A participative tool for sharing, annotating and archiving submarine video data

Yann Marcon*
Deep Sea Ecology and Technology
Alfred Wegener Institute
Helmholtz Center for Polar and Marine Research
Bremerhaven, Germany
Email: Yann.Marcon@awi.de

Renzo Kottmann*
Microbial Genomics and Bioinformatics Group
Max Planck Institute for Marine Microbiology
Bremen, Germany
Email: rkottman@mpi-bremen.de

Volker Ratmeyer
MARUM Center for Marine Environmental Sciences
University of Bremen
Bremen, Germany
Email: ratmeyer@marum.de

Antje Boetius
Deep Sea Ecology and Technology
Alfred Wegener Institute
Helmholtz Center for Polar and Marine Research
Bremerhaven, Germany
Email: Antje.Boetius@awi.de

Abstract— **The VIDLIB Deep-Sea Video Platform is an open access web-based tool for marine video data storage, streaming, sharing, and analysis. Using VIDLIB specialists can share, access and annotate the same videos from anywhere, thus accelerating the video annotation and analysis process. This way, scientists can share expert knowledge for video analysis (i.e. species identification) without the need to upload and download large video files. Moreover, the tool has the functionalities for participatory science, and science communication in that non-specialists can ask questions or comment on what they see, and get answers from scientists. VIDLIB is available at http://vidlib.marum.de. In this paper we describe the structure and workflow of the VIDLIB Deep-Sea Video Platform and present an example of analysis from video data in the Southwest Indian Ridge.**

Keywords— deep sea; video; database; open access; outreach; communication; video analysis

I. Introduction

Oceans cover more than 70 percent of the Earth's surface and are known to play an essential role for all of the Earth systems and element cycles. However, less than 5 percent of the ocean bottom has been explored so far and many aspects of the underwater world remain poorly known or understood – especially how it may influence anthropogenic activities or how it is affected by them. Therefore, increasing our ocean literacy is a necessity in order for specialists and non-specialists to better grasp the roles of the ocean on the Earth's system, its resources, as well as the chances and risks associated with human activities. Fortunately, manifold technological advances now permit marine scientists to access and study some of the deepest and most remote areas of the world's oceans, which has resulted in the discovery of new forms of habitats and life.

A key innovation for exploration and monitoring of deep ocean systems is the use of deep-sea robotics such as remotely operated vehicles (ROV), manned submersibles or towed camera systems, which produce ever-increasing quantities of video data. Especially with the use of HD cameras single dives with a towed camera commonly produce up to several hundreds of gigabytes of video data. However, using such data for research, science communication and public outreach purposes remains difficult as tools for accessing and sharing scientific videos are often lacking. Indeed, to date, there is no common solution for the storage, management, annotation and analysis of submarine video data. As a result, these unique data and documents of a vastly under-sampled ocean are scattered across multiple research institutes and/or working groups and it is difficult for scientists, as well as for the general public, to get an overview of the whereabouts of data, if they exist, how to access to them, how to query them, etc.

We present here the VIDLIB Deep-Sea Video Platform (VIDLIB), a web-based tool for archiving, streaming, sharing and annotating timecoded scientific deep-sea video data. VIDLIB provides an efficient way to ensure that marine video data remain easily accessible, query-able and retrievable. It also constitutes a participatory tool for scientists and laymen alike to work together on video data and exchange expert knowledge during the data analysis, as well as opportunities to document the scientific progress in investigating unknown life and habitats and invite non-specialists to participate.

In addition, VIDLIB is highly valuable in terms of scientific public outreach and popular science. Video data are by far the most efficient way to communicate scientific findings to a non-expert public. In this regard, video data of the deep-sea – a remote and barely accessible environment – may be particularly appealing to the general public. Moreover, we must consider that ocean and marine research is expensive and

* Both authors contributed equally to the work.

