Workshop on Visualisation in Environmental Sciences (EnvirVis) (2015)
A. Middel, G. Weber, and K. Rink (Editors)

Multi-Purpose Exploration of Uncertain Data for the Video Monitoring of Ecosystems

Emma Beauxis-Aussalet and Lynda Hardman

CWI, The Netherlands

Abstract

Computer Vision is a promising technique for in-situ monitoring of ecosystems. It is non-intrusive and cost-effective compared to sending human observers. Automatic animal detection and species recognition support the study of population dynamics and species composition, i.e., the evolution of species populations' size. Fixed cameras support continuous data collection, which can serve a variety of ecology research. Prior to in-depth data analysis, ecologists need to familiarise with the dataset, and with the limitations of video technologies. We propose an interactive visualization system for exploring the video data. It addresses user needs for i) eliciting information of interest for specific studies; and ii) identifying the uncertainty factors inherent to video technologies. We discuss generalisable interaction principes and illustrate them with screenshots of an online prototype.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1. Introduction

The Fish4Knowledge project developed computer vision technologies and data visualizations for the in-situ monitoring of fish populations [BHB*]. With 9 fixed underwater cameras continuously recording during 3 years, it supported innovative long-term studies of coral reef ecosystems. Video streams were split into 10-minute video samples, and computer vision performed fish detection and species recognition. Visualizations were developed for exploring the dataset and the uncertainties inherent to automatic video monitoring. Our specification of uncertainty factors and our visualization design are generalisable to other use cases using computer vision for ecology research [ACBCB*09,BCF*08, LMMZ*10,SMvO12], or needing multi-purpose data exploration.

2. Uncertainty Factors and Information Needs

From interviewing ecologists [BAAHVO13], we identified 2 core information needs: i) counting fish and their species over locations and time periods; ii) assessing the uncertainty of video data. Measuring fish body size was requested, but is not feasible with our technology. Uncertainty assessment requires both domain and technical expertise. From additional interviews of computer vision experts, we identified 10 uncertainty factors and their related metrics (Table 1-2). Ecologists may conduct a variety of research, focusing on specific species, locations or time periods. Their information needs depend on their specific research, both for the ecological (i) and uncertainty (ii) measurements to consider. Hence we designed a tool for data exploration addressing a wide range of user interests. It helps preparing further data analysis performed with specialised tools and methods. We generalised user tasks as: selecting datasets of interest, and exploring ecological and uncertainty *measurements* displayed over the multiple data *dimensions*.

The data *dimensions* are the location, time and image quality of video samples, the fish species and *certainty score* (indicating the quality of fish appearance, thus the chances of computer vision error). Datasets of interest are selected using the same dimensions, e.g., fish from specific species and time periods. The time dimension is decomposed into Year, Week of the Year, and Hour of the Day (e.g., to select fish occurring in the spring weeks, or the morning hours), as ecologists can study daily or yearly cycles, and compare populations across years (e.g., compare spring weeks' populations over several years).

Ecological *measurements* are the number of fish and species, and their mean and variance over *dimensions* of interest (e.g., mean per day or video sample). Using mean counts per video sample compensates for unbalanced numbers of videos due to *Fragmentary Processing* (missing or unprocessed videos). When comparing data subsets, the fewer the videos, the fewer the fish. But mean counts per 10-min video sample remain comparable.

Uncertainty *measurements* are the number of video samples, the fish scores, and the fish detection and species recognition errors, i.e., the number and rate of True Positives (TP), False Positives (FP) and False Negatives (FN). TP are correctly detected fish. FP are objects added to a category they do not belong to: non-fish added to the set fish, or fish from species A added to the set of fish from species B (denoted $FP_{B\leftarrow A}$). FN are objects missing from the category they truly belong to: undetected fish, or fish missing from species A since they are attributed to species B (denoted $FN_{A\rightarrow B}$). FP and FN are measured by comparing automatic and manual (ground-truth) fish classification. Users need to understand that FN for one species are FP for another (besides undetected fish). This creates biases, e.g., if species A increases, its FN increase, and species B may artificially increase too. Hence, to enable extrapolations of *Errors in Specific Output* under varying species proportions, rates of FN and FP are relative to the TP of their true class: $rate(FN_{A\rightarrow B}) = rate(FP_{B\leftarrow A}) = \frac{FN_{A\rightarrow B}}{TP_A}$. Contrarily to traditional metrics, $FP_{A\leftarrow X}$ must not be added to the denominator, as it depends on the sizes of other species populations.

