VIDEOMODULE TOWED SYSTEM: ACQUISITION AND ANALYSIS OF VIDEO IMAGING DATA FOR BENTHIC SURVEYS

By Ivan Anisimov, Andrey Lesin, Valeriya Muravya, Anna Zalota, and Maxim Zalota

SCIENTIFIC MOTIVATION

Ecological studies require two types of primary information: qualitative (what species are present in the study area) and quantitative (e.g., number, biomass, size structure). Such information is difficult to gather for underwater communities, especially in the deep sea, where scuba diving for census-taking is impossible. Traditionally, tools such as grabs and trawls have been used, but they have their limitations: grabs sample small areas, and the location of trawl samples is imprecise. In addition, these tools damage both the organisms and their habitats.

Still photography and video footage have been used for effectively decades to study underwater species (Mallet and Pelletier, 2014), as these tools are used to identify organisms that live above the substrate. Benthic video surveys are carried out with various underwater vehicles including remotely operated and autonomous underwater vehicles and towed camera systems that allow researchers to observe and document habitats and species in their natural settings without significant disturbance. Towed camera systems in particular offer advantages that enhance the effectiveness of benthic video surveys. Towed systems can efficiently cover extensive areas of the seafloor to provide a broad spatial context that is difficult to achieve with classic sampling methods or other types of vehicles. The connection of these towed systems to the vessel facilitates the estimation of transect distance and area covered. Towed systems operate at consistent speeds and depths, ensuring uniform data collection and reducing variability in the footage. Additionally, towed cameras are adaptable to various depth ranges, from shallow coastal waters to the deep sea, making them suitable for diverse marine environments. A comprehensive overview of such systems can be found in Durden et al. (2016).

Currently, numerous groups around the world are working to develop automatic identification systems for organisms using machine learning methods (ML; Li et al., 2023). Most ML efforts and error assessments are concerned with species identification: effectively finding an object in the footage and accurately identifying it and its taxonomic level. Very little attention is paid to the quantitative parameters. In machine learning, there are no standard technical or analysis parameters to account for the differences in underwater vehicles, such as in camera quality and distortion, water transparency, light level, and measurement methods, making comparative studies unreliable. Even multiple surveys of the same area using the same apparatus can result in significant differences in the quality of video footage due to, for example, different water transparency conditions. Therefore, the number of observed specimens may be undercounted because of actual changes in the assemblage or simply lower visibility in the water column.

We aim to develop an underwater vehicle and analysis tool that can minimize technical and environmental factors affecting video footage quality and quantify errors in the comparative quantitative analysis of underwater objects.

VIDEOMODULE TOWED SYSTEM

A towed system called *Videomodule* was designed at the Shirshov Institute of Oceanology for benthic research at depths of up to 6,000 m (Figure 1). *Videomodule* operates at a standard towing speed of approximately 0.5 knots and maintains an average height of 1.5 m above the seafloor. At this speed, the system is situated beneath the A-frame of the vessel, thereby facilitating system positioning. Eight 30 kg weights are distributed in a manner that minimizes pitch during towing. Stabilizers are mounted in the rear of the system to eliminate roll and course rotations.