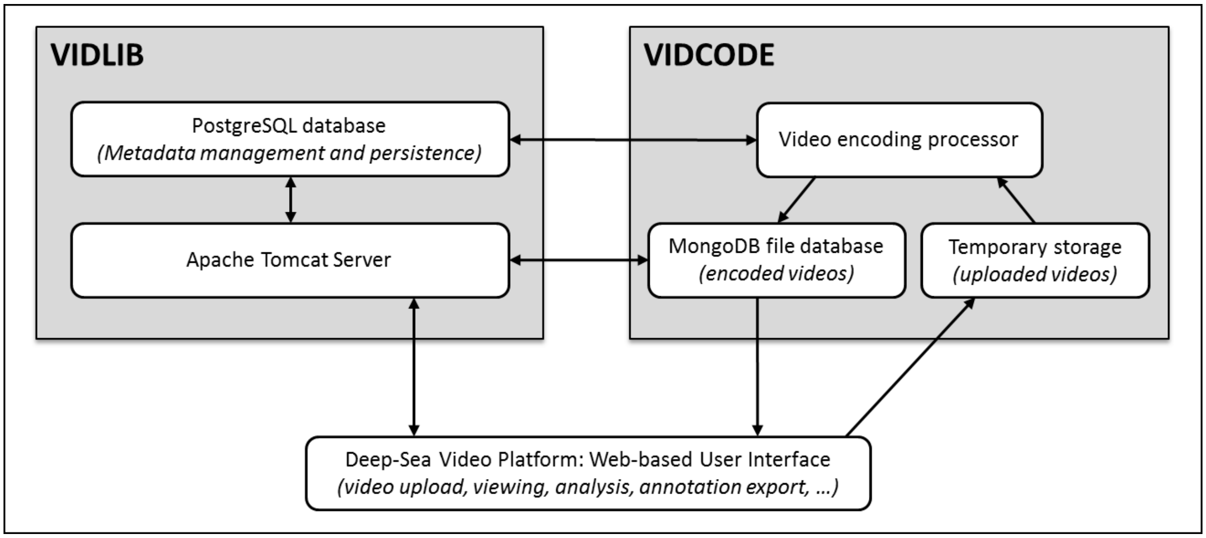

Fig. 1. Architecture of VIDLIB tool. Here, VIDLIB and VIDCODE are servers hosted at the MARUM. VIDLIB runs the meta-database and the web-based interface whereas VIDCODE stores the video files and runs the encoding processes.

that a large part of it is supported by public funds. Therefore, the results and products of such research should be easily accessible to the public and, thus, help raising awareness of remote habitats and increasing public interest on environmental protection issues. To this end, VIDLIB offers streaming capabilities in order to allow users to view the videos online from anywhere. The VIDLIB Deep-Sea Video Platform is available on the internet at http://vidlib.marum.de and can be accessed by personal computers, tablets and mobile phones with a recommended minimum bandwidth of 1+ Mbps.

Sections 2 and 3 of this paper present the structure of the servers and databases that constitute the VIDLIB Deep-Sea Video Platform and describe the workflow and the main features of VIDLIB. In section 4, we show an example of data analysis with VIDLIB, from the data pre-processing to the results.

II. VIDLIB ARCHITECTURE

The databases and web servers are currently entirely hosted at the MARUM institute and are comprised of two virtual servers, VIDLIB and VIDCODE (Fig. 1). The VIDLIB server runs the web-interface as well as the meta-database (PostgreSQL/PostGIS), whereas the VIDCODE server hosts the video file-database (MongoDB) and background processes such as the encoding of the video files into several suitable formats for video streaming supporting all major browsers (Internet Explorer/Edge, Firefox and Chrome).

The PostgreSQL meta-database manages all metadata associated with the video files, the annotations and the users. The simplified structure of the main PostgreSQL database is illustrated in Fig. 2. There are two types of metadata: dataset metadata and analysis metadata (Table 1).

TABLE I. MAIN TYPES METADATA USED BY VIDLIB

Dataset Metadata	Analysis Metadata
Expedition name Dive name Video name Date and time of acquisition Location (Latitude/Longitude) Description Country, etc.	Annotation Type Category Keyword Video ID Timecode Pixel position on the image frame Duration Comment Creation date Annotator's name

Dataset metadata refers to any information that is associated with the dataset but that does not result from the analysis itself. It includes particular information that is required for data querying and retrieval such as date and time of acquisition, expedition and dive numbers, etc. Dataset metadata are stored in the database in a set of tables related to the project (e.g. the expedition), the video set (e.g. the dive) and the video footage itself. These data are compulsory and must be provided with every video file upload.