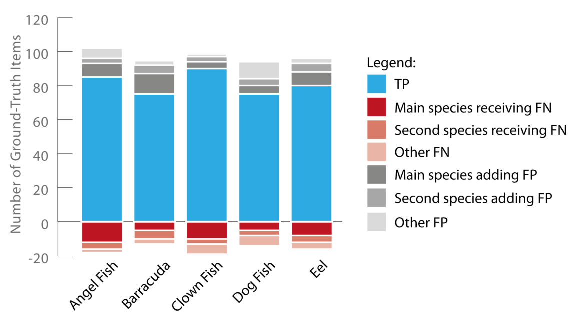

Figure 1: *Static visualization of computer vision errors from [BAH14c].*

3. Visualization and Interaction Principles

To reduce the complexity of exploring the multiple *dimensions* and *measurements*, we use simple graphs in multiple views. Multiple views introduce additional complexity (context switching) which can be balanced by the consistency between views (display same datasets or dimensions) [WBWK00]. We focused on ensuring views' consistency, while allowing flexible views' variation (Fig. 2). Zone A contains the main graph which is controlled with Zone B: users can select the type of graph (simple graph, stacked graph, or boxplot) and what *measurements* and *dimensions* are represented by its axes. While the Y axis displays a *measurement*, e.g., number of fish, the X axis can be swapped between *dimensions* (e.g., day, location, species). The Y axis can be swapped, e.g., for number of species, while keeping the same X axis. Incompatible options are disabled (e.g., number of species over species). Swapping axes is the core interaction principle for users to gradually navigate through the video data. Further multidimensional exploration is offered by stacked graph (Fig. 3) and boxplots (Fig. 4). Swapping graphs, our second interaction principle, preserves the X and Y axes' *dimensions* and *measurements*. For stacked charts and boxplots, users can swap the *dimension* (i.e., the Z axis) for which *measurements* are stacked, or summarized with mean and variance. Zone C contains filter widgets, one per *dimension*, for selecting datasets of interest. They are opened on-demand depending on user interests. They display histograms for each filterable value, which Y axis is the same *measurement* as for Zone A, thus ensuring multiple views consistency. Propagating Y axis *measurement* is our third interaction principle. It offers both overviewing (Zone C widgets) and detailed views (Zone A main graph). Early user feedback expressed enthusiasm for its flexibility and intuitiveness (e.g., *"I can display anything I want"*).

These interaction principles are applicable to studying ecological or uncertainty *measurements*. For the latter, numbers of video samples and fish scores can be visualised using the same graphs as in Fig. 2-4. Visualizing computer vision error requires specific graphs developed in [BAH14c, BAH14a] (Fig. 1). Although designed as a series of 4 static graphs (number and rate of errors for fish detection and species recognition), our interaction principles can be applied to provide integrated data exploration (Fig. 5). Error *measurements* (number and rate of errors) can made available for the Y axis. The X and Z axes can provide further options to decompose the errors over e.g., species, fish score, or image quality. But error *measurements* may not be available over all *dimensions*: ground-truth data spans over a limited range of time periods and locations, as otherwise it is too costly or useless (computer vision errors may not vary over years). Future work can investigate if it is preferable to use a distinct type graph, e.g., an *error graph* rather than a stacked graph.

For investigating uncertainty due to *Fields of View* and *Duplicated Individuals*, no measurements are available and users need to draw approximative estimations by inspecting the video footage. A video browser can be displayed in Zone A, as an option in the list of available graph.

