image/individual is displayed as a potential match, Discovery will assume, upon pressing Select Match, that you want to match the query to this one image/individual. The query image is not added immediately to the database upon pressing Select Match or New individual. An add image form will be displayed to specify additional parameters related to the image. After the user is satisfied with the information, they can then choose to add the new data to the existing catalog of individuals.

Quick Procedures:

1. Before matching, ensure "Preferences" have been configured for your matching needs.
2. On the main menu, click "Database" > "Match" to open the matching menu.
3. On the right side of the window you will see the existing database (for horizontal matching, bottom window for vertical matching).
4. Click "Load Query" button and select an image for matching.
5. The query image will be shown on the left/top. Click "Zoom" or "Mirror" to adjust the image.
6. Filter the existing database by using the "Filter" button on top right corner. Select criteria for filtering by checking appropriate boxes and click "Apply"
7. Check through the database by clicking "Previous" or "Next" at bottom right. User can also adjust the size of images in database by adjusting the resolution bar which is located to the left of "Previous" button.
8. Determining a New/Match:
    a. Match - If the query individual is matched, click "Select Match" and then click on the image of the matched individual in the database. Then fill in the appropriate information in the pop-up record sheet window and click "Add to Database" at the bottom right of the window.

b. New - If no match is found, click "New" and then fill in the appropriate information in the pop-up record sheet window; give a new ID (and new Name, if you want) for the new individual. Then click "Add to Database" at the bottom right of the window.

9. New individuals will be shown in the database immediately, the database is sorted by individual-IDs.

10. Start from step one until all new images and individuals are added to the database.

Note: User can double click on the images in database to view different images of the same individual. By clicking the "Type specimen" box of a selected image of an individual, this image will be shown in the database as the representative for matching purposes.

Note: User can Filter or Search through the database by "Database" > "Filter" and "Database" > "Find" functions in the menu bar of the Match menu.

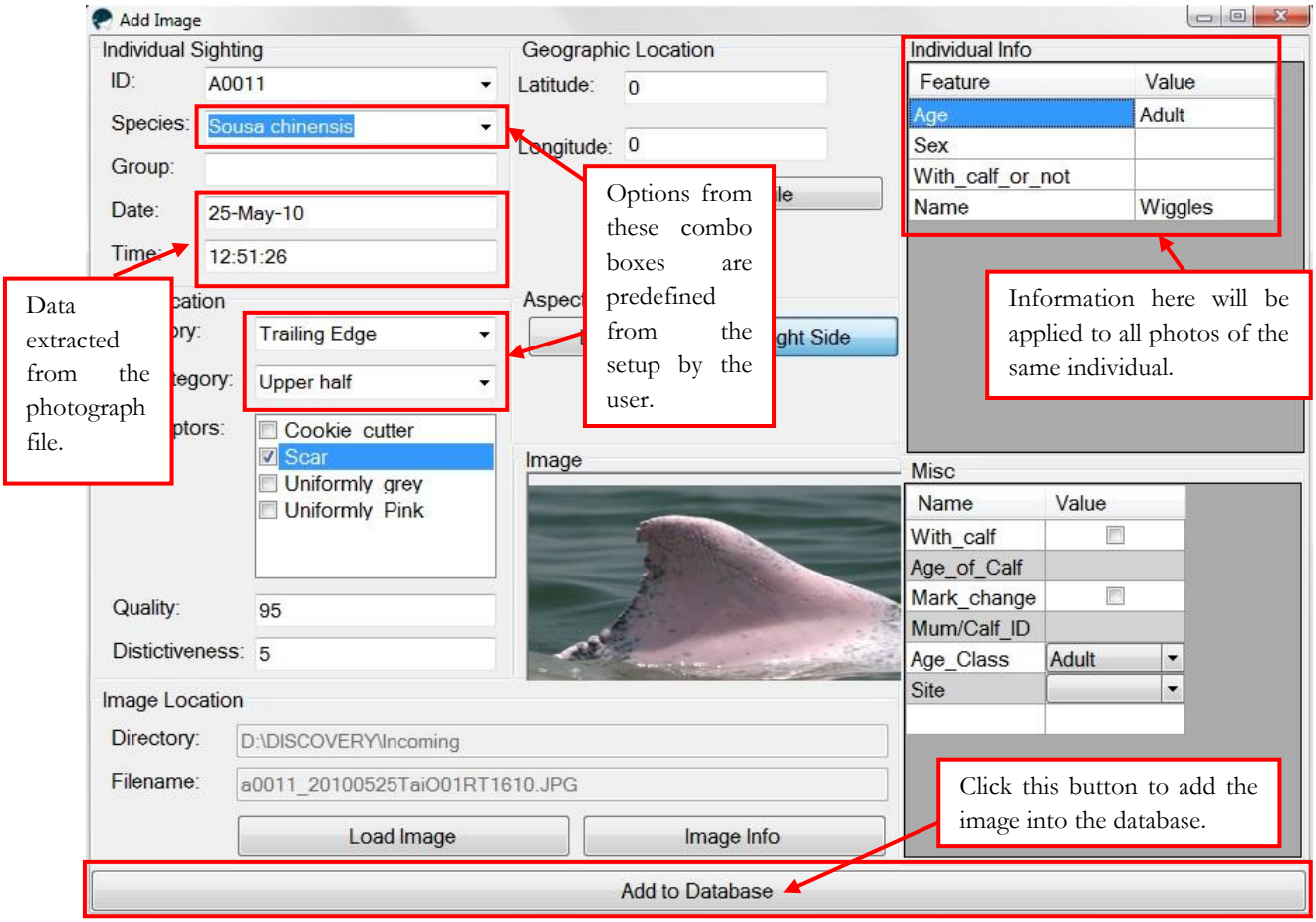


Figure 25. Adding individuals to Discovery's database.

Add Image to Database

Once an image is matched or determined to be a new individual in the catalog, the add image to database form opens to input various parameters. If geographic location information exists in the EXIF data of the image, Discovery will automatically extract the positional data and display it to the user. If no GPS position exists, the user can specify the GPS position by directly typing the latitude and longitude (note the format of GPS data as specified in Preferences) or by clicking the "from file" button and importing a GPX file and using either the waypoint, track or route data (see GPS Data Viewer and Extractor). Image quality and individual distinctiveness assessment, aspect, individual features, various categories and/or descriptors can be included with individual image as well (Figure 25).

ID Classifications

Discovery allows for both textual names and numeric values to be used as an individual identifier. However, it is *recommended* that the user uses numerical ID values to classify the individual as it simplifies potential further issues. We believe there is value in naming individuals (e.g. it is easier to memorize an identifiable feature of an individual by using a descriptive name, e.g. WhiteSpot or ScarFin), but from a database and classification scheme, it is often simpler to attribute numbers to individuals. To accommodate both options, one can setup Discovery to place a name as a feature (see *Note*), therefore the user has the primary classification scheme by numbers, while a feature called "Name" can be setup that allows the user to attribute a name to the individual in addition to the ID number. Both Name and Primary ID are searchable with Discovery. However, if a Name is primarily used to ID individuals, the user should not use special characters or include long names or multiple words separated by spacing (i.e. Six Notch Scar), as this will cause incompatibility issues when exporting data from the Discovery database to other software which may be used for analyses such as program Mark. A numerical ID will be compatible with Mark and less likely to result in conflicts later on.

Chapter

5

Sighting Information

Other than photographic images of individual animals, photo-identification surveys collect a variety of information about each sighting, such as behavior, group size, types of individuals, geographic positions, environmental data, video/photographs taken, associated comments, etc. Some information is usually standard for photo-identification surveys, while other parameters can be project-specific, species-specific or study-site-specific depending on the objectives of the study. As such, the sighting information application has some standard features as well as dynamic user-defined features to allow users to create a data form specific for their needs. The minimum standard information for a sighting is 1. Study Area, 2. Date, 3. Species, and 4. Group ID. These standard parameters are used as unique identifier for each sighting. Start/End times, Survey type, behavior, and group size (min-best-max *or* one group estimate) are provided but optional.

Other 'standard' sighting information are categorized as *Group Info* (i.e. # of Adults, # of Juveniles, # of Calves, etc., subject to change by users), *Environment*, *Comments*, *Geographic*, and *Photographic*. Additional data that are not considered in the 'standard' tabs can be recorded in the *Miscellaneous* tab. Most fields are user-specified and the data types (see Data Types) can also be classified by the user. For example, if the user specifies an environmental parameter such as Sea State (Beaufort) as numeric, then only numeric values can be entered. Date/time parameters are also validated to minimize user entry errors. Invalid entries (i.e. invalid date or time) are highlighted in red. This automatic error checking is designed to minimize data entry errors.

Viewing Sighting Information

To view the sighting information for a study area, the user can access the form from the main menu under the database menu (Database\Sighting menu). A window form will show a summary of the sighting information which can be partitioned by a combination of Study Area, Species, and/or Date. In addition to partitioning the data using the options in the boxes on the left side of the window, the rows can be sorted by column values by selecting the header of the column. For each row entry, the standard unique information is shown (Study Area, Species, Date, and Group ID) as well as the Start/Stop Time, Survey Type, Behavior, and Group Size information.

Buttons are displayed to access the Individual, Environmental, Comments, Geographic, Photographic, and user-defined Miscellaneous information (i.e. Genetics, for example see Fig. 24 below) for each sighting. Sighting entries can be copied to the clipboard, edited, or deleted by right clicking on the row and selecting the appropriate option.

Quick Procedures:

1. Go to Sightings menu: From the main menu, click "Database" > "Sightings".
2. All sightings records will be displayed in the right upper part of the window.
3. Lists of the Study Area(s) (top), Specie(s) (middle) and Date(s) (bottom) will be displayed on the left-hand side of the window. Click on the Study Area(s)/Species(es)/Date(s) to view the records of your choice. The list of records will be displayed instantly on the right according to the selected options on the left.

   Note: Study Area and Species can be edited in "Setup" > "Define" > "General".

4. To view details of certain records, double click the record.
5. Sighting records on the right can be sorted by clicking the header of interest.
6. User can also right click on the records to look for more options (e.g. delete record).
7. User can edit the sighting information by clicking on the tabs (e.g. "Geo" in Figure 26) and changing the records directly on the table.

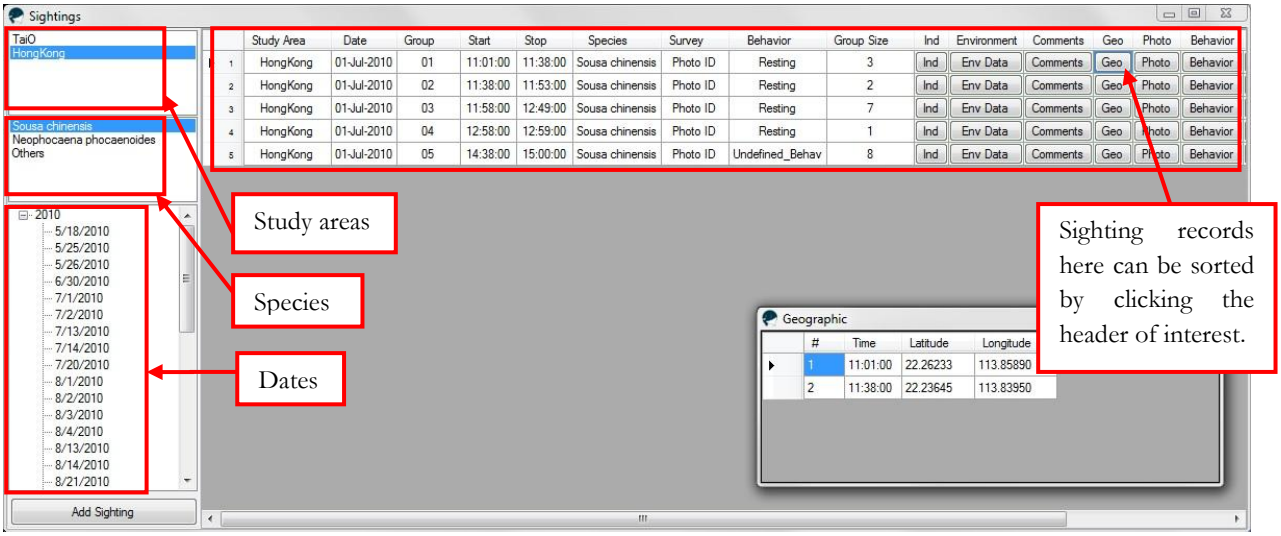


Figure 26. Summary of sighting information.

Note: "Study Area", "Date" and "Group ID" together are used as the ID for sighting information. It is not allowed to input sighting information under the same sighting ID (i.e. the same combination of the three parameters). However duplicate ID might still happen if users modify the data database in MS Access directly. In just a case, the row of sighting information with duplicate ID will be highlighted in **green** color, so as to remind users to check the data and make appropriate modifications.

Adding Sighting Information

To add a sighting, click on the "Add Sighting" button on the bottom left corner of the view sighting information form as described above. This will bring up a new window with multiple tabs for sighting information that the user can fill in. Figure 27 illustrates a basic summary of the sighting information form. The form consists of the Sighting ID, Standard Information, such as Geographic Information, Comments, etc., and user-defined miscellaneous information; each of them is described below.

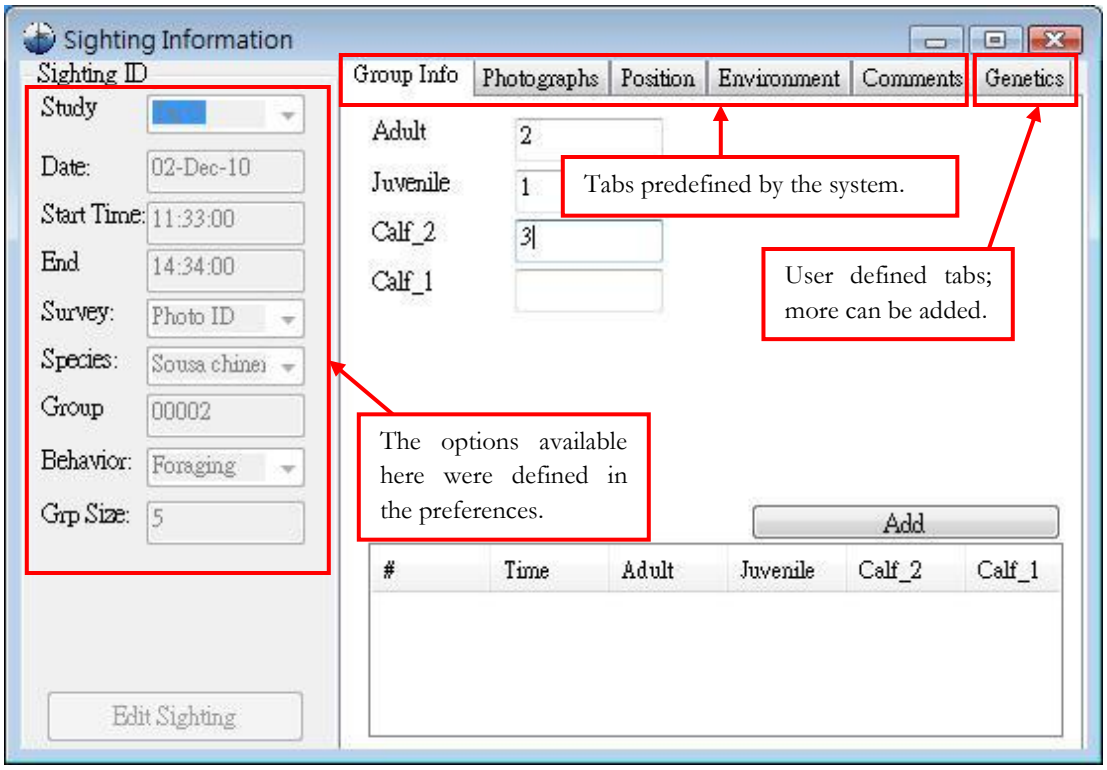


Figure 27. An illustration of the sighting information form.

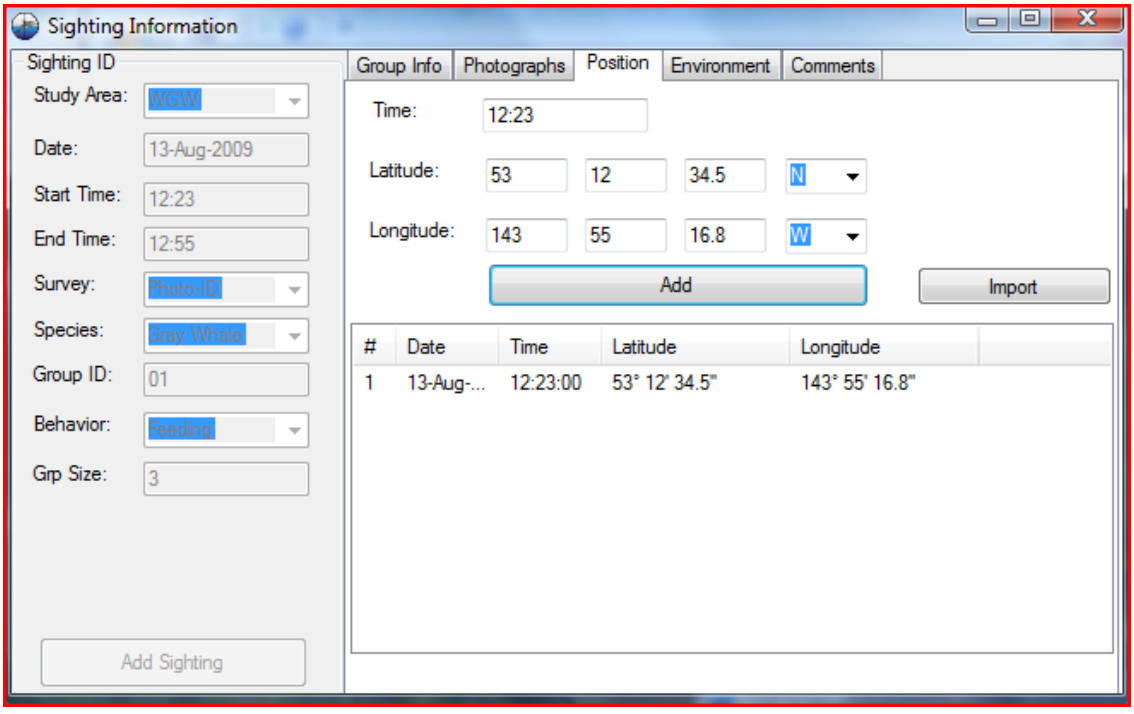


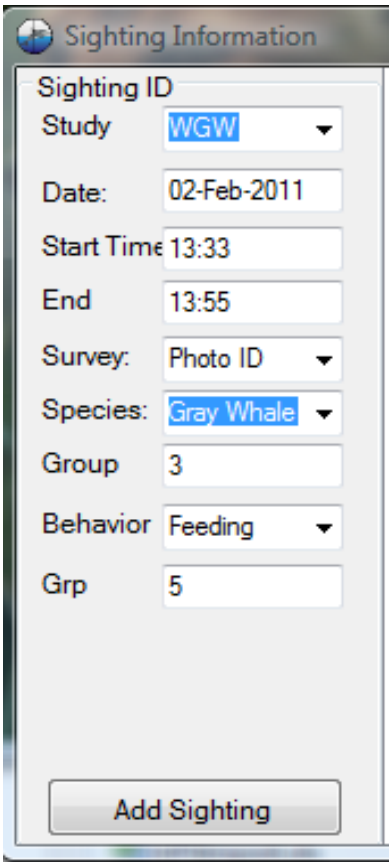
Figure 28. An illustration of how to enter geographic position in the sighting information form.

