An Approach to the Development of a Relational Database and GIS Applicable Scheme for the Analysis of Video-Based Surveys of Benthic Habitats

J. A. Sameoto, P. Lawton and M. B. Strong

Fisheries and Oceans Canada
Biological Station
531 Brandy Cove Road
St. Andrews, NB
E5B 2L9

2008

Canadian Technical Report of Fisheries and Aquatic Sciences 2818

Fisheries and Oceans Canada   Pêches et Océans Canada

Canada

Canadian Technical Report of
Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of the Department of Fisheries and Oceans, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in *Aquatic Sciences and Fisheries Abstracts* and indexed in the Department's annual index to scientific and technical publications.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and the Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page. Out-of-stock reports will be supplied for a fee by commercial agents.

Rapport technique canadien des
sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techni-ques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques du ministère des Pêches et des Océans, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications complètes. Le titre exact parait au-dessus du résumé de chaque rapport. Les rapports techniques sont résumé dans la revue *Résumés des sciences aquatiques et halieutiques*, et ils sont classés dans l'index annual des publications scientifiques et techniques du Ministère.

Les numéros 1 à 456 de cette série ont été publiés à titre de rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de rapports techniques de la Direction générale de la recherche et du développe-ment, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre. Les rapports épuisés seront fournis contre rétribution par des agents commerciaux.

Canadian Technical Report of
Fisheries and Aquatic Sciences 2818

2008

An Approach to the Development of a Relational Database and GIS Applicable Scheme for the Analysis of Video-Based Surveys of Benthic Habitats

by

J. A. Sameoto, P. Lawton and M. B. Strong

Fisheries and Oceans Canada
Biological Station, 531 Brandy Cove Road
St. Andrews, New Brunswick, Canada, E5B 2L9

This is the two hundred and eighty first Technical Report
of the Biological Station, St. Andrews, NB

Cat. No.Fs97-6/2818E ISSN 0706-6457

Correct citation for this publication:

Sameoto, J. A., Lawton, P., and Strong, M. B. 2008. An approach to the development of a relational database and GIS applicable scheme for the analysis of video-based surveys of benthic habitats. Can. Tech. Rep. Fish. Aquat. Sci. 2818: iv + 34 p.

TABLE OF CONTENTS

ABSTRACT

Sameoto, J. A., Lawton, P., and Strong, M. B. 2008. An approach to the development of a relational database and GIS applicable scheme for the analysis of video-based surveys of benthic habitats. Can. Tech. Rep. Fish. Aquat. Sci. 2818: iv + 34p.

An adaptation of an existing marine benthic habitat classification scheme is presented to facilitate the incorporation of video-based survey data into marine benthic habitat characterization. The present scheme is designed to describe seafloor environments through both qualitative and quantitative descriptors and is based on the classification system of marine sublittoral habitats proposed by Valentine et al. (2005). We discuss the importance of including information on video quality, since this can strongly influence the interpretation of video data, and present a database structure for our proposed classification scheme that allows easy migration of habitat data into Geographic Information Systems (GIS) for analysis and display. We illustrate our approach based on the analysis of video collected from a survey of deep seafloor habitats in the Discovery Corridor[1], Gulf of Maine, using the Canadian underwater research vehicle ROPOS. Although this scheme is based on observations taken of deep seafloor habitats in Northeastern North America, its structure is intended to be easily adapted to various seabed environments.

RÉSUMÉ

Sameoto, J. A., Lawton, P., and Strong, M. B. 2008. An approach to the development of a relational database and GIS applicable scheme for the analysis of video-based surveys of benthic habitats. Can. Tech. Rep. Fish. Aquat. Sci. 2818: iv + 34p.

Une adaptation d’un système de classification d'habitat benthique marin déjà établi, est présentée pour faciliter l'incorporation des données de sondage à base de vidéo dans la caractérisation d'habitat benthique marin. Le présent système est conçu pour décrire l’environnement du fond de la mer avec des descripteurs tant qualitatifs que quantitatifs et est basé sur le système de classification d'habitats sous littoraux marins proposé par Valentine et al. (2005). Nous discutons l'importance d'inclure l'information sur la qualité de la vidéo, puisque cela peut fortement influencer l'interprétation des données vidéo et nous présentons une structure de base de données pour notre système de classification proposé qui permet la migration facile des données d'habitat dans des Systèmes d'Information Géographiques (GIS) pour l'analyse et la visualisation. Nous illustrons notre approche basée sur l'analyse de vidéo rassemblée durant un sondage d'habitats des fonds profonds de la mer dans le Couloir de Découverte et le Golfe du Maine[1], en utilisant le véhicule canadien de recherche sous marine ROPOS. Bien que ce système soit basé sur des observations d'habitats prises dans les fonds profonds du Nord-est de l’Amérique du Nord, sa structure est conçue pour être facilement adapté aux environnements variés du fond des mers.

---

[1] Appendix A

INTRODUCTION

The demand to classify and identify benthic marine habitats has increased with the move by Fisheries and Oceans Canada (DFO) to adopt an ecosystem approach to ocean resource management and with increased concern over the conservation and protection of biodiversity and characteristic features of marine environments. Benthic habitat classification and mapping are valuable since they provide a means to understand the distributional patterns of populations, communities, and habitats. Currently, there is a concerted international effort to develop structured frameworks whereby various technical approaches to marine classification, mapping and biodiversity analysis can be linked to explicit ocean management objectives (recently reviewed by Cogan and Noji 2007).

There are a variety of approaches to benthic habitat classification and what constitutes a resulting habitat map can vary greatly from regional scale geophysical maps to detailed descriptions of biota. This diversity in benthic mapping products reflects the variability in survey technologies and classification schemes used. Currently, there is no single adopted classification scheme for the marine environment.

The use of a standard and systematic scheme would help facilitate scientific workflow, interoperability, and long-term interpretation of data. Variability in the organization of habitat classification schemes used to describe data can cause problems when comparing habitat information. Classification systems are often hierarchical (e.g. Cowardin et al. 1979, Allee et al. 2000, Connor et al. 2003) or partly hierarchical (e.g. Greene et al. 1999, Valentine et al. 2005, Greene et al. 2007), and the same habitat type or habitat descriptor may be at different levels across multiple schemes. Comparisons of classifications and cross study analysis between habitat maps are also difficult when the focus (e.g. geo-physical versus biological descriptors) and level of detail of the schemes vary.

Recently, Valentine et al. (2005) proposed a classification scheme designed to be applicable across various marine environments. It addresses the need for a holistic ecosystem-based approach to benthic characterization. The classification scheme recognizes and describes the main geological, biological, and oceanographic attributes of a habitat, as well as seabed modifiers such as natural and anthropogenic disturbances. Unlike many other schemes, the organization proposed by Valentine et al. (2005) is only partly hierarchical. Each main class is a unique formal unit that can be applied to a range of different environments. The scheme proposes a suite of descriptors which can be recorded independently of one another. Then, based on a defined suite of characteristics, these descriptors can be grouped to produce a set of habitat types.

This type of habitat classification system has numerous benefits. Firstly, this approach enables a user to modify the habitat descriptors such that they apply to their region of interest. Secondly, by having the lowest unit of classification a ‘descriptor’, it enables data to be collected and stored in relatively high detail. Descriptors can then be aggregated into higher level groups or classes during subsequent data analysis. This permits increased user flexibility when determining the set of descriptors that best defines the observed habitat types. Thirdly, whereas the majority of habitat classifications lack quantitative descriptors (for examples, see review by Green et al. 1996), the scheme by Valentine et al. (2005) defines how percent cover of seabed structures can be interpreted as habitat structural complexity. In addition, quantitative information associated with qualitative descriptors aids data interpretation and permits increased options for data analyses. Lastly, the scheme by Valentine et al. (2005) can be easily adapted as a database template from which data can be queried. The storage of data in databases makes data sharing easier and facilitates the migration of data into Geographic Information Systems (GIS) for analysis and display. The organization of the Valentine et al. (2005) classification scheme makes it an ideal candidate template for standardized use across various environments.

For any survey technique used to collect information on the benthic environment, two levels of data exist. Habitat attribute data is captured by the habitat classification scheme, and information on the survey technology and processes is captured by metadata. However, for data interpreted from video-based survey techniques, additional information needs to be recorded. Information on the quality of the video and its sampling resolution are required for these data to be interpreted properly. During the transcription phase, when visual information on tape or disk is converted into qualitative and quantitative data, information on survey processes that influence the quality of the video, from which habitat information is interpreted, should be recorded and associated with each habitat record. This supplementary data on video quality, here defined as 'secondary survey information', improves the value of the data in the scientific workflow by helping future data users recognize limitations and potential biases of the dataset by identifying the spatial scale at which data were collected. This information affects the ecological processes that can be studied and the analyses that can be conducted on the data (Fortin and Dale 2005, Jones et al. 2006).

In this paper we apply a modified version of the benthic habitat classification scheme developed by Valentine et al. (2005) to video data. We 1) present a method for classifying habitat attribute data, 2) discuss the value of secondary survey information and how it can be applied to video data, and 3) apply the modified Valentine et al. (2005) scheme as a database template.

VIDEO DATA SOURCE

The initial application of this work was to classify the offshore benthic environments of the Discovery Corridor (Appendix A), Gulf of Maine, investigated using video survey equipment on the Remotely Operated Vehicle (ROV) ROPOS, operated by the Canadian Scientific Submersible Facility (CSSF). This research mission was conducted over a 2 week period in July of 2006 aboard the CCGS Hudson. In adapting the Valentine et al. (2005) scheme, we focused on the development of a generic and systematic approach to the analysis and classification of video-based survey data of benthic habitats. Although this scheme is based on observations taken of deep seafloor habitats in Northeastern North America using an ROV, its structure is intended to be easily adapted to numerous seabed environments and survey instrumentation.

HABITAT CLASSIFICATION STRUCTURE

The Valentine et al. (2005) scheme identifies eight seabed themes to characterize marine habitat: Topographical Setting, Seabed Dynamics and Currents, Seabed Texture Hardness and Layering, Seabed Grain Size, Seabed Roughness, Fauna and Flora, Habitat Association and Usage, and Habitat Recovery from Disturbance. It was developed using a combination of acoustic (e.g. multibeam and side-scan sonar) techniques, video systems, and direct ecological sampling (e.g. cores and grabs). Combined, these survey techniques provide detailed information regarding each of the eight seabed themes. However, individual technologies vary in their sampling resolution (for review, see Diaz et al. 2004), thus for a single survey technique, a unique combination of classification units from Valentine et al. (2005) will apply. For example, acoustic data is capable of providing broad-scale physical habitat features, but it is often incapable of identifying specific biological variables (Smith et al. 2001).