4. Conclusions

Our initial work investigated interactive data exploration [BAH14b] and simplified visualization of computer vision errors [BAH14c]. We introduce here an application of the interaction principles to the visualization of computer vision errors. We argue that the interaction principles (swapping graphs and axes, propagating Y axis) are generalisable, and that the visualization system is extensible. It can be extended with further *measurements* for the Y axis (e.g., growth rate of populations, fish body size) ; with *dimensions* for the X and Z axes (e.g., subsample over random splits for boxplots); and with other types of graph. For further reuse, an initial prototype [F4Kb] and open source code are available [F4Ka].

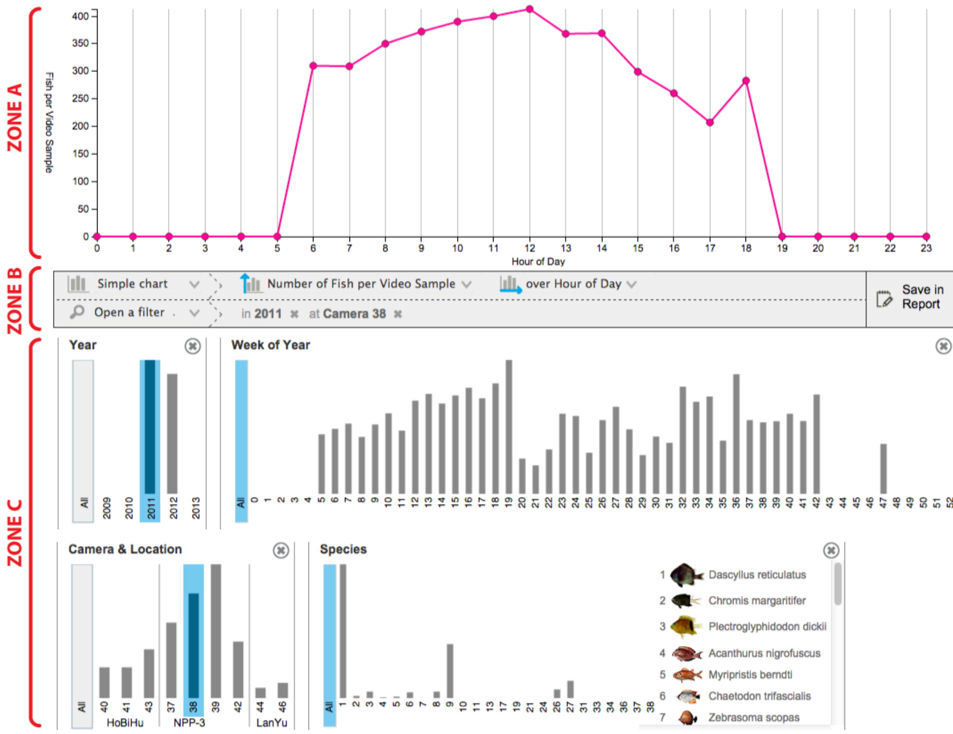

Figure 2: *Multi-purpose visualization system.*

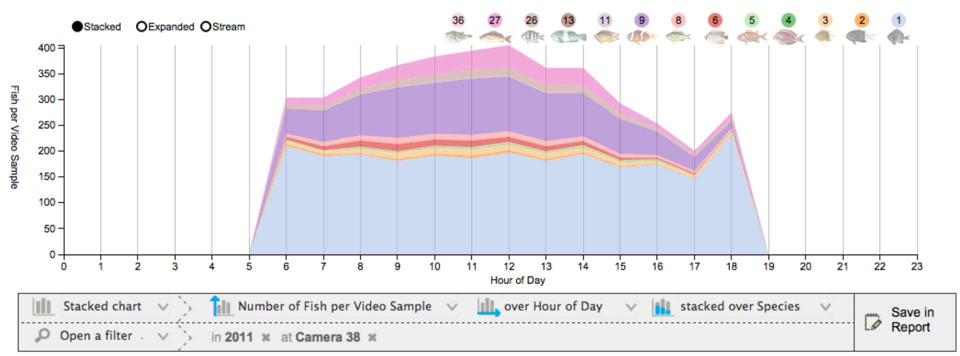

Figure 3: *Multi-purpose stacked visualization (omitting Zone C).*

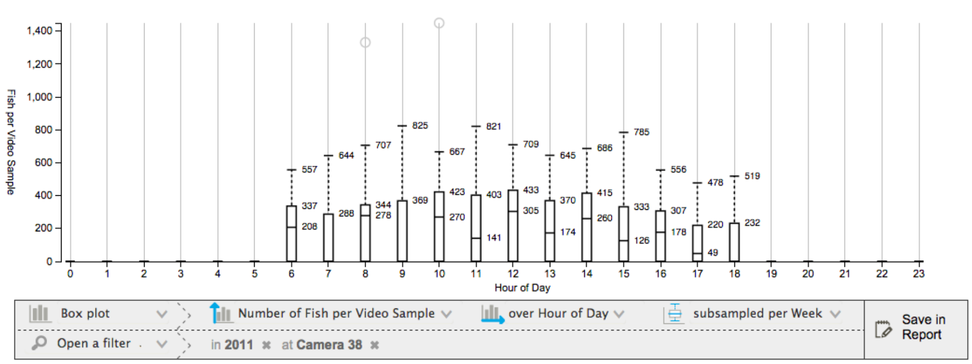

Figure 4: *Multi-purpose boxplot visualization.*

Factor	*Description*
Uncertainty due to computer vision algorithms	
Ground-Truth Quality	Ground-truth items may be scarce, represent the wrong objects, or odd fish appearances.
Fish Detection Errors	Some fish may be undetected, and non-fish objects may be detected as fish.
Species Recognition Errors	Some species may not be recognized, or confused with another.
Image Quality	Lighting, contrast, resolution or fuzziness may impact the magnitude of computer vision errors.
Uncertainty due to in-situ system deployment	
Field of View	Cameras may observe heterogeneous ecosystems, and over- or under-represent some species.
Fragmentary Processing	Some videos may be yet unprocessed, missing, or unusable (e.g., encoding errors).
Duplicated Individuals	Fish swimming back and forth are repeatedly recorded. Rates of duplication vary among species behaviour (e.g., sheltering in coral head) and *Fields of view* (e.g., open sea or coral head).
