Interactive Visualization of Video Data for Fish Population Monitoring

Emma Beauxis-Aussalet
CWI
Science Park 123, Amsterdam
emmanuelle.beauxis-aussalet@cwi.nl

Lynda Hardman
CWI
Science Park 123, Amsterdam
lynda.hardman@cwi.nl

ABSTRACT
The recent use of computer vision techniques for monitoring ecosystems has opened new perspectives for marine ecology research. These techniques can extract information about fish populations from in-situ cameras, without requiring ecologists to watch the videos. However, they inherently introduce uncertainty since automatic information extraction is imperfect. To be accepted for scientific use, video analysis tools must support the analysis of the extracted information and of their uncertainty. Another challenge concerns the diversity of scientific interests. Ecologists have diverse research goals and information needs, for instance specific species, time periods, or locations. We present a visualization interface addressing these two challenges: providing information about fish populations as well as computer vision uncertainty; and enabling the exploration of specific subsets of the video data depending on user needs.

Author Keywords
Data Visualization; Uncertainty; Interaction Design; User Trust; Computer Vision Application.

ACM Classification Keywords
H.5.2. [User Interface]: Graphical user interfaces;

INTRODUCTION
The Fish4Knowledge project (fish4knowledge.eu) has continuously recorded video footage of coral reef fish from 9 underwater cameras during 3 years. This collection motivated the use of computer vision for automatically recognizing fish from different species, and monitoring the population dynamics. The original video collection is processed in 3 steps: sequencing of continuous video streams into 10-minute clips (for storage purposes), identification of fish amongst other objects, and recognition of fish species.

From an ecologist’s perspective, each information processing step potentially introduces errors: from data collection (recording videos) and processing (recognizing fish), to interpretation (deriving facts from automatic fish counts). We developed visualizations and interaction designs for exploring these uncertainties. Our interface discloses the data collection and data processing methods and their possible biases. It lets users walk through explanations of the video analysis processes and their uncertainty. The interface offers interactive visualizations of fish populations, e.g., to explore numbers of fish from specific species, locations or time periods. The uncertainties of the observed fish populations are also visualizable. Finally, users can annotate and share their findings.

ELICITING AND ADDRESSING USER NEEDS
Ecologists are not experts in computer vision systems and their technical concepts. The information about provenance and uncertainty needs to be comprehensive and sufficiently detailed, while remaining understandable. We investigated user information needs by interviewing ecologists and computer vision experts [1-3]. Regularly collecting user feedback along design stages revealed user needs throughout all stages, depending on users’ expectations and knowledge of computer vision. User-Centered Design had limitations for introducing this novel technology: non-experts did not foresee all uncertainties inherent to the technology, nor appropriate metrics for interpreting them (e.g., ground-truth evaluation). Requirements from non-experts only would produce incoherent design as [5] notifies. Interviews of computer vision experts complemented the information requirements. Computer vision experts were particularly needed for supplying uncertainty metrics, and specifying Uncertainty and Error Modeling [4]: leveraging user support depended on experts’ uncertainty evaluation. We identified 5 main information needs and 10 underlying uncertainty factors (Tables 1-2).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright is held by the author(s).


1 The prototype is online: f4k.project.cwi.nl
truth evaluation for non-experts (Figure 3), and an interaction design for exploring multiple uncertainty factors.

<table>
<thead>
<tr>
<th>Information Need</th>
<th>Source</th>
<th>UI Tab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watch Videos</td>
<td>Users</td>
<td>Video tab</td>
</tr>
<tr>
<td>Fish and species counts</td>
<td>Users</td>
<td>Visualization tab</td>
</tr>
<tr>
<td>Trends and Correlations</td>
<td>Users</td>
<td>Visualization, Report tabs</td>
</tr>
<tr>
<td>Uncertainty of Computer</td>
<td>Users, Experts</td>
<td>Video, Video Analysis, Visualization tabs</td>
</tr>
<tr>
<td>Vision system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertainty of Computer</td>
<td>Users, Experts</td>
<td>Visualization tab</td>
</tr>
<tr>
<td>Vision results</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. The main information needs.

EXPLORING VIDEO PROCESSING STEPS

The user interface tabs deliver manageable units of information and reflect the information processing sequence: data collection (Video tab), data processing (Video Analysis, Extracted Data tabs) and data interpretation (Visualization, Report tabs).

The **Video** tab supports video browsing (Figure 2). It contains filtering functionalities for specifying the videos of interest (e.g., at specific location or time periods). Users can control the data collection conditions: which ecosystems are observed, with which field of view and image quality (e.g., lens biofouling, water turbidity).

The **Video Analysis** tab provides explanations of the video processing steps and visualizations of their uncertainty. It exposes the technical concepts needed for understanding computer vision uncertainty. The **Overview** sub-tab provides explanations of the main video processing steps. The **Fish Detection**, and **Species Recognition** sub-tabs provide visualizations of ground-truth evaluations (Figure 3). The **Workflow** sub-tab provides on-demand video processing. Users can request the analysis of specific videos (from user-defined time periods and cameras) with specific component versions (e.g., with the best accuracy for the species of interest). It serves either for processing videos that were not yet analyzed, or for experimenting with different versions of the video analysis components (e.g., to check robustness of observations).

