A TOWED BODY STEREO-VIDEO SYSTEM FOR DEEP WATER BENTHIC HABITAT SURVEYS

Shortis, M. R.[1], Seager, J. W.[2], Williams, A.[3], Barker, B. A.[3] and Sherlock, M.[3]
[1]RMIT University, Australia
[2]University of Western Australia (now with SeaGIS P/L, Australia)
[3]CSIRO Marine and Atmospheric Research, Hobart, Australia.
mark.shortis@rmit.edu.au

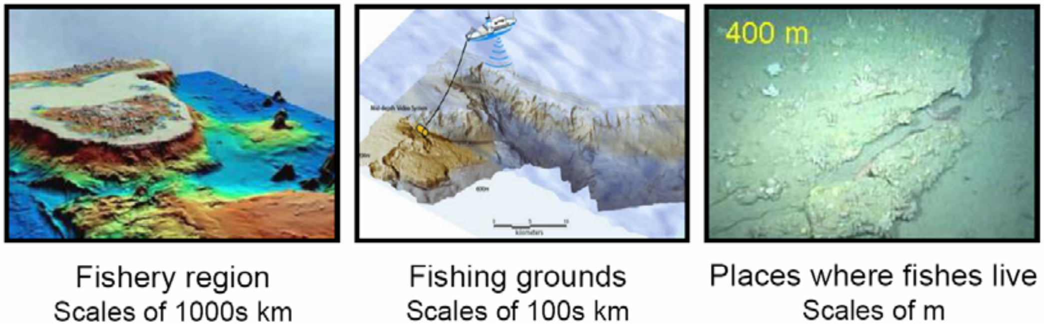
KEY WORDS Stereo-video, towed body, deep water, fisheries, marine habitat

ABSTRACT

Towed body systems of various configurations have been used for many years to map the seabed. Prior to the last several years, single video camera systems were widely used to gather qualitative data, or collect low accuracy quantitative data using laser dot patterns projected into the field of view. The introduction of stereo-video systems has enabled the capture of high accuracy spatial information. CSIRO has recently adopted stereo-video on a towed body system used for habitat mapping and biodiversity survey work. This paper provides an overview of the research context, examples of habitats being mapped in the south east Australia fishery region, describes the towed body system, and reports on the status of the development project.

1. INTRODUCTION

Over the last few decades there has been a rapidly increasing global focus on marine benthic biodiversity conservation, especially through the implementation of Marine Protected Areas (MPAs), that has generated the need for multi-scale maps of seabed habitat (Williams *et al.*, 2005). In addition, the acknowledgement that fisheries need to be managed for ecological sustainability, rather than simply on the basis of regulating catch or effort, has generated the need to understand and quantify the interactions of fishing gear with the benthic environment (Hobday *et al.*, 2006).



Fishery region
Scales of 1000s km

Fishing grounds
Scales of 100s km

Places where fishes live
Scales of m

Figure 1. Multi-scale mapping of the SE fishery region.

Australia is constructing a national network of MPAs in offshore waters (DEWR, 2007) where all waters are deeper than SCUBA diving depths (>50 m). As well, a large proportion of total fisheries catches in Australia are taken below this depth. As a consequence, deep benthic habitats in the South East Region of Australia are being mapped to support the development of an integrated and ecosystem-based approach to planning and management for all ocean uses. Information from surveys is being integrated to produce habitat maps at various scales of resolution so that the multi-scale structure of benthic habitats (see figure 1) can be understood and natural regions can be identified as planning units.

Research information for the spatial management of deep seabed habitat has to be gathered remotely, and towed camera systems are an integral part of that capability. Video sequences are used to verify habitats mapped at coarser scale by multi-beam sonar. The methodology overlays fine-scale detail from video transects onto the intermediate-scale detail provided by hydro-acoustics in order to understand the broad scale issues across the fishery region as a whole.