Sampling Coverage	The numbers of video samples may not suffice for software outputs to be statistically representative.
Uncertainty due to both computer vision algorithms and deployment conditions	
Biases Emerging from Noise	Errors may be random (noise) or systematic (bias). Biases may emerge from a combinaison of factors (*Image Quality, Field of View, Duplicated Individuals, Fish Detection* & *Species Recognition*).
Uncertainty in Specific Output	Uncertainty in specific set of outputs may be extrapolated from errors measured in test conditions, compared to the specific characteristics of the output (e.g., fewer low quality images).

Table 1: *Uncertainty factors introduced by computer vision software and in-situ system deployment.*

Factor	*Metrics*	*Visualization*
Uncertainty due to computer vision algorithms		
Ground-Truth Quality	Number of ground-truth items over species and image quality.	Fig. 5
Fish Detection Errors	Number and rate of TP, FN and FP.	Fig. 5
Species Recognition Errors	Number and rate of TP, FN and FP.	Fig. 5
Image Quality	Number of video samples from each type of image quality.	Fig. 2-4
Uncertainty due to in-situ system deployment		
Sampling Coverage & Fragmentary Processing	Number of video samples over time and location. Average fish count per video.	Fig. 2-4
Field of View	No existing metric. Manual inspection of video footage.	
Duplicated Individuals	No existing metric. Manual inspection of video footage.	
Uncertainty due to both computer vision algorithms and deployment conditions		
Biases Emerging from Noise	Estimation of significantly higher *Duplicated Individuals*, *Fish Detection* and *Species Recognition Errors* over species, *Image Quality* and *Field of View*.	Fig. 2-5
Uncertainty in Specific Output	Fish *certainty score*. Correlate *Biases emerging from Noise* with dataset characteristics (Species, *Image Quality, Field of View,* and certainty score distributions).	Fig. 2-5

Table 2: *Metrics addressing the uncertainty factors of Table 1 and corresponding visualizations.*