Videomodule is equipped with a video camera, a still camera, a side-scan sonar, and various auxiliary sensors, including a depth gauge, an altimeter, and an inclinometer. The system uses six 30 W LED lights and a 200 W strobe light for optimal illumination. Two parallel red lasers fixed 20 cm apart are mounted next to the camera to provide a scale for the frame. The video camera is a 4K IP surveillance camera with a Sony Exmor R sensor that provides 6 MP resolution and a 76° field of view underwater. The down-looking camera captures a bird's-eye view of the seafloor to identify the best possibilities for quantitative analysis of examined species. At a standard height above the seabed, organisms approximately 1 cm in size can be distinguished in the video. The photo system is based on a consumer-grade Sony ILCE 7RM2 digital photo camera with 24.3 MP resolution and a 56° field of view under water, enabling the study of smaller organisms (less than 1 mm). For operational purposes, two analog TV cameras, one down-looking and one tilted, transmit low latency video onboard, allowing operators to avoid obstacles on the seafloor. The side-scan sonar operates at a frequency of 240 kHz, enabling high-resolution imaging of a 300 m wide strip of seabed.
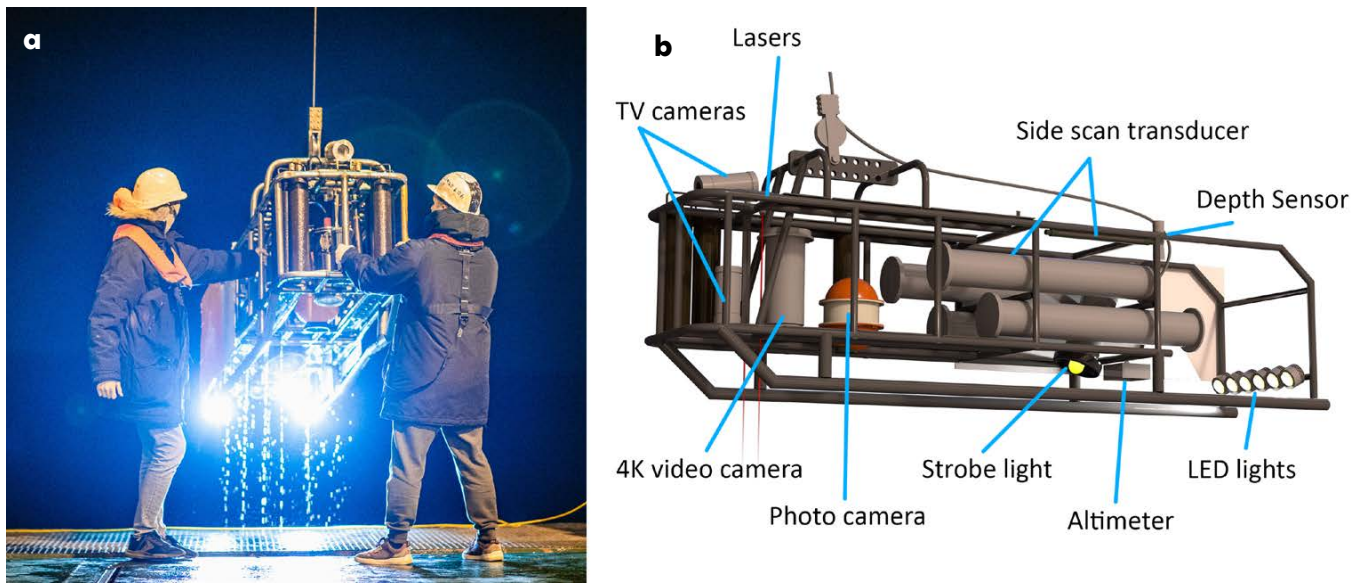
FIGURE 1. *Videomodule* towed system designed at the Shirshov Institute of Oceanology. (a) *Videomodule* is recovered here during the 85th expedition of R/V *Akademik Mstislav Keldysh* in 2021. *Photo credit: Ivan Potylicin* (b) Schematic shows *Videomodule* equipment including cameras, side-scan sonar, and auxiliary sensors.

Videomodule is connected to the research vessel via a fiber-optic cable, allowing real-time data transmission to a computer on board the vessel. The system is powered by two LiFePO4 batteries that provide sufficient power for surveys lasting up to 12 hours at full load without recharging.

Surveys conducted with *Videomodule* are typically organized into 500–600 m transects. Each transect yields approximately 40 minutes of video footage, 80–100 photographs, and a single side-scan stripe image. Comprehensive metadata, including system depth, distance from the seafloor, pitch, roll, heading, and vessel coordinates, are recorded in a log file at one-second intervals. These metadata are also embedded in the video's captions and photos to facilitate analysis.

IMAGE ANALYSIS

Two main groups of software tools that have been developed for analysis of underwater video data (1) generate three-dimensional models of the seafloor based on photos and video sequences, and (2) provide tools for direct analysis of the images themselves using laser marks as references. The latter group is simpler and faster but is typically better suited for analyzing images of relatively flat seafloor areas. Another issue to consider when using such methods is the need for constant camera tilt to avoid perspective distortions (Istenič et al., 2020).

Image distortion can reduce the accuracy of species measurements. To address this issue, the camera's radial distortion coefficients were evaluated and corrections were made. In addition, we developed a novel method of perspective correction. Although similar methods are widely used in computer vision systems (e.g., self-driving cars), as far as we are aware, this method has not yet been applied to underwater image analysis. Our approach involves estimating the correspondence between consecutive images, triangulating three-dimensional points, and approximating them with a plane. This plane is then virtually reoriented to align with the image plane, thereby reducing measurement errors. At 1.5 m above the seafloor, we achieve an average measurement uncertainty of 6 mm (Anisimov, 2023).