For video-based survey techniques, data is typically collected on macro- (1 to 10 meters) and micro-scales (10s centimeters but ≤ 1 meter) and allows direct observations and quantitative estimations of geo-physical and biological seafloor features. Classification units were therefore chosen from the Valentine et al. (2005) scheme that matched the data types provided by direct underwater video observations. Data derived by video survey fell under 4 main themes: Seabed Texture Hardness and Layering, Seabed Grain Size, Seabed Roughness, Fauna and Flora. To facilitate data collection and data entry into a relational database format, the information associated with these themes was reorganized into three modified themes (Figure 1). Information for all other themes from Valentine et al. (2005) that cannot be directly obtained from video can likely be determined or calculated from supplementary information on the region of interest. However, for the purpose of this project, we examined only those themes for which we could obtain information directly from the video survey data.

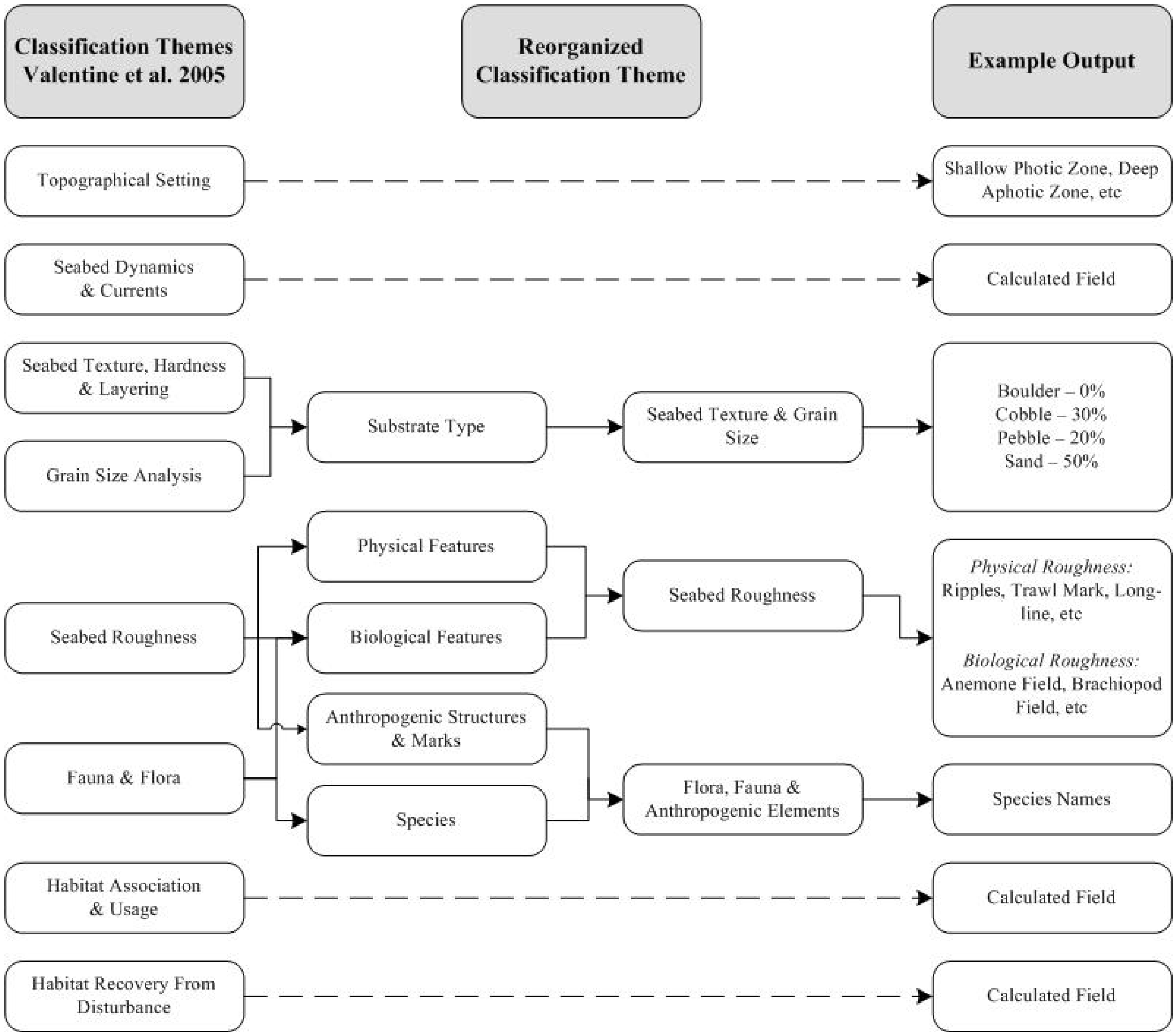


Figure 1. Reorganization of Valentine et al. (2005) habitat classification scheme for video-based data entry into a relational database. Solid lines indicate information on themes that are collected directly from video. Broken lines indicate information obtained from sources other than video.

Reorganized Classification Scheme Themes

1. *Seabed Texture and Grain Size*

This classification provides the results of analysis of visible substrate types. The majority of substrate entries are determined based on the Wentworth scale in accordance with Valentine et al. (2005). However, the Wentworth scale does not encompass all possible substrate types, therefore our classification scheme (Table 1) was designed to permit additional substrate types to be recorded. For example, under the Wentworth scale, an outcrop of bedrock is not distinguishable from a boulder.

All substrate types are recorded as percent cover as determined from an area of ~1-2 m^2 (Appendix D).

Table 1. Example of substrate and sediment descriptors defined for benthic habitat classification of the Discovery Corridor, Gulf of Maine.

Substrate	*Size (mm)*
Bedrock	NA*
Boulder	X > 256
Cobble	64 < X < 256
Pebble	4 < X < 64
Sand	0.0625 < X < 4
Silt	0.0039 < X < 0.0625
Clay	X < 0.0039
Fine grained sediment**	X < 0.25

*Not applicable
**Used in situations where the sampling resolution is too low to distinguish between sediments that are < 0.25 mm in diameter

2. *Seabed Roughness*

This classification provides information on the contribution of physical and biological features to the three-dimensional structure of the seabed surface. Physical roughness elements are geological or anthropogenic in nature and augment the complexity of the underlying substrate and sediments. Biogenic roughness elements are created by biological factors and add structure to the underlying abiotic environment, often providing unique environments for other organisms. Biogenic structures include both attached and emergent epifauna as well as sedentary organisms. The degree to which physical or biological features contribute to forming a roughness element is dependent on the abundance or dominance of that feature in its environment.

How a feature contributes to the structure and complexity of a habitat, is dependent on its coverage of the seafloor. Feature entries are quantitatively defined when possible, and estimates of percent cover or abundance are included in the descriptions of the features. Quantitatively describing habitat attributes is important as it allows an index of structural complexity to be applied to the data. For example, Valentine et al. (2005) classify features that cover >25% of the substrate as adding a 'high' degree of structural complexity, while features that cover 5–10% of the seabed add a 'low' degree of complexity. The ability to quantitatively summarize roughness entries enables flexibility when choosing data analyses options.

When roughness features are determined and their quantitative values assigned (for the derivation processes of features refer to Appendix D), the spatial resolution at which the observations are made need to be recorded. The ability of the researcher to determine the presence of features and their contribution to the structural complexity of the seafloor is scale dependent. Therefore, entries in the database are inherently biased by the sampling technique employed. For example, acoustic techniques can identify a sea mount as a single unique feature, however, from the perspective of a video survey, which typically cover a swath of ~ 1–10 m^2, the side of a sea mount would be recorded as a slope. Since the spatial scale of the sampling influences the cameras perception of the seafloor, information on the cameras footprint area should be included in the description of the roughness features. For the analysis of the video data of the Discovery Corridor, Gulf of Maine survey, roughness features are defined based on an area of ~1–2 m^2 (Tables 2, 3).

Table 2. Examples of physical roughness elements defined for benthic habitat classification of the Discovery Corridor, Gulf of Maine.

Entry	*Descriptor*
Ripples	Undulations of non-cohesive sediment. Distance between peaks ~ 5 to 10 cm. Length of ripples > 100 cm (i.e. > largest length dimension of field of view). Visible on m^2 scale.
Dunes	Low mounds of fine-grained sediment. Likely accumulated as a result of sediment transport (physical process).
Ledge	A shelf-like projection or projecting ridge, from the face of a steep declivity, a rocky outcrop, or reef (Howell 1957).
Vertical face	A surface vertical in orientation. It is part of a substrate feature (e.g. an outcrop) that is larger than the field of view (~1–2 m^2).
Outcrop	Part of stratum that appears, i.e. is exposed, at the surface of the ground (Howell 1957).
Cliff	A steep high face, usually formed of bedrock. This descriptor is used when the field of view is large enough to view both ledge and vertical face simultaneously.

Table 3. Examples of biogenic roughness elements defined for benthic habitat classification of the Discovery Corridor, Gulf of Maine.

Entry	*Descriptor*
Brachiopod dominated	> 200 per m^2
Anemone dominated	> 200 per m^2, 2 distinct size classes: ~1–4 cm dia. and ~10–20 cm dia.
Coral thicket	≥ 0.2–0.5 per m^2, high density aggregations of hard corals consisting mainly of *Primnoa* sp. and *Paragorgia* sp.
Tube worm dominated	> 200 per m^2
Brittle star dominated	> 25% cover
Sponge dominated	Sponges > 40 per m^2
Bioturbation	Prevalent (> 25% cover) biogenic modifications of the seabed: burrows, mounds, tracks, depressions, on 1–2 m^2 scale.

3. Fauna, Flora and Anthropogenic Elements

This field enumerates biological and anthropogenic elements present in a habitat. Elements are recorded in this field as discrete events if they are present in the habitat but do not meet classification requirements for a biogenic or physical roughness element. At the sampling resolution of the survey (< 1 m^2 to 15 m^2), these elements are usually not numerically dominant and do not define the habitat type. However, these elements can provide important information on species-habitat relationships, ecosystem function, as well as the stability, and susceptibility of the habitat to disturbance. The proposed classification system can accommodate these elements.

For our analyses of the Discovery Corridor video, only anthropogenic elements were recorded for this field. However, the marine biodiversity research to be undertaken in the Discovery Corridor, Gulf of Maine, was envisaged to be collaborative in implementation and long-term in outlook (Appendix A). Our research collaborators have already been analysing the distribution and abundance of a range of benthic fauna observed during the 2006 research mission using the same benthic video imagery, along with its consolidated navigation data. We anticipate a number of independent and joint statistical analyses to be forthcoming regarding benthic organism-habitat relationships and the ecological processes and patterns that occur within the surveyed area. A major benefit to the current organization of our modified scheme is its application in this type of evolving scientific workflow.