Figure 2. The Video tab.

The **Extracted Data** tab provides an overview of the available video data and their properties (i.e., their dimensions). It shows all the characteristics of fish extracted from the video footage. It also explains the 4 main metrics provided for describing fish populations and their uncertainty: Number of Fish, Number of Video Samples (to check for missing videos), Number of Fish per Video Sample (to compensate for missing videos), and Number of Species. It helps understanding how fish populations can be monitored, and identifying the information relevant for particular studies.

VISUALIZING COMPUTER VISION UNCERTAINTY

The computer vision community uses well-accepted methods for evaluating uncertainty. They basically rely on a ground-truth: a set of images with manually identified objects, to which automatically identified objects are compared. A variety of metrics and visualizations can be derived from ground-truth evaluation. They are easily understood by experts but not by non-experts [1]. We adapted their visualization for non-experts. We reduced complexity by avoiding the use of advanced metrics (e.g., rates such as Precision/Recall), thresholds (e.g., ROC curves) and True Negatives (i.e., errors automatically discarded and having no impact on fish counts). [1] highlights that providing more details is likely to overwhelm users. The visualization exposes the proportions of items in the ground-truth, which is often omitted in more traditional evaluations. It indicates potential biases: the fewer the items, the higher the chance of error. The **Video Analysis** tab provides uncertainty visualizations (Figure 4) for each video analysis steps (fish detection and species recognition). Uncertainty is detailed for each video quality and each species, since different levels of errors indicate potential biases in the video data.

VISUALIZING COMPUTER VISION RESULTS

The **Visualization** tab (Figure 4) provides means to explore the video data, and the uncertainty due to missing videos, video quality or fish appearance quality. Videos can be missing due to camera maintenance, encoding errors, or unfinished processing queues. The quality of each fish appearance is measured using a certainty score. It indicates how much fish look like the fish model for its species. The higher the score, the more certain is the species recognition.
In Figure 4, Zone A contains the main graph. Zone B supports the adaptation of the main graph to specific user needs. Users can specify what the axes of the main graph represent. For instance, while the y-axis represents numbers of fish, the x-axis can represent their distribution over weeks of the year or hours of the day. Users can also select other types of graph (stacked chart or box plot). They provide additional information about the visualized fish population: e.g., the proportion of each species, or the variance of fish abundance. The selection of stacked charts or boxplot leads to the display of dedicated menus for adapting further the visualization. For instance fish counts can be stacked by species or by camera. Zone C contains filter widgets for both selecting datasets of interest, and overviewing datasets over several dimensions. Filter widgets are displayed on-demand. There are widgets for each dimension of the data, namely: Year, Week of Year and Hour of Day of fish occurrence, Camera, Species, Certainty Score, Video Quality and Software Version. A summary of the filters applied is provided in Zone B. To limit information overload, the default filters (e.g., all species, all cameras) are not mentioned in the summary. The widgets’ histograms display the same metric as the main graph, and applied to the same dataset. E.g., in Figure 6 both the graph of Zone A and the histograms of Zone C display numbers of fish per video sample. Both use a dataset of fish detected by software D50-R52, occurring in 2011 at Camera 38, and belonging to all species, certainty score, video quality, week of year and hour of day. The Camera widget uses a dataset from all cameras, and highlights in blue which camera is selected.

The Report tab supports manual grouping and annotation of graphs created in the Visualization tab. Graphs can be added to and removed from a report. They can be annotated with a title and a comment. Users can download and upload reports in the form of text file containing a list of parameters. Reports can be saved and shared any file.

Our interaction design let users specify which data dimensions are relevant for their goal. Information of interest is displayed on-demand (open widgets in Zone C, change graph axes, display details in stacked charts and boxplots). Irrelevant information is not displayed (close widgets, switch back to simple graph). It provides both overviewing (widgets in Zone C) and detailed views (main graph in Zone A). It supports a wide range of data analysis goals, while limiting display cluttering and information overload. It addresses our design context with ecologists pursuing a variety of research goals, while being unfamiliar with video data.

CONCLUSION

Our interface allows ecologists to monitor fish populations using novel computer vision techniques. It allows the exploration of multiple uncertainty factors and aspects of the data. Early user feedback shows we achieved intuitive interactivity and easy to understand visualizations, although the dataset is unfamiliar to users. It allows preliminary data exploration for a variety of user goals, and the identification of uncertainties that may impact further information processing. This design informs and inspires other use cases dealing with open-ended data exploration, familiarization with novel data, or visualization of machine learning uncertainty.

REFERENCES

3. Fish4Knowledge Project Deliverables: D2.1 User Information Needs; D2.3 Component-based Prototype and Evaluation; D6.6 Public Query interface. http://groups.inf.ed.ac.uk/f4k/deliverables.htm
Table 2. The uncertainty factors, identified by users all along design stages and complemented by system experts.