For quantitative analysis of images aided by software applications, researchers primarily work with still images rather than continuous video data (Gomes-Pereira et al., 2016). The usual method for generating sets of still images from video involves extracting frames at fixed time intervals. However, this approach can result in gaps between or overlaps with subsequent frames, which can cause errors in the analysis. Some organisms may be counted multiple times, leading to inaccuracies in biomass and quantity estimates. To address this issue, we developed a Python-based program (available on GitHub) that analyzes the movement of objects in video sequences and calculates the displacement between two frames, referred to as "drift." The program can estimate the optimal time intervals for capturing frames to place seafloor images side by side without overlapping by determining the drift between frames. These non-overlapping frames are referred to as "seafloor fragments." Afterward, each seafloor fragment's distance between red laser marks is calculated to provide a scale. There is an option to manually adjust the drifts and laser mark positions if the program evaluates them incorrectly.

The program provides tools for marking different organisms (Figure 1a) and calculating their linear dimensions (Figure 2b). Every marked organism is associated with the corresponding seafloor fragment. Because the analysis is done using video rather than independent seafloor fragments, users can switch between video frames to select the best lighting and perspective for estimating dimensions without changing the seafloor fragment with which the organism is associated. As a result, we can increase the accuracy of identification and measurements of the organisms in a fixed area of the seafloor. These areas are stacked next to each other, forming the entire transect.

APPLICATION IN THE KARA SEA: SPREAD OF INVASIVE SNOW CRABS

Benthic surveys using *Videomodule* were conducted in the Arctic's Kara Sea, among other polar regions. The software was developed for and used extensively to study the spread of the large predatory snow crab *Chionoecetes opilio* into the Kara Sea (Zalota et al., 2019, 2020; Udalov et al., 2024). Unlike still images, long video transects supplied information on the size structure dynamics of the growing invasive crab population and its density on the seafloor in different regions of the Kara Sea. The software allowed us to use all video footage for the analysis rather than just random fragments, increasing the accuracy of identification by providing views of objects from different angles as the camera passes over, and increasing the accuracy of objects' measurements by using images with the most accurate angle of view in addition to distortion and pitch and tilt corrections. We observed that crabs rarely notice the vehicle and generally do not flee from it. The size structure of crab assemblages calculated from the video is less detailed than that of physical specimens caught by bottom trawls. This is due

to measurement errors, which we strive to minimize, and biological factors such as sexual dimorphism that cannot be accounted for in the video. However, the dominant size groups are easily identified from video data and correspond to those observed from trawl data (Zalota et al., 2019). Unlike bottom trawls, video data have allowed us to calculate crab densities in different regions of the sea and observe changes in these densities as the invasion progresses. In addition, no apparent clustering of crabs has been observed. Rather, they are evenly distributed on the flat Kara seafloor unless it has been disturbed. After bottom trawling, crabs gather around the trench created by it, possibly to consume unearthed burrowing organisms.

Videomodule has also been used to investigate the dynamics of benthic communities in response to the invasion of this active predator (Udalov et al., 2024). Research on this topic has been ongoing for more than 10 years, resulting in the collection and analysis of a substantial amount of data on benthic assemblages. The towed camera system has allowed diverse and long-term monitoring of large areas of the seafloor, revealing significant changes in the ecosystem due to the snow crab invasion.

FUTURE WORK

In the absence of standardized methodologies for evaluating the quality of images captured via benthic video surveys, there is a significant gap in our ability to accurately identify organisms. Variations in light conditions, water turbidity, the height of the camera above the seafloor, sensor resolution, and image compression can substantially influence the visibility of organisms. The size, texture, and color of the organisms play crucial roles in this regard. Our objective is to develop criteria that ensure the reliability and comparability of benthic organism analysis under diverse conditions
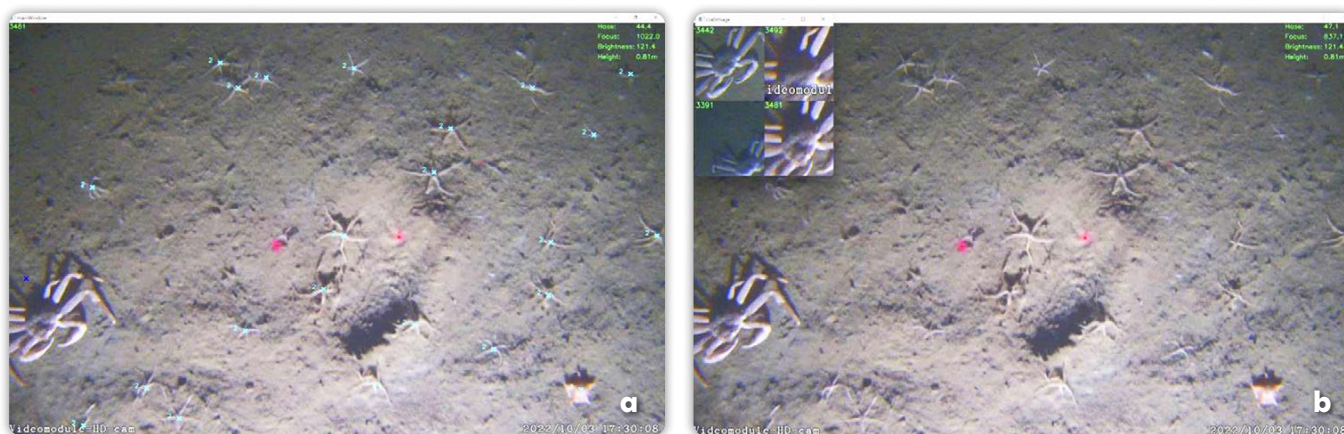

