Long-term underwater camera surveillance for monitoring and analysis of fish populations

Bastiaan J. Boom[1], Phoenix X. Huang[1], Cigdem Beyan[1], Concetto Spampinato[3]
Simone Palazzo[3], Jiyin He[2], Emmanuelle Beauxis-Aussalet[2], Sun-In Lin[4]
Hsiu-Mei Chou[4], Gayathri Nadarajan[1], Yun-Heh Chen-Burger[1], Jacco van Ossenbruggen[2]
Daniela Giordano[3], Lynda Hardman[2], Fang-Pang Lin[4], Robert B. Fisher[1]
[1]*School of Informatics, University of Edinburgh*, [2]*Interactive Information Access, CWI, Amsterdam*
[3]*Department of Electrical, Electronics and Computer Engineering, University of Catania*
[4]*National Applied Research Laboratories, Taiwan*

Abstract

Long-term monitoring of the underwater environment is still labour intensive work. Using underwater surveillance cameras to monitor this environment has the potential advantage to make the task become less labour intensive. Also, the obtained data can be stored making the research reproducible. In this work, a system to analyse long-term underwater camera footage (more than 3 years of 12 hours a day underwater camera footage from 10 cameras) is described. This system uses video processing software to detect and recognise fish species. This footage is processed on supercomputers, which allow marine biologists to request automatic processing on these videos and afterwards analyse the results using a web-interface that allows them to display counts of fish species in the camera footage.

1. Introduction

In order to study the effects that climate change and pollution has on the environment, long-term monitoring of the environment is necessary. One of the most important natural environments on earth are the coral reefs, however monitoring the fish population and biodiversity is still a challenging task. Data collection in this kind of environment is labour intensive, requiring divers to count the fish species in a certain area [7]. In recent years, digital video recording has become much cheaper which makes underwater cameras a good alternative for data collection. Furthermore, automatic video processing and pattern recognition is able to process this kind of data. This paper describes an entire system which is been developed to allow marine biologists to analyse large amounts of video data for long-term monitoring purposes.

To give an indication of the challenges in this project, a summary is given of the amount of data that the system is expected to process. At the moment, around 10 cameras record 12 hours a day (daylight) where some have already recording over 4 years. The estimated amount of raw video data is at the moment 112 Terabytes. By processing this data using automatic video processing software, we expect to find around 10^{10} fish, which will also be categorized by species (or family). All this data will be stored in a database, which is expected to take up to 500 Gigabytes. Besides processing all this data, it will also be a challenge to present this data to marine biologists in usable manner.

The goal of this paper is to present the structure of the system. This system has three main challenges:

1. Processing the data, both the raw videos processing and querying based data extracted from these videos
2. Dealing with the uncertainty in the data. This uncertainty is caused by mistakes during detecting and recognizing the fish.
3. Visualization of the data, where the results are statistically correct, verifiable and reproducible.

2. Literature Review

Several studies [7] have been performed in marine biology where the fish population is counted using volunteer divers. This has several drawbacks in comparison with underwater camera surveillance. Firstly, some fish species will hide themselves if divers appear, while, in the case of camera surveillance, the fish get used to the cameras. Secondly, the observation performed by

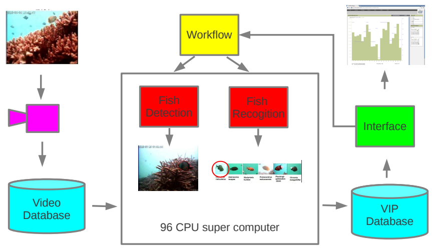


Figure 1: A schematic overview of the system software components and the two databases

Figure 2: An example of fish detection and tracking

the volunteers is not always very stable, because both the region covered and ability to recognise species can vary. Finally, the observations of the volunteers are not reproducible.
In recent years, underwater cameras have been used to study coral reefs [8]. Several approach have been developed to detect and recognise fish using underwater camera footage, but this work is mainly restricted to more constrained environments [13, 9, 5, 4]. This paper uses an improved version of the methods developed by Spampinato et al [10], where both fish detection, tracking and recognition are performed on video footage recorded in corals. In [10], one of the largest experiments used 320 fish image to verify the performance of their methods. Since our goal is to run on 10^{10} fish, a larger dataset is obtained to verify our automatic video processing software.

3. Overview of System

The system can be divided into four software components and two databases as shown in Figure 1. The software components are the fish detection, fish recognition, user interface and workflow. These software components run on supercomputers at NARL, which provides the architecture to record the videos and store and process the data. At the moment, both the fish detection and recognition is running distributed on a 96 CPU supercomputer in Taiwan and the VIP (Video/Image Processing) database is filled with fish which are located in the video and afterwards the species of the fish is determined. The user interface is web-based, connection to the database allowing the user to, for instance, visualize the fish count at certain time intervals. In the following section, more details of the different software components are given.