SURVEY CLASSIFICATION STRUCTURE

Video-based surveys provide vast amounts of information; however the quantitative interpretation of this data is often difficult. Methods of scaling, footprint determination, and standards of data quality can vary within a survey, between remote video systems, and are defined by project objectives. Information on video quality is critical to the proper interpretation of associated habitat data but is often not recorded during video-based analyses. This typically occurs when the researchers that conduct the primary analysis of video data are involved in collecting the data and consequently, they are aware of the limitations imposed by their survey and sampling design. They use this inherent knowledge when interpreting and analyzing their data but often do not document it with the survey data.

Increasingly, information is being integrated across sampling techniques and disciplines, allowing investigators to observe patterns and ask questions that would not be possible from individual datasets (Jones et al. 2006; Cogan and Noji 2007). Data users are no longer inherently associated with the process of deriving datasets yet they still wish to ask questions of these data. However, without knowledge of the data's quality, secondary data users cannot easily identify potential limitations of the datasets they wish to use.

Secondary Survey Information

Benthic habitat classification and mapping projects often involve the collaboration of numerous researchers, science organisations, and scientific disciplines. The inclusion of information on survey processes that influence data quality, and therefore data analysis and interpretation, is critical to data-users that are not directly involved with the initial data collection. It allows them to identify and determine the ecological processes that can and cannot be studied, and the appropriate spatial and statistical analyses for the data (Fortin and Dale 2005). Data related to the survey process that influences data quality and interpretation is here defined as 'secondary survey information'.

The importance of including metadata (e.g. survey tool user, operational protocols, navigation source, etc) and having metadata standards for ecological surveys has long been established. However, secondary survey data differs from metadata in that it is recorded on the same temporal resolution as the habitat classification data. For each habitat record (data on substrate type, physical and biological roughness elements), there is associated secondary survey data that is used to indicate the quality of the video at that moment in time. The video quality information reflects the quality of the data and can be used to determine potential biases in the dataset during analyses.

For data to be properly interpreted and synthesized into the scientific workflow, proper documentation and standards are critical. However, for video-based data interpretation, basic standards do not yet exist. To assist in the interpretation of video-based survey data, we propose the survey classification fields of Survey Mode and View that provide an indication of the survey technique being employed, and the video sampling resolution, respectively. (For the derivation processes of survey mode and view fields, refer to Appendix D).

Survey Mode

In a database, for a section of records, if all the entries across the various habitat attribute fields were null, a database user may interpret this as 'missing' data. This would likely result in the user attempting to track down the source data or documentation to determine if these data exist. This problem could be eliminated if there was a field that detailed the survey equipment process that was associated with each habitat record. For example, if the equipment was pulled off the bottom, perhaps due to unexpected ship motion, then this section of video could be recorded as 'off bottom' in the survey mode field. A detailed description of what the entry 'off bottom' means would be included in an associated table that defined the attribute entries for the data field 'Survey Equipment Mode' (Table 4). For the entry 'off bottom', the description would inform a data user that no habitat records exist for these sections of the survey since the bottom was not visible on the video at this time.

Survey mode records the survey technique being employed by the equipment. Entries are classified as either directed or non-directed sampling (Table 4). Directed sampling is defined as the structured, intentional collection of data. Non-directed sampling is the opportunistic collection of data.

An additional benefit of the attribute 'Survey mode' is that it allows the data user to summarize the types of activities performed by the survey equipment during a dive. For example, Table 5 breaks down the total data collected in 'Direct' versus 'Non-Direct' survey modes on the 2006 Discovery Corridor research mission, allowing the data user to identify the efficiency of data collection for each dive.

Table 4. Survey modes of video-based benthic habitat sampling equipment.

Entry	*Descriptor*
Sampling	Taking/removing a physical sample from the environment. Equipment is typically stationary. Directed sampling.
Investigation	In-depth exploration of an area. This is a non-transect mode and the survey instrument is usually relatively stationary (e.g. examining an organism, bedform, etc). Directed sampling.
Transect	Transecting e.g. moving video survey of area. Directed sampling.
Inactive	Equipment is not engaged in scientific activity. Survey equipment is usually relatively stationary. e.g. stopped to check gauges, checking/placing equipment, etc. Non-directed sampling.
Transiting	Moving between sampling sites. Substrate is usually some-what visible. Non-directed sampling.
Off bottom	Equipment off bottom/in water column and substrate is usually not visible. e.g., pulled off due to currents, ship motion, etc. Non-directed sampling.

Table 5. Percent of Data Collected in Direct and Non-Direct Sampling Survey Mode.

Dive	*Directed (%)*	*Non-Directed (%)*
R967	81.42	18.51
R968	79.32	20.68
R969	94.01	5.99
R970	95.73	4.27
R971	98.73	1.27
R972	75.22	24.78
R974	89.64	10.35
R975	97.48	2.52
R976	97.15	2.85
R977	79.66	20.34
R978	73.74	26.26
R979	91.61	8.35
R980	89.30	10.70
R981	98.28	1.72
R982	96.73	3.27

View

Video data goes through a number of processing steps before its information is in a usable format. The primary step of video analysis is viewing the video and converting the visual information to a text or numeric format that is recorded in a table within a spreadsheet or database. Secondly, this information is extracted from the table to be analyzed for patterns, correlations, etc. When this second phase of analysis is conducted, caution must be taken when interpreting the spatial distribution of features.

Unlike other data collection techniques, the sampling resolution of video can vary quite dramatically within a single sample, e.g. a transect (see Figure 2) and can result in the misinterpretation of data. To avoid drawing erroneous conclusions, the spatial distribution of features should be determined using a combination of the information on the presence and absence of a feature and the sampling resolution of the video.

The sampling resolution of video is a function of the camera's field of view and its limit of resolution. The combination of these parameters determines the size and level of detail of the sample and, consequently, affects the identification of physical and biological features. For example, if there is a transect section where a feature is absent in the database, the spatial resolution limit of the feature should be compared with the resolution limit of the video. If the resolution of the video is lower than the resolution needed to identify the feature, then it is possible that the feature was present but that the camera setup did not allow it to be observed. This could inherently bias the results of an analysis on the spatial distribution of that feature (see Figure 2). Informed inferences on the absence of features should only be made after examining associated changes in the sampling resolution of the video.

For the proposed classification, the sampling resolution of the video is recorded in the 'view' field and acts as an index of the quality and usability of the data. Each habitat record derived from the analysis of the video has an entry that defines the video resolution at that moment in time. The entries used in the 'view' field are derived from a previous overview of the video data (for derivation process see Appendix D) and each entry is defined by the lowest size limit of object discrimination (Table 6). The entries in this field are an indication and function of the speed of the survey equipment over the substrate, the focus and zoom of the camera, and the lighting quality. For example, if the camera is moving relatively quickly across the substrate, the limit of resolution may be low due to motion blur. If the smallest object that can be distinguished from this section of video is 5 cm, then it is assigned a view of 2 (Table 6). This indicates to future data users that they cannot draw definitive conclusions regarding the presence or absence of features that are < 5 cm in diameter. (Analyses of video-based survey data collected using the Remotely Operated Platform for Ocean Sciences (ROPOS) in the Discovery Corridor, July 2006, resulted in View categories as defined in Table 6.)

The 'view' field can also provide an indication of the camera's footprint of the substrate and define the ecological processes that can be interpreted from a section of analyzed video. Typically, when the field of view increases, the resolution decreases. By examining the video data, this relationship can be defined. For the current application, the relationship between the limit of resolution, as reflected by the 'view' index, and the field of view is presented in Table 7 (for derivation process see Appendix D).

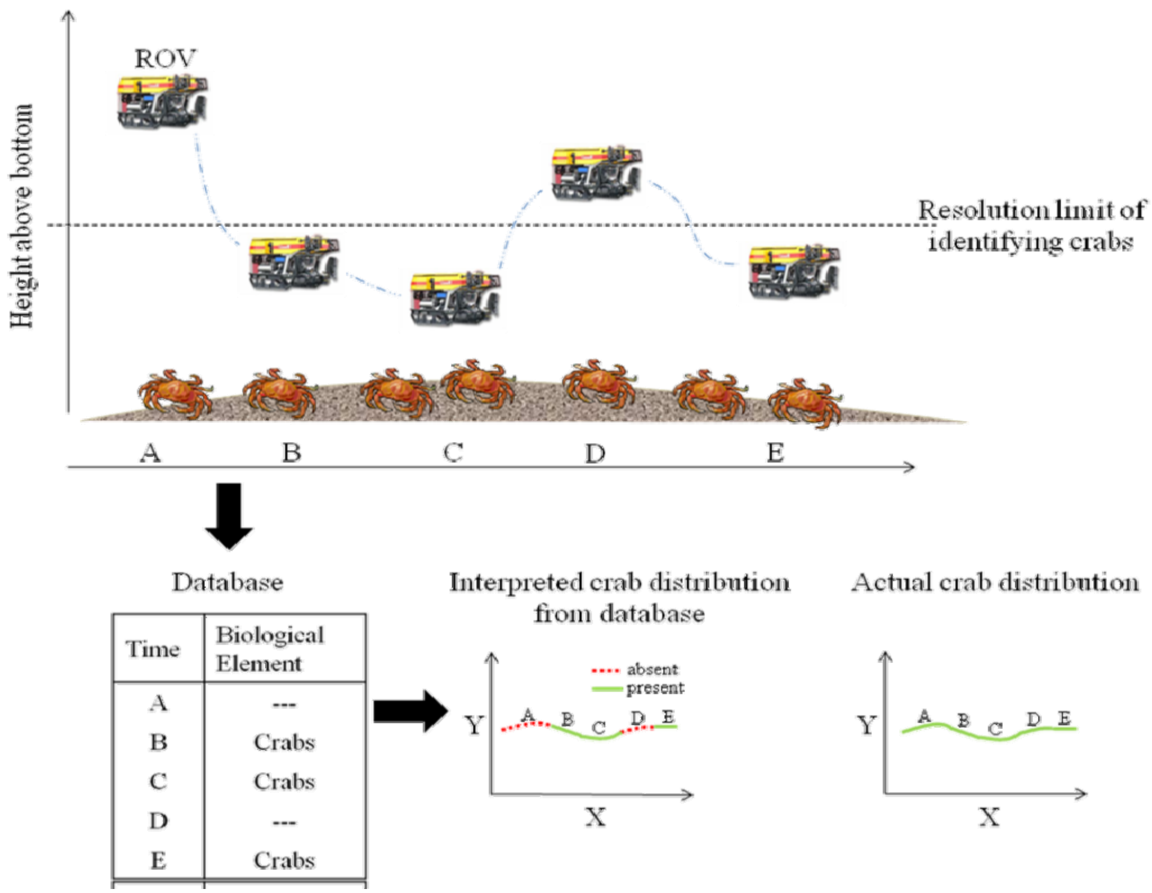


Figure 2. Example of misinterpreted distribution of features from video survey data collected and analyzed without secondary survey information.