FIGURE 2. Two windows show the interface of *Videomodule*'s Python-based program for video data analysis. (a) The window for marking species in the frame. Light blue dots indicate marked species. Two red dots in the center are laser marks 20 cm apart. (b) The window for carrying out measurements. A dialog box at top right allows selection of optimal measurements for various organisms.

by assessing the impacts of these variables on image quality. At present, we can only reliably compare large organisms under different conditions; however, with such criteria, we could understand size limitations of organisms studied and the associated range of errors.

REFERENCES

Anisimov, I.M. 2023. Improving the accuracy of measuring seabed objects using video images with a laser scaler. [In Russian] *Underwater Investigations and Robotics* 4(46):16–28.

Durden, J.M., T. Schoening, F. Althaus, A. Friedman, R. Garcia, A.G. Glover, J. Greinert, N.J. Stout, D.O.B. Jones, A. Jordt, and others. 2016. Perspectives in visual imaging for marine biology and ecology: From acquisition to understanding. *Oceanography and Marine Biology: An Annual Review* 54:9–80, https://doi.org/10.1201/9781315368597.

Gomes-Pereira, J.N., V. Auger, K. Beisiegel, R. Benjamin, M. Bergmann, D. Bowden, P. Buhl-Mortensen, F.C. De Leo, G. Dionísio, J.M. Durden, and others. 2016. Current and future trends in marine image annotation software. *Progress in Oceanography* 149:106–120, https://doi.org/10.1016/j.pocean.2016.07.005.

Istenič, K., N. Gracias, A. Arnaubec, J. Escartín, and R. Garcia. 2020. Automatic scale estimation of structure from motion based 3D models using laser scalers in underwater scenarios. *ISPRS Journal of Photogrammetry and Remote Sensing* 159:13–25, https://doi.org/10.1016/j.isprsjprs.2019.10.007.

Li, J., W. Xu, L. Deng, Y. Xiao, Z. Han, and H. Zheng. 2023. Deep learning for visual recognition and detection of aquatic animals: A review. *Reviews in Aquaculture* 15(2):409–433, https://doi.org/10.1111/raq.12726.

Mallet, D., and D. Pelletier. 2014. Underwater video techniques for observing coastal marine biodiversity: A review of sixty years of publications (1952–2012). *Fisheries Research* 154:44–62, https://doi.org/10.1016/j.fishres.2014.01.019.

Udalov, A.A., I.M. Anisimov, A.B. Basin, G.V. Borisenko, S.V. Galkin, V.L. Syomin, S.A. Shchuka, M.I. Simakov, A.K. Zalota, and M.V. Chikina. 2024. Changes in benthic communities in Blagopoluchiya Bay (Novaya Zemlya, Kara Sea): The influence of the snow crab. *Biological Invasions* 26:3,455–3,473, https://doi.org/10.1007/s10530-024-03388-1.

Zalota, A.K., O.L. Zimina, and V.A. Spiridonov. 2019. Combining data from different sampling methods to study the development of an alien crab *Chionoecetes opilio* invasion in the remote and pristine Arctic Kara Sea. *PeerJ* 7:e7952, https://doi.org/10.7717/peerj.7952.

Zalota, A.K., V.A. Spiridonov, S. Galkin and A.A. Pronin. 2020. Population structure of alien snow crabs (*Chionoecetes opilio*) in the Kara Sea (trawl and video sampling). *Oceanology* 60(1):83–88, https://doi.org/10.1134/S0001437020010257.

AUTHORS

Ivan Anisimov (oceanbreak@gmail.com), **Andrey Lesin**, **Valeriya Muravya**, and **Anna Zalota**, Shirshov Institute of Oceanology of the Russian Academy of Sciences, Moscow, Russia. **Maxim Zalota**, WhoTalks.app, Westbrook, CT, USA.