APPLICATION

Cpw Photo Warehouse: a custom database to facilitate archiving, identifying, summarizing and managing photo data collected from camera traps

Jacob S. Ivan* and Eric S. Newkirk

Colorado Parks and Wildlife, Fort Collins, CO 80526, USA

Summary

1. Contemporary methods for sampling wildlife populations include the use of remotely triggered wildlife cameras (i.e., camera traps). Such methods often result in the collection of hundreds of thousands of photos that must be identified, archived, and transformed into data formats required for statistical analyses.

2. Cpw Photo Warehouse is a freely available software based in Microsoft Access® that has been customized for this purpose using Visual Basic® for Applications (VBA) code. Users navigate a series of point-and-click menu items that allow them to input information from camera deployments, automatically import photos (and image data stored within the photos) related to those deployments, and store data within a relational database. Images are seamlessly incorporated into the database windows, but are stored separately from the database.

3. The database includes menu options that (i) facilitate identification of species within the images, (ii) allow users to view and filter any subset of the database on study area, species, season, etc., and (iii) produce input files for common analyses such as occupancy, abundance, density and activity patterns using Programs MARK, PRESENCE, DENSITY and the r packages ‘secr’ and ‘overlap’.

4. Our database makes explicit use of multiple observers, which greatly enhances the efficiency and accuracy with which a large number of photos can be identified. Modular subsets of the data can be distributed to an unlimited number of observers on or off site for identification. Modules are then re-incorporated into the database using a custom import function.

Key-words: camera traps, database, multiple observers, photos, wildlife cameras

Introduction

Ecologists have employed remotely triggered cameras (camera traps) to sample wildlife for many decades (Kucera & Barrett 2011). In recent years, however, significant technological advances have been made in battery life, sensor quality, overall reliability and memory card capacity (Kucera & Barrett 2011) such that camera traps have become a very efficient, even preferred, means of collecting information on a broad array of taxa including mammals, birds and herpetofauna. Such information can be used to inform conservation and management decisions (Kucera & Barrett 2011). Accordingly, use of these devices has erupted. A search of published literature in the Web of Science database (Thompson Reuters 2015) returned only six citations for “camera trap” in 2005, 37 in 2010, and 92 in 2014.

A common characteristic of studies that employ camera traps is the generation of a large quantity of data. Even single surveys or sampling sessions routinely generate hundreds of thousands of photos which must be processed and archived before data can be extracted from them for analyses. At least five general approaches have been authored for the sole purpose of handling the large quantity of photographic data generated from research that employs camera traps. These range from manual organization of photos into file structures that can be used by DOS programs to perform summaries (Renamer + CamTrap, Harris et al. 2010), to spreadsheet-based applications (Photospread, Kandel et al. 2007; Sundaresan, Riginos & Abelson 2011), to tailored use of commercial photographic archive systems (e.g., Picasa, Sundaresan, Riginos & Abelson 2011), and custom relational desktop databases such as Aardwolf (Krishnappa & Turner 2014) and Camera Base (Tobler 2014). Additionally, sophisticated server-based databases exist that automatically archive data from numerous projects across multiple jurisdictions (e.g., eMammal, http://emammal.si.edu/; DeskTEAM, Fegraus et al. 2011). Until recently, these latter developments required a project to be part of a larger, exclusive monitoring effort. Both now offer versions of their software that are available for independent research (the publically available version of DeskTEAM is known as Wild.ID; http://www.teamnetwork.org/wildlife-monitoring-solutions).

*Correspondence author. E-mail: Jake.Ivan@state.co.us

© 2015 The Authors. Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
Table 1. Comparison of software available for archiving and managing photo data collected from independent camera trap projects

| Operating System | PhotoSpread | Renamer + CamTrap | Camera Base | Aardwolf | Wild.ID | eMammal | Crw Photo Warehouse |
|------------------|-------------|-------------------|-------------|----------|---------|---------|---------------------|
| Single, relational platform | Yes | Windows, MacOS | Windows, MacOS | Windows, MacOS | Windows, MacOS | Windows, MacOS | Windows |
| Storage capacity | Unlimited | Unlimited | Unlimited | c. 2,000,000 | Unlimited | Unlimited | c. 800,000–2,000,000 |
| Automatic import | No | No | No | Yes | Yes | No | No |
| Assign multiple species | Yes | No | No | No | No | Yes | Yes |
| Record individuals | No | No | Sex Only | Yes | No | No | Yes |
| Record ancillary data | No | No | No | Yes | No | Yes | Yes |
| Double observer | No | No | No | No | No | Yes | Yes |
| Batch ID | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Crowd source ID | Yes | No | No | No | No | Yes | Yes |
| Record active days | No | Yes | Yes | No | Yes | No | Yes |
| Filter, query data | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Auto-generate input files | No | Yes | Yes | No | Yes | Yes | Yes |
| Auto-generate reports | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Help | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