Quick Procedures:

1. From the main menu, click "Database" > "Sightings" to go to the sightings database menu.
2. To add new sighting record > Select a Study Area and Species, then click the "Add Sightings" button.

Sighting ID

The Sighting ID consists of both mandatory and optional information. The mandatory sighting information is the Study Area, Date, Species, and Group ID. At a minimum, this information needs to be specified prior to adding new sighting to the database.



Optional Sighting ID information consists of the Start/End Time of the sighting, the Survey Type, the Behavioral state, and Group Size (there are two different input types for group size based on Preferences). Survey Type and the Behavioral state can be either typed in upon entry or selected from predefined categories the user defined in the "Sightings Information" from in Parameter in Setup (see Sightings).

Once the basic Sighting ID information has been added, the user can click on "Add Sighting" to add the information to the database as well as add additional details (Individual Information, Geographic Position, Comments, etc.) in relation to the current sighting. These details can only be added once the Sighting ID has been defined and added as a sighting. Once the sighting ID has been added, the sighting ID information cannot be edited as the additional details depend on the unique sighting ID. To edit/delete an entry, see Viewing Sighting Information.

Quick Procedures:

1. From the main menu, click "Database" > "Sightings" to go to the sightings database menu.
2. To add new sighting record, select a Study Area and Species and click "Add Sightings".
3. Fill in / Select the appropriate information for the Sighting ID on the left.
4. Study Area, Date, Species and Group ID are required parameters.
5. Behavioral state can be either entered or chosen from the predefined list. The list can be defined by user in "Setup" > "Define" > "Sightings" > "Behavior". (see Behavior)
6. Click "Add Sighting" button to proceed with recording other sighting details.

Group Information

The Group Information tab in the sighting information form is intended to store additional details about a group, such as group membership. For example, you may want to record the "Number of Adults", "Number of Calves", "Number of Females/Males", etc. Each parameter is user-defined and consists of the user's selected Data Types for that parameter. To setup group information parameters, see Parameter Setup. The entry type will depend on the data type. For True/False or Yes/No Data Types, a button or check box will be provided for data entry. For numeric data types, a text box will be provided, but only numeric values can be entered. Once the Group Information is entered, the information can be added to the database by pressing the add button. You will be asked to input a time value for the entry. The time value is needed to separate multiple entries in case the user wants to incorporate multiple group information such as changes in membership over time.

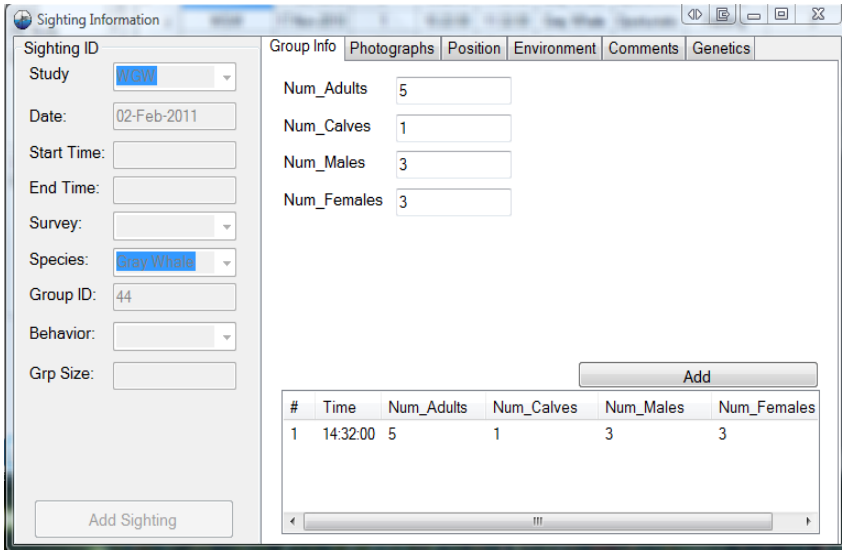


Figure 29. Adding a new entry of group information.

Quick Procedures:

1. Fill in the user-defined entry boxes, then click "Add".
2. Type the time value of the information entered.
3. A new record displaying the information will be shown in the table at the bottom of the tabbed window.
4. Repeat steps 1 to 3 until all data are entered.
5. To delete a record, right click on the record and select the delete option.
6. Entry boxes can be added, edited or deleted in "Setup" > "Define" > "Sightings" > "Group info". Entry boxes type can also be changed in the same place (see Group Info).

Photographic Information

The Photographic tab for sightings contains information about photographic records, video, or other media records that were collected for the sighting. Data entries for photographic info consist of the Recorder (person making the recording), the Media (camera type, video type, voice recorder, etc.), Card /Tape #, and Frames/Time. The name of the Recorder and type of Media can be predefined in

the Parameter Setup or simply typed in upon entry. Once the appropriate information has been entered, simply click on the Add button on the tab to store the photographic data in the database. Multiple entries can be entered for each recording.

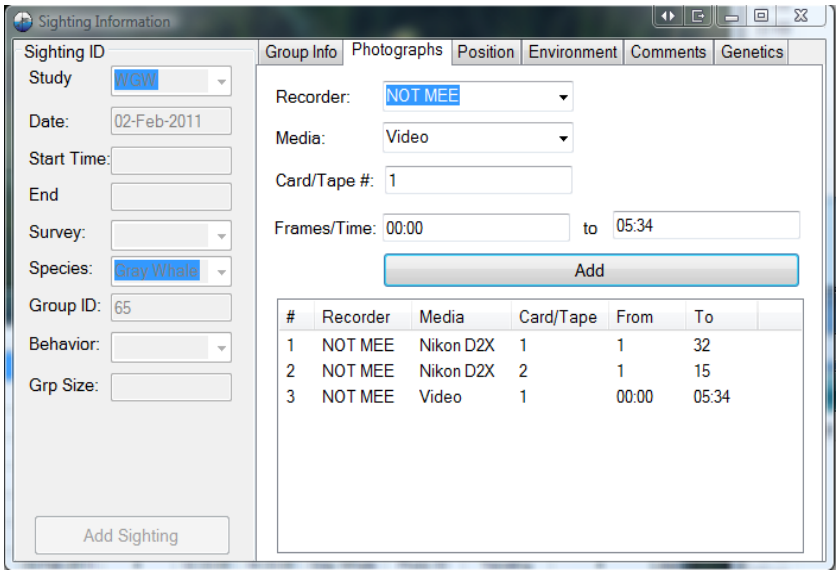


Figure 30. Adding photographs information.

Quick Procedures:

1. Fill in the entry boxes, then click "Add".
2. Type the time value of the information entered.
3. A new record of the information will be shown in the table in the lower half of the window.
4. Repeat steps 1 to 3 until all information is entered.
5. To delete a record, right click on the record and select the delete option.
6. Entry boxes can be added, edited or deleted in "Setup" > "Define" > "Sightings" > "Photographs". Entry boxes type can also be changed in the same place (see Photographs).
7. Recorder and media can be edited/defined in "Setup" > "Define" > "Sightings" > "Photographs" > "Recorder" / "Media".

Geographic Position

The geographic location is important for records of each sighting. Geographic position(s) can be added in the Position tab. These can be manually entered with the Time, Latitude, and Longitude values. Alternatively, geographic positions can be imported automatically from a GPS file(s), such as a GPS Exchange Format (GPX) file, that can be downloaded from a GPS device. EasyGPS software (http://www.easygps.com) is a freely available tool that downloads GPS data and stores GPS information into a GPX file format. Excel and Access formats can also be used to store GPS information and import into the positions tab (see GPS Data Viewer and Extractor).

Latitude and Longitude formats will be displayed based on the user's preferences defined in "Setup" > "Preferences". Despite the format selection, the user can input other formats and have them converted to the current format. For example, if the user selects degrees, minutes, seconds in preferences, but enters decimal degrees in the degrees box, then the minutes and seconds will be automatically converted to the preference values. All positional data are stored in decimal degrees in the database with negative values for southern (latitude) and western (longitude) hemispheres.

Once the user has entered the appropriate information, simply click the "Add" button to store the geographic position of the sighting. Multiple entries can be added for each sighting.

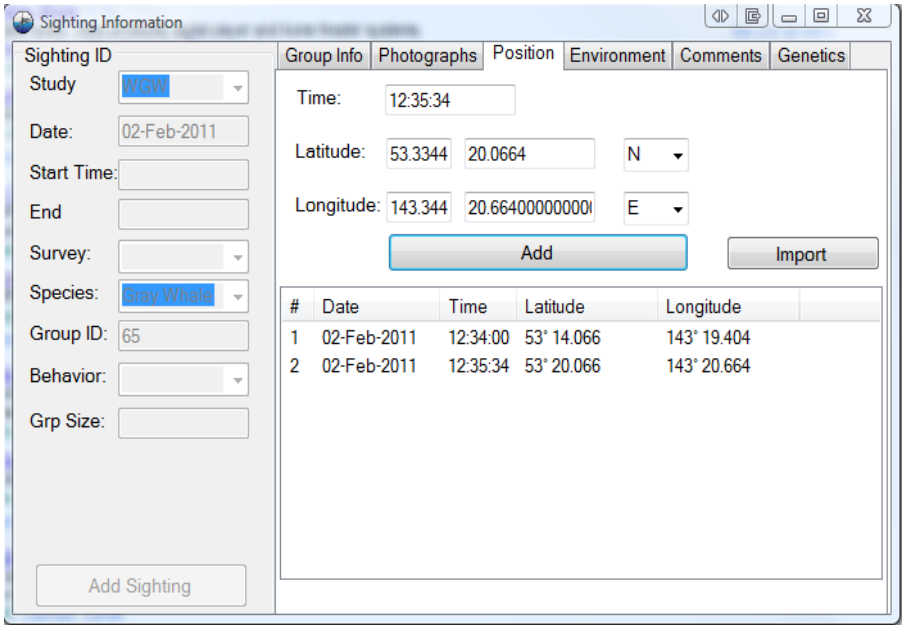


Figure 31. Adding geographical information records.

Quick Procedures:

1. Fill in the GPS information in the entry boxes, then click "Add".
2. GPS information inputted manually will be converted to the format defined by user in "Setup" > "Preferences".

or

3. New records of GPS information will be imported from a GPS file downloaded to the computer (in GPX format).
4. To import a GPS file, click "Import" and select the GPS file.
5. New records will be shown in the table that lists the geographic position data.
6. To delete a record, right click on the record and select the delete option.

Environment

Environmental data are commonly recorded for each sighting. However, the type of parameters recorded can vary among different studies. Discovery allows users to define their own environmental parameters in Parameter Setup and these user-defined parameters are shown in the environmental tab for the sighting entry form. In the above example, Sea State (Beaufort), Visibility, Cloud_Cover, Salinity, and Depth were defined as numeric values with Habitat_Type as a combo data type (with Sandy & Rocky as categories of the combo) and Fog was defined as a Yes/No data type.

Once all values for the environmental records are entered, click the "Add" button; the program will then request a Time for the recording to be stored with the environmental parameters entered. Multiple environmental recordings can be stored for each sighting.

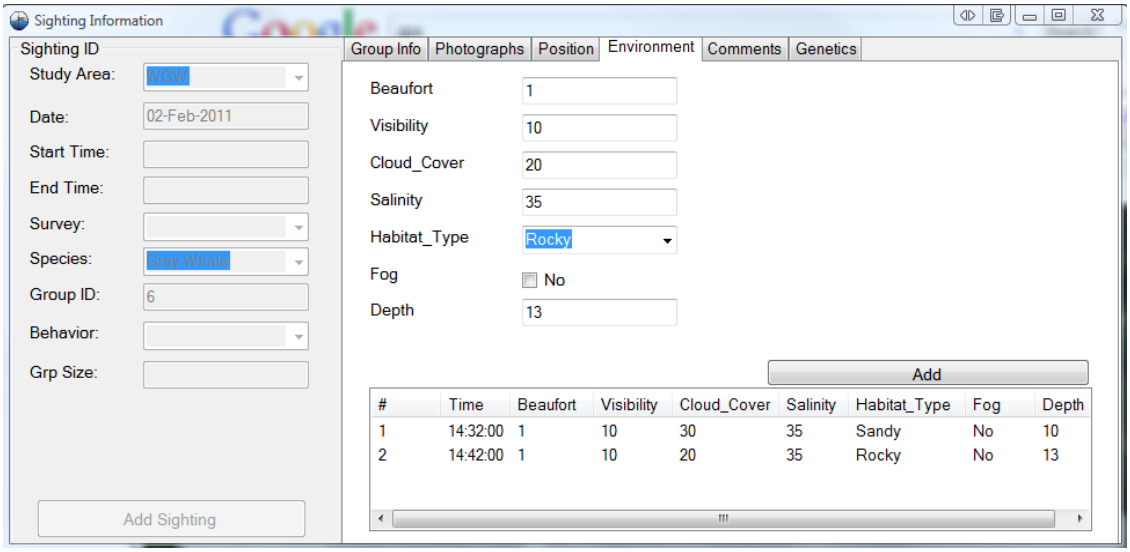


Figure 32. Adding new environmental information records.

Quick Procedures:

1. Fill in the user-defined entry boxes, then click "Add".
2. Type the time value of the information entered.
3. New records will be shown in the table at the bottom right of the window.
4. Repeat steps 1 to 3 until all information for the sighting are entered.
5. To delete a record, right click on the record and select the delete option.
6. Entry boxes can be added, edited or deleted in "Setup" > "Define" > "Sightings" > "Environment". Entry boxes type can also be changed in the same place (see Environment).

Comments

A comment tab is provided to store any additional description or information about the sighting. The comment requires a time and the comment. Simply enter in the time and the comment and click the "Add" button to store a comment in the database. Multiple comments can be stored for each sighting.

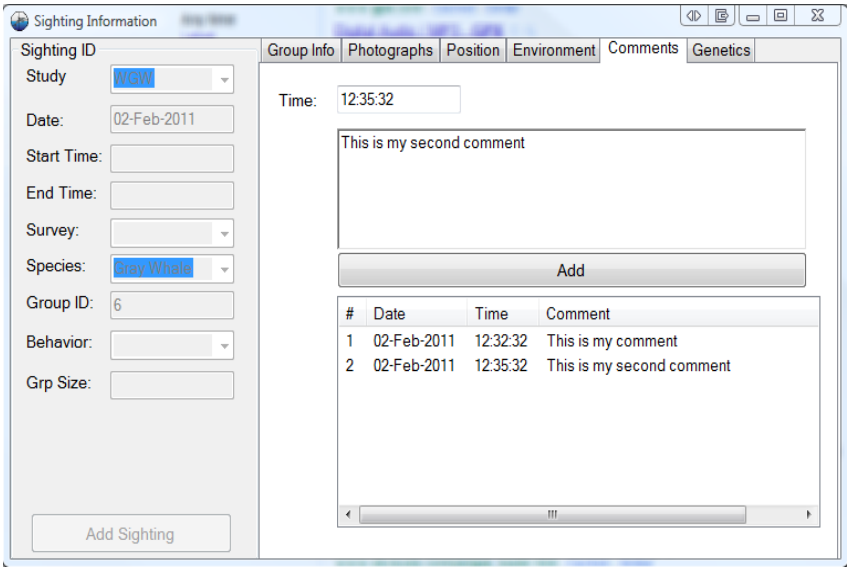


Figure 33. Adding comments into sightings record.

Quick Procedures:

1. Fill in the time and comment box, then click "Add".
2. New record of the information will be shown in the table listing the data.
3. Repeat steps 1 to 2 until all comments are entered.
4. To delete a record, right click on the record and select the delete option.

Miscellaneous Information

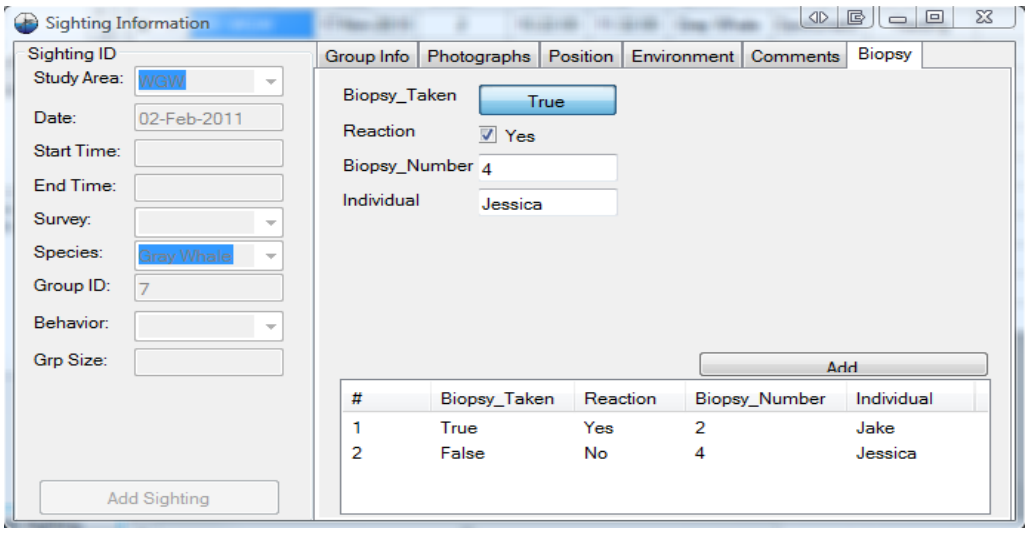


Figure 34. Editing user-defined sightings parameter(s).

Beyond the 'standard' sighting information, additional tabs can be created by adding parameters in Parameter Setup window. For example, collecting biopsies for genetic research maybe part of a study and a Miscellaneous tab called "Biopsy" can be created with categories such as "Biopsy Taken" (Yes/No data type), "Reaction" (Yes/No or Strong/Moderate/Mild/Negligible/None), "Biopsy number" (Numeric), "Individual" (Text), etc. Since many studies have different objectives and collect different information, the Miscellaneous tab offers a means to provide an area to consider these data.

Miscellaneous tabs do not automatically record a time with the entry. If time is an important variable, the user can add it as a time variable in the Miscellaneous setup in Parameter Setup.

Quick Procedures:

1. Additional sighting information forms can be added by user in "Setup" > "Define" > "Sightings" > "Miscellaneous" (see Miscellaneous).
2. Fill in the user-defined entry boxes, then click "Add".
3. New records of the information will be shown in the table on the bottom right.
4. Repeat steps 2 to 3 until all information is entered.
5. To delete a record, right click on the record and select the delete option.