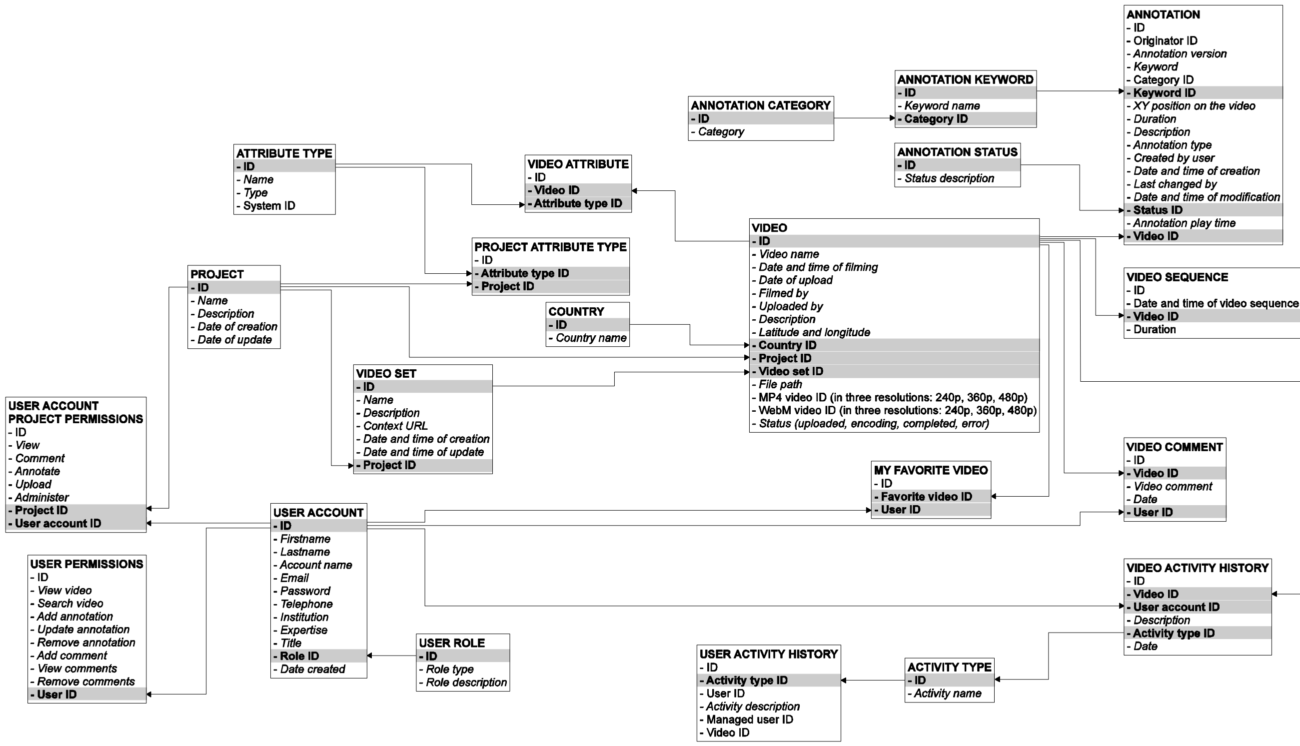

Fig. 2. Relational model of the PostgreSQL meta-database.

In addition to the dataset metadata, the meta-database also stores the analysis metadata, which includes all results of analyses that are conducted with the VIDLIB interface, e.g. the annotations and their position in the video, the annotator's name and so on. Analysis metadata can be found in the tables related to the actual annotations or categories (e.g. marine geology, marine biology) and keywords (e.g. sea cucumber, starfish).

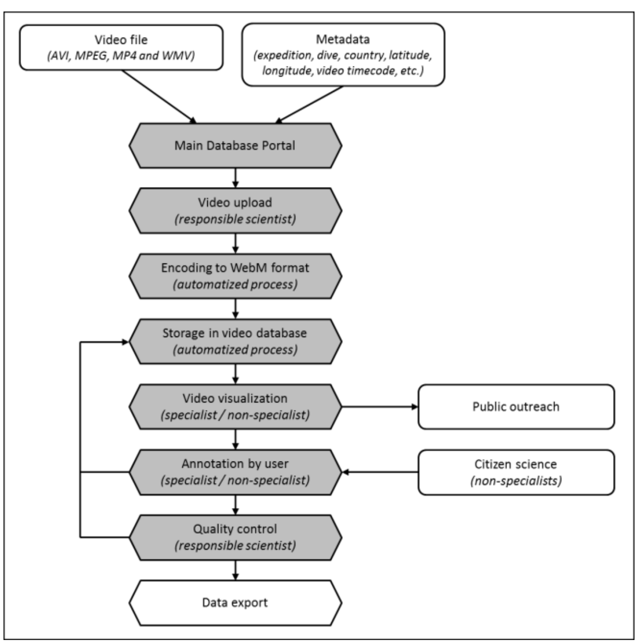

Fig. 3. Flow-chart showing the main steps of the VIDLIB workflow.

Finally, another set of tables are used to keep track of the different users, as well as their role (user, annotator, uploader, and administrator), permissions, activity history, favorite video, etc.

III. WORKFLOW DESCRIPTION

The different steps of the VIDLIB workflow are illustrated in Fig. 3. All steps can be performed through the main web-interface of VIDLIB. However, to ensure consistency and correctness in the data, some of the steps (e.g. data upload, quality control) must be performed by the scientist responsible for the dataset.