1 Assumed based on its use of Microsoft Access, which has a 2,000,000-record limit per table.
2 Depending on whether the user employs double observers for each photo (c. 800,000) or only a single observer (c. 2,000,000).
3 Photos and associated metadata can be imported to database structure automatically without the need to manually enter or manipulate data.
4 Multiple species or number of individuals can be assigned to each photo if the user copies photos to multiple folders.
5 Multiple species can be assigned to each photo if the user imports photos multiple times, once for each species present.
6 Includes ability to record custom details for each photo such as sex, behaviors, radio collars, tags, etc.
7 Photo data can be auto-filled within a single photo burst or within a certain period of time of the focal photo but the user cannot auto-fill across bursts or longer (user-defined) periods of time.
8 Authors state that porting the software to a web interface would make this possible—unclear if the functionality is currently available and implemented.
9 Allows users to record and/or manage the days over which each camera was active and operating properly.
10 Software produces input files for Program PRESENCE; limited to a single occasion length (10 days).
11 Software produces input files for Programs MARK (closed capture), CAPTURE, PRESENCE, R ‘Mark’ (occupancy), DENSITY, and ESTIMATES.
12 Software produces input files for Programs PRESENCE and R ‘unmarked’ and produces output graphs from R ‘overlap’, and R ‘vegan’.
13 Software produces input files for Programs MARK (closed capture, occupancy), PRESENCE, DENSITY, R ‘secr’, and R ‘overlap’.
14 Basic reports available (total photos, number of photos per camera, total species). More advanced reports require SQL coding.
15 User manual available for download.
16 User manual available for download; context specific help available within each form.
We suggest that software used in camera trap research should at a minimum provide the user with means of (i) automatically importing photos and photo metadata into a relational database setting to minimize data entry errors, (ii) facilitating efficient and accurate identification of species in a potentially overwhelming number of photos including the ability to batch identify a series of photos, crowd-source the identification process to maximize efficiency, and make use of double observers to maximize accuracy, (iii) quickly sorting and querying photographic information once identification is complete, (iv) auto-generating summary reports, and (v) auto-generating input files for further analysis in software such as Program MARK (White & Burnham 1999) to minimize errors and time required to complete analyses.

Each of the current software programs mentioned above offers useful solutions that cover at least some of the desired functionality (Table 1). Thus each has relative strengths subject to user preference. However, to our knowledge, none of them offer the full array of properties and capacities necessary to maximize efficiency with which camera trap data can be utilized in a research context. Here, we describe a new application (CPW Photo Warehouse, CPW) that combines all of these features into a single solution (Table 1). The application and manual are freely available for download from Colorado Parks and Wildlife (thus the CPW moniker and recursive acronym) at: http://cpw.state.co.us/learn/Pages/ResearchMammalsSoftware.aspx.

Specifications

CPW is a Microsoft Access® database that uses extensive Visual Basic® for Applications (VBA) code to bring the complex functionality of the database to the user in the form of an intuitive, easily accessible graphical user interface. The current application operates on any modern Windows® operating system (XP or newer), with Microsoft® Access 2007, 2010 or 2013. Windows users who do not own an Access® license can use the database after downloading free Access® Runtime software. Mac users have successfully used CPW to manage their photos by installing Windows in a virtual machine environment, and the same solution will work with Linux operating systems. Because the database is distributed as an .accdb file, all of the design elements of the database (including VBA code) can be accessed by closing the startup form. Thus, if necessary, intermediate users can make modifications to suit the needs of a specific project by leveraging the startup form. Advanced users can further modify the software by altering the VBA or SQL code we provide.

Photos are stored separately from CPW but link seamlessly into forms and windows when called. Links to photo locations can be updated easily from within the database if photos are moved to a new location. Depending on whether the user opts to make use of the double-observer functionality (see below), the database can store data from...
c. 800,000 (double observer) to c. 2,000,000 (single observer) photos when operating as a single file on a personal computer. For large, recurring projects CPW can be modified so that it functions as a “front-end” with tables stored in a more robust enterprise-level database such as SQL Server®. In this case, storage of photo data is unlimited.