Table 6. View modes of video-based survey equipment

Entry	*Lowest limit of object discrimination*
view 1	< 3 cm
view 2	3–5 cm
view 3	10–15 cm

Table 7. View modes and associated camera footprint of Discovery Corridor Survey using ROPOS, July 2006.

Entry	*Lowest limit of object discrimination*	*Camera Footprint of Substrate*
view 1	< 3 cm	$<1\ m^2$
view 2	3–5 cm	$\sim1–2\ m^2$
view 3	10–15 cm	$\geq10\ m^2$

Combining Survey Mode and View

By combining Survey Mode and View fields, video and data quality can be summarized. For example, by querying those records where the survey mode is in a directed sampling mode (i.e. 'sampling', 'transect' and 'investigation'; Table 4) and View is view 2 (Table 6), you can determine the percent of records where there was structured data collection and the camera footprint of the substrate was ~ 1–2 m^2 (Table 8).

Table 8. Percent of Data in Direct Sampling Survey Mode where the camera footprint is ~1–2 m^2.

Dive	*Footprint ~ 1–2m^2 (%)*
R967	44.63
R968	53.53
R969	70.57
R970	74.54
R971	67.40
R972	48.40
R974	29.16
R975	64.64
R976	52.73
R977	48.44
R978	43.52
R979	68.50
R980	61.02
R981	76.93
R982	85.64

DATALOGGING

Data was logged using the software ClassAct Mapper. ClassAct Mapper is a utility developed within DFO to assist in the post-processing and analysis of video data (Appendix B). This software was developed to enable researchers to interactively log information related to classifying habitat types as well as recording discrete events that occur on the video. The main advantage of this program is that it allows for almost real-time analyses of video data. Previously, video data were analyzed manually, with data being transcribed by the viewer and then manually entered into a database. ClassAct Mapper eliminates the transcription phase of the analysis by enabling direct-to-database recording of data as the video is being viewed. This software builds on previously developed geo-referencing software called Class-Event that was also developed within DFO and has been used in similar applications by Strong and Lawton (2004).

The navigation data associated with the video survey equipment are recorded to the field on the audio track of the digital video tape using an acoustic modem (GeoStamp®Audio;

Intuitive Circuits, LLC, Troy, Michigan, USA). When the video data are analyzed in the laboratory, the audio track is demodulated using the same modem and then fed through a serial connection to the computer. The ClassAct Mapper utility records the navigation data (latitude, longitude, time) directly to a Microsoft Access® database and enables the user to interactively log additional data to the corresponding navigation records by using the ClassAct Mapper GUI.

There are two data types that are logged by ClassAct Mapper: Class Data and Event Data. Class data are all habitat attributes that contribute significantly to the structure of the seafloor environment. For the purpose of our proposed classification scheme, this includes seabed texture and grain size descriptors and seabed roughness element descriptors. Since survey mode and view data are designed to accompany habitat records on the same temporal scale, these data were also recorded as 'class data'. To record class data, when an attribute entry is selected (e.g. sand 100%) that entry is logged automatically to the database at each navigation update (2 or 3 s intervals). This automatic logging occurs until the attribute for that classification data field is changed by the operator (e.g. sand changed to cobble). This provides for the quasi-continuous geo-referencing of habitat type. Event data are discrete occurrences observed for a seafloor environment and includes descriptors that fall under the theme of fauna, flora and anthropogenic elements. For discrete events that occur on video, such as passing by a fish or over a cable, the event and its associated navigation information, (latitude, longitude, time) are logged as a discrete record when the user enters the event code.

This software was modified to accommodate the habitat and survey fields as defined in the current classification scheme. In total, over 100 hours of high quality video with audio encoded navigation from the 2006 survey of the Discovery Corridor, Gulf of Maine was analyzed using this program. However, since the software logs data to a database normalized on classification type and not on time, the data were reorganized into a relational database format normalized on time (resolution of seconds). Further details of the software are provided in Appendix B.

DATABASE DESIGN

Analysis of benthic habitat attributes, in particular, the attribute data collected from video imagery, results in large volumes of data. For example, for the current analysis of video collected from the Discovery Corridor, Gulf of Maine, over 280 000 habitat records were recorded for ~ 100 hours of video. To store this amount of data, a relational database format was chosen. The benefits of storing data in a relational database include expedited information flow, ease of data access using SQL queries, data integrity rules, and the ability to integrate data with analysis programs such as Geographic Information Systems (GIS). Habitat records are geographically referenced and their storage in a relational database format enables the user to import attribute tables into GIS and create thematic maps based on classification attributes. The Valentine et al. (2005) classification scheme was chosen as a candidate database template since it is based on individual attributes. The logic structure for this classification scheme also builds on prior relational database designs for remote video surveys and SCUBA-diver-based transect survey approaches (M. B Strong and P. Lawton, unpubl. data; contact address on title page).

A relational database was created in Microsoft Access® to store all information related to the analysis of the video data collected from the ROV ROPOS during the survey of the Discovery Corridor, Gulf of Maine, July 2006. The database consists of 2 main components: 1) attribute tables, and 2) attribute definition tables. Attribute tables are here defined as containing spatially relevant information – that is data that has an associated geographic location, whereas attribute definition tables contain descriptive information that define attribute entries.

Four attribute tables were designed to store information related to the habitat classification information derived from analyzing the video survey data. Attribute tables were normalized on date-time (YYYY-MM-dd HH:mm:ss) and designed to minimize the redundant storage of information, yet be comprehensive enough to facilitate making simple queries to extract summaries of data. The organization of records was based on time to facilitate ordering and sorting the data and all date-time fields were stored as the data-type date-time.

Attribute tables are joined based on their unique date-time stamp (YYYY-MM-dd HH:mi:ss) (Figure 3).

To facilitate the migration of data from the database into Geographic Information Systems (GIS), in particular ArcGIS® for display and analysis, all data field headings were created as single strings and all fields that could be quantitatively categorized and displayed using GIS were converted to the data-type 'number'. For example, all data fields that store percent cover data on substrate type are of the data-type 'number'.

Attribute Tables:

1) *Mission Table* – contains data on the mission, dives and tapes.
2) *Habitat Classification Table* – contains fields described under Seabed Texture and Grain Size, Seabed Roughness, Survey Mode, View, and information on the geographic location of the habitat features.
3) *Fauna, Flora and Anthropogenic Elements Table* – contains biological and anthropogenic element data present in a habitat.
4) *Survey Equipment Table* – contains fields related to the function and position of the video-based survey equipment. This includes information related to the navigation and function of the equipment.

Four attribute definition tables were designed to store information that described attribute table entries. Storage of attribute table entry descriptions (qualitative and quantitative) within the database consolidates information used to interpret and analyze video data with its corresponding transcribed tabular data. Single location storage helps avoid misinterpretation of data by future database users.

Attribute definition tables are joined to attribute tables based on matched attribute entry names (Figure 3).

Attribute Definition Tables:

1) *Biological Roughness Elements* – contains qualitative and quantitative descriptions of the entries found in the Biological Roughness Element field of the Habitat Classification Table.
2) *Physical Roughness Elements* – contains qualitative and quantitative descriptions of the entries found in the Physical Roughness Element field of the Habitat Classification Table.
3) *Survey Mode* – contains descriptions of the entries in the survey mode field of the Habitat Classification Table.
4) *View* – contains descriptions of the entries in the view field of the Habitat Classification Table.

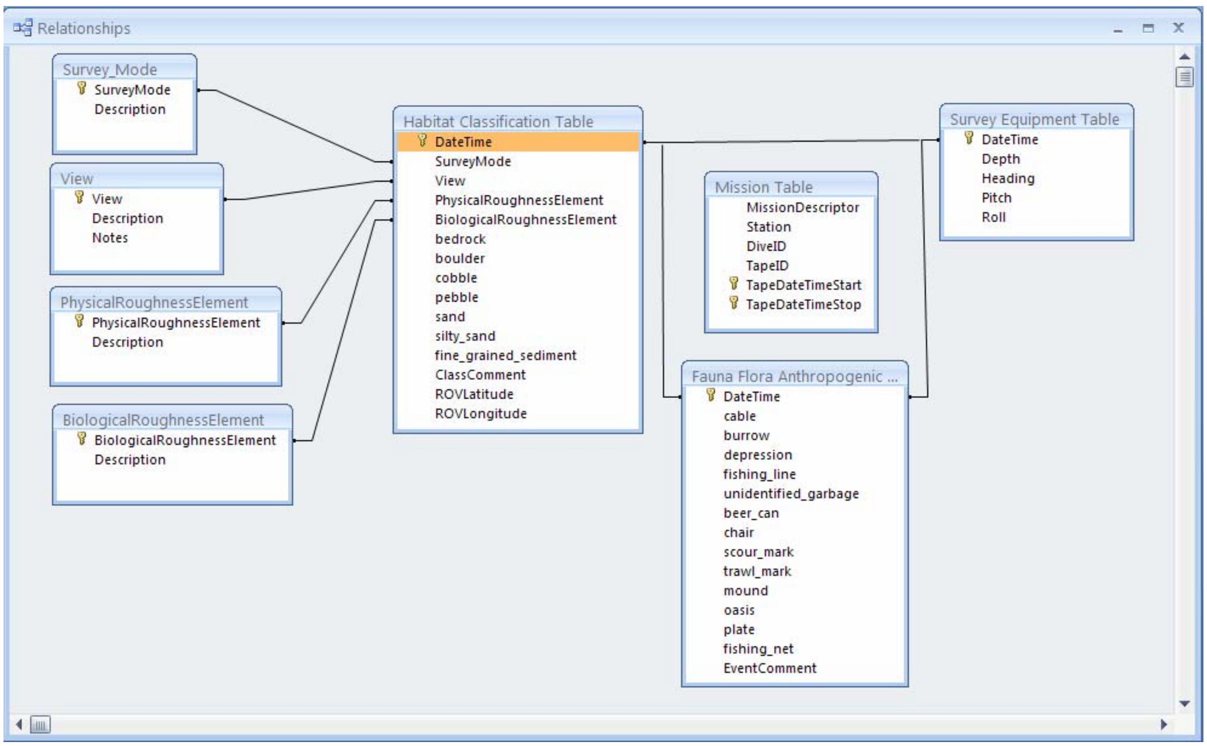


Figure 3. Relationships between tables of the Discovery Corridor 2006 survey Database.