A. Data upload

The VIDLIB Deep-Sea Video Platform is compatible with a number of video types. For instance, video files in different container formats such as the AVI, MPEG, MP4 and WMV formats can be uploaded to the database. When uploading a video file, users must provide meta-data that includes the latitude and longitude where the video data was acquired as well as the exact date and time of the start of the video file. This information can later be used to query and retrieve the data as well as timestamp all annotations. If desired, a watermark logo can be provided that is later superimposed on the video to honor copyright.

Upon successful upload onto the database the new video files are automatically re-encoded in the WebM format, regardless of input format. This permits rapid online streaming. Other advantages of the WebM format are that it is a royalty-free format and it is natively supported by most internet-browsers, meaning no external plugins should be installed to stream videos. Videos are encoded into three different resolutions with progressive scanning (240p, 360p and 480p) in order to ensure fluidity and minimize the amount of artifacts when streaming the videos. The encoding step may take several hours depending on the duration and resolution of the original video. Therefore, if several videos are simultaneously uploaded, the encoding processes may get queued in order to avoid a processing overload on the VIDCODE server. This process is entirely automated and videos are directly accessible for the VIDLIB server once the encoding is complete. Videos are then available to users through the main web-based interface for viewing and analysis (Fig. 4).

B. *Data visualization, analysis and fusion*

At this point, users can freely stream the videos. While visualizing the videos, users with annotator rights can access the annotation interface. This interface provides tools to mark and comment all observations on the video image itself (Fig. 4). Annotation comments can contain either free text or selectable keywords. The list of keywords can be customized for the different video files depending on the needs. Annotations are classified into different categories (e.g. marine geology, marine biology, etc.) in order to allow viewers to select the different types of annotations while visualizing a video file. If selected, annotations will appear on the video when their starting timestamp is reached and then disappear when the annotation process is completed.

After the data analysis is complete, the results can be exported into both text and Excel files. Users can also export annotation results for a batch of videos into a single file. The exported annotations include additional information, such as the identity of the annotator, the position of the annotation in the image frame, and the absolute date and time of the video frame related to the annotation. The annotator's identity is stored as metadata with every annotation in order to facilitate quality control. The main purpose of time-stamping every single annotation is to permit the merging of the video analysis results with the underwater navigation data in order to geo-reference and display the spatial distribution of all observations. Time-stamped annotations can also be fused with additional dive metadata such as physicochemical measurements from in-situ sensor packages (e.g. temperature, turbidity, salinity, chemical concentrations, etc.).

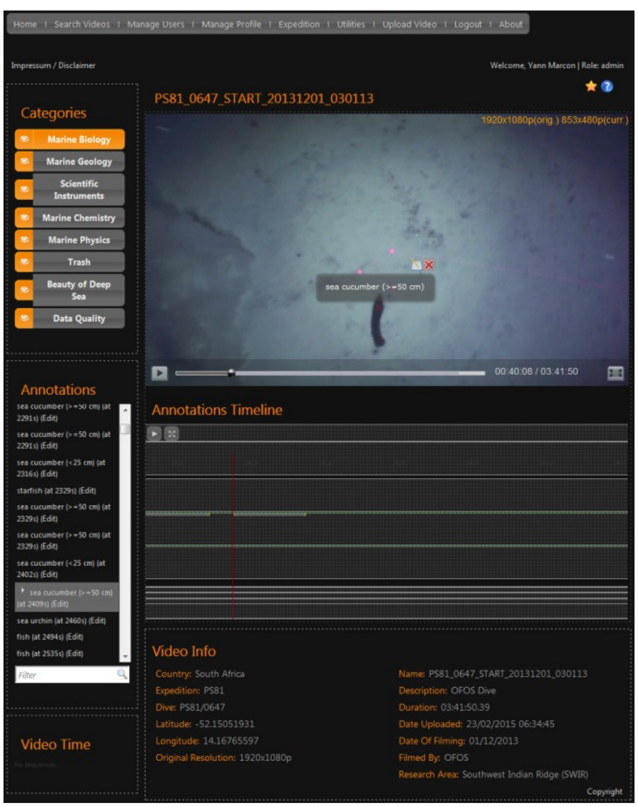

Fig. 4. Visualizing and annotation interface of the VIDLIB Deep-Sea Video Platform.