Survey Effort

Survey effort can be recorded with GPS the track data downloaded from a GPS device. This serves as an overall spatial and temporal effort for the number of surveys conducted for each study area and can be used/summarized/visualized later on.

Note: For the survey track in effort data, *only a reference to the GPS/GPX file location is recorded.* This information is used for later information and geographic displays. Users are therefore highly recommended to keep their GPS/GPX files centralized in one location; the track info will not be read and displayed if the files cannot be found under the recorded entry. Users can change the directory information in the Survey Effort window or View Database window, if need be.

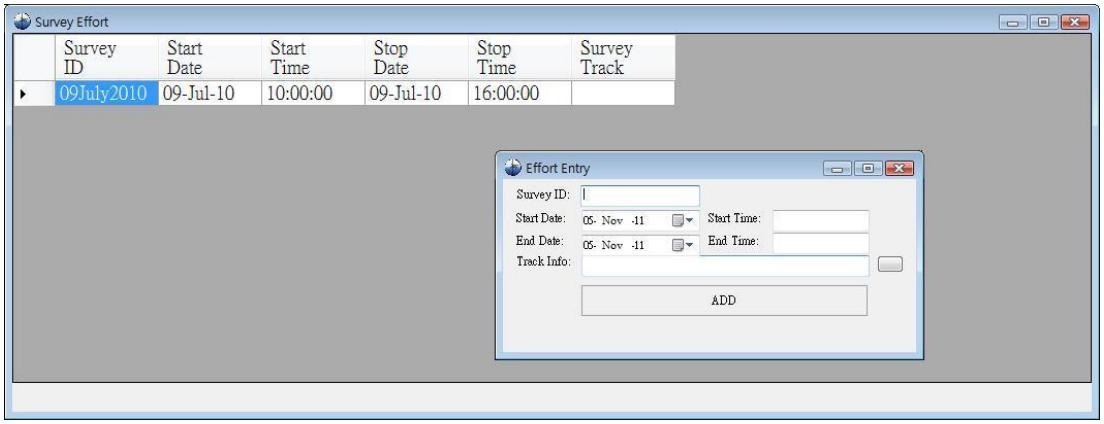


Figure 35. Survey effort can be recorded and viewed in a single table.

Quick Procedures:

1. On the main menu, click "Database" > "Survey Effort".
2. Right click on empty space to add / edit or delete survey effort data.
3. To add a new entry, click "Add" on the right-click menu.
4. Input Survey ID, Start, and End time.
5. User can also incorporate track of the survey from a GPS file by clicking the "…" button at the end of "Track Info" row, and select the track info file. File formats can be found in Chapter 8.
6. Click "ADD" to finish the procedure.
7. New record will be shown in the table.

Other

This provides a highly dynamic tool to record miscellaneous information that may not be included in the process of matching images (see Matching in Chapter 4), sighting and/or survey information records (see Sighting Information and Survey Effort in Chapter 5). For example, strandings of cetacean species could be recorded here since these records do not share the same fields of parameters of live encounters of the animals.

Given the flexibility of recording various information with almost no restrictions, users are required to define the "Type" of parameters, the "Parameter Name" and the "Data Type" (see Other in Parameter Setup) before having the option appearing on the main menu (Figure 36). Each new "Type" defined in Setup will create a separate table in data database, and thus a new tap under the "Database">"Other" menu.

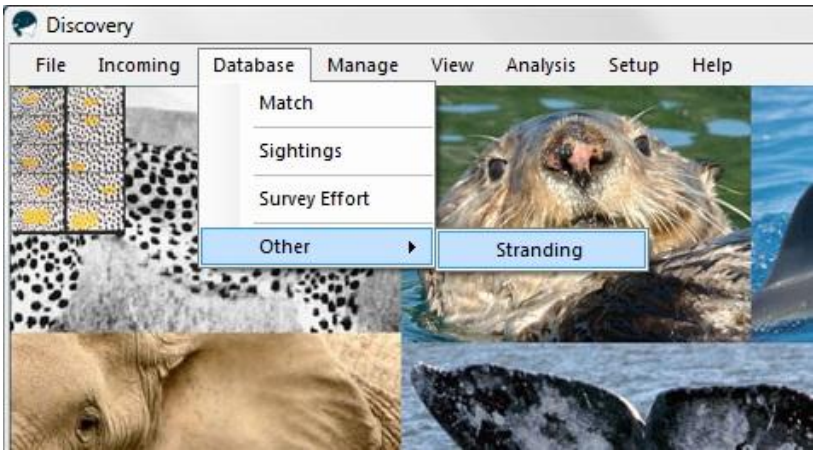


Figure 36. The tool "Other" appears after defining in Setup.

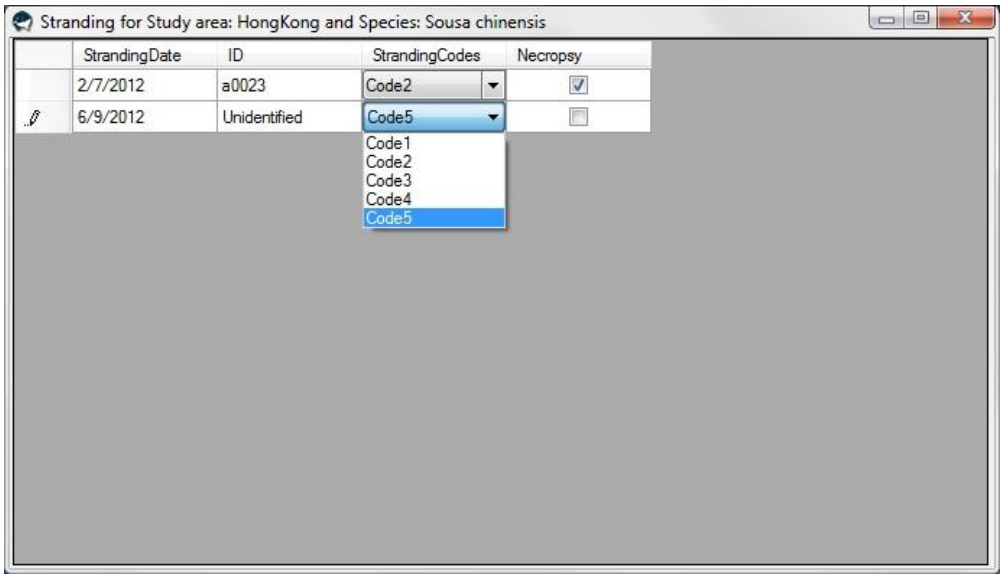


Figure 37. Example of a "Stranding" dataset as an Other dataset.

Quick Procedures:

1. In the main menu, click "Database" > "Other" > "Stranding" (or other tabs defined in Parameter Setup).

2. To add a new entry, right click on empty space and click "Add row" on the right-click menu.

3. Input / Select the appropriate data in the relevant boxes.

4. To quickly import filenames and/or directories of relevant files (e.g. photos of stranded cetacean animals), left-click to select a box, then right-click and click "Get" on the menu and choose from "Filename", "Directory + Filename" or "Directory".

5. New record will be saved to data database once finished typing or selecting. Repeat steps 2 to 4 to add more entries.

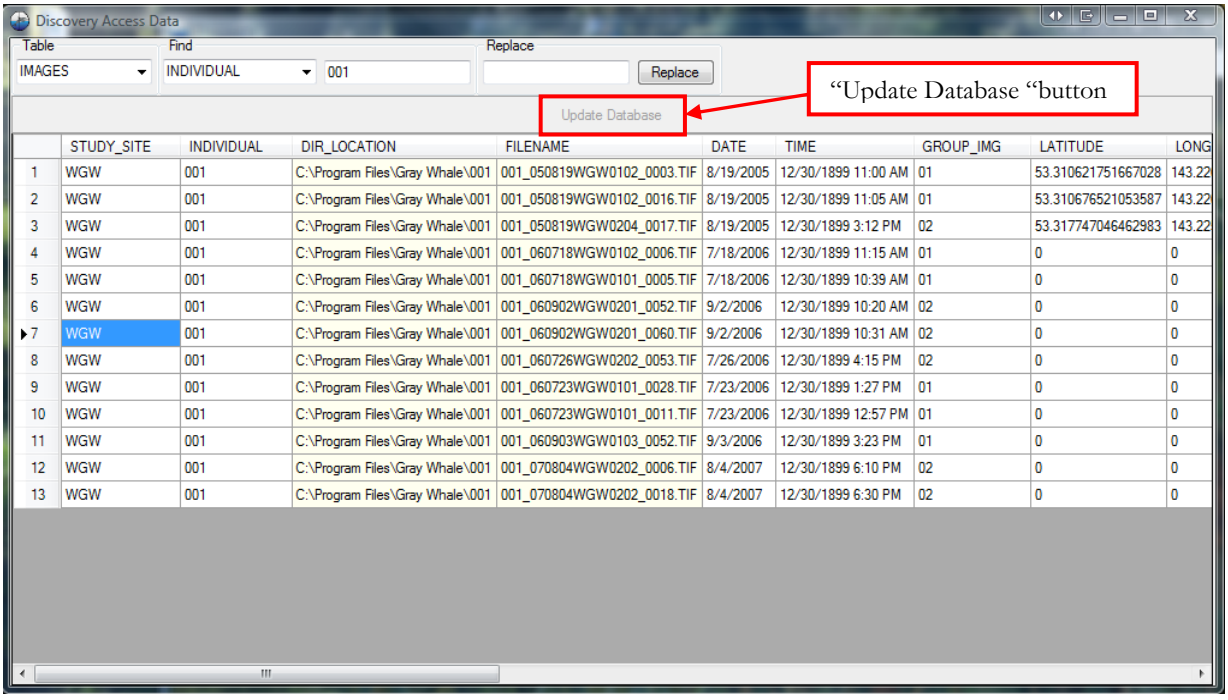
Chapter

6

Managing Data

Discovery stores all data into a Microsoft Access database. If the user has Access installed on their computer, they can view the database directly or use the File\Open in Access function to open either the system or data file. Microsoft Access has various tools to assist with database management. However, Microsoft Access is not required, but highly recommended, to use Discovery.

If a user does not have Microsoft Access and wishes to view the data, they can use the View\Database option. This window provides a direct link to the database and allows the user to partition, search, and manipulate their data directly without the need to open the Access database.



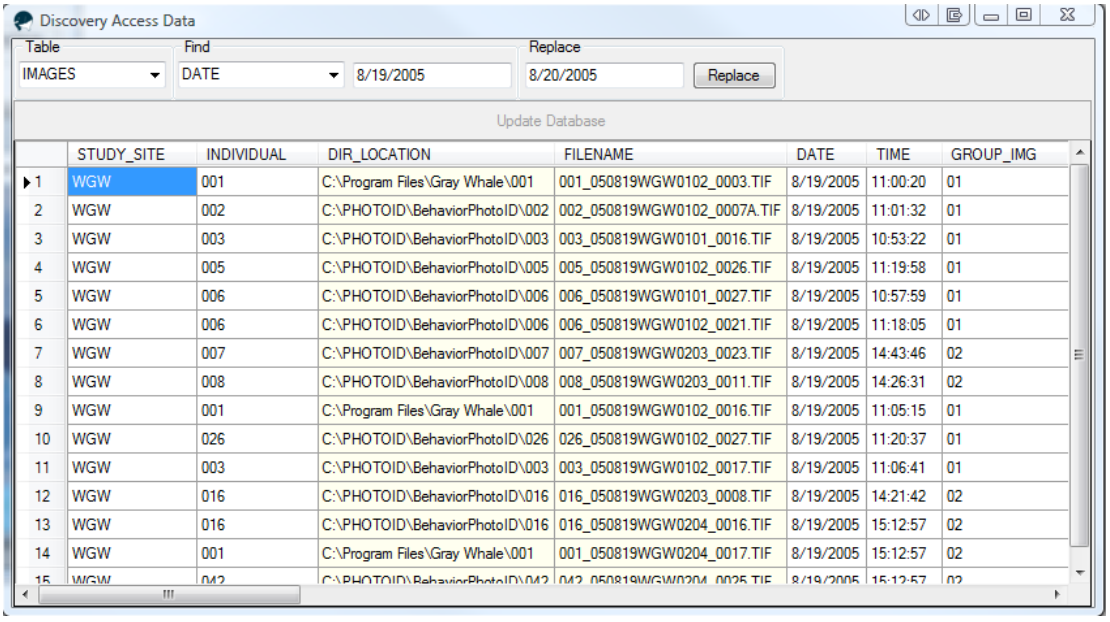
	STUDY_SITE	INDIVIDUAL	DIR_LOCATION	FILENAME	DATE	TIME	GROUP_IMG	LATITUDE	LONG
1	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0102_0003.TIF	8/19/2005	12/30/1899 11:00 AM	01	53.310621751667028	143.22
2	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0102_0016.TIF	8/19/2005	12/30/1899 11:05 AM	01	53.310676521053587	143.22
3	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0204_0017.TIF	8/19/2005	12/30/1899 3:12 PM	02	53.317747046462983	143.22
4	WGW	001	C:\Program Files\Gray Whale\001	001_060718WGW0102_0006.TIF	7/18/2006	12/30/1899 11:15 AM	01	0	0
5	WGW	001	C:\Program Files\Gray Whale\001	001_060718WGW0101_0005.TIF	7/18/2006	12/30/1899 10:39 AM	01	0	0
6	WGW	001	C:\Program Files\Gray Whale\001	001_060902WGW0201_0052.TIF	9/2/2006	12/30/1899 10:20 AM	02	0	0
7	WGW	001	C:\Program Files\Gray Whale\001	001_060902WGW0201_0060.TIF	9/2/2006	12/30/1899 10:31 AM	02	0	0
8	WGW	001	C:\Program Files\Gray Whale\001	001_060726WGW0202_0053.TIF	7/26/2006	12/30/1899 4:15 PM	02	0	0
9	WGW	001	C:\Program Files\Gray Whale\001	001_060723WGW0101_0028.TIF	7/23/2006	12/30/1899 1:27 PM	01	0	0
10	WGW	001	C:\Program Files\Gray Whale\001	001_060723WGW0101_0011.TIF	7/23/2006	12/30/1899 12:57 PM	01	0	0
11	WGW	001	C:\Program Files\Gray Whale\001	001_060903WGW0103_0052.TIF	9/3/2006	12/30/1899 3:23 PM	01	0	0
12	WGW	001	C:\Program Files\Gray Whale\001	001_070804WGW0202_0006.TIF	8/4/2007	12/30/1899 6:10 PM	02	0	0
13	WGW	001	C:\Program Files\Gray Whale\001	001_070804WGW0202_0018.TIF	8/4/2007	12/30/1899 6:30 PM	02	0	0

Figure 38. Illustration of viewing the current Access data database in Discovery.

To view the data, the user has to select the table of interest. An explanation of the table types are outlined in Table 3. For example, if the user wants to view the image catalog, they can open the table named "Images". This table should illustrate the study site, individual, directory location (dir_location), filename, date, time, group ID, etc. The Image Table view has a routine to check the filename locations to see if the file exists on the computer. If the fields are red, then the system could not find the file, otherwise the directory and filename fields are ivory if the file exists at the specified location.

Data can be directly manipulated by selecting a cell and typing the new values. However, the *new information will not update the database values until the user presses the "*Update Database*" button.*

Find and Replace functions are provided to assist with data manipulation. To find an entry in a column, select the column and type in the item you want to find. The find function works dynamically to find values that are similar to the values entered, therefore the entire entry may not be required. Once the item is selected, the user can also replace values in the find box with another value. This can be a very useful tool to alter directory locations. For example, if you moved your catalog images from "C:\My_Catalog\Spinners" to "C:\Photo_ID\Spinners". You can find the entry "C:\My_Catalog" and replace them with "C:\Photo_ID" to update the location of the images. Again, ensure you press the "Update Database" button to apply the changes to the actual database.



	STUDY_SITE	INDIVIDUAL	DIR_LOCATION	FILENAME	DATE	TIME	GROUP_IMG
1	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0102_0003.TIF	8/19/2005	11:00:20	01
2	WGW	002	C:\PHOTOID\BehaviorPhotoID\002	002_050819WGW0102_0007A.TIF	8/19/2005	11:01:32	01
3	WGW	003	C:\PHOTOID\BehaviorPhotoID\003	003_050819WGW0101_0016.TIF	8/19/2005	10:53:22	01
4	WGW	005	C:\PHOTOID\BehaviorPhotoID\005	005_050819WGW0102_0026.TIF	8/19/2005	11:19:58	01
5	WGW	006	C:\PHOTOID\BehaviorPhotoID\006	006_050819WGW0101_0027.TIF	8/19/2005	10:57:59	01
6	WGW	006	C:\PHOTOID\BehaviorPhotoID\006	006_050819WGW0102_0021.TIF	8/19/2005	11:18:05	01
7	WGW	007	C:\PHOTOID\BehaviorPhotoID\007	007_050819WGW0203_0023.TIF	8/19/2005	14:43:46	02
8	WGW	008	C:\PHOTOID\BehaviorPhotoID\008	008_050819WGW0203_0011.TIF	8/19/2005	14:26:31	02
9	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0102_0016.TIF	8/19/2005	11:05:15	01
10	WGW	026	C:\PHOTOID\BehaviorPhotoID\026	026_050819WGW0102_0027.TIF	8/19/2005	11:20:37	01
11	WGW	003	C:\PHOTOID\BehaviorPhotoID\003	003_050819WGW0102_0017.TIF	8/19/2005	11:06:41	01
12	WGW	016	C:\PHOTOID\BehaviorPhotoID\016	016_050819WGW0203_0008.TIF	8/19/2005	14:21:42	02
13	WGW	016	C:\PHOTOID\BehaviorPhotoID\016	016_050819WGW0204_0016.TIF	8/19/2005	15:12:57	02
14	WGW	001	C:\Program Files\Gray Whale\001	001_050819WGW0204_0017.TIF	8/19/2005	15:12:57	02
15	WGW	042	C:\PHOTOID\BehaviorPhotoID\042	042_050819WGW0204_0025.TIF	8/19/2005	15:12:57	02

Figure 39. Replacing certain records (or part of) from the current loaded database in Discovery using the Find and Replace functions.

Note: For each table there is a field called "PKey". This is a unique number that is auto-incremented. This identifier is normally used for database purposes. It is *highly recommended* that the user does not change or alter this field.

Quick Procedures:

1. From the main menu, click "View" > "Database" to open the database window.
2. On the top left, select a table that you want to view. Explanation of Table Types is given in Table 3.
3. User can click on any cell and edit the table.

   Note: Database is not modified until step 4.

4. Click on the "Update Database" button to save the edits.
5. To *find* an entry:
   a. Select the column and type in the item you want to find in the "Find" box.
   b. Select the cell.
   c. Enter value in the "Replace" box and click on the "Replace" button, the value in the selected cell will be replaced by the value in the Replace box.