EXAMPLE APPLICATION

A key attribute of this benthic habitat classification system and its relational database design is that various components of the scheme can be visualized and interpreted separately as well as through various combinations of attributes depending on the specific benthic habitat mapping application or habitat utilization question being addressed. Appendix E presents several aspects of the survey information captured from a video transect performed using ROPOS to survey benthic habitat in the Discovery Corridor, Gulf of Maine in July 2006.

DISCUSSION

The habitat classification structure we proposed is designed for analyzing video-based survey data of marine benthic environments and is based on the classification scheme for marine sublittorial habitats developed by Valentine et al. (2005). Our approach is designed to describe seafloor environments through both qualitative and quantitative descriptors and emphasizes the importance of geo-physical and biological features that contribute to habitat complexity.

The Valentine et al. (2005) scheme proved to be robust and easily adaptable to classifying habitats sampled using video surveys. Not all habitat themes and classes proposed by Valentine et al. (2005) applied to our data, however, since the lowest level of their scheme is based on habitat attributes, the classification structure was easily modified to fit the attributes found in the environments we surveyed. The Valentine et al. (2005) scheme is just one of a number of extant marine habitat classification schemes (McDougall et al. 2007; see also Greene et al. 2007) and this report does not attempt to arbitrate on the merits of the various classification schemes in terms of their utility for general application. However, our experience in the translation of the Valentine et al (2005) scheme into a database template usable for video analysis indicates that the essential design elements for partly hierarchical classification are highly adaptable to GIS-based display and analysis.

In a recent review of the latest application development with a long-established marine habitat classification scheme, Greene et al. (2007) identify specific expansions to their schema to take advantage of new spatial analysis approaches and also make specific recommendations on developing attribute codes to use habitat classification data within GIS. Dealing with collection techniques for documenting seafloor characteristics at macro- (1–10 meters) and micro-habitat (10s centimeters but ≤ 1 meter) scales, Greene et al. (2007) acknowledge the importance of various optical imaging approaches, but identify that the current use of categorical classification codes for *in situ* observations represents a limitation in analysing habitat associations. With multiple attributes (their examples being depth, substrate type, slope, rugosity or the roughness of the seafloor) reduced to categorical codes there is a significant likelihood that individual attributes within a code may have separate potential associations (their examples being sediment waves and flat or volcanic rock and high rugosity) which could affect or limit statistical analyses. They recommend the collection of habitat variables (their examples being temperature, depth, substrate type) continuously along a dive or transect but their conclusion was that this was currently beyond the capabilities of modern technology.

However, we have shown that by developing a relational database approach for benthic habitat classification, and using software to partly automate the classification of video at the level of individual geographic position and time records from telemetry, we have achieved an initial approach to “deconstruct” benthic habitat classification variables. We are now poised to conduct statistical analyses at macro- and micro-habitat scales for various organism-habitat associations within the Discovery Corridor. These will be based either on internalized analyses from our own benthic habitat data set, or will be conducted in conjunction with research collaborators that have identified distribution and abundance variables for specific benthic fauna from the benthic video surveys conducted on the 2006 Discovery Corridor research mission. In addition we anticipate working with marine geologists to “reconstruct” benthic habitat variables in various combinations and at different spatial scales to provide ground-truthing information that can be used in the interpretation of seafloor features from acoustic mapping techniques (multibeam sonar).

A further important component of the classification scheme we have developed is its ability to provide information on the quality of the video data from which habitat features are determined. We propose that 'secondary survey information' should be routinely recorded at the same temporal resolution as habitat descriptors when analyzing video-based data. This supplementary information acts to preserve the value of video survey data since it allows limitations in the dataset to be identified by future data users.

With increased use of video-based techniques to survey the seafloor, the processes by which video data is analyzed will likely increase and diversify. Here we have presented an approach and organizational scheme from which future analyses of benthic video-data may be modeled. Although this scheme is based on observations taken of deep seafloor habitats in the Gulf of Maine, Northeastern North America using the ROV ROPOS, its structure is intended to be easily adapted to various seabed environments and imagery-based survey equipment.

ACKNOWLEDGEMENTS

We are very grateful to Robert Benjamin for his assistance and expertise regarding the adaptation of his ClassAct Mapper software for the above mentioned benthic habitat classification scheme. Also to Robert, and to Vince Auger (Canadian Scientific Submersible Facility) thanks for their assistance with marine navigation data issues and other aspects of relational database design. Funding for the 2006 research mission by the Canadian Coast Guard Ship (CCGS) Hudson in the offshore portion of the Gulf of Maine Discovery Corridor was provided jointly by DFO and the Natural Sciences and Engineering Research Council of Canada. We thank Drs. Ellen Kenchington, Anna Metaxas, and Paul Snelgrove, for their academic collaborations in defining the research objectives, securing funding, and also for their participation in the 2006 CCGS Hudson mission. The collection of over 100 hours of high quality benthic video data in challenging offshore operational environments (survey depths up to 2600 m, high bottom currents, surface weather factors) is a testament to the professional operation of ROPOS by staff of the Canadian Scientific Submersible Facility, and also to the dedication of the captain and crew of the CCGS Hudson in making that first mobilization of ROPOS on Hudson such a success. Financial support to complete the development and documentation of the classification scheme was provided by DFO and by the Sloan Foundation (through the Gulf of Maine Area Program of the Census of Marine Life).

REFERENCES

Allee, R. J., Dethier, M., Brown, D., Deegan, L., Ford, R. G., Hourigan, T. F., Maragos, J., Schoch, C., Sealey, K., Twilley, R., Weinstein, M. P., and Yoklavich, M. 2000. Marine and estuarine ecosystem and habitat classification. NOAA Technical Memorandum NMFS-F/SPO-43.

Cogan, C. B., and Noji, T. T. 2007. Marine classification, mapping, and biodiversity analysis. *In* Mapping the seafloor for habitat characterization. Edited by B.J. Todd and H. G. Greene. Geological Association of Canada Special Paper 47: pp. 129-139.

Connor, D. W., Allen, J. H., Golding, N., Howell, K. L., Lieberknecht, L. M., Northen, K. O., and Reker, J. B. 2003. The marine habitat classification for Britain and Ireland, version 04.05. Joint Nature Conservation Committee (JNCC), UK.

Cowardin, L. M., Carter, V., Golet, F. C., and LaRoe, E. T. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Fish and Wildlife Service FWS/OBS-79/31. 108p.

Diaz, R. J., Solan, M., and Valente, R. M. 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. J. Environ. Manag. 73: 165–181.

Fortin, M. J., and Dale, M. 2005. Spatial analysis: a guide for ecologists. Cambridge University Press, Cambridge. 378p.

Green, E. P., Mumby, P. J., Edwards, A. J., Clark, C. D. 1996. A review of remote sensing for the assessment and management of tropical coastal resources. Coastal Management. 24: 1–40.

Greene, H. G., Yoklavich, M. M., Starr, R. M., O'Connell, V. M., Wakefield, W. W., Sullivan, D. E., McRea Jr., J. E., and Cailliet, G. M. 1999. A classification scheme for deep seafloor habitats. Oceanologica Acta. 22: 663–678.

Greene, H. G., Bizzarro, J. J., O'Connell, V. M., and Brylinsky, C. K. 2007. Construction of digital potential marine benthic habitat maps using a coded classification scheme and its application. *In* Mapping the seafloor for habitat characterization. Edited by B.J. Todd and H. G. Greene. Geological Association of Canada Special Paper 47: pp. 141-155.

Howell, J.V. (COORD). 1957. Glossary of geology and related sciences. NAS-NRC Publication, USA. 335p.

Jones, M. B, Schildhauer, M. P., Reichman, O. J., and Bowers, S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. Annu. Rev. Ecol. Evol. Syst. 37: 519–544.

McDougall, P. T., Janowicz, M., and Franks-Taylor, R. 2007. Habitat classification in the Gulf of Maine: A review of schemes and a discussion of related regional issues. Gulf of Maine Council on the Marine Environment. www.gulfofmaine.org/habitatclassification. 15 p.

Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. J. Theor. Biol. 13: 131–144.

Smith, G. F., Bruce, D. G., and Roach, E. B. 2001. Remote acoustic habitat assessment techniques used to characterize the quality and extent of oyster bottom in the Chesapeake Bay. Marine Geodesy. 24: 171–189.

Strong, M. B., and Lawton, P. 2004. URCHIN – Manually-deployed geo-referenced video system for underwater reconnaissance and coastal habitat inventory. Can. Tech. Rep. Fish Aquat. Sci. 2553. 32p.

Valentine, P. C., Todd, B. J., and Kostylev, V. E. 2005. Classification of marine sublittoral habitats, with application to the Northeastern North America Region. Am. Fish. Soc. Symp. 41: 183–200.

Zwanenburg, K. C. T., Querbach, K., Kenchington, E., and Frank, K. (editors). 2003. Three Oceans of Biodiversity. Development of a Science Plan for Marine Biodiversity in Canada. Can. Tech. Rep. Fish. Aquat. Sci. 2432: 80p.

APPENDIX A: THE DISCOVERY CORRIDOR INITIATIVE

The Discovery Corridor Initiative is an international collaborative research effort that commenced in 2004 to explore marine biodiversity across a particular swath of the Gulf of Maine. The project was initially conceived as a way in which to move forward with one of five key recommendations contained within Canadian national reports on status of knowledge and priorities for research on marine biodiversity (*Three Oceans of Biodiversity: A Canadian National Plan 2004-2009* available online from the Centre for Marine Biodiversity (www.marinebiodiversity.ca); see also Zwanenburg et al. 2003).

The specific recommendation was to establish reference sites in each of Canada's three oceans (the North Atlantic, the Arctic, and the North Pacific). The Gulf of Maine Discovery Corridor was established as the pilot reference site for the North Atlantic. The concept of discovery corridors includes making a census of marine life and furthering understanding of the processes that contribute to biodiversity. Discovery corridors are also anticipated to provide for the development of education and outreach resources to engage public and stakeholder interest in marine biodiversity and marine conservation. Through the research conducted within discovery corridors, technical approaches to, and implementation of, marine conservation should be improved, although their delineation is not meant to imply that the full discovery corridor area should become a marine conservation area.

The Gulf of Maine Discovery Corridor extends from the lower Bay of Fundy across the northern Gulf of Maine, over the northeast tip of Georges Bank and the Northeast Channel, and out to the New England Seamount Chain (see Figure E1, Appendix E for the corridor extent within the Gulf of Maine and some recent offshore research locations). It was situated with two objectives in mind: 1) maximize known information, and 2) include a variety of habitats.