IV. EXAMPLE OF DATA PROCESSING, ANALYSIS AND GIS EXPORT

The VIDLIB database was successfully used to analyze the megafauna density in video surveys acquired with the Ocean Floor Observation System (OFOS) at the Southwest Indian Ridge (SWIR) during the PS81 expedition with RV Polarstern in November 2013 [1]. The OFOS system is a towed-camera system equipped with a high-definition video camera (Sony FCB-H11) and a photo camera (Canon EOS-1Ds Mark III). In addition to the cameras, the OFOS system includes an ultra-short baseline (USBL) positioning system (IXBLUE USBL-BOX) with a POSIDONIA acoustic antenna [2], an altimeter and three lasers (positioned 50 cm from each other) for scaling purposes. In this example we show the analysis workflow, from the data pre-processing to the final results.

A. *Video pre-processing*

To prevent any error or system instability that may arise during the acquisition, the OFOS generates several smaller video files of a few minutes in duration instead of one large single file for the entire dive. When uploading the video to VIDLIB for analysis, we concatenated all individual video files to obtain one single video file per OFOS dive. This resulted in two video files of 3h30 (7.7 GB) and 5h40 (14.3 GB) for the

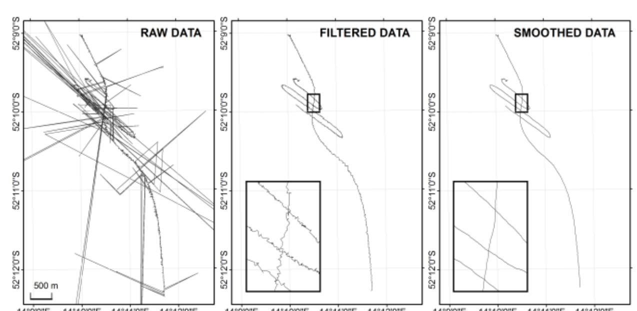

Fig. 5. Steps of the USBL data processing workflow. The raw data contain many outliers, which are filtered out with a speed filter. The remaining high-frequency noise of the data is reduced with a moving average filter.

OFOS deployments PS81/0347 and PS81/0350 respectively. The video upload carried out on the MARUM intranet only took a few minutes. However, the automatic video encoding took approximately 24 hours for the longer dive.

B. Navigation data processing

The performance and accuracy of USBL positioning systems decreases greatly with increasing depth. Because of the considerable depth of the study area (3600-3800 m), the recurrence of the USBL transponder (i.e. the time-delay between two positions) had to be set to 10 seconds. Furthermore, the navigation data provided by the USBL positioning system were impacted by a number of erroneous and inaccurate positions, mainly caused by the high depth of the area and the roughness of the sea state at those latitudes.

We first applied a speed filter (1 knot) to the navigation data to correct for the inaccuracies of the USBL data and remove large outliers. Next, the high-frequency noise of the dive track was smoothed with a moving average filter (Fig. 5). Finally, the dive track was spline-interpolated using a piecewise cubic polynomial for every second of the dive.

C. Video analysis and merging with navigation data

The video analysis was done by annotating every benthic organism that could be identified, such as sea cucumbers, starfish and shrimp. It is important to be aware of the resolution of object sizes; e.g. in this study approximately 10 cm, so that smaller organisms are not reported. Here we focused on sea cucumbers, which have recognizable shapes and can be easily identified, also in areas where the video quality decreased. Additionally, we could group the sea cucumbers into three categories based on their length using the laser points: small (<25 cm), medium-size (25-50 cm) and large (>50 cm).