3.1. Workflow

The workflow component has to dynamically generate the commands to invoke the fish detection and recognition software based on requests from marine biologists (called workflow composition) and schedule them on the distributed platform/supercomputer using a resource scheduler. It is common for computer vision methods to have some parameters which can be tuned for different environments/fish species. The workflow performs reasoning based on semantics [6] to find the correct parameter set. In addition, some parameters cannot be set until the execution of another task is complete, causing the workflow to have to separately manage the composition and scheduling phases in order to control task and data dependencies. The system is still under development so newer versions of the software are sometimes released, however to keep the result verifiable and reproducible older versions of the software are still supported. In the interface, marine biologists will be able to request the processing on different videos and the workflow gives an estimate of the running time. At the moment, the workflow is used to process historical (previously recorded) videos.

3.2. Fish Detection and Tracking

The fish detection component both finds the fish in the image and follows the fish as time progresses using tracking algorithms. The first version of this software is described in [10]. The current system has multiple background subtraction methods, which can deal with different underwater conditions. After the detection, tracking of the fish is performed (both shown in Figure 2) using covariance based fish tracking [12], which gives the trajectory for each fish throughout the video. Assuming that behavior is related to the trajectories of the fish, a study has analysed the behaviour of the fish during typhoons [11].

3.3. Fish Recognition and Clustering

The fish recognition component assigns to each fish a species label [2]. At the moment, a vector of descriptors is extracted using the color, contour and texture information of the fish. In order to classify the fish based on the descriptors, we are using a hierarchical classifier.

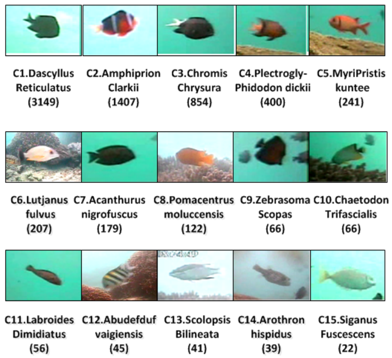


Figure 3: 15 most common fish species in the underwater camera footage

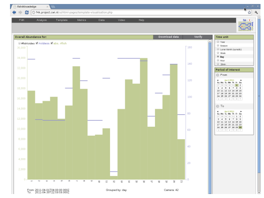
Fish species are organised in a hierarchical manner by biologists and we use a classifier with a similar structure. Forward feature selection is used to select a subset of all features, while support vector machines are used as the classifier at each node of the tree. At the moment, the fish recognition software is able to recognise the 15 fish species shown in Figure 3.

3.4. Interface

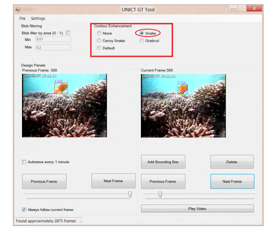
The interface allows the marine biologist to ask certain queries to the system. Based on several interviews conducted with marine biologists, a list of twenty question has been developed with typical questions which can be expected from a marine biologist. Examples of these question are: "What populations live in area X?", "What predator-prey relationship can be observed in area X?", etc. Lots of these questions are related to counting the fish species in the camera footage. A typical interface for this is shown in Figure 4(a). However it is also important that the marine biologist can verify the results of the different components and can change some of the settings based on observations of the data. For this reason, the interface allows users to focus on specific aspects of the data, and check which components are responsible for processing the video and for describing the fish appearing the video (e.g., species or trajectory of fish). The interface also allows marine biologists to share the information with each other, allowing other marine biologists to check the data.

4. Verification of the System

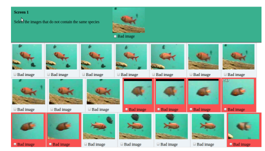
The verification of the system is very important and the user interface helps marine biologists to perform visual verification, however it is also important to measure the performance of automatic video processing software. The basic idea is that if we have a measure of accuracy in fish detection and recognition on a couple videos, we can extrapolate this accuracy on all videos. In order to determine the accuracy, the data needs to be annotated by humans. The annotation process in the case of fish detection is very different from that of the fish recognition. Both processes are briefly discussed in the following sections.



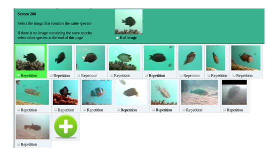
(a) Interface showing the fish population for the month of April



(b) Interface to annotate the fish location, contour and trajectories in a video



(c) The first interface to remove images that do not belogn to the clusters



(d) The second interface to link the image in the top row to a label

Figure 4: Interfaces

4.1. Annotations for fish detection and tracking

For fish detection, three important aspects can be evaluated: the boundingbox, contour and trajectory of the fish. A tool was developed for this kind of groundtruth annotation [3], shown in Figure 4(b). Automatic algorithms are used to give suggestions about the boundingbox, contour and tracking, however the user can change these suggestions. For the boundingbox, the background subtraction methods are used. The initial contour is also given by the background subtraction method, but other automatic methods like grabcut and snakes can be used to improve the contour if incorrect. The tool allows the user to easily show consecutive frames which makes it possible to verify the trajectory of the fish.