Primary research funding from 2004 to 2008 has been provided by DFO, the Natural Sciences and Engineering Research Council of Canada, and the Sloan Foundation (through the International Census of Marine Life). The Discovery Corridor Initiative is a Canadian contribution to the Gulf of Maine Area Census of Marine Life, coordinated through the Centre for Marine Biodiversity (www.marinebiodiversity.ca). Further details can be obtained from http://www.marinebiodiversity.ca/cmb/research/discovery-corridor.

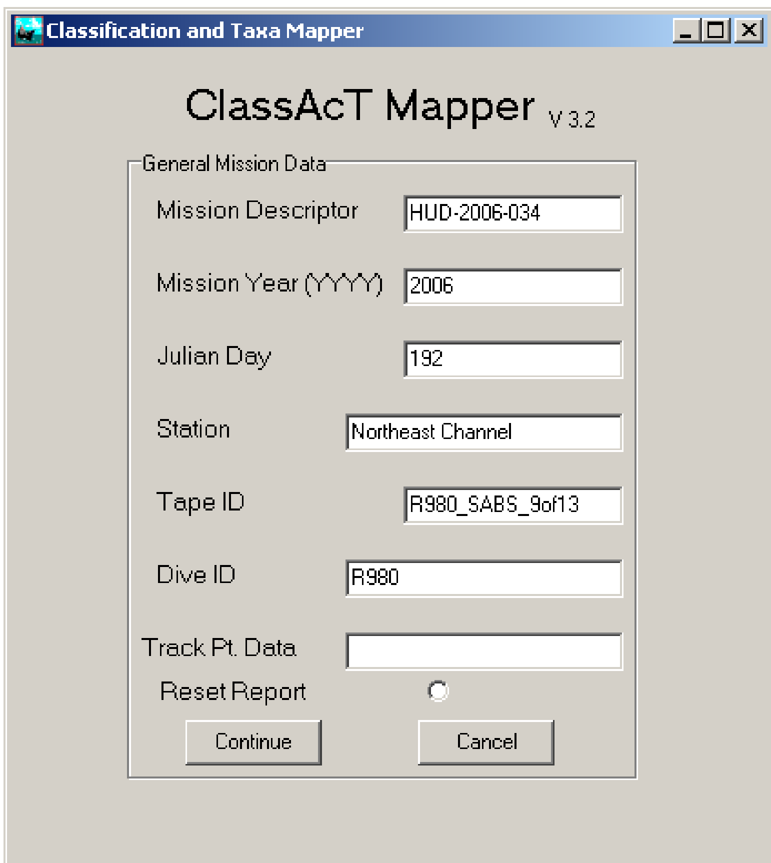
APPENDIX B: DATALOGGING SOFTWARE

ClassAct Mapper Overview:

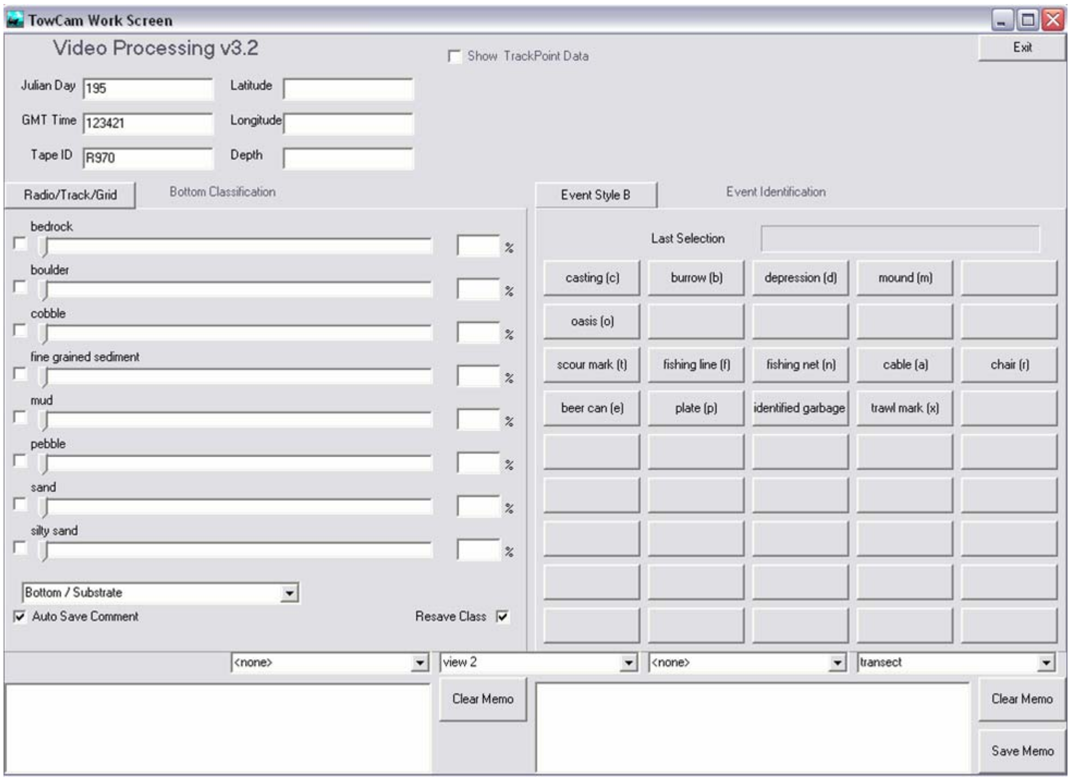
ClassActMapper v. 3.2 software is authored for DFO by Robert Benjamin at the Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada.

This program is used to create database files in Microsoft Access® of time and position (latitude and longitude) stamped bottom classification descriptions and discrete events while observing underwater video in real-time. The database is normalized on classification type and not on time. Therefore, customized report generation is required after the data is logged to reorganize the database into a relational database format normalized on date-time (resolution of seconds).

The general mission data window appears as below.



The main window for the ClassAct Mapper logging software appears below. In this case, the classification and discrete event variables have been defined by the user as habitat classification variables (substrate variables) and anthropogenic variables, respectively.



APPENDIX C: DISCOVERY CORRIDOR CLASSIFICATION SCHEME

Attribute Tables:

Table Name	Descriptor
MissionData	General mission and analysis metadata
Classification	Habitat classification data
ROVData	Data pertaining to the condition of the survey equipment used to collect the data being analyzed
Events	Features that are part of, or contribute to the habitat on a spatial scale below that of our analysis

Attribute Definition Tables:

Table Name	Descriptor
SurveyMode	Contains descriptors of entries that define the survey technique being employed
View	Contains descriptors of entries that help define the resolution/quality/usability of the video
PhysicalRoughnessElements	Contains descriptors of physical roughness element entries
BiologicalRoughnessElements	Contains descriptors of biologic roughness element entries

Attribute Table 1:
Name: MissionData

Field Name	Description	Data Type
MissionDescriptor	Cruise name: e.g. Hudson-2006-034	Text
Station	Name of location e.g. Northeast Channel	Text
DiveID	Dive Identification Name/Number	Text
TapeID	Tape Identification Name/Number	Text
TapeDayTimeStart	Calculated Field, Min(DateTime of portion of dive recorded on associated tape)	Date/Time
TapeDateTimeStop	Calculated Field, Max(DateTime of portion of dive recorded on associated tape)	Date/Time

Attribute Table 2:
Name: Classification

Field Name	Descriptor	Data Type
DateTime	Combined date-time: YYYY-MM-dd HH:mm:ss	Date/Time
SurveyMode	Indication of the survey technique being employed by the equipment. The categories are designed for ROPOS but can also be used for other devices and other categories can be added for techniques unique to other equipment.	Text
View	Index of the resolution limit of video. Can be used in combination with Camera mode to indicate what types of analysis can be conducted on different sections of video. Often an indication of camera footprint, the zoom of the camera, and video quality (focus, lighting, etc).	Text
PhysicalRoughnessElements	Physical structures that add to the three-dimensionality of the seabed surface. Includes bedforms. A feature must contribute substantially to the seabed. Entries are designed to be mutually exclusive; however, when overlap is unavoidable, the most dominant feature relative to the area of the field of view is recorded.	Text
BiogenicRoughnessElements	Biogenic structures that add to the three-dimensionality of the seabed surface. Includes attached and non-attached epifauna. Must contribute substantially to the seabed.	Text
bedrock	Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
boulder	X > 256 mm Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
cobble	64 mm < X < 256 mm Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
pebble	4 mm < X < 64 mm Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
sand	0.0625 mm (63 µm) < X < 4mm Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
silty_sand	Appears: Silt < X < Sand (for analysis, can be grouped into Fine grained sediment classification). Typically non-cohesive. Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number

fine_grained_sediment	Sediment < 0.25 mm Use in areas where we cannot definitively distinguish between fine sand, silt, mud, and clay, etc. Quantitative field indicating % cover of seabed. % cover determined from area of ~ 1–2 m^2	Number
ClassComment	Logger's comments/data entry that is not applicable or possible in other fields. (Comment field from ClassAct Mapper) Common keywords and phrases used for consistency in database to ensure accurate queries.	Memo
ROVLongitude	Longitude of Remotely Operated Vehicle	Number
ROVLatitude	Latitude of Remotely Operated Vehicle	Number

Attribute Table 3:
Name: ROVData

Field Name	Descriptor	Datatype
DateTime	Combined date-time: YYYY-MM-dd HH:mi:ss	Date/Time
Depth	Depth of ROV from surface. Unit = meters	Number
Heading	The direction which the equipment is moving measured as the horizontal angle between a ground line and true north. Unit = degrees	Number
Pitch	Pitch of survey equipment. Unit = degrees	Number
Roll	Roll of survey equipment. Unit = degrees	Number

Attribute Table 4:
Name: Events

Field Name	Descriptor	Datatype
GMTtime	Combined date-time: YYYY-MM-dd HH:mm:ss	Date/Time
Event 1	Discrete element/event	Text
Event 2	Discrete element/event	Text
. .	Discrete element/event	Text
Event n^e	Discrete element/event	Text
Event Comment	Logger's comments/data entry that is not applicable or possible in other fields. (EComment field from ClassAct Mapper)	Memo

Attribute Definition Table Structure:
All attribute definition tables consist of 2 main fields:

Field Name	Description	Data Type
Field Heading		Text
Description		Memo

Attribute Definition Table Entries:
Attribute Definition Table 1: SurveyMode

Entries are mutually exclusive. Directed sampling is structured, intentional collection of data while non-directed sampling is opportunistic collection of data. This field is used in combination with View entry to determine quality of video and the approximate field of view.