Verifying Individuals

To ensure the ID and/or other image information (e.g. category/sub-category, quality, distinctiveness, etc) have been correctly entered to the database, all newly-cataloged images, either as new or previously matched individuals, should be checked and verified. An ID is considered a "new individual" if no images for a defined ID have been verified; or it is considered as a "matched individual" if at least one image has been verified. Therefore users should first verify the earliest image of all new IDs, which then become matched; and the remaining images will be verified together with the previously matched individuals.

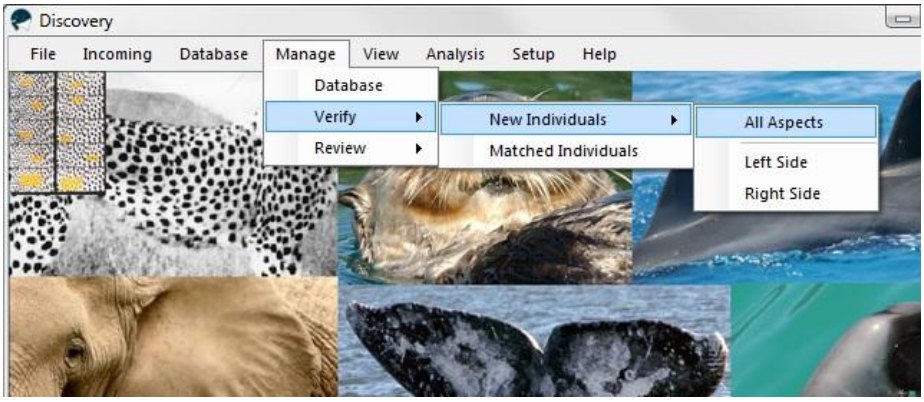


Figure 40. Illustration of the Verify menu in Discovery.

Verifying New Individuals

While verifying "new individuals" (those that have not been verified previously), users need to consider whether various aspects (if there are multiple aspects) would affect the certainty of identification in the study. For example, in photo-ID study of spinner dolphins using notches on dorsal fins as the primary feature, the notches have the same appearance on the left or right side; however in the study of gray whales which primarily use body markings for identification, left and right aspects have no relationship. Therefore users are provided with options to verify new IDs by particular aspect or all combined together (Figure 40).

After selecting a particular aspect or all aspects, the first images (in term of date and time) of the new IDs will be automatically selected and listed (Figure 41). Users could go through the existing verified catalog to ensure that there are no matches to the proposed new IDs (Figure 42). Users could change the ID if the query image is found to be a match in the catalog. The image details could also be changed.

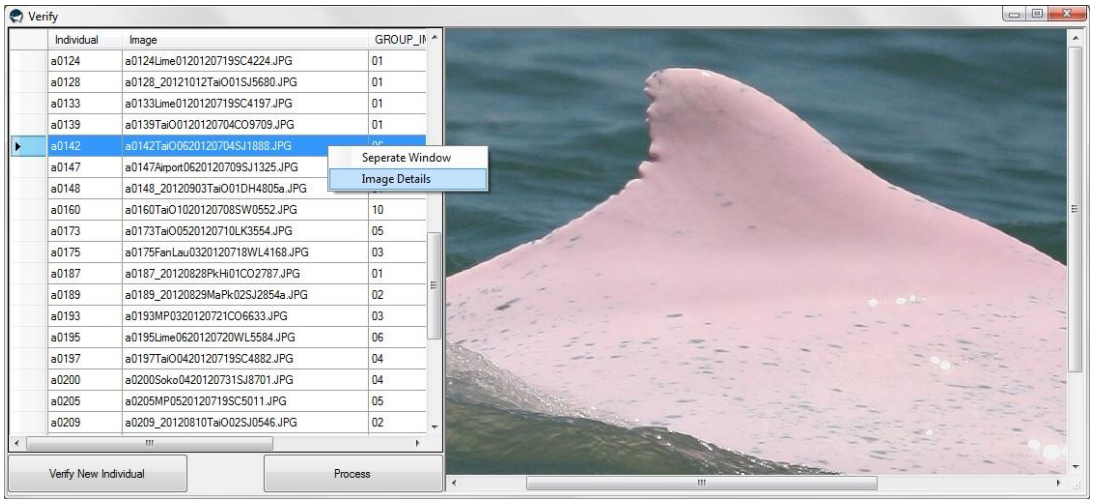


Figure 41. Illustration of Verifying New Individual form.

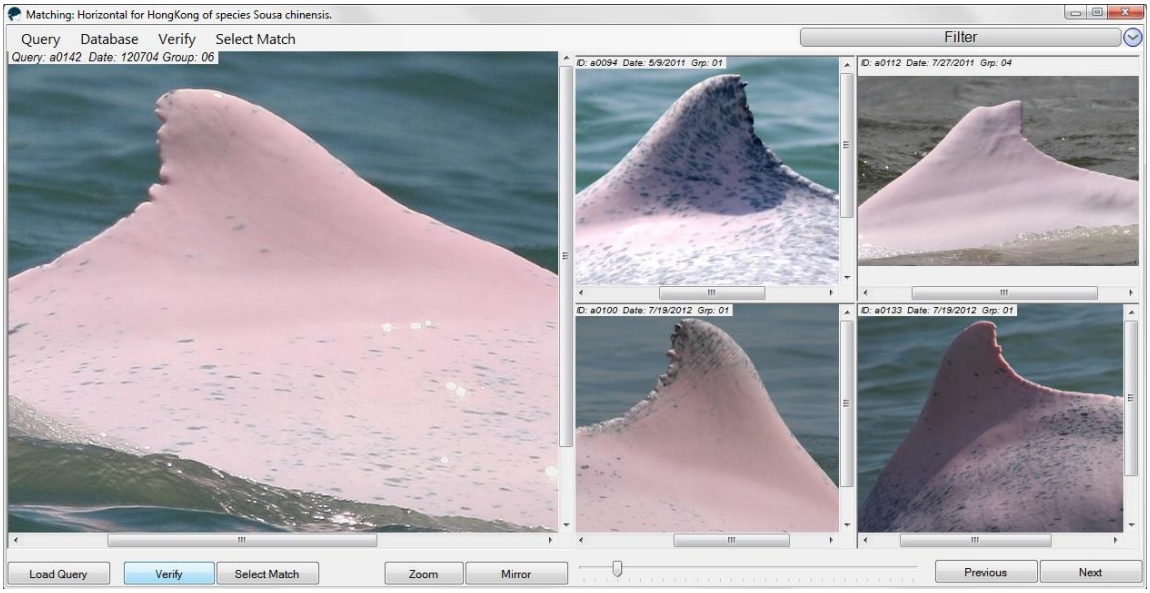


Figure 42. Illustration of the process of comparing new individual to the existing verified catalog.

Quick Procedures:

1. From the main menu, click "Manage" > "Verify" > "New Individuals". Choose "All aspects" or a particular aspect.

2. The first images of new IDs with selected aspect(s) are listed on the left. Right-click on the ID to be verified. Choose "Separate Window" to show the image in a separate window. Choose "Image Details" to view/edit the image parameters. A new form similar to matching image will appear. Change the image details and click "Edit" to save the changes.

3. Move the horizontal scroll bar to the right and click on the "Marginal" and "Temp ID" checkbox if appropriate.

4. Click "Verify New Individual" at the left bottom corner to quickly verify the selected new ID, or click "Process" to compare the selected image with the existing verified catalog.

5. The "Process" form functions similarly to the matching form (see Figure 22 in Matching).If no matches are found, click "Verify" to verify the selected ID; if there is a match, click "Select Match" to change the ID of the query image.

6. Repeat step 2 to 5 until all new individuals are verified or matched to existing catalog.

Verifying Matched Individuals

The Verify Match form allows users to verify the newly-matched images of the IDs which are verified in the existing catalog. The query images are compared with the type specimen image(s) of that particular individual. Users can either verify or reject the match (Figure 43). If the match is rejected, the ID of the query image should be changed to a new ID and the image will be moved to the new ID folder and the ID, filename and directory updated in data database, so that the rejected image would enter the procedure of verifying new individuals, which might result in a new individual or another match to the verified catalog.

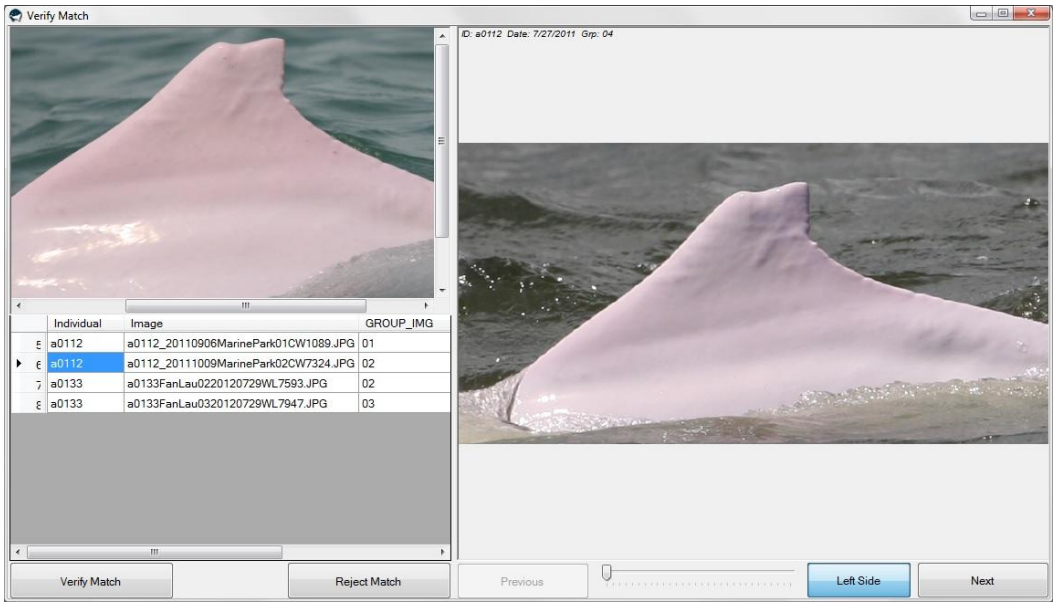


Figure 43. Illustration of Verifying Matched Individuals form.

Quick Procedures:

1. From the main menu, click "Manage" > "Verify" > "Matched Individuals".

2. All non-verified images will be listed on the left. Right-click on the ID to be verified for the following options: 1. Choose "Separate Window" to show the image in a separate window, 2. Choose "Image Details" to view/edit the image parameters. A new form similar to matching image will appear. Change the image details and click "Edit" to save the changes.

3. Move the horizontal scroll bar to the right and click on the "Marginal" and "Temp ID" checkbox if needed.

4. Type specimen image(s) of the proposed ID of the selected query image are displayed on the right. Click the aspect button to switch between "All Aspects" and the aspect of the query image.

5. Click either 1. "Verify Match" at the left bottom corner to verify the query image or 2. "Reject Match" to reject the proposed ID.

6. If the query image is rejected, the query ID should be changed to another ID existing in the verified catalog or a new ID so that the rejected image would enter the procedure of verifying new individuals to find another match or enter as a new ID.

7. Repeat step 2 to 7 until all query images are verified or rejected.

Review

The verified catalog should be regularly reviewed to prevent and correct mistakes in identifications. Users can also use the review process to determine the best representative image ("Type Specimen") per aspect. The choice of type specimen images should be updated to the most recent high quality representative image of that individual for each aspect. In the Review page, the type specimen images are highlighted in **green** color (Figure 44). Right click on the image and select "Change to Type Specimen" to update the representative image.

Quick Procedures:

1. From the main menu, click "Manage" > "Review" and choose "All Aspects" or a user-defined aspect.
2. All verified individuals are listed on the left. Double-click on the ID or click on the "+" symbol next to the ID to expand the list of images of the selected ID.
3. Click on an image to display it on the right. Right-click on the image and select "Separate Windows" to display in a separate window.
4. To display multiple images, left-click at empty space and move the mouse pointer to include multiple images in the colored drag box or hold "Ctrl" key while choosing multiple images.
5. If an incorrect match of ID is identified, right-click on the image and select "Image Details" in the context menu. Change the ID, and other image details if needed, and click "Edit". Users will also be asked if they want to move the image to the new ID folder.
6. To update the type specimen image, right-click on the most recent high quality image of an individual and select "Change to Type Specimen". The previous type specimen image of the same aspect will be un-selected, while the updated one will be highlighted.
7. Repeat step 2 to 6 to review all individuals in the catalog.

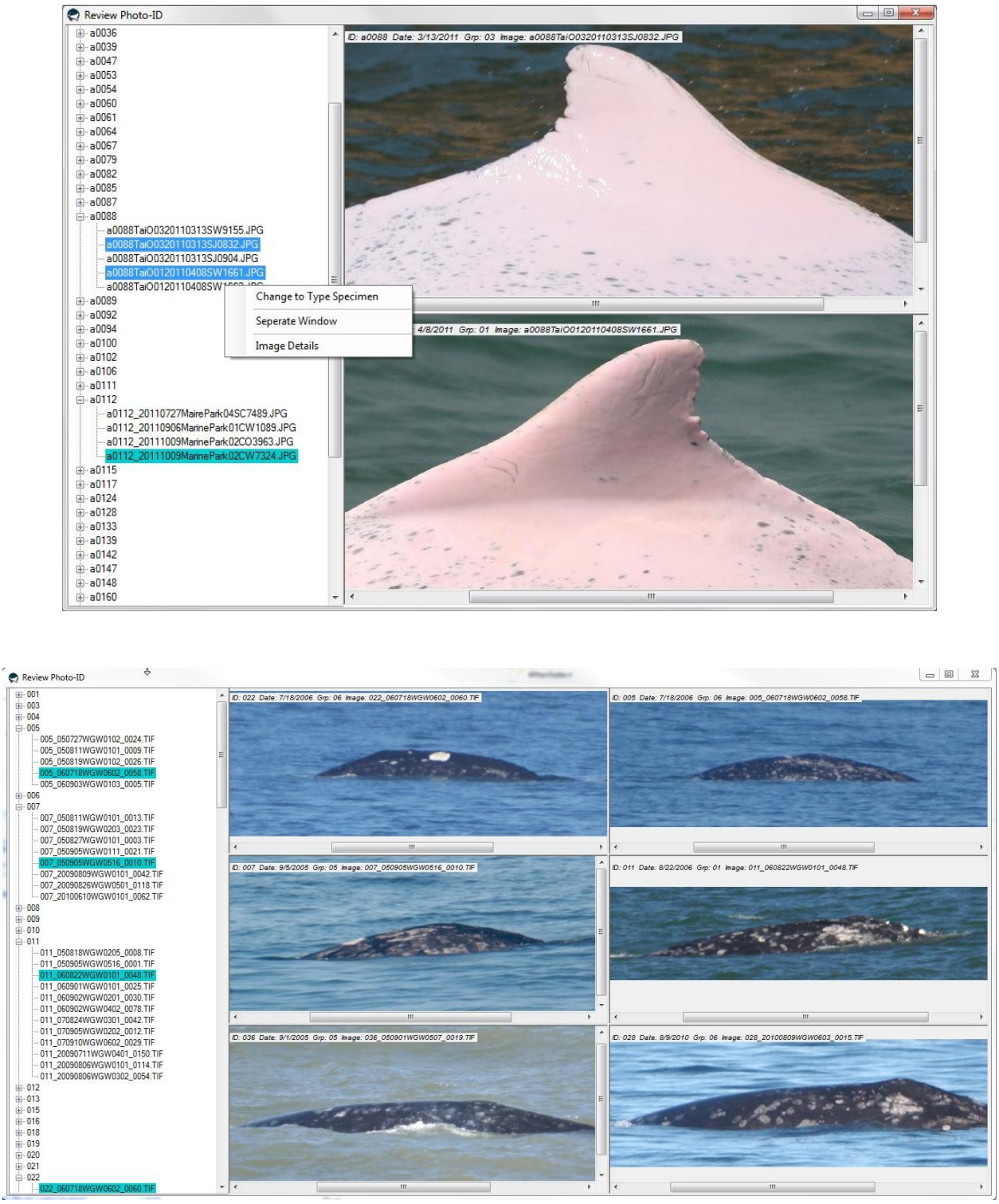


Figure 44. Illustration of Review form.

Chapter

7

Viewing Individuals

The View Individuals form provides an overview of the individuals in the existing catalog. To view the individuals in the catalog, click "View" in the main menu and then select "Individual". Select an individual in the left-hand table (catalog of individuals) and all the images for that individual along with the details will be displayed in the bottom right-hand table. Information pertaining to the individual ID can be edited by double clicking on the cell to be modified. In the Image list for an ID, the user can select which image will be the "Type Specimen" for each aspect by ticking the box (see the indication of Type specimen in the screen capture below; Figure 45). In most instances, "Type Specimen" is likely to be the most recent highest quality representative image of an individual. Only one "Type Specimen" can be selected per Aspect of the individual. "Type Specimen" is used as the primary image shown as a potential match in the catalog.

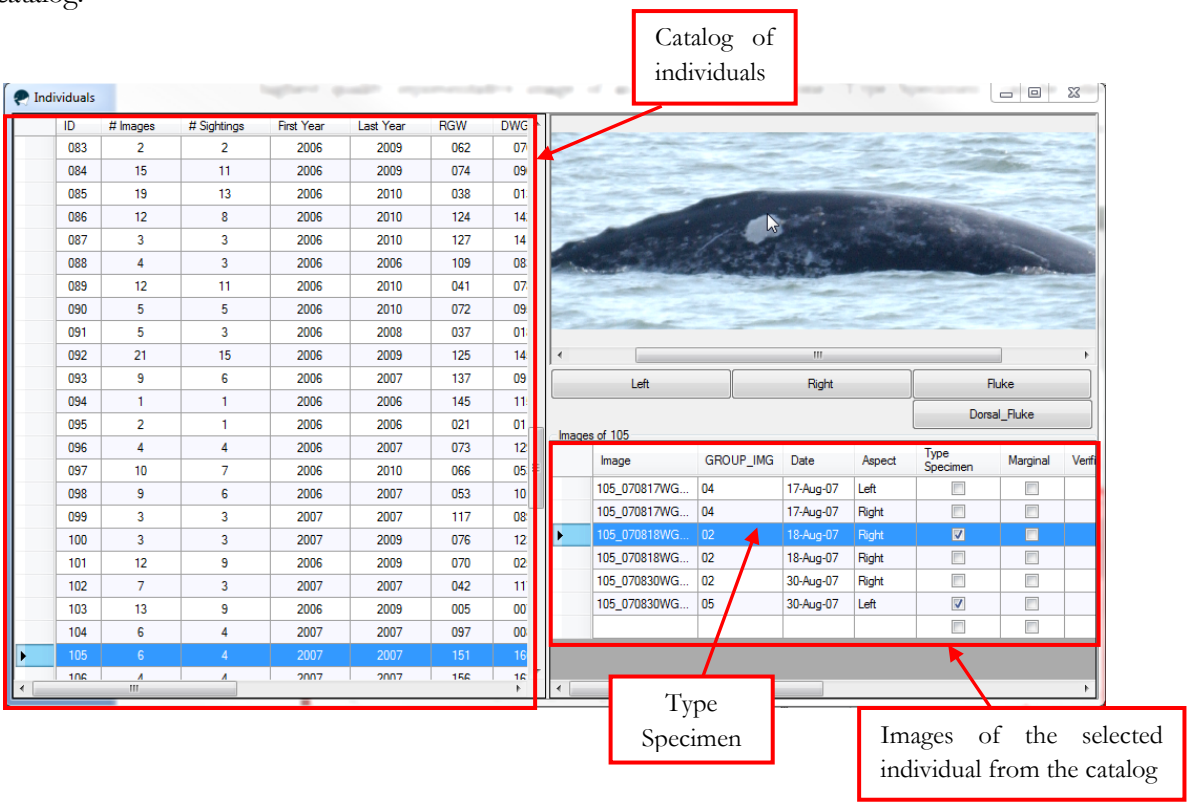


Figure 45. Form illustrating View Individuals in an existing gray whale catalog.