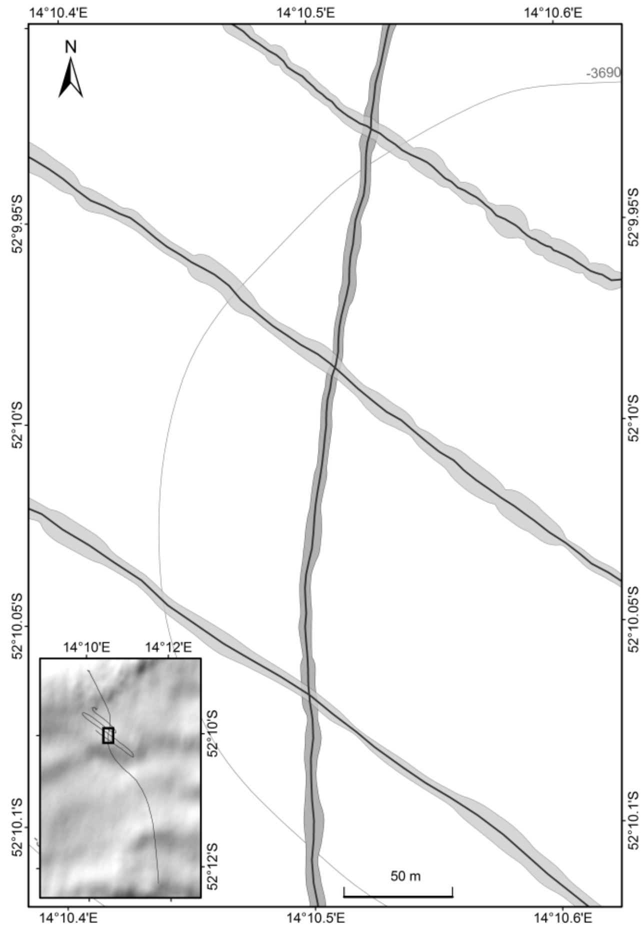

Fig. 6. The field of view of the camera on the sea floor is a function of the distance between the camera and the bottom. The variations of the camera altitude are mainly caused by the swell at the ocean surface, which makes it difficult to maintain the OFOS at a constant altitude.

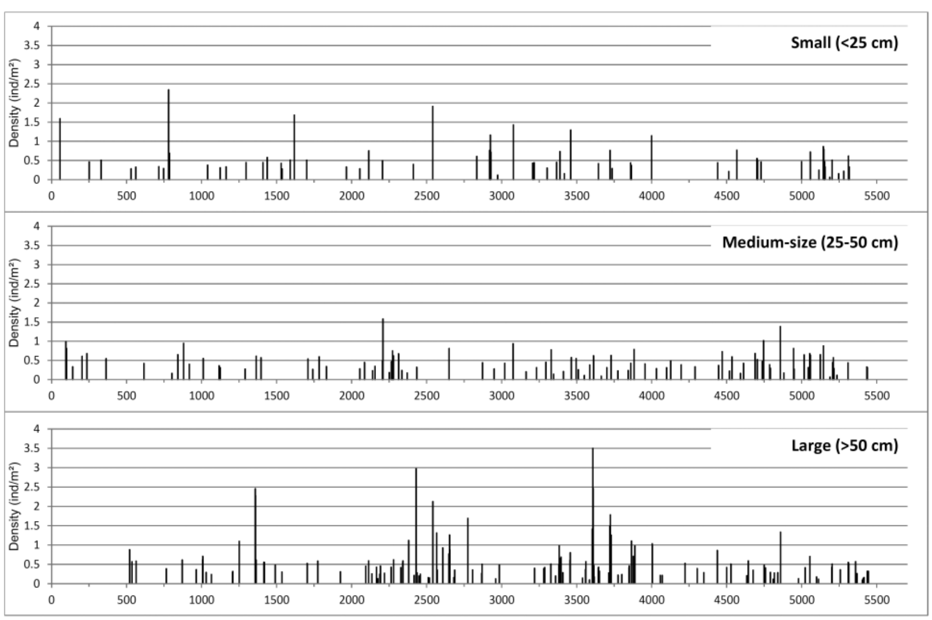

Fig. 7. Plots showing the variations of the sea cucumber density for every 1 m segment of the bottom track of the PS81/0347 OFOS deployment. The horizontal axis is expressed in meters from the start of the bottom track. Only the density distribution of the large sea cucumber (> 50cm) seems to show a trend, with elevated densities towards the middle of the deployment.