4.2. Annotations for fish recognition

Annotation of the fish species based on the detected fish is more difficult. The reason is only marine biologists might know the species name given a fish image. However, the marine biologists are not going to annotate thousands of fish images. To overcome this problem the methodology in [1] was used to perform annotation based two interfaces (Figure 4(c) and 4(d)). Firstly clustering is performed obtaining clusters of similar fish images, where we have one representative image for each cluster (mean of cluster for instance). The first interface in Figure 4(c) allows users to label if the fish belongs to the same species as the representative image. The second interface (Figure 4(d)) allows users to link this representative image to a species, where each species is represent by a typical image of that species. Marine biologists are then only asked about new species. This annotation framework has the advantage that it is faster and requires almost no domain knowledge.

5. Current Status

The system is not yet fully operational, but so far we have detected fish on approximately 958 hours ($2\frac{1}{2}$ months) of video, resulting in 2819257 detected and tracked fish and 205 hours of video are processed by fish recognition software. Performance analysis using a ground truth of 35935 detected fish shows a detection and tracking rate of 79.8% with a 11.8% false detection rate. 20074 images of fish have been annotated by at least 2 persons, however because of low resolution and murky water only a smaller subset of 6874 annotated fish containing 15 species is used to evaluate the fish recognition algorithm that achieves an average recall of $83.15\%\pm7.10$. The user interface is still under development, but is already able to run some simple statistics on the fish stored in the VIP database (see Figure 1). The system currently uses 96 processors and is processing previously recorded videos of a single camera at the rate of ±4 hours per day.

6. Conclusion

A research tool is described in this paper to monitor the fish population in a coral reef environment by analysing massive amounts of video footage. This is achieved by automatic video processing software and the supercomputer facilities in Taiwan. Groundtruth data is in place to measure the performance of automatic video processing software. The user interface is being developed to allowing marine biologists to access the data and reprocess video with different parameter settings in order to verify hypotheses.

Acknowledgements: This research was funded by the European Commission (FP7 grant 257024) and the Taiwan National Science Council (grant NSC100-2933-I-492-001) and undertaken in the Fish4Knowledge project (www.fish4knowledge.eu).

References

[1] B. J. Boom, P. X. Huang, J. He, and R. B. Fisher. Supporting ground-truth annotation of image datasets using clustering. ICPR 2012, to appear.

[2] P. X. Huang, B. J. Boom, J. He, and R. B. Fisher. Underwater live fish recognition using balance-guaranteed optimized tree. ACCV 2012, to appear.

[3] I. Kavasidis, S. Palazzo, R. Di Salvo, D. Giordano, and C. Spampinato. A semi-automatic tool for detection and tracking ground truth generation in videos. VIGTA 2012.

[4] R. Larsen, H. Olafsdottir, and B. K. Ersbøll. Shape and texture based classification of fish species. SCIA 2009, pages 745–749.

[5] D. Lee, J. Archibald, R. Schoenberger, A. Dennis, and D. Shiozawa. Contour matching for fish species recognition and migration monitoring. In *Applications of Computational Intelligence in Biology*, pages 183–207, 2008.

[6] G. Nadarajan, Y.-H. Chen-Burger, R. Fisher, and C. Spampinato. A flexible system for automated composition of intelligent video analysis. In *ISPA 2011*, pages 259 –264.

[7] C. V. Pattengill-Semmens and B. X. Semmens. Conservation and management applications of the reef volunteer fish monitoring program. *Environmental Monitoring and Assessment*, 81:43–50, 2003.

[8] D. Pelletier, K. Leleu, D. Mallet, G. Mou-Tham, G. Herv, M. Boureau, and N. Guilpart. Remote high-definition rotating video enables fast spatial survey of marine underwater macrofauna and habitats. *PLoS ONE*, 7(2), 02 2012.

[9] A. Rova, G. Mori, and L. M. Dill. One fish, two fish, butterfish, trumpeter: Recognizing fish in underwater video. In *In IAPR Conference on Machine Vision Applications*, pages 404–407, 2007.

[10] C. Spampinato, D. Giordano, R. Di Salvo, Y.-H. J. Chen-Burger, R. B. Fisher, and G. Nadarajan. Automatic fish classification for underwater species behavior understanding. ARTEMIS '10, pages 45–50, 2010.

[11] C. Spampinato, S. Palazzo, B. Boom, J. Ossenbruggen, I. Kavasidis, R. Salvo, F.-P. Lin, D. Giordano, L. Hardman, and R. Fisher. Understanding fish behavior during typhoon events in real-life underwater environments. *Multimedia Tools and Applications*, pages 1–38, 2012.

[12] C. Spampinato, S. Palazzo, D. Giordano, I. Kavasidis, F.-P. Lin, and Y.-T. Lin. Covariance based fish tracking in real-life underwater environment. In *VISAPP 2012*, 2012.

[13] N. J. C. Strachan. Recognition of fish species by colour and shape. *Image Vision Comput.*, 11(1):2–10, Jan. 1993.