The left table contains the individual IDs, the first year seen, last year seen, number of images, number of sightings, individual features (see below), and crossmatching numbers (optional, see below). The user can perform updates and maintenance of the catalog from this form and classify "Type Specimens", "Marginal" images, "Temporary" IDs, etc.

Individual Features

The user can specify individual features, such as sex, age, name, etc. Such information usually pertains to the individual and not to a particular sighting or photograph. For each individual, the individual features are shown. To modify the individual information, simply select the cell and type the desired value for the individual feature parameter.

Crossmatching

A crossmatching function is available to centralize different individual catalogs (see CrossMatch in Preferences). For example, the work involving western gray whales has three separate photo-identification catalogs from different teams. Each catalog has different identification numbers and all catalogs can be centralized but kept separate so the user can see the images from another catalog compared to their catalog. These catalogs are maintained as separate by assigning different Study Area in Setup. Only individuals of the default study area are listed in the left table, while crossmatched images from other catalogs and/or other study areas will be displayed with all images of a selected individual on the bottom right.

The initial left individual table can be changed to display images by survey date that illustrates the date, number of individuals and number of groups in the left table, and all images (of all individuals) on the specific date will be displayed in the bottom right table (Figure 46).

Quick Procedures:

1. From main menu, click "View">"Individuals" to open the individuals window.
2. Select an individual in the left table; all images in the database for that individual will be displayed in the table on the bottom right side of the window. Right-click anywhere on the left grid form, the options "Resighting" (showing all resighting of the selected individual), "Delete" (deleting the individual record) and "Change ID" (quickly changing ID of all images of the selected individual) are available.

3. To edit the information in a table on the cell, double click and type the desired value.

4. In the table on the bottom right side of the window, select which image is the "Type Specimen", "Marginal" or "Temporary ID" by ticking the appropriate box.

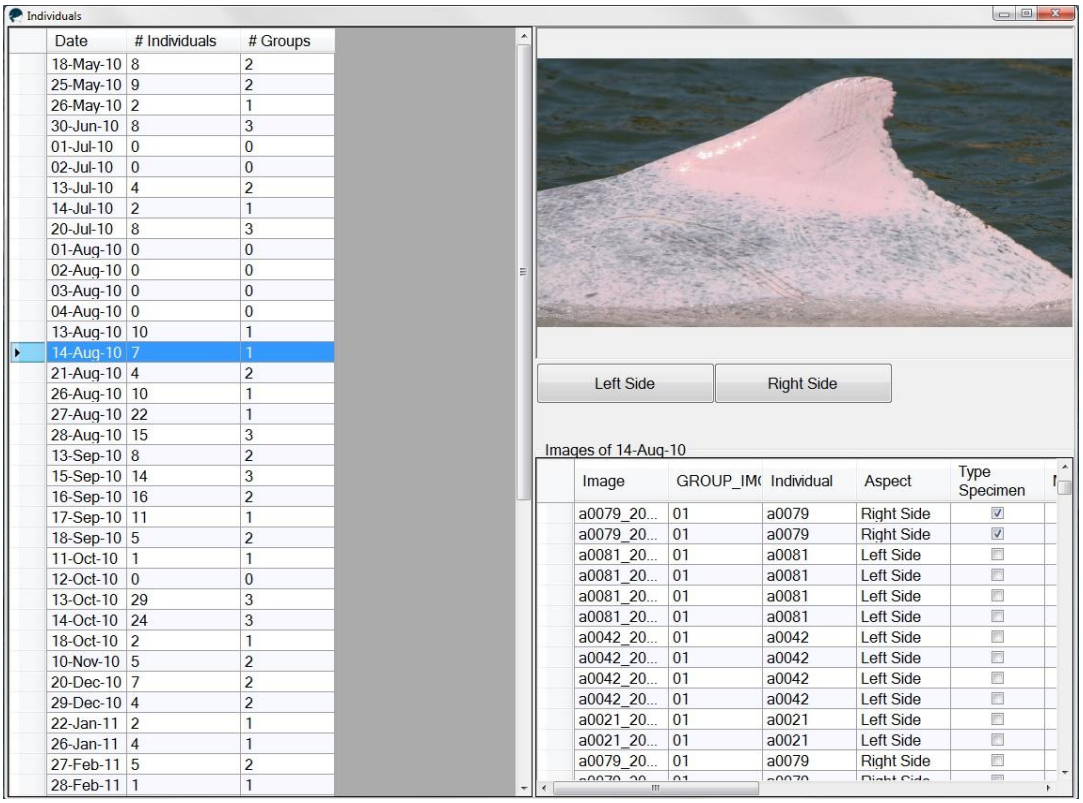
5. In the left side table, individual features can be entered by clicking the corresponding cell.

6. Individual features can be added, edited or deleted in "Setup" > "Define" > "Individual ID" > "Individual Info" (see Individual Info).

7. Individual features can also be used as a searching key in Matching (see Matching).

8. Clicking on the aspect below the displayed image (e.g. Left side) will by default display the type specimen image for that aspect only.

9. Right-click on any space in the left table and select "Show Date" to display the list of survey dates instead of individuals; right-click again and select "Show Individual" to switch back to the list of individuals.

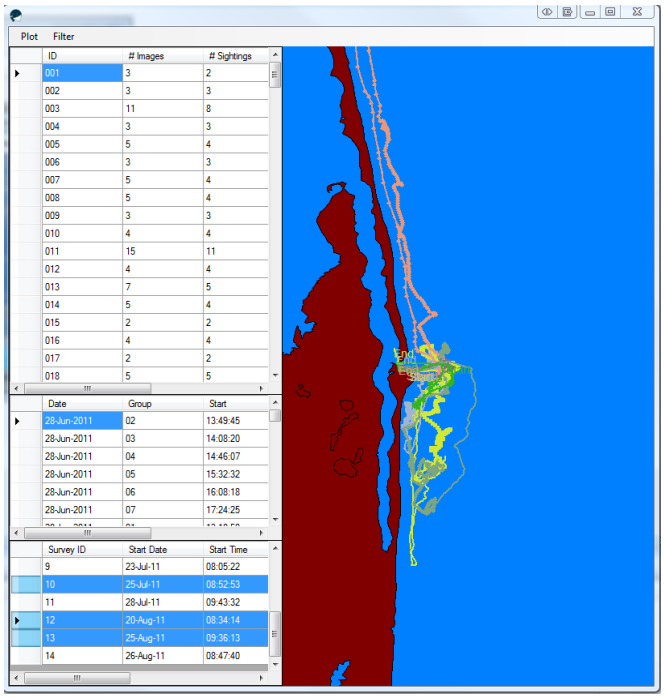


Date	# Individuals	# Groups
18-May-10	8	2
25-May-10	9	2
26-May-10	2	1
30-Jun-10	8	3
01-Jul-10	0	0
02-Jul-10	0	0
13-Jul-10	4	2
14-Jul-10	2	1
20-Jul-10	8	3
01-Aug-10	0	0
02-Aug-10	0	0
03-Aug-10	0	0
04-Aug-10	0	0
13-Aug-10	10	1
14-Aug-10	7	1
21-Aug-10	4	2
26-Aug-10	10	1
27-Aug-10	22	1
28-Aug-10	15	3
13-Sep-10	8	2
15-Sep-10	14	3
16-Sep-10	16	2
17-Sep-10	11	1
18-Sep-10	5	2
11-Oct-10	1	1
12-Oct-10	0	0
13-Oct-10	29	3
14-Oct-10	24	3
18-Oct-10	2	1
10-Nov-10	5	2
20-Dec-10	7	2
29-Dec-10	4	2
22-Jan-11	2	1
26-Jan-11	4	1
27-Feb-11	5	2
28-Feb-11	1	1

Image	GROUP_IM(	Individual	Aspect	Type Specimen
a0079_20...	01	a0079	Right Side	☑
a0079_20...	01	a0079	Right Side	☑
a0081_20...	01	a0081	Left Side	☐
a0081_20...	01	a0081	Left Side	☐
a0081_20...	01	a0081	Left Side	☐
a0081_20...	01	a0081	Left Side	☐
a0042_20...	01	a0042	Left Side	☐
a0042_20...	01	a0042	Left Side	☐
a0042_20...	01	a0042	Left Side	☐
a0042_20...	01	a0042	Left Side	☐
a0021_20...	01	a0021	Left Side	☐
a0021_20...	01	a0021	Left Side	☐
a0079_20...	01	a0079	Right Side	☐

Figure 46. View database form illustrating survey days in the left grid for Indo-Pacific humpback dolphin photo-ID surveys.

GIS Viewer



A mapping application in Discovery allows users to geographically plot sightings of individuals, group sightings, and effort information. ESRI shapefiles can be incorporated to include the study area, bathymetry, etc., to be displayed along with the individual sightings. To add shapefiles, see Map.

The top left table lists all individuals in the catalog for the default selected study area. Click on the individuals to view all sightings of that individual. The middle left table shows all sighting records. Select the group sightings to see positional information in relation to the group. The bottom table illustrates effort data and can display track information of the vessel for each survey. Press the "Ctrl" button to select more than one record to be displayed on map.

The "Plot" menu box allows you to view all sightings of all individuals ("All Individuals") as well as place lines between sequential sightings of the individual(s) ("Lines"). The sightings are color coded with the earliest sightings in red and later sighting dates in colors deviating further from red, based on the differences in dates between sightings.

Sightings can be filtered to include/exclude marginal, temporary, and non-verified images in the database (from the "Filter" menu box). To view the sighting information on the map, simply hover the mouse near the sighting and the date and group information will be displayed. Left-click / right-click on the map to zoom-in/out the map.

Chapter

8

Data Summary & Analysis

Several charts and basic analytical summaries are provided based on the centralized database; the data are given by year and total. The following are provided: the number of individuals, new individuals, images in database, # marginal, # temporary, # non-verified, resightings (min, max, and mean), and percentage (%) of aspects (left, right, fluke, etc), discovery curve, and individual sighting histories by date.

Sighting information is summarized by # of days, # of groups, average group size, and number of records for environmental, geographic, comment, and individual records. Frequency distribution of group sizes is provided as well as survey type, behavior, and species, and can be displayed as pie or frequency charts. Charts summarizing the information are displayed in excel-like grids that can be copied/paste to other applications to further display of the data. Data can be filtered to exclude/include study areas, species, types of behaviors, types of surveys, marginal images, temporary IDs, and non-verified images.

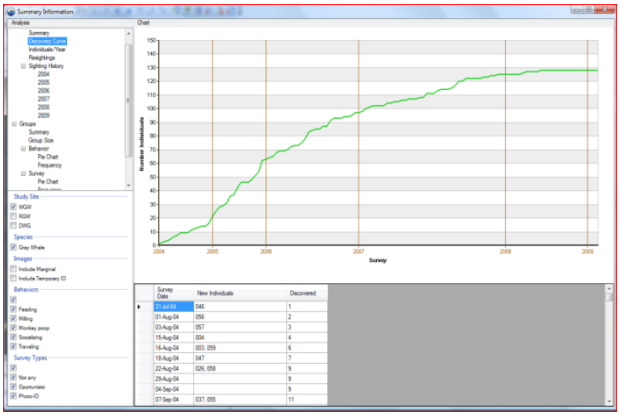
Image Database Summary

A summary tool is provided that summarizes aspects of the database for each year of the data. Most of the summary information is rather basic and self-explanatory.

Parameter	2004	2005	2006	2007	2008	Total
# Individuals	16	52	68	76	21	128
New Individuals	16	47	39	26	0	128
# Images in DB	40	146	205	327	35	753
# Marginal Images in DB	4	8	11	1	0	24
# TempID Images in DB	0	0	0	0	0	0
# Non-Verfied Images in DB	40	146	205	327	35	753
Min Resightings	1	1	1	1	1	1
Mean Resightings	1.81	1.62	1.90	2.43	1.38	3.56
Max Resightings	7	5	5	8	3	12
Left (%)	62.5	73.0	81.4	86.7	86.7	86.7
Right (%)	56.3	79.4	76.5	85.9	87.5	87.5
Fluke (%)	6.3	9.5	9.8	10.2	10.9	10.9
Dorsal_Fluke (%)	6.3	7.9	8.8	10.9	10.9	10.9

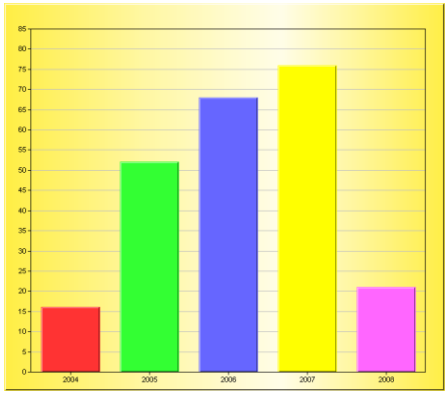
Figure 47. Summary of various data parameters from the catalog of identified individuals.

Discovery Curve



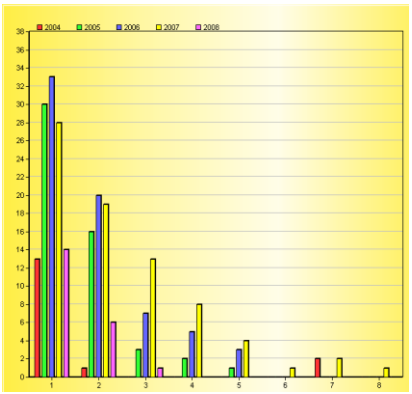
A discovery curve is an illustration of the cumulative number of individuals identified as a function of time and the cumulative effort expended. For Discovery, the effort consists of survey days. The number of survey effort can change per year, hence the different width of the years in the example graph. With each new individual identified and included in the database, the discovery curve increases for the value of one. If no new individuals have been identified during the survey, the discovery curve remains at the current number of identified individuals.

Number of Identified Individuals



The total number of individuals identified per year of survey effort can be graphed using the Individuals/Year. This simple bar graph illustrates the number of individuals on the y-axis and the year on the x-axis. A summary table is provided with the number of individuals per year of survey effort.

Number of Resightings



A resighting plot indicates how many times an individual was sighted. In Discovery, the resighting plot is constructed per each year of survey effort. The number of individuals is plotted on the y-axis and the number of resightings is plotted on the x-axis. In the example given on the right, a total of 33 individuals were sighted one time in 2006 and two individuals were resighted seven times in 2007 and 2004. A table indicating each individual and the number of resightings for each year is also provided along with the plot.

Sighting History

A plot of all sightings of all individuals in the database across time is provided. The y-axis shows individuals while the x-axis indicates each survey date. Red triangles represent a sighting of the individual, red hash marks represents within year sightings and blue bars represent between year sightings. This plot provides a complete history of each individual's sighting records. Sighting histories can also be selected for a given year of the survey data.

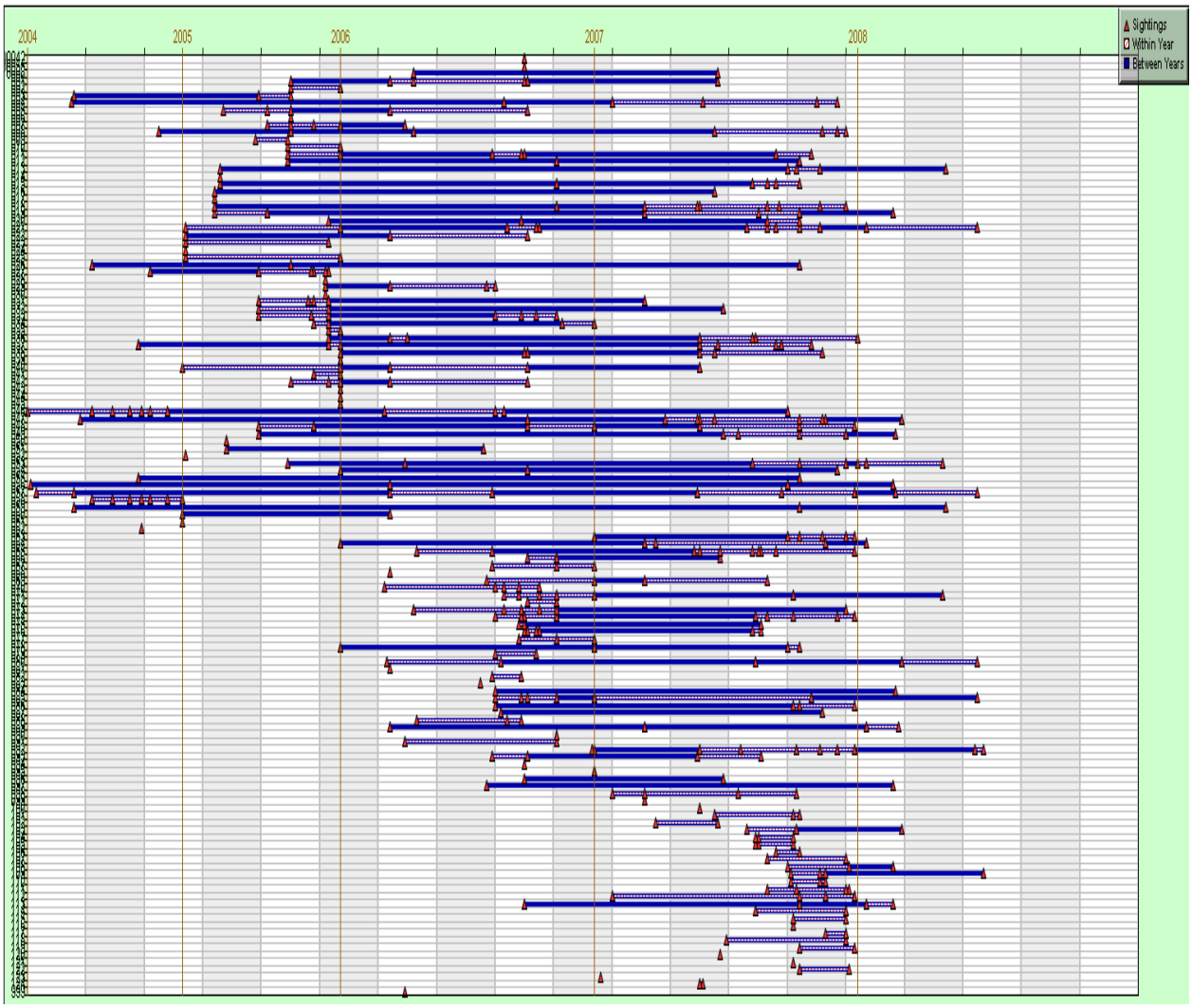


Figure 48. Graphical display of sighting histories for different individuals.