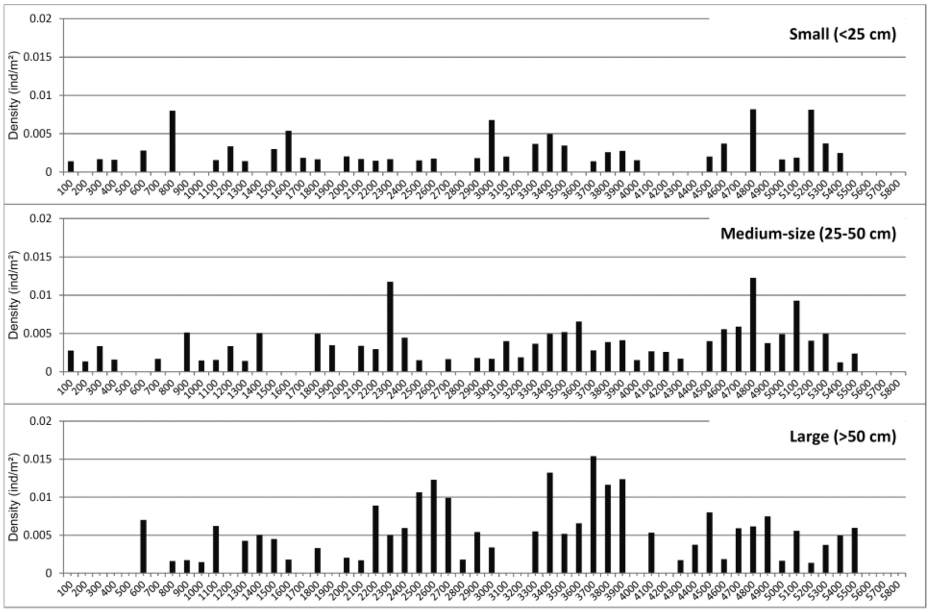

Fig. 8. Plots showing the variations of the sea cucumber density for each 100 m segment of the bottom track of the PS81/0347 OFOS deployment. The horizontal axis is expressed in meters from the start of the bottom track. These results show more clearly than Fig. 6 that the density of large sea cucumber is more elevated near the middle of the dive track.

After the video analysis, all annotations were exported to an Excel file and georeferenced with the 1-second interpolated navigation data. The densities of sea cucumbers were calculated for every 1 m and 100 m segments of the dive tracks. The areas used for the density calculations were estimated from the altitude data and the camera calibration parameters, which were used to compute the along-track field-of-view of the camera (Fig. 6). This provides density estimations for the entire transects, without altitude-induced biases (Fig. 7-8).

The results allow us to quickly detect areas with higher or lower faunal densities (Fig. 7-9). This information is crucial in order to understand what other factors may explain these distributions and guide further sampling in this study area. In our example, the zones of high sea-cucumbers densities may be partly related to some high reliefs or slopes on the sea floor (Fig. 9), which may be exposed to a higher energy supply. However, more data is needed to confirm this hypothesis.

V. CONCLUSION AND PERSPECTIVES

Currently, in comparison to similar databases used in other countries [3]–[5], the VIDLIB video platform is still at its early stages. However, it already contains more than 330 videos of the deep-sea that were acquired with the ROV QUEST 4000 and the OFOS during several German research cruises to hydrothermal vents, cold seeps, ultraslow spreading ridges and arctic areas [1], [6]–[8]. The videos are freely accessible and the platform can be used for data analysis as well as outreach purposes. As an example, it was used to detect and extract all sequences containing deep-sea fishes to provide these as footage to science media and television.

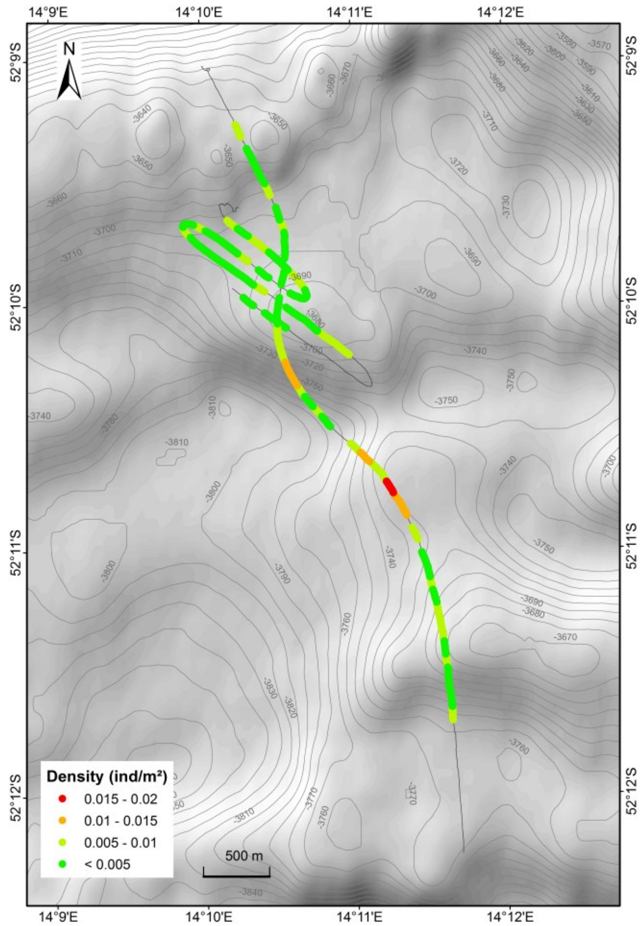

Fig. 9. Density distribution of the large sea cucumbers (> 50cm) along bottom tracks of the PS81/0347 and PS81/0347 OFOS deployments. The black lines indicate the parts of the dive tracks for which the quality of the video data was too low (e.g. elevated turbidity, altitude of the camera too high) to reliably be analyzed. According to the bathymetry, the zones of high sea cucumber density may coincide with highs and lows in the bathymetry. Those areas may be exposed to higher energy supply.