Entry	Description
Sampling	Taking/removing a physical sample from the environment. Equipment is typically stationary. Directed sampling.
Investigation	In-depth exploration of an area. This is a non-transect mode and the survey instrument is usually relatively stationary (e.g. examining an organism, bedform, etc). Directed sampling.
Transect	Transecting e.g. moving video survey of area. Viewing area (range: < 1 to > 10 m^2) can be approximated using this entry in combination with the View entry. Directed sampling.
Inactive	Stationary but not sampling or in investigation mode. Non-directed sampling e.g. stopped to check gauges, checking/placing equipment, etc.
Transiting	Moving between sampling sites. Not in survey mode. Non-directed sampling. Substrate is usually visible.
Off bottom	Equipment off bottom/in water column and substrate is not visible. e.g. pull off due to currents, ship motion, etc. Non-directed sampling.

Attribute Definition Table 2: View

The view category is a function of the field of view/camera footprint, the speed of the survey equipment over the substrate, and the focus of the camera which contribute to the resolution limit of the video.

Entry	Description
View 1	Lower limit of object discrimination: < 3 cm dia. If visibility is good and Camera is moving slowly across substrate: Field of view <1 m^2
View 2	Lower limit of object discrimination: > 3–5 cm dia. If visibility is good and Camera is moving slowly across substrate: Field of view ~1-2 m^2
View 3	Lower limit of object discrimination: > 10–15 cm dia. If visibility is good and Camera is moving slowly across substrate: Field of view ≥10 m^2

Attribute Definition Table 3: PhysicalRoughnessElements

Entry	Descriptor
Megaripples	Large ripples of non-cohesive sediment Height 3-9 m, length 125–1300 m. Are large bedforms, visible on scales > 10s m^2, this is likely a greater scale than visible by sampling with ROPOS.
Ripples	Undulations of non-cohesive sediment. Distance between peaks ~ 5 to 10 cm. Length of ripples > 100 cm (i.e. > largest length dimension of field of view) Visible on m^2 scale
Dunes	Low mounds of fine-grained sediment. Likely accumulated as a result of sediment transport (physical process).
Slump	Slump = material that has slid down from a slope (Howell 1957)
Ledge	A shelf-like projection or projecting ridge, from the face of a steep declivity, a rocky outcrop, or reef. It is part of a substrate feature (e.g. an outcrop, clay canyon, etc) that is larger than the field of view (~1–2 m^2)
Vertical face	A surface vertical in orientation. It is part of a substrate feature (e.g. an outcrop) that is larger than the field of view (~1–2 m^2)
Outcrop	Part of stratum that appears, i.e. is exposed, at the surface of the ground.
Shell fragments	Broken pieces of shells. Contribute to habitat complexity by cover >25% of an area ~1–2 m^2.
Cliff	A steep high face, usually formed of bedrock. This descriptor is used when the field is view is large enough to view both ledge and vertical face simultaneously.

Attribute Definition Table 4: BiogenicRoughnessElements

Features that are created by biogenic factors and are important elements of a habitat. They add structure to the underlying abiotic environment and provide unique environments for other organisms. The degree to which a biogenic feature contributes to forming a biogenic habitat is dependent on the density of that feature. Density is approximated by abundance or percent cover. The following features are recorded when their presence reaches or exceeds the level defined in the descriptor.

Biological Roughness Element	**Description**
Brachiopod dominated field	> 200 per m^2
Anemone dominated field	> 200 per m^2, 2 distinct size classes: ~1–4 cm dia and ~10–20 cm dia
Coral thicket	≥ 0.5 per m^2, o High density aggregations of hard corals consisting mainly of Primnoa sp. and Paragorgia sp.
Tube worm dominated	> 200 per m^2
Biogenic turf	Upper surface layer of fine sediment dominated by biogenic structures: burrows, anemones, hydrozoans, brittle stars. It is associated with high-relief areas, e.g. along vertical and high sloped faces.
Brittle star dominated	> 25% cover
Sponge dominated field	Sponges > 40 per m^2
Bioturbation	High abundance > 25% cover of burrows, mounds, tracks, depressions on 1–2 m^2 scale. The result of, and indicators of, sediment reworking by biological organisms. Biogenic modifications of the seabed. Includes burrows, mounds, depressions, tracks.

APPENDIX D: DERIVATION OF QUANTITATIVE DESCRIPTORS

An important aspect in the above proposed classification scheme is the derivation of quantitative descriptors. Here we discuss the method by which quantitative descriptors were determined.

The ROV ROPOS has scaling lasers which are used as a sizing reference for video data. Two green laser dots spaced 10 cm apart are visible for all video and images (Figure D1). A calibration factor was calculated for each analyzed image using the calibration tool in Image Pro Plus® and used to determine the real-world size of observed features.

Roughness Elements

A preliminary review of the ROPOS video footage was undertaken in the laboratory by the data analyst and habitat characteristics that required quantitative descriptors were identified. For each characteristic, a minimum of 4 images which contained that characteristic were selected. Selected images were temporally and spatially independent and, when possible, images of the same feature were selected from different dives. Images were then imported into the software Image Pro Plus® for analysis.

The percent cover or abundance of the target feature was then determined using the measurement toolbox and values were standardized to 1 m^2. The mean and standard deviation for each feature was calculated and used as a reference to set the quantitative description used to define the habitat descriptor. For an example, the quantitative description of Coral Thicket is >0.5 per m^2 and the derivation of this value can be seen in Table D1.

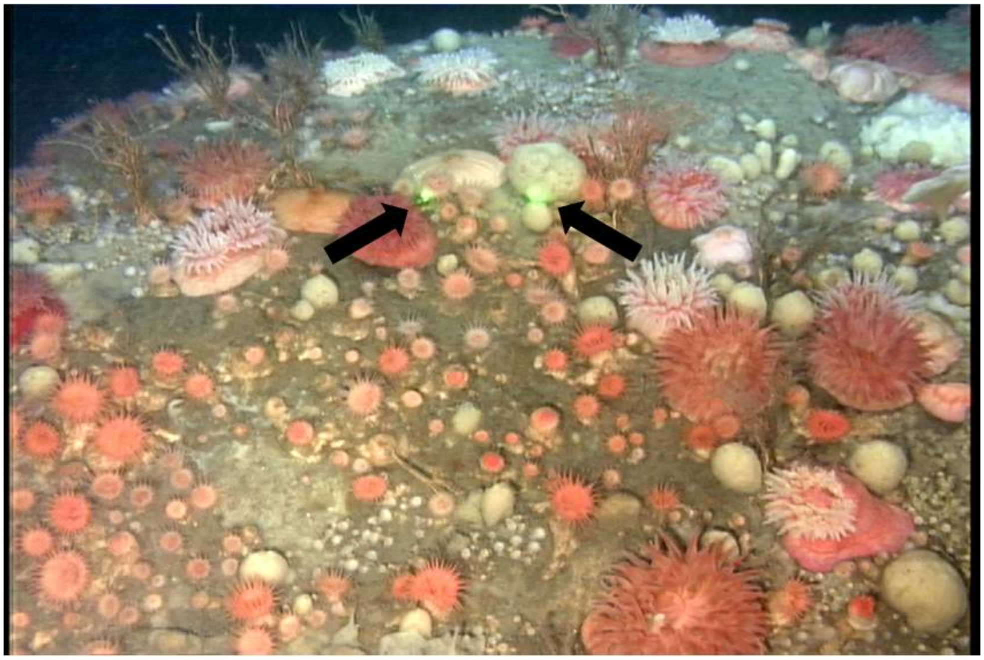


Figure D1. Example image taken from a video transect using the ROV ROPOS. Note green laser dots spaced 10 cm apart.

Table D1 Quantitative derivation of the biological roughness element 'Coral Thicket'

Entry	Qualitative Description	Image Number	Count	Area Sampled	Coral per m^2
Coral Thicket	High density aggregations of hard corals consisting mainly of Primnoa and/or Paragorgia sp.	1	9	11.80	0.76
		2	8	15.29	0.52
		3	6	13.29	0.45
		4	11	10.54	1.04
				Mean	0.70
				StDev	0.27

View Derivation

A preliminary review of the video footage was undertaken and categories of 'view' were established by the data analyst. The analyst then arbitrarily selected at least 4 images from each view category and determined the area of the camera's substrate footprint and the size of the smallest identifiable features using the image analysis program Image Pro Plus® (Figure D2).

For the majority of video, the lowest limit of object discrimination was proportional to the camera's footprint. The average minimal resolution per view category was used to quantitatively define the view level.

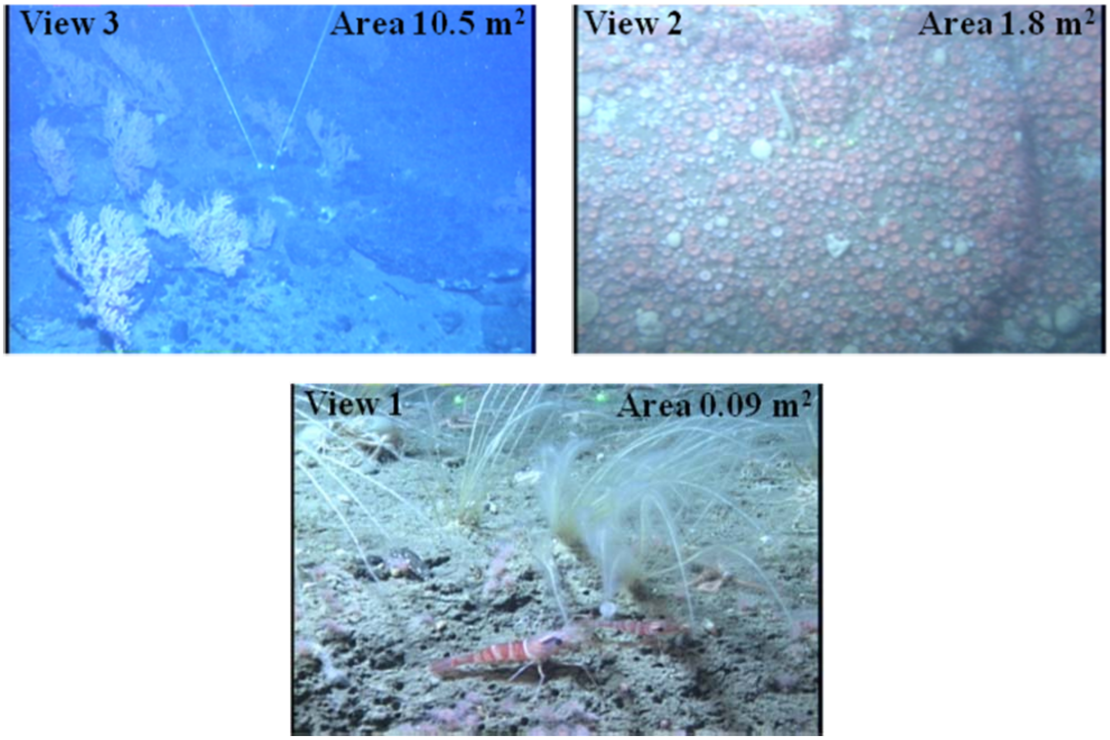


Figure D2. View fields for analysis of the video survey data from the Discovery Corridor, Gulf of Maine, 2006.