Group Summary

For group sighting information, a basic summary of the database information can be viewed. The summary consists of the "number of survey days", "number of groups encountered", the "mean group size", and the "number of records" for the individual, "environmental", "comment", "geographic", "photographic", and "miscellaneous information" stored in the database.

Group Size Plot

A frequency distribution plot is provided to examine the group size data given in the group sighting information stored in the database. For min, best, max estimates, the plot is based on the best group size estimates. A summary table is provided for each group size estimate in the database.

Behavioral State Plots (Pie, Frequency)

The behavioral state information for the sightings in the database can be summarized by either a pie or frequency distribution chart. Discovery will evaluate the number of occasions different behavioral states were observed and plot these by the desired plot. The numbers of occurrences are also shown in a table by year and total.

Survey Plots (Pie, Frequency)

Survey plots provide a summary of the different survey types conducted. These consider the different categories defined and provide a frequency distribution or pie chart illustrating the survey types for each sighting in the database.

Species Plot (Pie, Frequency)

Discovery's database can maintain data on a variety of species and studies. To examine the sightings for each species, a summary plot of the different species can be displayed by either a frequency distribution or a pie chart.

Chapter

9

Import

Import Images

A pre-existing database of images can be imported into Discovery with the import images option. The user can initially select a directory where the images are stored. This directory may contain subdirectories that have images of individuals in a subfolder or the entire directory may have all the images. Once the user selects the directory to import, the subfolder(s) of that directory will be illustrated. If the user selects a subfolder, the images in that subfolder will be displayed.

When importing images, it is handy to incorporate certain parameters:

1. *Aspect* - The aspect of the image (left, right, fluke, etc); the user can choose one of the following:

   A. Default - The user is asked to specify one aspect which is given to all images that are imported.

   B. Filename (Standard*)* - The user has a standardized naming format where the aspect is defined in the filename structure and therefore can be extracted out of the filename and placed as an aspect when imported.

   C. Filename (Search) - The user has aspect as part of the filename structure of the images, but the position of the aspect name in the filename is not always standard and therefore needs a search and match routine to extract the necessary information and record it with the image when imported.

2. *Individual* - Information pertaining to the individual identification ID to be incorporated with the imported image. Two options are available:

   A. Folder Name - the folder name in which the images are located is the name of the individual ID (can be either numbers or names).

   B. Filename (Standard) - The individual ID is placed in the filename with Discovery's format.

3. *Image* - Image options allow bitmap images to be converted to either TIFF or JPG formats. This conversion is convenient and allows the images to store EXIF data.

4. *Species* - The name of the Species of the imported images.

Once the user has selected all appropriate parameters, then he or she can simply click the "Start Import" button to include the images into the *Discovery Photo-Identification System.*

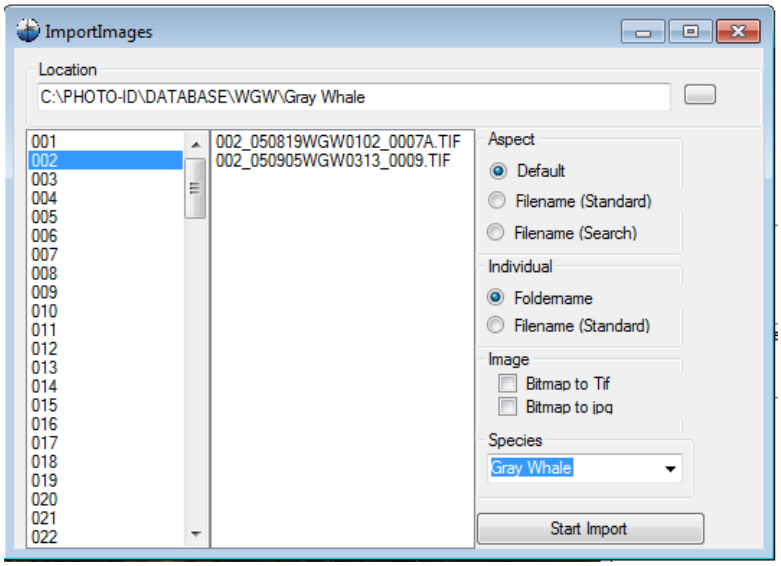


Figure 49. Illustration of importing images into the database.

Quick Procedures:

1. On the main menu, click "File" > "Import" > "Images" to open the Import Image interface.
2. Click on the button at top right to select the folder containing the images that are to be imported.
3. Subfolders will then be shown on the left box, click on the name of subfolder to be imported.
4. The images inside the folder will be shown in the right box.
5. Select appropriate mode of "Aspect", "Individual" on the right. This allows Discovery

to recognize the aspect and individual ID of each image automatically.

6. Select one of the two options under "Image" on the right of the window to convert bitmap (.bmp) images into JPEG or TIFF format, which allows EXIF data to be stored and used in the future.

7. Choose the "Species" in the drop down box at bottom right and click "Start Import" to import the images.

Import Images II

Similar to ***Import Images***, this function imports images from already existing database; however, ***Import Images II*** provides a more advanced and user-defined interface to input images. Besides extraction from filenames, users can choose to extract data from EXIF info, to fill the entire column with a single value, or type each cell directly.

1. *From EXIF* - User can extract any EXIF data to any parameters of image imported. The left drop-down menu lists all the fields of EXIF data to be extracted; while the right drop-down menu lists all parameters of image records.

2. *From Filename* - Information embedded in filenames can be extracted and input to any fields of the importing images. The start and end positions are specified under the same rationale as described in Incoming Renaming, with the first digit as 0 (zero).

3. *Fill Column* - This provides the tool to fill one field of all importing images with a single value. It is useful when inputting features that are shared among all images, e.g. study area, species, etc.

4. *Options* - If the user previously processed the image with Discovery and the image contains information in the EXIF that Discovery placed in the image (such as study area, group ID, individual ID name/number), then the information can be extracted automatically by selecting the "Discovery EXIF" option. Another option is available ("Move to Catalog") that will move the imported image to the defined catalog structure.

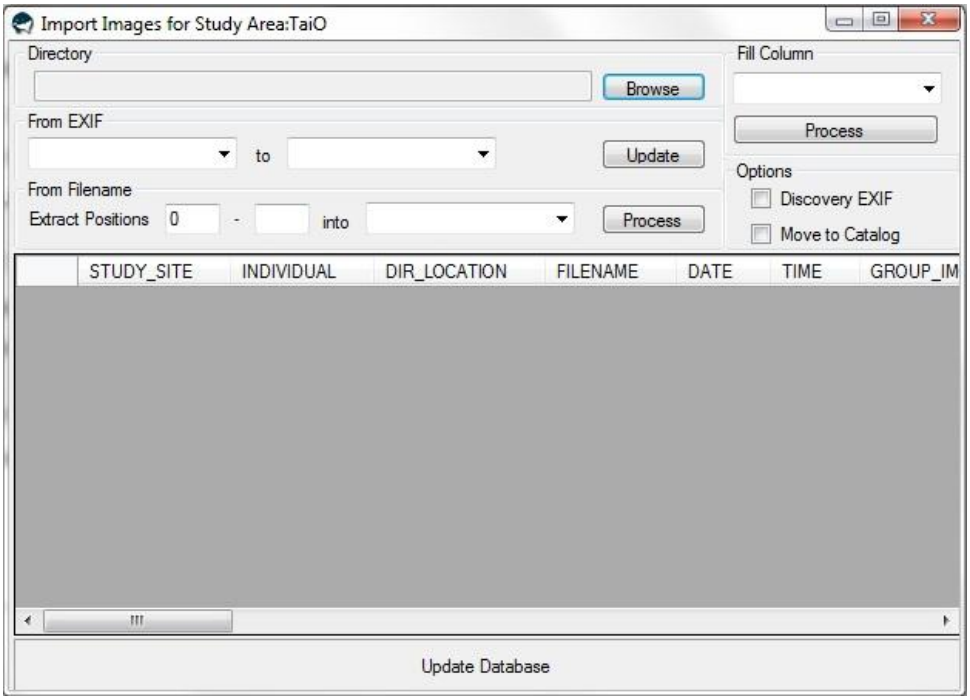


Figure 50. Illustration of importing images into database (Import Image II).

Quick Procedures:

1. On the main menu, click "File" > "Import" > "Images II" to open the Import Image II interface.

2. Click on the "Browse" button at the end of the row "Directory" to select the folder containing images to be imported.

3. Click the left drop-down menu under "From EXIF" and choose the field of EXIF data to be extracted, and click the right drop-down menu to select the parameter to be updated. Press the "Update" button to confirm. Repeat this for another extraction if needed.

4. Under the words "From Filename", input the start and end position in the first two boxes, and select the parameter to be updated by clicking the drop-down menu. Press "Process" button to proceed. Repeat this process for other extractions as needed.

5. Click the drop-down menu under the words "Fill Column" to select the parameter to be filled. Press "Process" button and then enter the value (or text). Repeat this process for other fields as needed.

6. Check the boxes “Discovery EXIF” and “Move to Catalog” if needed.

7. Click the "Update Database" button to import the image records to database.

Note: No records are saved before this step.

Import Databases

Existing data can be imported from Excel spreadsheets or Access databases. The system is designed to provide a flexible method to enter existing data into the application. You can import databases by going to File\Import\Database. Once the form is loaded, the user selects the file to be imported.

Select Table - Once the database is selected, the user will need to indicate which table of their database they wish to enter into Discovery's data database. For Excel spreadsheets, a table would be viewed as a worksheet (i.e. Sheet1, Sheet2, etc.).

Select Import Type - The import type refers to the various tables in the data database. These tables are:

Table 3. Definition of the various import types and their associated fields to facilitate importing data from various data sources.

Import Type	Description	Fields
Comments	Comments recorded during a sighting.	Date Time Study Group Species Comment
Crossmatch	Data that link two different individual IDs from two separate studies, that are actually the same individual	Study1 Individual1 Species Study2 Individual2

Environment	Environmental data that were recorded during a sighting. Environmental data can have varying fields dependent on the user defined parameters. Therefore, if the user wants to import certain types of fields, they must setup this option in setup (see Chapter 1) prior to importing their data.	Study Date Time Group user-defined parameters
Features	Unique features that apply to an individual. For example, sex, age, name, etc. Features can have varying fields dependent on the user defined parameters. Therefore, if the user wants to import certain types of fields, they must setup this option in setup (see Chapter 1) prior to importing their data	Study individual user-defined features
GeoPos	Geographic positions recorded during a sighting	Study Date Time Group Species Latitude Longitude
Images	Images of each sighting recorded in the database. These images are considered to be part of the catalog. Image descriptors can have varying fields dependent on the user defined parameters. Therefore, if the user wants to import certain types of fields, they must setup this option in setup (see Chapter 1) prior to importing their data	Study Individual Directory Location Filename Date Time Group Latitude Longitude Type Specimen Species Quality Distinctiveness Category Subcategory Aspect Marginal Temporary ID Verified User-defined Descriptors
IndSight	Individual information (such #Adults, #Calves, etc) that were recorded with the sighting of a group. IndSight can have varying fields dependent on the user defined parameters. Therefore, if the user wants to import certain types of fields, they must setup this option in setup (see Chapter 1) prior to importing their data	Study Date Time Group Species User-defined parameters

Photos	Information pertaining to the recorder, media, and photographs taken for each group sighting.	Study Date Recorder Group Media Card Species From To Photo	 Photo
Sightings	Information pertaining to a unique sighting of a group. Note: If you have only one group size estimate, place your group size estimate in Grp_Best and ensure you select the correct preference (see Chapter 1).	Study Date Start Stop Group Survey Species Grp_Min Grp_Best Grp_Max	 Time Time Type
Miscellaneous	Miscellaneous tables will be named based on the user-defined data type in the Setup. The fields will also be user-defined based on the setup. However, some standard fields do apply	Study Date Species Group User-defined parameters	

Once you select the table from your Access Table or Excel spreadsheet and the Import Type, a list of the fields (as noted above) displays. The field name, data type, and required designation are shown. The required designation refers to a value that either must be placed in the field or not. A True value means the field is required; a False value means the data is not required. If the user does not have data in their database, but need to enter it into Discovery's database, they can place a default value in the value field. For example, if the user did not specify Species in their data since it was all one species, they can enter a default species name to be recorded with each of their imported entries. The import field reads each field in the Access database and assumes the first row in the Excel Spreadsheet is the field name.

Once the user have matched up their fields to Discovery's fields for an Import Type or at least placed default values, the user can simply click on the import button to transfer the data into Discovery's database.

Note: The Import Database does not check if the input data are duplicate records in the database. Therefore, if identical parameters are being imported into a Discovery database then multiple records will exist. It is recommended to use the "Import Discovery database" method (see Import Discovery Database).

Quick Procedures:

1. In the main menu, click "File" > "Import" > "Database" to open the Import Database interface.

2. Click on the "Browse" button and choose the Excel (.xls) or Access database (.mdb) file.

3. Choose the appropriate table to be imported in the drop down button in "Table" on left.

4. Choose the appropriate "Import Type" from the drop down button (refer to Table 3 for description of the import types).

5. Once an import type is selected, the details of the import type will be shown in the grey box at the bottom.

6. In each row:

   a. Fill in the "Value" field by clicking on the field and then typing. This value will be applied to every entry in the importing table.

   or

   b. Select from the drop down box in the "Import Field", which tells Discovery which field from the source table is equal to the field being entered in the destination table (shown on the left of the row).

   or

   c. Leave the row unchanged.

   Note: If the field under column "Required" is "True", the option (c) above cannot be used; the user must either type "Value", see step 6(a) above, or select from the Import Field, see step 6(b) above.

7. Finish the Import procedure by clicking the "Import" button on middle right.

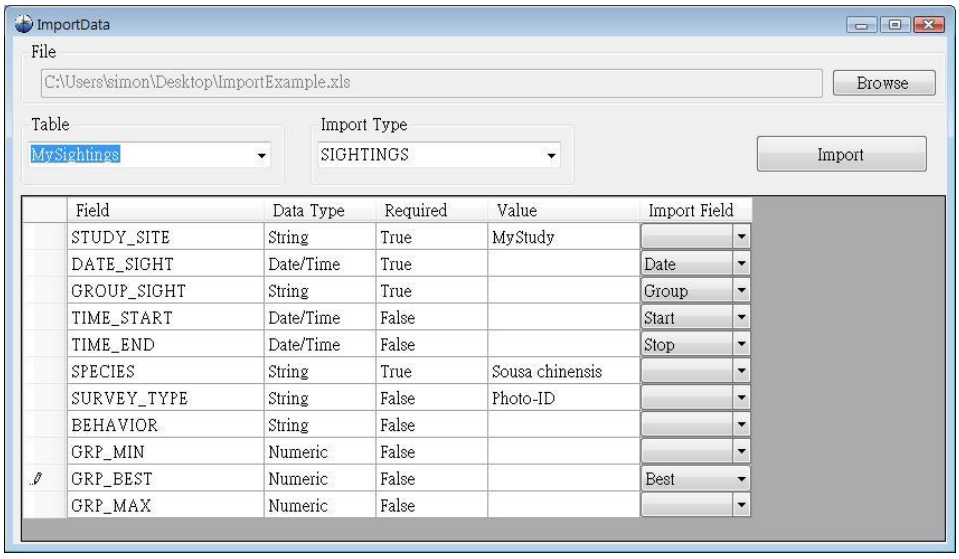


Figure 51. A form illustrating importing data from one Access database or Excel spreadsheet into Discovery's database.

Importing Sightings (General)

This option allows the user to import sighting information from Excel where the first row header cells match Discovery's database field columns names. This procedure will minimize selection of various data fields as noted above.

Quick Procedures:

1. Click "File" > "Import" > "Sightings (General)" on the main menu to open the Import Sightings interface.

2. Select the file to be imported.

	A	B	C	D	E	F	G	H	I	J	K
1	STUDY_SITE	DATE_SIGHT	TIME_START	TIME_END	GROUP_SIGHT	SPECIES	SURVEY_TYPE	BEHAVIOR	GRP_MIN	GRP_BEST	GRP_MAX
2	Me Study	5/5/2009	12:22:00	13:23:00	3	Pickle heads	ds	fds	0	0	0
3	Me Study	5/6/2009	12:22:00	13:23:00	4	Pickle heads	ds	fds	1	2	3
4	Me Study	5/7/2009	12:22:00	13:23:00	5	Pickle heads	ds	fds	2	4	6
5	Me Study	5/8/2009	12:22:00	13:23:00	6	Pickle heads	ds	fds	3	6	9
6	Me Study	5/9/2009	12:22:00	13:23:00	7	Pickle heads	ds	fds	4	8	12
7	Me Study	5/10/2009	12:22:00	13:23:00	8	Pickle heads	ds	fds	5	10	15
8	Me Study	5/11/2009	12:22:00	13:23:00	9	Pickle heads	ds	fds	6	12	18
9	Me Study	5/12/2009	12:22:00	13:23:00	10	Pickle heads	ds	fds	7	14	21
10	Me Study	5/13/2009	12:22:00	13:23:00	11	Pickle heads	ds	fds	8	16	24
11	Me Study	5/14/2009	12:22:00	13:23:00	12	Pickle heads	ds	fds	9	18	27
12	Me Study	5/15/2009	12:22:00	13:23:00	13	Pickle heads	ds	fds	10	20	30
13	Me Study	5/16/2009	12:22:00	13:23:00	14	Pickle heads	ds	fds	11	22	33
14	Me Study	5/17/2009	12:22:00	13:23:00	15	Pickle heads	ds	fds	12	24	36
15	Me Study	5/18/2009	12:22:00	13:23:00	16	Pickle heads	ds	fds	13	26	39
16	Me Study	5/19/2009	12:22:00	13:23:00	17	Pickle heads	ds	fds	14	28	42
17	Me Study	5/20/2009	12:22:00	13:23:00	18	Pickle heads	ds	fds	15	30	45
18	Me Study	5/21/2009	12:22:00	13:23:00	19	Pickle heads	ds	fds	16	32	48
19	Me Study	5/22/2009	12:22:00	13:23:00	20	Pickle heads	ds	fds	17	34	51
20	Me Study	5/23/2009	12:22:00	13:23:00	21	Pickle heads	ds	fds	18	36	54
21	Me Study	5/24/2009	12:22:00	13:23:00	22	Pickle heads	ds	fds	19	38	57
22	Me Study	5/25/2009	12:22:00	13:23:00	23	Pickle heads	ds	fds	20	40	60
23	Me Study	5/26/2009	12:22:00	13:23:00	24	Pickle heads	ds	fds	21	42	63
24	Me Study	5/27/2009	12:22:00	13:23:00	25	Pickle heads	ds	fds	22	44	66
25	Me Study	5/28/2009	12:22:00	13:23:00	26	Pickle heads	ds	fds	23	46	69
26	Me Study	5/29/2009	12:22:00	13:23:00	27	Pickle heads	ds	fds	24	48	72
27	Me Study	5/30/2009	12:22:00	13:23:00	28	Pickle heads	ds	fds	25	50	75
28	Me Study	5/31/2009	12:22:00	13:23:00	29	Pickle heads	ds	fds	26	52	78

Figure 52. An example Excel spreadsheet illustrating columns of data for importing sighting information.