The VIDLIB Deep-Sea Video Platform has been designed in a modular and flexible way that should facilitate future linking with national and international scientific data repositories. Linking VIDLIB to existing databases and repositories would help maximizing performances by decentralizing data archives and optimizing the storage capacity. Indeed, it would allow users to use VIDLIB as a front-end tool to visualize and annotate data that are hosted on different repositories.

Furthermore, by allowing the use of both free text and keywords for annotating videos, VIDLIB makes the data analysis both flexible and fast. Nevertheless, in order to further improve the consistency of annotations as well as the searching and querying possibilities, we may consider implementing the use of a hierarchical knowledgebase [3], [9] in future developments of this platform.

ACKNOWLEDGMENT

We would like to thank Goce Mihailovski, Mile Stefanovski, Marjan Sterjev, Mathias Weinrebe and Paul Wintersteller and for their help and support. VIDLIB was developed within the framework of a cross-institute cooperation that involved the MARUM at Bremen University, the Max Planck Institute for Marine Microbiology in Bremen, and the Alfred Wegener Institute in Bremerhaven, Germany. Its development is funded through the European FP7 EUROFLEET project and the Helmholtz Alliance "Robotic Exploration of Extreme Environments – ROBEX".

REFERENCES

[1] V. Schlindwein, "The expedition of the research vessel 'Polarstern' to the Antarctic in 2013 (ANT-XXIX/8) / Ed. by Vera Schlindwein with contributions of the participants," Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Berichte zur Polar- und Meeresforschung = Reports on polar and marine research 672, 2014.

[2] J. P. Peyronnet, R. Person, and F. Rybicki, "POSIDONIA 6000: a new long range highly accurate ultra short base line positioning system," in *OCEANS'98 Conference Proceedings*, 2002, vol. 3, pp. 1721–1727.

[3] B. M. Schlining and N. J. Stout, "MBARI's Video Annotation and Reference System," in *OCEANS 2006*, Boston, MA, 2006, pp. 1–5.

[4] M. Leslie, N. Scott, E. Guillemot, and V. Auger, "Video acquisition, archiving, annotation and analysis: NEPTUNE Canada's real-time georeferenced library of deep sea video," in *OCEANS 2010*, Seattle, WA, 2010, pp. 1–9.

[5] Y. Hanafusa, H. Saito, and Y. Abe, "The JAMSTEC Metadata Publication and Search System," *Data Sci. J.*, vol. 12, no. 0, pp. WDS221–WDS224, 2013.

[6] G. Bohrmann, A. Bahr, F. Brinkmann, M. Brüning, S. Buhmann, V. Diekamp, K. Enneking, D. Fischer, A. Gassner, G. Von Halem, D. Huettich, S. Kasten, S. A. Klapp, M. Nasir, N. Nowald, W. T. Ochsenhirt, T. Pape, V. Ratmeyer, R. Rehage, J. Rethemeyer, M. Reuter, P. Rossel, M. Saleem, W. Schmidt, C. Seiter, S. Stephan, K. Thomanek, N. Wittenberg, M. Yoshinaga, and K. Zonnefeld, "Report and Preliminary Results of R/V Meteor Cruise M74/3, Fujairah-Male, 30 October-28 November, 2007: Cold Seeps of the Makran Subduction Zone (continental Margin of Pakistan)," Geowissenschaften, Universität Bremen, Bremen, Cruise report 266, 2008.

[7] W. Bach, N. Dubilier, C. Borowski, C. Breuer, B. Brunner, P. Franke, O. Herschelmann, S. Hourdez, L. Jonda, N. Jöns, S. Klar, K. Koloa, A. Mai, A. Meyerdierks, I. Müller, S. Petersen, P. Pjevac, V. Ratmeyer, E. Reeves, R. Rehage, C. Reuter, A. Schaen, L. Shu, J. Thal, and M. Zarrouk, "Report and preliminary results of RV SONNE Cruise SO-216, Townsville (Australia)-Makassar (Indonesia), June 14-July 23, 2011," Geowissenschaften, Universität Bremen, Bremen, Cruise report 280, 2011.

[8] A. Boetius, "The expedition PS86 of the research vessel POLARSTERN to the Arctic Ocean in 2014," Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Berichte zur Polar- und Meeresforschung = Reports on polar and marine research 685, 2015.

[9] J. A. Harrison, M. J. Wyatt, D. Marrable, M. Gray, and L. Edwards, "Collaborative and Annotation Tools for Analysis of Marine Imagery and video (CATAMI)," in *2012 IEEE 8th International Conference on eScience*, Chicago, IL, 2012.