Data Analysis Quality

The data analyst was trained to identify habitat features prior to conducting the video analysis with ClassAct Mapper. To identify substrate types and percent cover of substrate, the analyst selected images from the video data and conducted quantitative analyses to determine the size of substrate features and their associated percent cover with Image Pro Plus®. The analyst was trained to identify biological roughness elements by studying images that had been used to obtain the quantitative descriptors for the biological roughness elements. The images quantitatively analyzed to determine the view categories were also studied by the analyst to help them consistently identify the view of the camera.

The same person conducted all preliminary viewing of the video, quantitative derivation of features, and data analysis to reduce the potential of a learning curve effect on the analysis of the video data.

APPENDIX E: EXAMPLE APPLICATION

Detailed analysis of linkages between physical and biotic habitat structural elements and the distribution and abundance of benthic organisms within the Discovery Corridor will be reported separately. Examples of data representation available as a consequence of adopting a relational database design and GIS-applicable scheme for benthic habitat classification are presented for a portion of a dive conducted with the ROV ROPOS in Northeast Channel during the 2006 CCGS Hudson research mission.

The variables depth, physical roughness, and biological roughness are recorded directly in the relational database, while substrate majority and substrate evenness are derived from the percentages of different substrates recorded within the database at each geo-referenced time.

Substrate evenness is a diversity index that quantifies how equal the representation of six substrate categories (silty sand, sand, pebble, cobble, bedrock, boulder) is within a sample (defined as the percent cover observations at each given time step). This index follows the general form of Pielou's evenness index (Pielou 1966):

$E = \frac{H'}{H'_{max}}$ where,

$H' \; (Shannon's \; Diversity \; index) = -\sum_{i=1}^{s} \mathrm{p_i} \ln \mathrm{p_i}$ where,

$\mathrm{p_i} = \frac{n_i}{N}$

n_i = proportion of each substrate in each category and $N = 1$

s = total number of substrate categories

$H' \max = \ln S$

For a segment of dive R976 (Figure E2), the substrate majority changes from bedrock dominanted to pebble dominated with increasing depth. Substrate evenness has higher values, indicating higher substrate diversity and evenness, in the shallower portions of the dive segment. Biological roughness elements derived from brittlestar fields, sponge dominated fields, and coral thickets are associated with these shallower bedrock-dominated areas that have high substrate evenness. A region of bioturbation occurs at intermediate depth within this segement of dive R976.

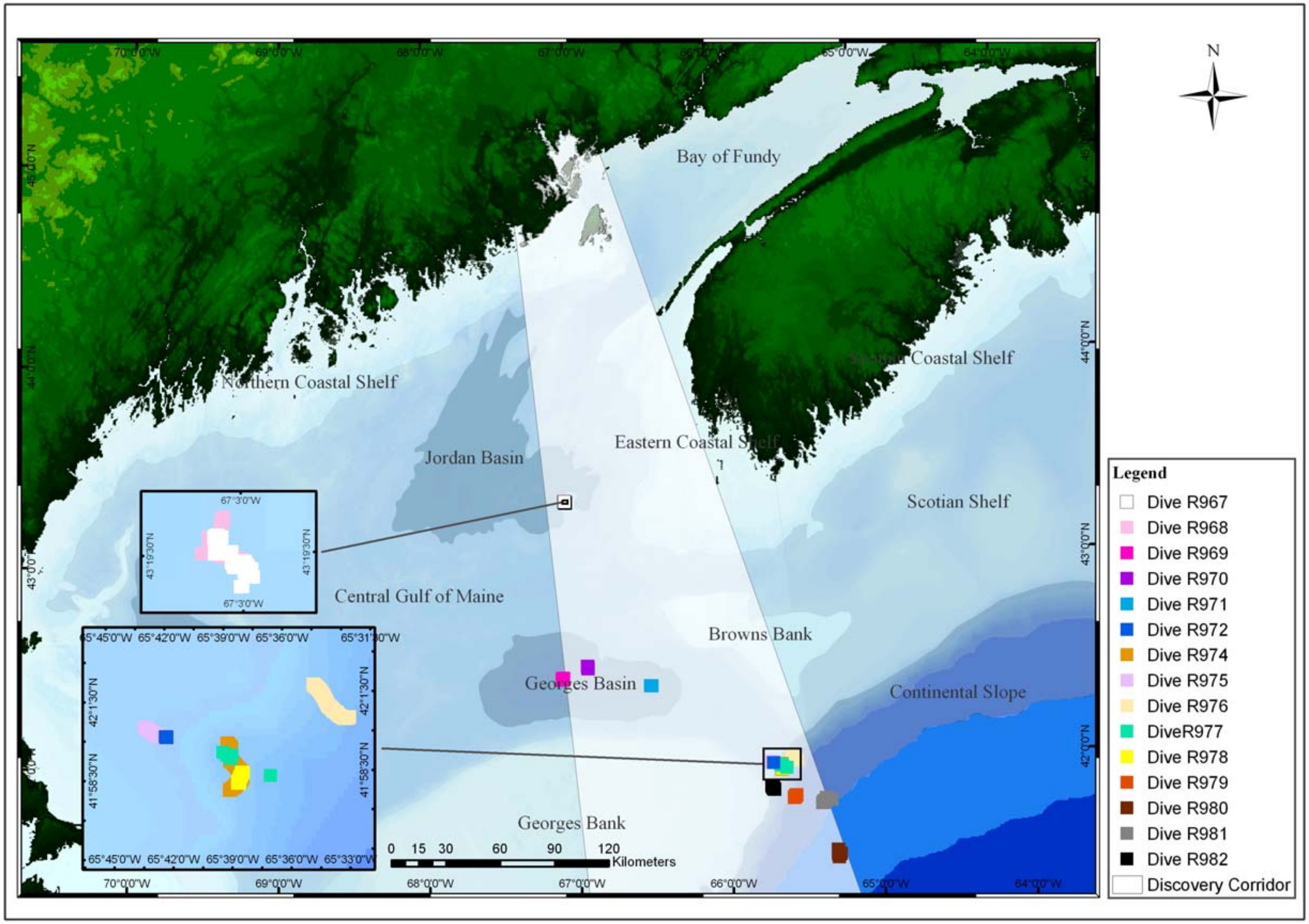


Figure E1. General overview of the locations of 15 ROPOS dives conducted in the Discovery Corridor in 2006

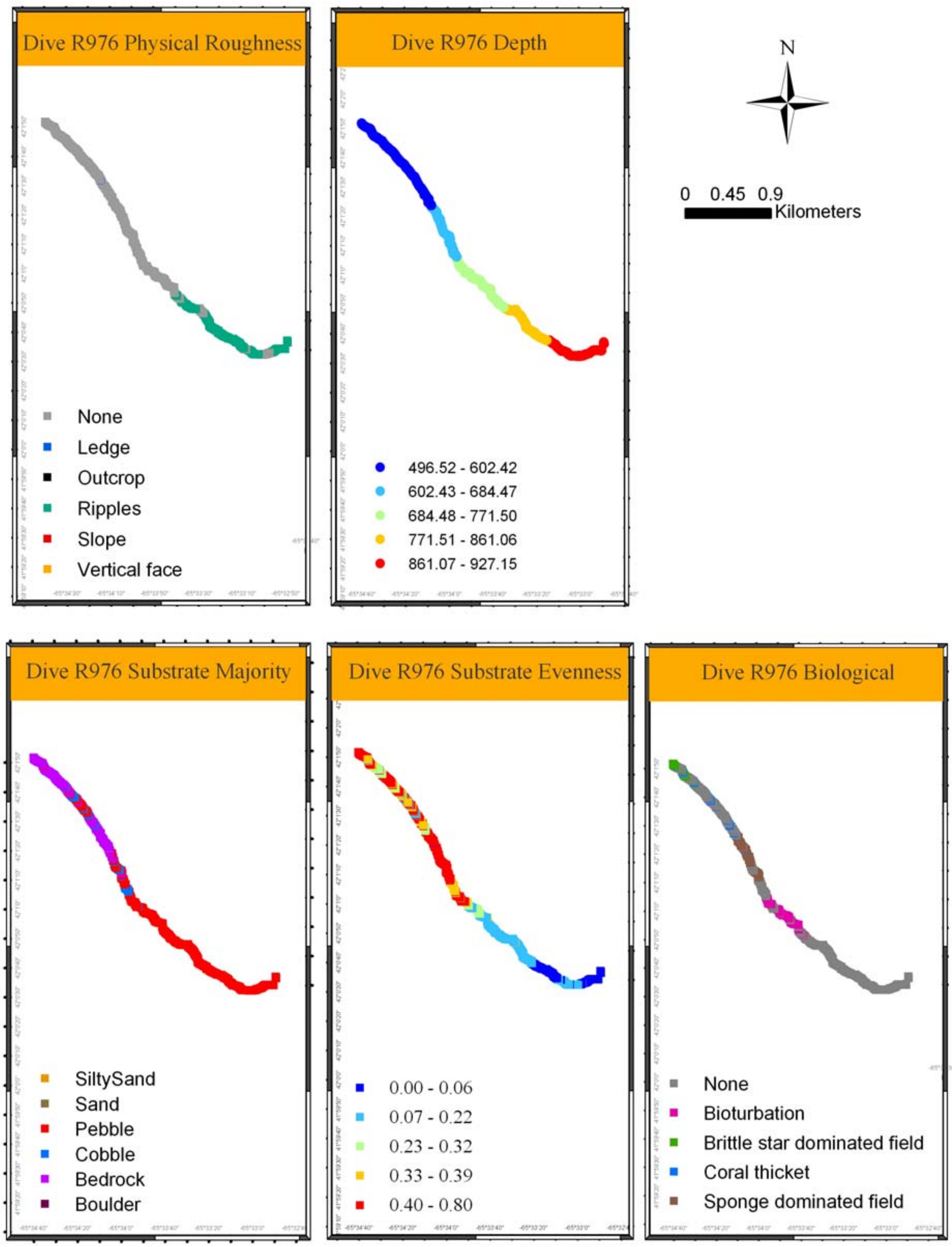


Figure E2. Segment of Dive R976 covering a depth transition from approximately 500 to 900 m over a distance of 5 km. Five separate panels show different benthic habitat variables