Import Discovery Database

This option allows the user to import other Discovery database(s) into the current Discovery database being used by the system (i.e. external Discovery database into the "master" database). Compared to other import options, it acts as a quick way to import a Discovery database. This import routine, includes data "as is" without checks for duplicate records. In addition, only user-defined parameters of the master database (the default database loaded by Discovery) will be imported. For example, if the import database has a field name "Field1" and this field is not defined in the master database, Discovery will not import "Field1". Therefore, no additional fields or tables will be created/imported with this option. Users should check the consistency between the two databases prior to importing and create the additional fields in setup, if needed.

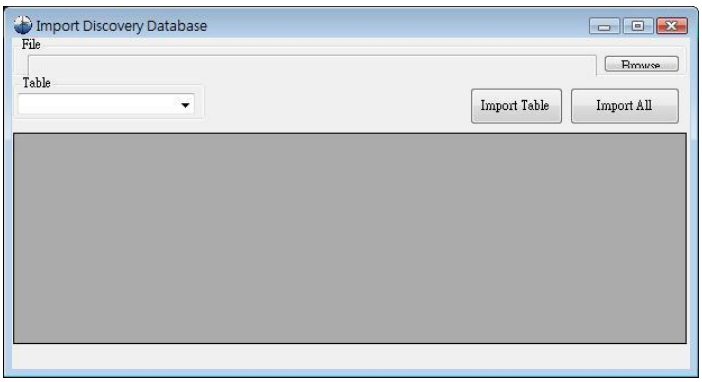


Figure 53. Interface of importing a Discovery database.

Quick Procedures:

1. On the main menu, click "File" > "Import" > "Discovery Database" to open the Import Discovery Database interface.

2. Click on the "Browse" button and choose the access database (.mdb) file.

3. Click "Import All" button on the middle right to import all tables from the database file.

   or

   Choose the appropriate table of the import file to be imported in the drop down button in "Table" on left.

4. Click "Import Table" button to import the table.

Export

Currently, all data is stored in a Microsoft Access database since most users are familiar with Access. Export routines are intended to export data into Excel and other applications such as SOCPROG, MARK, ARC GIS, and other applications.

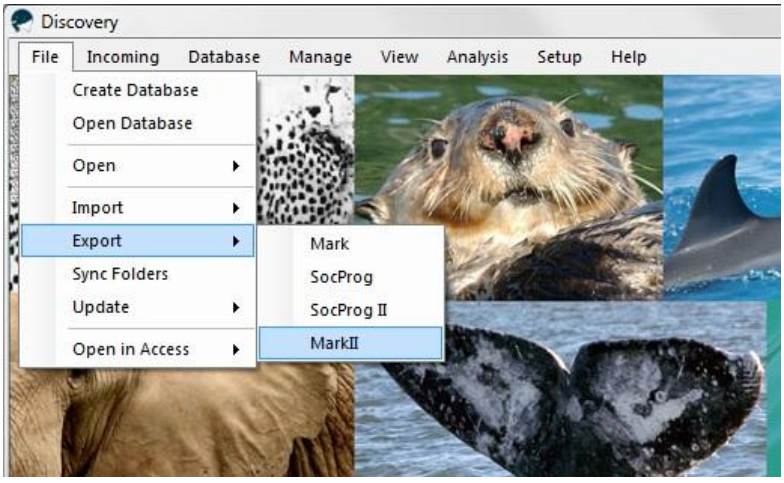


Figure 54. Exporting database information into an output file for program MARK or SocProg.

MARK

INP files (.inp) can be generated based on the images entered into the Discovery database. Filtering for Marginal, Temporary ID, and Verified images can be performed when prompted at the time the output file is generated. Two export methods of MARK output files are available:

1. *Simplified (Mark)* - Provides a table of individual sighting histories according to user-defined length of occasion.
2. *User-defined (Mark II)* - Provides a platform for users to specify the output file to fit closed, open, multi-state or robust design models.

Note: The smallest sampling unit for both Mark export methods is by survey day - i.e. each survey is considered as one sampling occasion in program MARK. Longer intervals can be included.

Quick Procedures:

1. In the main menu, click "File" > "Export" > "Mark" or click "File" > "Export" > "Mark II" to start exporting database information.
2. If "Mark" is chosen, user will be asked to specific the survey day interval (which is equivalent to pooling data by regular interval in "Mark II" described below). User will also be asked if Marginal, Temporary ID or non-Verified images are to be exported. A window will appear after successfully creating the export file.
3. If "Mark II" is selected, a window will appear for users to specify the output files.
4. Click the checkboxes under "Options" to include/exclude Marginal, non-Verified and Temporary ID images.
5. The "Model" option can be either Single State or Multi-State, which disable/enable the additional states below.
6. The "Variables/States" option allows users to add variables (which is used as covariates, groups or mixtures according to users' setting in program MARK) and states (which is used in multi-state models). Click "Table" and "Field" to select the parameters in database to be exported as variables or states.

7. In the "Pool Data By" option:

   a. "Regular Intervals" - The day interval can be specified and all data in database will be pooled by regular interval.

   b. "Irregular Intervals" - The right upper side of the window will be changed to the list of survey days. This option allows users to manually pool data by irregular interval. Left click on a survey day, press and hold "Ctrl" or "Shift" button, and then left click on other survey day(s) to select more than 1 survey. Right click on the selected survey(s) and select "Add Occasion", and a row of occasion will be added to the right lower table. Repeat to add more occasions.

   c. "Robust" - The right upper side of the window will become a list of survey days. This option allows users to manually pool data to primary and secondary occasions, which are used in robust design models in program MARK.

      Left click on a survey day, press and hold "Ctrl" or "Shift" button, and then left click on other survey day(s) to select more than 1 survey. Right click on the selected survey(s) and select "Add Occasion", and a primary occasion composing of several secondary occasions (surveys selected above) will be added to the right lower table.

      4 columns will appear in the table:

      1. Occasion - Primary occasion, which can also be distinguished by the colors.

      2. Date - Surveys pooled as secondary occasions within the corresponding primary occasion.

      3. Secondary Occasion - Number of secondary occasions.

      4. Day interval - Number of days between two adjacent primary occasions, which is calculated using the mid-point of primary occasions. For example, 12th April is the mid-point of 11th-13th April (3 surveys in one primary occasion) (see Figure 57).

   Repeat to add more occasions.

8. Press "Process" button to export database data. A preview of the export file will appear on the right side of the window. Note that the export data have not been output at this stage. Press the "Save" button to save the INP file.

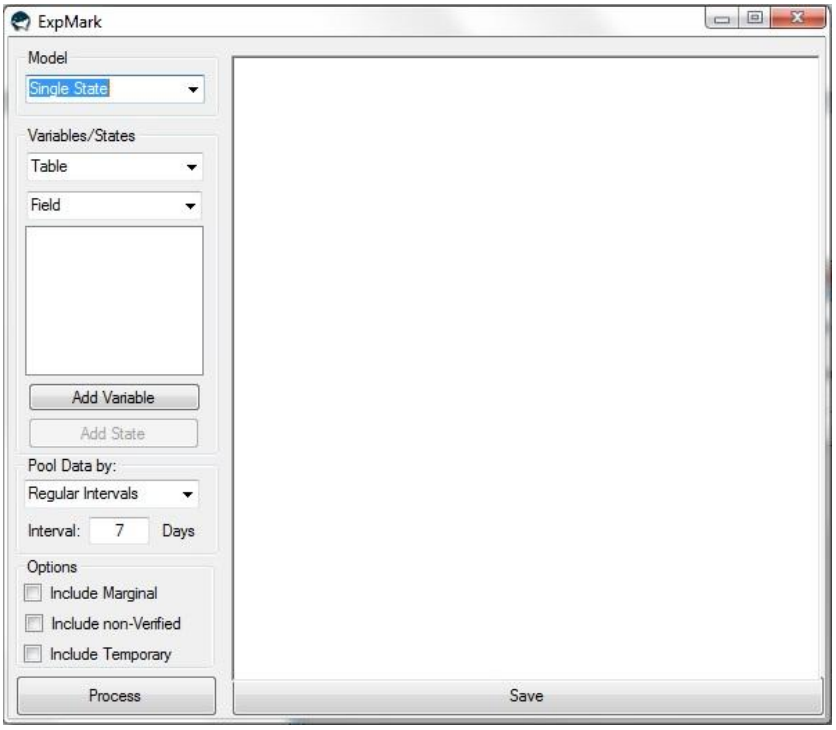


Figure 55. Mark II form for exporting data for program MARK.

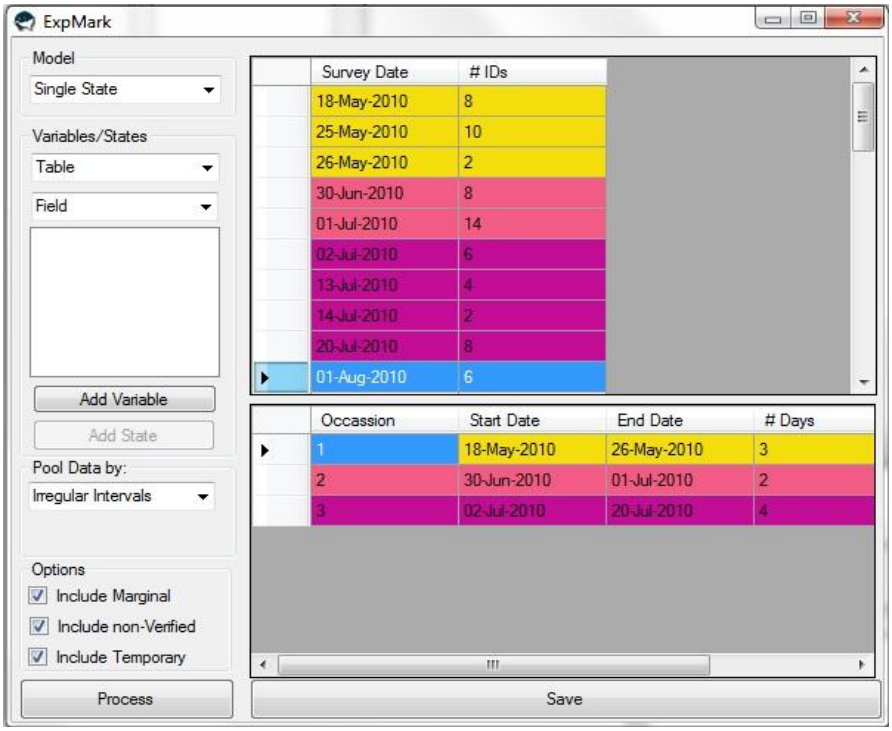


Figure 56. Mark II form illustrating pooling survey days for MARK output files by user-defined irregular intervals.

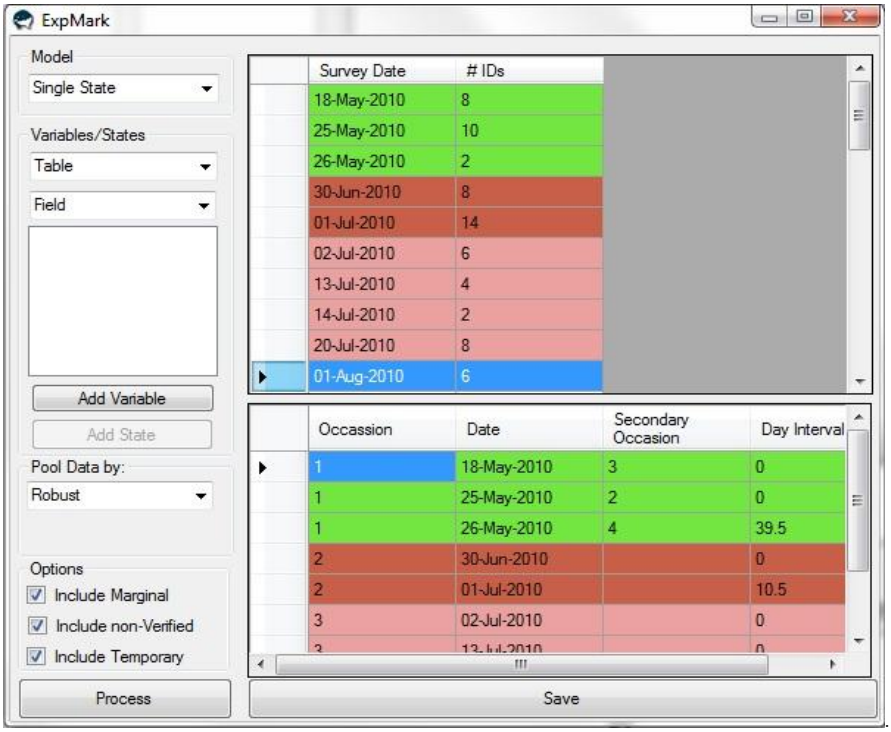


Figure 57. Discovery's Mark II form for pooling data to primary and secondary occasions to export in MARK files for Robust design models.

SOCPROG

Microsoft Excel files[1] can be generated based on the images entered into the Discovery database. Similarly as with the output file for MARK, filtering for Marginal, Temporary ID, and Verified images can be done when prompted at the time the output file is generated. Two export methods for SOCPROG output files are possible:

1. *Simplified (SocProg)* - Provides the Date, Group Number, Group Membership (group defined as 'in group together'), Quality and Distinctiveness.
2. *User-Defined (SocProg II)* - User-specify dynamic output where user can select what and how data are included in the primary and supplementary output file(s).

[1] Only Microsoft Excel 2003-2007 (*.xls) is currently supported. New formats, such as Excel 2010; .xlsx, is not currently supported.

Quick Procedures:

1. In the main menu, click "File" > "Export" > "SocProg" or click "File" > "Export" > "SocProg II" to start exporting the information.

2. If "SocProg" is chosen, options will appear to ask the user if Marginal, Temporary ID or non-Verified images are to be exported. Once this is confirmed, the export can proceed.

3. If the user chooses "SocProg II", a window with several options will appear for the user to specify what and how data are to be exported.

4. The "Output Type" for SocProg can be chosen as either Linear, Group or Supplementary.

5. Click "Source" and "Field" to select the "Variables" to be exported. Note that the choice of "Variables" changes under different "Output type".

6. If the database has more than one value for the same variable (e.g. 2 images, thus 2 values of Quality and Distinctiveness for the same individual in the same sighting), the user can go to "Multiple values" and select "First", "Mean" or "Last" to specify which value to be exported. Users can select different methods for different variables (see Figure 59 for an example).

7. Press "Add" to include the "Variables" in the method of "Multiple values" previously selected.

8. Click the checkboxes under "Filters" to include the Marginal, Temporary IDs or Non-Verified records, which are then included in addition to the default values (i.e. as the default, with no boxes checked, only the best quality records are included).

9. Press "Export" to proceed with the output file.

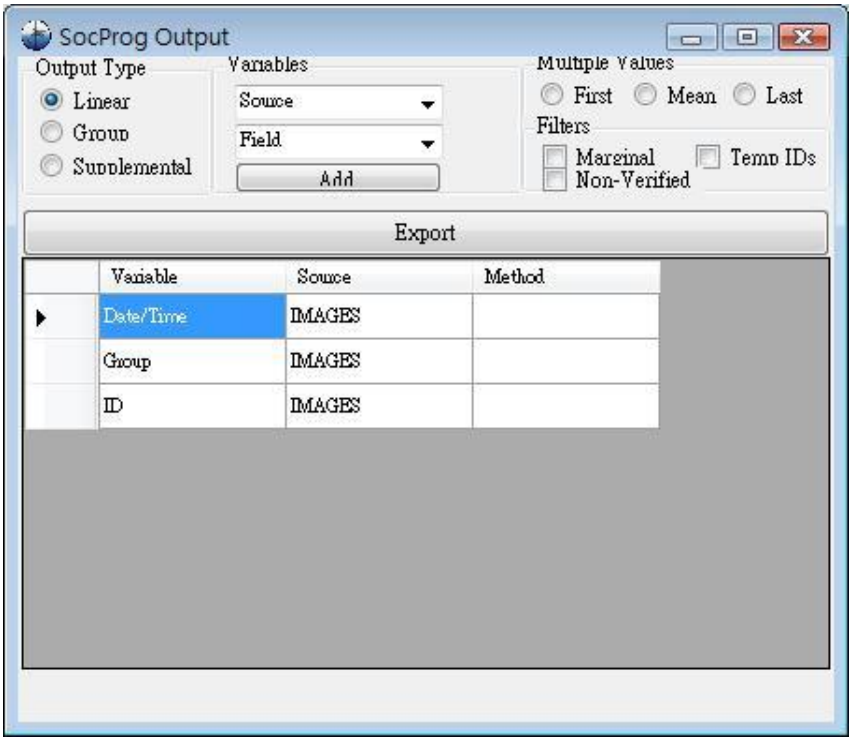


Figure 58. Defining output file for SOCPROG in function SocProg Output.

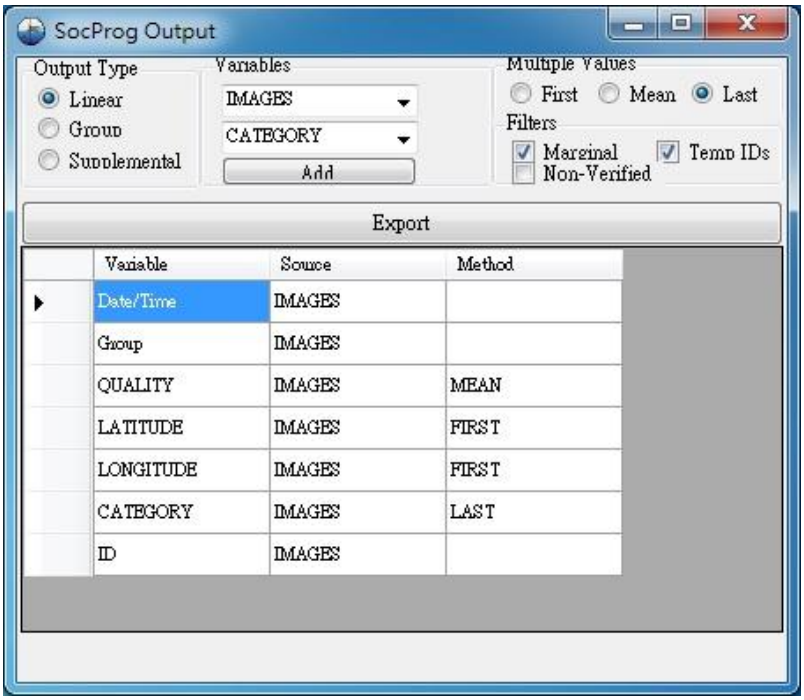


Figure 59. Exporting selected and user-defined database information into an output file for program SOCPROG using Discovery (with user-defined SocProg II).