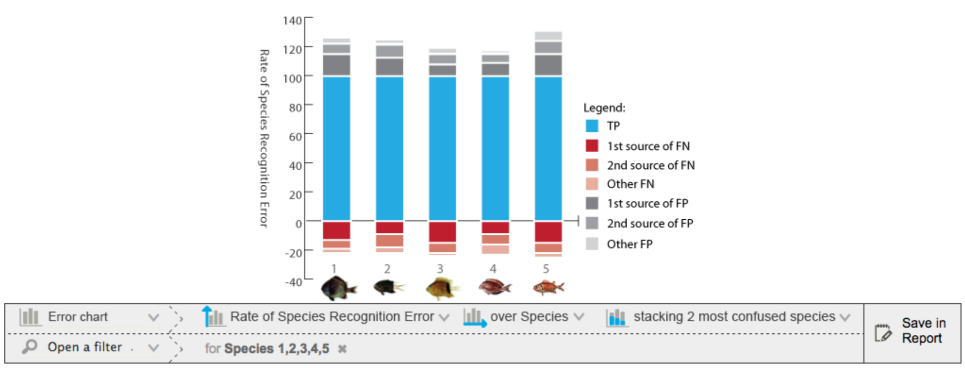

Figure 5: *Multi-purpose visualization of computer vision errors, using the interaction principles of Fig. 2-4 for Fig. 1.*

References

[ACBCB*09] ACEVEDO M. A., CORRADA-BRAVO C. J., CORRADA-BRAVO H., VILLANUEVA-RIVERA L. J., AIDE T. M.: Automated classification of bird and amphibian calls using machine learning: A comparison of methods. *Ecological Informatics 4*, 4 (2009), 206–214. 1

[BAAHVO13] BEAUXIS-AUSSALET E., ARSLANOVA E., HARDMAN L., VAN OSSENBRUGGEN J.: A case study of trust issues in scientific video collections. In *Proceedings of the 2nd ACM international workshop on Multimedia analysis for ecological data* (2013), ACM, pp. 41–46. 1

[BAH14a] BEAUXIS-AUSSALET E., HARDMAN L.: Simplifying the visualization of confusion matrix. BNAIC. 2

[BAH14b] BEAUXIS-AUSSALET E., HARDMAN L.: Uncertainty-aware visualization of fish populations. In *Demo at the International Working Conference on Advanced Visual Interfaces. ACM* (2014). 2

[BAH14c] BEAUXIS-AUSSALET E., HARDMAN L.: Visualization of confusion matrix for non-expert users. In *Poster at the IEEE Symposium on Information Visualization* (2014). 2

[BCF*08] BELHUMEUR P. N., CHEN D., FEINER S., JACOBS D. W., KRESS W. J., LING H., LOPEZ I., RAMAMOORTHI R., SHEOREY S., WHITE S., ET AL.: Searching the world's herbaria: A system for visual identification of plant species. In *Computer Vision–ECCV 2008*. Springer, 2008, pp. 116–129. 1

[BHB*] BOOM B. J., HUANG P. X., BEYAN C., SPAMPINATO C., PALAZZO S., HE J., BEAUXIS-AUSSALET E., LIN S.-I., CHOU H.-M., NADARAJAN G., ET AL.: Long-term underwater camera surveillance for monitoring and analysis of fish populations. 1

[F4Ka] Source code. URL: http://sourceforge.net/projects/fish4knowledgesourcecode/. 2

[F4Kb] Visualization prototype. URL: http://f4k.project.cwi.nl/demo/ui/. 2

[LMMZ*10] LYTLE D. A., MARTÍNEZ-MUÑOZ G., ZHANG W., LARIOS N., SHAPIRO L., PAASCH R., MOLDENKE A., MORTENSEN E. N., TODOROVIC S., DIETTERICH T. G.: Automated processing and identification of benthic invertebrate samples. *Journal of the North American Benthological Society 29*, 3 (2010), 867–874. 1

[SMvO12] SPAMPINATO C., MEZARIS V., VAN OSSENBRUGGEN J.: Multimedia analysis for ecological data. In *Proceedings of the 20th ACM international conference on Multimedia* (2012), ACM, pp. 1507–1508. 1

[WBWK00] WANG BALDONADO M. Q., WOODRUFF A., KUCHINSKY A.: Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces* (2000), ACM, pp. 110–119. 2