**Photo import**

The structure of CPW follows a logical hierarchy similar to that reported by previous authors (e.g., Krishnappa & Turner 2014). Cameras are organized first according to study areas, which are groups of cameras that the user may later want to summarize or compare to other groups based on attributes such as geographic areas, treatments and controls, or habitat strata. Within study areas, locations are specified for each camera (or set of cameras if they are deployed in pairs, as is often the case with mark-recapture studies); within locations, information is stored from each time a camera was deployed, checked or retrieved. Once design information regarding study area(s), locations and visits has been entered, photos can be added using the *Import Photos* form. Users specify which visit the photos should be associated with, which folder contains the photos, and whether and how to copy/rename the photos. This automated import process eliminates the need for manually creating a spreadsheet row or database record for each photo, thus minimizing opportunity for mistakes.

**Photo identification**

Once photos have been imported, the user logs in to the *Photo ID* form and navigates through the images in sequential order, providing identifications (ID) of the species in each photo. Photos and data fields appear together in the form and the user can type in identifiers, select them from a menu, or use shortcut keys (including one to repeat all information from the previous photo) to quickly populate the fields (Fig. 1). Each user can customize keyboard shortcuts to suit them, and filters can be employed to focus on photos from a specific location and/or season and/or only those photos that yet need identification. Users can quickly highlight images of particular importance so they can be easily retrieved later for use in reports and presentations. Double-clicking on an image opens it in an external program so users can edit or zoom as needed to assist in identification. The form can record any number of species and individuals present in a given image. A batch ID mode is also available such that a particular ID can be applied to an unlimited series of photos in a single stroke.

**Double observers to improve accuracy**

The database is designed such that each photo is viewed by two observers, which improves accuracy of the resulting data. Thus, when a user clicks the filter to only view those images that need ID, the program shows those photos that have not yet been identified by that user and that need identification.
from at least one more user. Once photos have been subjected to two observers the database manager or a third observer can then log in to the Compare ID form, which displays those photos in which the two observers disagreed. The manager decides which ID was correct and deletes the other. The database tracks each observer’s IDs, even those that were deemed incorrect and deleted, so that a complete history of the observations is maintained. Note that two observers are not strictly required; if the database manager determines that a single observer per photo is sufficient the database will still function properly.

Crowd-sourcing identification

To expedite the photo ID process, the database manager can use the Photo Viewer form (Fig. 2) to quickly filter the data to a manageable subset, which can then be exported to a modular version of the Photo ID form and shipped to observers on or off site. These modules contain the same species list as the master database, and observers identify photos as described above (including logging in to the form to register who is assigning the IDs), return the module, and the manager then uses the Import Module form to import identification data into the master database. As before, mismatched IDs will be flagged and can then be reconciled by the database manager. The ability to parcel out photographs for ID across an unlimited number of observers vastly improves the efficiency with which photos can be identified, thus accelerating the most tedious and time-consuming aspect of research involving camera traps. For example, recently we used the module function within CPW to identify species in 197 247 photos collected from 150 camera traps during summer 2014. Modules were distributed simultaneously across a team of 10 observers; each photo was examined by two observers and discrepancies were rectified by the database manager. The entire identification process consumed approximately 15 days. Note that it is the user’s responsibility to independently track how photos have been distributed to observers, which can be easily done with a simple spreadsheet.

Creating reports

The View or Print a Report option provides a means for users to easily generate various summary reports for a camera trap project. For example, the ‘General Camera Summary’ includes total cameras, total photos, average photos/camera, effort summaries (camera days), and number of photos of each species. Such a report is available at both the project and location level, and by year. Users can also choose to print a summary of detection data for individual animals, a photo ID summary (list of the number of photos by location along with the number identified by one or two observers and their initials), reports comparing performance of each observer who identified photos, or lists of cameras currently in deployment, including coordinates and access notes for each.

Fig. 3. 

© 2015 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society, *Methods in Ecology and Evolution*, 7, 499–504.
Creating input files

CPW provides forms to easily generate input files for occupancy, abundance, density, or activity pattern analyses performed in Programs MARK (White & Burnham 1999), PRESENCE (Hines 2015), DENSITY (Efford, Dawson & Robbins 2004), and the r (R Development Core Team 2015) packages ‘secr’ (Efford 2015) and ‘overlap’ (Meredith & Ridout 2015; Fig. 3). The user can define which species the files should be built for, and has complete control over the number and length of occasions generated. Options are available to allow the user to specify an absolute start date for a survey (e.g., June 1), or a start date that is relative to the set date for each camera, and the user can specify a buffer between when each camera was set and when the survey (and input file) actually begins. Finally, filters are available to create input files for a particular study area, year, or spatial extent. The form provides a preview of the input file and when satisfied, the user can export in a format specific to software mentioned above, or as an Excel® spreadsheet or Access® query for further modification.