Update

This function provides an option for users to quickly synchronize image-related information to/from the database.

Exif Based on DB (Database)

This function updates the EXIF data information based on the Discovery database. Information changed in the database does not automatically update the EXIF data. Such information could be related to the geographic position, study area, group number, date, etc. To keep the image information up-to-date with the database, users can use this function to quickly update the image's EXIF details.

Quick Procedures:

1. On the main menu, click "File" > "Update" > "Exif based on DB".
2. A green progress bar will appear, illustrating the completion of the update process.
3. When the progress bar reaches 100%, a window showing the number of images updated will appear.

GPS Effort to DB (Database)

To quickly update the GPS locations (latitude and longitude) of ALL cataloged image records (that falls within the timeframe of a GPX file) in the database, users can select a GPX file and choose to update the GPS locations based on the track, waypoint and/or route data from a GPX file.

Quick Procedures:

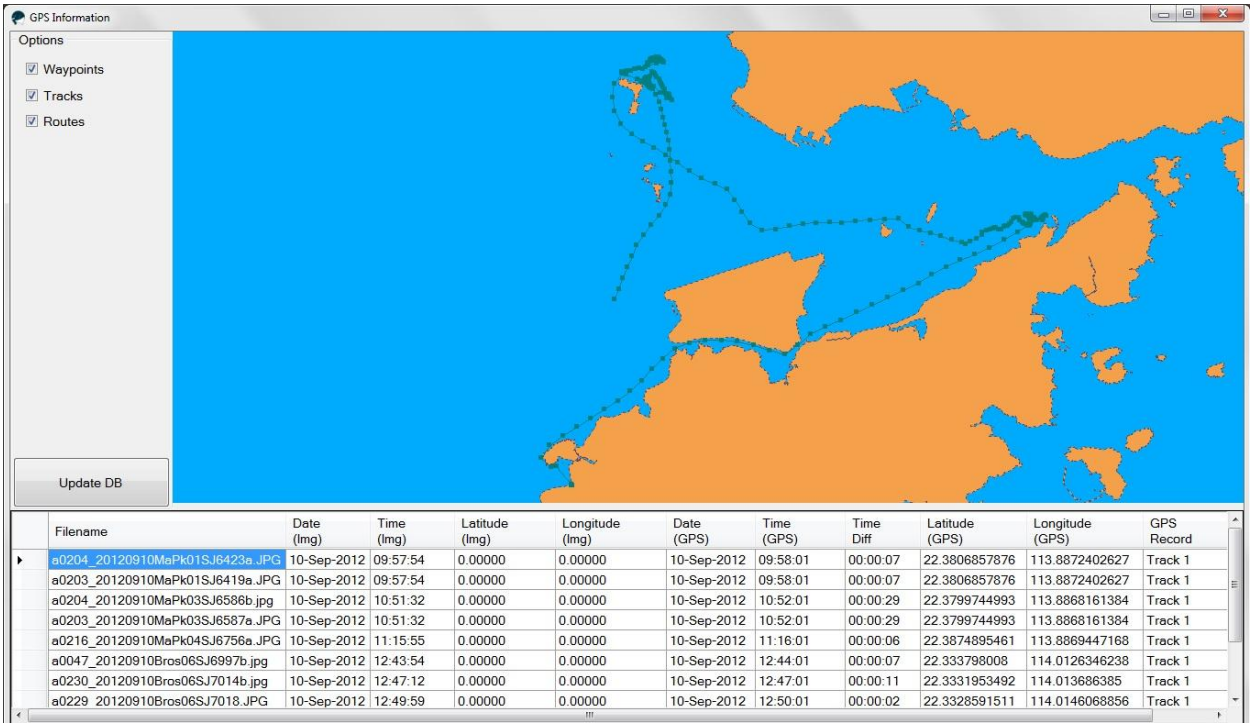
1. On the main menu, click "File" > "Update" > "GPS effort to DB".
2. User will be asked to find and select a GPX file.

3. A window with a map showing the study area and a list of images will appear. The images are automatically selected based on the time. For example, suppose the selected GPX file stores the GPS information of only one survey day (e.g. 25-Mar-2013), all cataloged images that are taken on that day will be automatically filtered from the database, and the records will appear on the list.

4. Under "Option" of the left upper corner, user can check the boxes "Tracks", "Waypoints" and "Routes" to include/exclude these information in the update.

5. In the list of records at the bottom, the columns "Latitude (GPS)" and "Longitude (GPS)" represent the GPS location extracted from the GPX file, while the column "GPS record" indicates the origin of the particular GPS location (e.g. the track number). The column "Time Diff" gives the difference between the image time and the time of the closest point of GPS location. The last column on the right "Add to DB" allows user to choose the image records to be updated.

6. Press the "Update DB" button for the update to proceed.



Filename	Date (Img)	Time (Img)	Latitude (Img)	Longitude (Img)	Date (GPS)	Time (GPS)	Time Diff	Latitude (GPS)	Longitude (GPS)	GPS Record
a0204_20120910MaPk01SJ6423a.JPG	10-Sep-2012	09:57:54	0.00000	0.00000	10-Sep-2012	09:58:01	00:00:07	22.3806857876	113.8872402627	Track 1
a0203_20120910MaPk01SJ6419a.JPG	10-Sep-2012	09:57:54	0.00000	0.00000	10-Sep-2012	09:58:01	00:00:07	22.3806857876	113.8872402627	Track 1
a0204_20120910MaPk03SJ6586b.jpg	10-Sep-2012	10:51:32	0.00000	0.00000	10-Sep-2012	10:52:01	00:00:29	22.3799744993	113.8868161384	Track 1
a0203_20120910MaPk03SJ6587a.JPG	10-Sep-2012	10:51:32	0.00000	0.00000	10-Sep-2012	10:52:01	00:00:29	22.3799744993	113.8868161384	Track 1
a0216_20120910MaPk04SJ6756a.JPG	10-Sep-2012	11:15:55	0.00000	0.00000	10-Sep-2012	11:16:01	00:00:06	22.3874895461	113.8869447168	Track 1
a0047_20120910Bros06SJ6997b.jpg	10-Sep-2012	12:43:54	0.00000	0.00000	10-Sep-2012	12:44:01	00:00:07	22.333798008	114.0126346238	Track 1
a0230_20120910Bros06SJ7014b.jpg	10-Sep-2012	12:47:12	0.00000	0.00000	10-Sep-2012	12:47:01	00:00:11	22.3331953492	114.013686385	Track 1
a0229_20120910Bros06SJ7018.JPG	10-Sep-2012	12:49:59	0.00000	0.00000	10-Sep-2012	12:50:01	00:00:02	22.3328591511	114.0146068856	Track 1

Figure 60. GPS waypoint, track, and route information for photo-ID surveys in Hong Kong illustrating the GPX file import function in Discovery's "GPS effort to DB" form.

Chapter

10

Image Formats

The imaging component of the application is an independent (third-party) component that was incorporated into the system. The component continues to be updated and as such it supports the latest image file formats of the majority of digital camera manufactures. Table 4 outlines the current supported formats that can be displayed and processed with the Discovery photo-identification system.

Table 4. Discovery's supported image formats.

File Type(s)	Image Format
AWD	Microsoft FAX
BMP, DIB	Windows/OS2 Bitmap
BRK, BRT	Brooktrout FAX
BW, RGB, RGBA	SGI Format
CAL, CALS	CALS
CAM	Casio QVxx
CRW, CR2, TIF	Canon Raw
DB	Windows XP Thumbnails
DCR, TIF, TIFF	Kodak Professional
DCX	DCX
DNG	Adobe Digital NeGative
EMF, WMF	Windows Metafile
ERF	Epson ERF
FPX	FlashPix
FXD, FXM, FXR, FXS	WinFax
GIF	Graphics Interchange Format
HDR	Radiance Picture File Format
ICO	Windows Icon
IFF	Commodore-Amiga IFF
IMG, LV	LaserView Fax
J2C, J2K, JP2, JPC, JPF, JPX	JPEG 2000
JPE, JPG, JPEG, THM	JPEG
KDC	Kodak Digital Camera

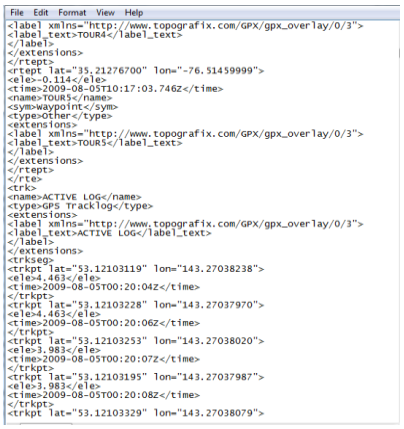
MBM	Psion MultiBitMap
MRW	Konica Minolta Raw
NEF	Nikon Raw
ORF	Olympus Raw
PBM	Portable Bitmap
PCD	Photo-CD
PCT, PIC, PICT	Apple Picture Format
PCX	ZSoft PCX
PDD, PSD	Adobe Photoshop Format
PEF	Pentax Raw
PGM	Portable Greymap
PMP	Sony Image Format
PNG	Portable Network Graphics
PNM, PPM	Portable Pixelmap
PSP, PSPIMAGE	Paint Shop Pro Format
RAF	FujiFilms Raw
RAS	Sun Raster Format
RAW	Leica and Panasonic Raw
SFW	Seattle Filmworks Format
SRF	Sony Raw
TGA	Truevision TARGA
TIF, TIFF	Tagged Interchange File Format
WBM, WBMP	WAP WBMP
X3F	Foveon Image Format
XBM	X Bitmap
XPM	X PixMap
XWD	X Windows Dump

EXIF (EXchangeable Image File format)

EXIF constitutes the information that is stored in the header of digital images. Many cameras store information into each image that is captured in JPEG, RAW or TIFF formats. Furthermore, EXIF data usually has additional specific metadata tags associated with it. EXIF data usually consist of the date and time information, camera settings, a thumbnail image, descriptions and copyright information, and geo-location (latitude/longitude) information. Discovery makes use of these data at every opportunity to streamline data entry as well as store useful information such as study area, group ID, species, and geo-location information.

GPS Data Viewer and Extractor

Discovery can view and extract Waypoints, Tracks, and Route GPS information from data stored in common file formats. There are a variety of tools freely available to download GPS data from a GPS device. EasyGPS (http://www.easygps.com/) is one common GPS download software that is freely available. Discovery currently supports importing the following file formats:



1. *GPX (the GPS Exchange Format)* - Waypoint, Track, and Route information is stored in an XML format for the interchange of GPS data between applications and web services on the internet. Discovery supports GPX 1.1 schema.

2. *Microsoft Access and Excel files* - Access and Excel files can be used to import GPS information. The format for these files needs to have standard header or column name information. The names of the header or column should be similar to example illustrated. However, not all values need to be included as illustrated in the example with missing Date_UTC and Time_UTC. A minimum of a date, time, latitude, and longitude is needed.

	A	B	C	D	E	F	G	H
1	Name	Date_Local	Time_Local	Date_UTC	Time_UTC	Latitude	Longitude	Elevation
2	MyDate	12/1/2010	11:30			53.233232	143.23323	10
3	MyDate	12/1/2010	11:31			53.233232	143.23323	11
4	MyDate	12/1/2010	11:32			53.233232	143.23323	12
5	MyDate	12/1/2010	11:33			53.233232	143.23323	13
6	MyDate	12/1/2010	11:34			53.233232	143.23323	5
7	MyDate	12/1/2010	11:35			53.233232	143.23323	8
8	MyDate	12/1/2010	0:12			53.233232	143.23323	9
9	MyDate	12/1/2010	0:13			53.233232	143.23323	6
10	MyDate	12/1/2010	0:14			53.233232	143.23323	4
11	MyDate	12/1/2010	0:15			53.233232	143.23323	3

Discovery's GPS Viewer and Extractor is offered in a number of areas to streamline data entry:

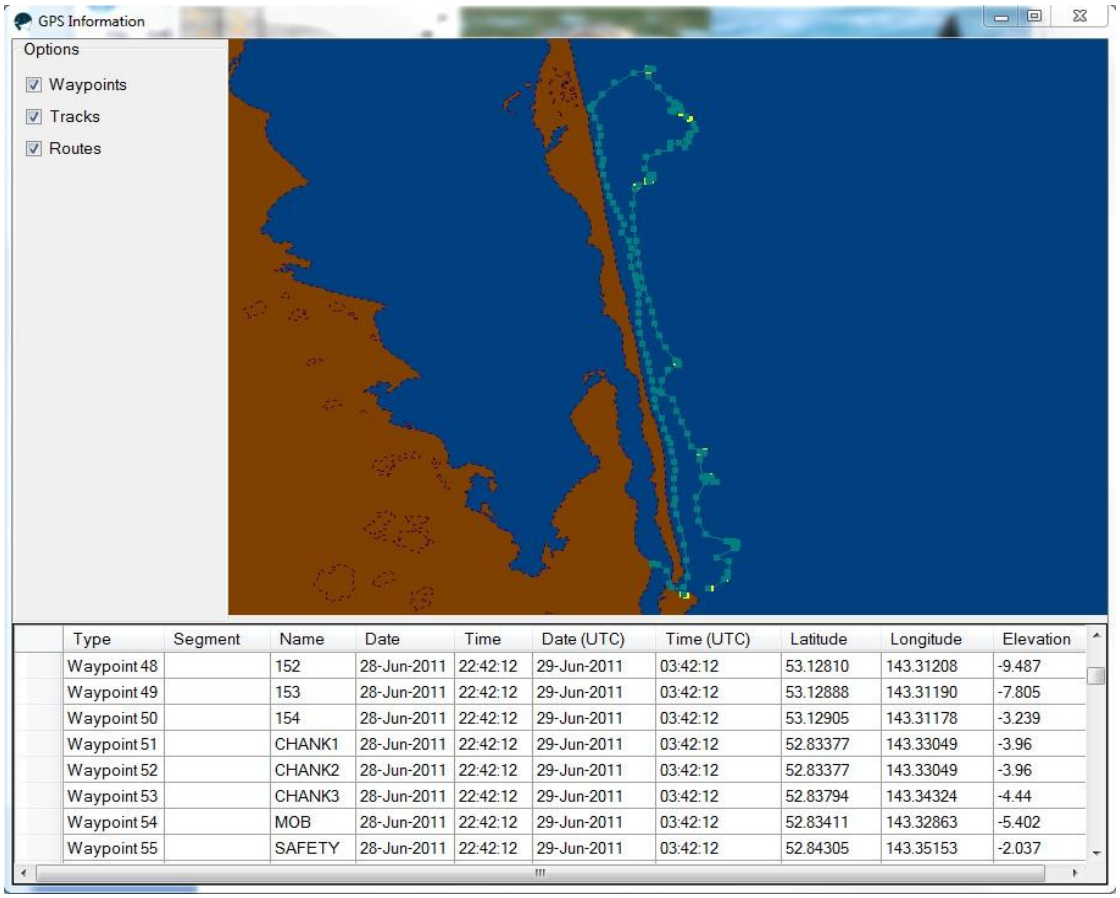
1. *View GPS File* - to view your GPS data, you can open the file from the main menu under File\Open\GPS File. Select the desired file format and file to view the GPS data.

2. *Extract GPS Information for Images* - Data collected from the GPS device can be used to incorporate the geo-location into the EXIF of the image as well as include into the position in the database. The EXIF date/time values can be compared to the GPS data and the closest time within that day is then extrapolated and incorporated into the EXIF of the image. The user can evaluate the time difference between the image time and the GPS time and decide if they want to include the geographic location as part of the image location. If no geographic position is found within that

date, then the program will return a "NA" (Not Any) position. You can incorporate the GPS data to the images when renaming (see Renaming) or when adding an image to the database (see Add Image to Database).

3. *Import GPS Data for Group Sighting* - Geographic positions recorded during a sighting can be easily added to the sighting information. In the Position tab in the sighting form, you can click on Import to extract the positions. The specified date, start time, and end time will be evaluated and the geographic locations during those times will be suggested to include as geographic coordinates information for that sighting. If the GPS data does not include positions during that time, no geographic positions will be given.

Note: The GPS Data Viewer's mapping system uses ESRI Shapefiles that the user specified in Parameter Setup as background images to plot the GPS points. If a shapefile is not defined for the area, the GPS Viewer may be unable to plot the GPS points. It is *recommended* to have a default shapefile for your study area.



Type	Segment	Name	Date	Time	Date (UTC)	Time (UTC)	Latitude	Longitude	Elevation
Waypoint 48		152	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	53.12810	143.31208	-9.487
Waypoint 49		153	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	53.12888	143.31190	-7.805
Waypoint 50		154	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	53.12905	143.31178	-3.239
Waypoint 51		CHANK1	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	52.83377	143.33049	-3.96
Waypoint 52		CHANK2	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	52.83377	143.33049	-3.96
Waypoint 53		CHANK3	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	52.83794	143.34324	-4.44
Waypoint 54		MOB	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	52.83411	143.32863	-5.402
Waypoint 55		SAFETY	28-Jun-2011	22:42:12	29-Jun-2011	03:42:12	52.84305	143.35153	-2.037

Figure 61. Viewing GPS information and other map related data for a gray whale photo-id survey off Sakhalin Island, Russia.

Quick Procedures:

Viewing GPS data

1. Click "File" on the main menu
2. Click "Open > GPS File"
3. Choose the appropriate file to view

Viewing GPS data in database

1. Define a shape file as the default map of the software (see Map).
2. On the main menu of Discovery, click "View" > "Map"

Import GPS data from GPS device into EXIF of images

1. See Renaming in Parameter Setup, choose "GPS from file" in the box "Change EXIF" at bottom right.

or

2. See Add Image to Database in Chapter 3; in the provided information form, before adding an image into the database, click on the button "From File" on the top middle part of the form.

Import GPS data from GPS device into Group sighting information

See Geographic Position in Adding Sighting Information in Chapter 5.

Note: The local date and time of the imported GPS data are corrected based upon the computer's time zone UTC offset. Users who potentially collect their data in one area and process in another (i.e. different time zones), make sure the computer's time zone is the same as the one collected to ensure proper conversion of the date/time GPS information.

Acknowledgements

We gratefully acknowledge the financial support received from the National Research Foundation (NRF) South Africa, Research Grants Council (RGC) Hong Kong and Ocean Park Conservation Foundation Hong Kong (OPCFHK) as well as the facility support provided by The Swire Institute of Marine Science and The University of Hong Kong.