

## Fish4Knowledge Publishable Summary - Year 1

### 1 Project Context

The study of marine ecosystems is vital for understanding environmental effects, such as climate change and the effects of pollution, but is extremely difficult because of the inaccessibility of data. Undersea video data is usable but is tedious to analyse (for both raw video analysis and abstraction over massive sets of observations), and is mainly done by hand or with hand-crafted computational tools. Fish4Knowledge will allow a major increase in the ability to analyse this data: 1) Video analysis will automatically extract information about the observed marine animals which is recorded in an observation database. 2) Interfaces will be designed to allow researchers to formulate and answer higher level questions over that database.

The project will investigate: information abstraction and storage methods for reducing the massive amount of video data (from  $10E+15$  pixels to  $10E+12$  units of information), machine and human vocabularies for describing fish, flexible process architectures to process the data and scientific queries and effective specialised user query interfaces. A combination of computer vision, database storage, workflow and human computer interaction methods will be used to achieve this.

### 2 Publishable Project Summary

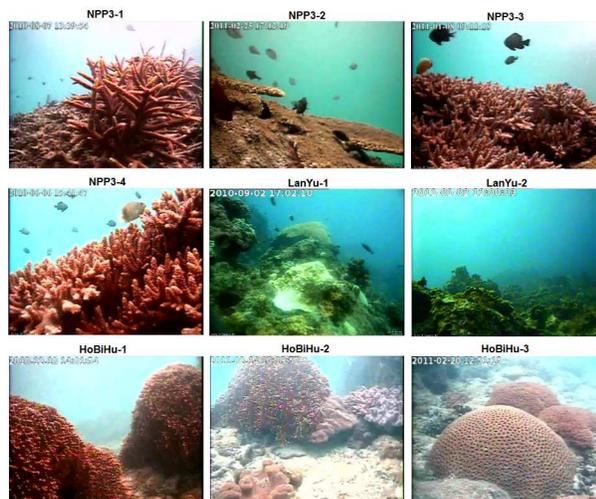
Based on visits to and discussions with several marine biologists, we have identified a set of 20+ questions to focus on, which has led to an analysis of their and the system's information needs and from this we have an initial design for the user interface, a portion of which is shown here:



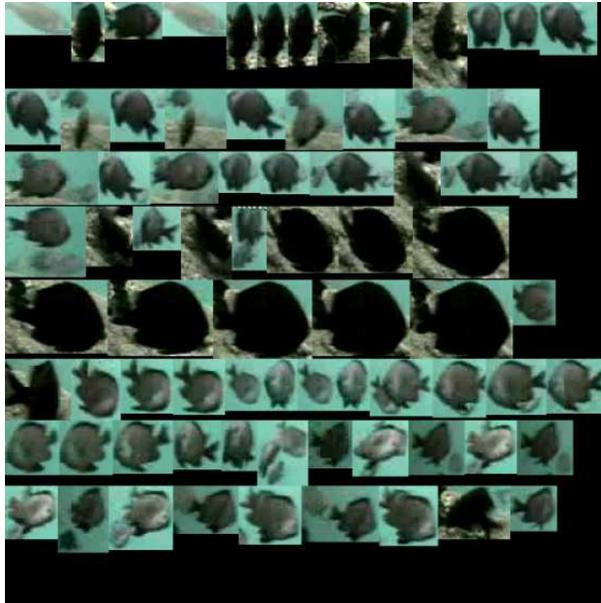
This has led to the design of the software data collection and analysis architecture, consisting roughly of the stages of video capture, fish detection, tracking and description, after which fish species recognition occurs.

The data sources are in place: 9 cameras (10th not operational at the moment) at 4 different locations, for which we have approximately 4000 hours of video now recorded. Here are views from the 9 cameras. To the right is the 48 processor system which has been assembled at NARL

for executing the project code.



The first stage of processing of the video is fish detection and tracking, for which we have fairly effective processes running. These have been applied to 471 hours of video, resulting in 3869473 detections of 456622 tracked fish. The detection and tracking data is being recorded in an SQL datafile currently hosted at Catania. Some example fish detections are:



From these, we have extracted about 3000 detections for initial work on species recognition, resulting in a cross-validated ground truth over about 25 species. Here are the top 11 species:



Current the main workflow is the production data collection and analysis, which is expected to be task farmed across the 48 processors. In addition, the workflow team has completed a data, capability and user needs ontology for the project, which will be part of the foundation for the dynamic workflow needed for query answering.

A Second Life virtual world has been constructed as a public virtual showcase for the project. Here are images of the upper (academic showcase) floor:



and the lower (public interest) floor:



### 3 Expected final results

The project will use live video feeds from 10 underwater cameras as a testbed for investigating more generally applicable methods for capture, storage, analysis and querying of multiple video streams. We will collate a public database from 2 years containing video summaries of the observed fish and associated descriptors. Expert web-based interfaces will be developed for use by the marine researchers themselves, allowing unprecedented access to live and previously stored videos, or previously extracted information. The marine researcher interface will also allow easy formulation of new queries. Extensive user community evaluations will be carried out to provide information on the accuracy, ease and speed of retrieval of information.

### 4 Project public website

[www.Fish4Knowledge.eu](http://www.Fish4Knowledge.eu)