Discussion

The practice of sampling wildlife with cameras has increased tremendously in recent years, a trend that will likely persist into the future as technological advances continue to improve the efficiency and effectiveness of these devices. This methodology, however, produces large quantities of raw data that must be archived and processed before analyses can be performed. **Photo Warehouse** provides researchers with a useful tool for achieving this task. Our program provides essential utility collectively available in existing software programs (e.g., maximized automation of photo and metadata input to minimize data entry mistakes, ability to auto-generate reports and export data to input files for analysis with other software) but also introduces important advancements that allow database managers to formally utilize and track multiple observers during the photo identification process. Such functionality greatly increases efficiency of photo identification (undeniably the most time-consuming and tedious aspect of research associated with camera traps) as an unlimited team of observers work simultaneously to identify photos. Additionally, this new multiple observer functionality allows the database manager to require that each photo be identified by two observers, which improves accuracy of analyses based on photo data. CPW is under continuous development to accommodate common uses of photo data from camera traps and should be a helpful tool for scientists who require the functionality encompassed here but do not have the time or expertise to design and program such utility themselves.

Acknowledgements

Funding for this project was provided by Colorado Parks and Wildlife and Colorado State University. We thank M. Schuette for developing critical aspects of the initial Lynx Occupancy database upon which parts of this application are based. We thank J. Lewis, K. Crooks, K. Blecha, S. Breck, R. Conrey, J. Halseth, R. Much, L. Stinson, M. Strauser and C. Wait for testing, feedback and patience. D. Johnston and K. Logan provided helpful reviews on early drafts. Microsoft®, Access®, SQL Server® and Excel® are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. Use of these names does not imply sponsorship, affiliation or endorsement of or by Colorado Parks and Wildlife.

Data accessibility

No data were displayed, analysed or summarized in this publication.

References

Efford, M.G. (2015) Package ‘secr’: Spatially explicit capture-recapture. Version 2.9.3. http://cran.r-project.org/web/packages/secr/secr.pdf.

Efford, M.G., Dawson, D.K. & Robbins, C.S. (2004) DENSITY: software for analysing capture-recapture data from passive detector arrays. *Animal Biodiversity and Conservation*, 27, 1–12.

Fegraus, E.H., Kai, L., Jorge, A.A., Chaitan, B., Sundaresan, S.R. & Abelson, E.S. (2011) Management and analysis of camera trap data. *Bulletin of the Ecological Society of America*, 91, 352–360.

Hines, J.E. (2015) PRESENCE 8.3: Software to estimate patch occupancy and related parameters. http://www.mbr-pwrc.usgs.gov/software/presence.html.

Kandel, S., Abelson, E., Gairea-Molina, H., Papecke, A. & Theobald, M. (2007) **PhotoSpread**: A Spreadsheet for Managing Photos. CHI 2008. Florence, Italy.

Krishnappa, Y.S. & Turner, W.C. (2014) Software for minimalistic data management in large camera trap studies. *Ecological Informatics*, 24, 11–16.

Lynx Occupancy database upon which parts of this application are based.

Meredith, M. & Ridout, M.S. (2015) Package ‘overlap’: Estimates of coefficient of overlapping for animal activity patterns. Version 0.2.4. http://cran.r-project.org/web/packages/overlap/overlap.pdf.

R Development Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Sundaesaran, S.R., Riggins, C. & Abelson, E.S. (2011) Management and analysis of camera trap data: alternative approaches (Response to Harris et al. 2010). *Bulletin of the Ecological Society of America*, 92, 188–195.

Thompson Reuters. (2015). Web of Science. http://thomsonreuters.com/en/products-services/scholarly-scientific-research/scholarly-search-and-discovery/web-of-science.html

Tobler, M. (2014) Camera Base 1.6.1. http://www.atrium-biodiversity.org/tools/camerabase.

White, G.C. & Burnham, K.P. (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study*, 46(Suppl), 120–138.

Received 25 August 2015; accepted 27 October 2015

Handling Editor: Patrick Jansen