2. DATA COLLECTION

The primary survey tool is a towed camera platform that records continuous, medium resolution stereo-video sequences and intermittent high-resolution digital still images along transects. The platform operates to depths of 1,500 metres and is connected to the vessel via a 3,200 metre steel-armoured cable containing fibre-optic and conducting wires. Two PAL video cameras, configured as a stereo-pair, transmit live video sequences that are recorded on time-coded DV tape. The recordings are indexed to vessel differential global positioning system (DGPS) navigation data and to ultra short baseline (USBL) tracking beacon data on the towed body, so that imagery can be accurately geo-located. Geo-location of sampling on the seabed is critical to relate the sampled area to environmental co-variates extracted from hydro-acoustic and other sensors. Calibration and data processing requirements for the various sensors is described in Williams *et al.* (2007).

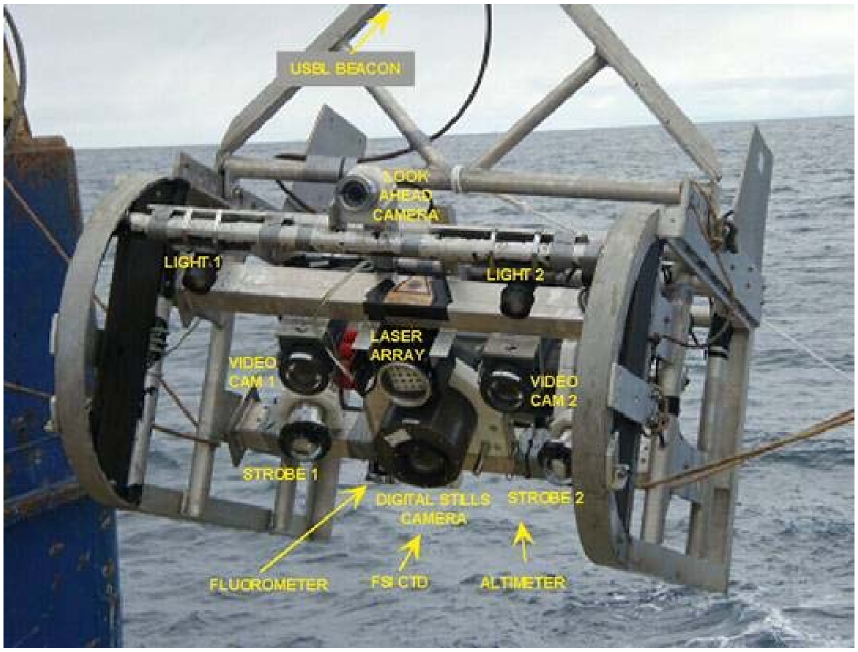


Figure 2. CSIRO towed body platform

A separate forward-looking camera provides an additional view for navigation and obstacle avoidance. Additional sensors record altitude, pressure, pitch, roll, water temperature, conductivity and fluorescence. All sensor data is captured to a log file and combined with vessel DGPS and USBL information. Several sources of incoming data are displayed graphically on a custom-made LabView “console” on an onboard PC screen to provide feedback to the pilots for control of the system. The console is also the switching interface for components. AC power is supplied to the system from the ship. Two 250 Watt incandescent lights provide illumination for the video cameras. Strobes provide illumination for the digital still imagery.

The camera system is deployed over the stern of the vessel using a gantry. The optimum tow speed is 1-1.5 knots and the pilot operates the winch to “fly” the platform just above bottom so that the cameras view the sea floor obliquely from 1-3 metres above the seabed. Deployments are typically 30-60 minutes duration, producing transects of 1-3 kilometres in length, but, if required, the body can be towed continuously for several hours.

The resolution of the video images is a limitation, so high resolution digital still images enable qualitative analysis at a greater level of detail. The digital still camera is remotely triggered by the operator or programmed to fire at set intervals. Images are captured to the internal storage of the camera and later uploaded to the logging computer. As a further measure to overcome the limitation of PAL video resolution, high-resolution (1392 x 1040 pixel) progressive scan cameras are under evaluation for the stereo-video imaging. Synchronized image pairs are transmitted via fibre-optic transmission to the surface and recorded at 12 to 15 image pairs per second. Whilst the fibre-optic bandwidth and other factors limit the frame rate compared to standard video, there are several advantages to this approach. The high resolution images improve the measurement accuracy from the stereo image pairs, are accurately synchronized and are recorded direct-to-disk in readiness for analysis.

3. SYSTEM CALIBRATION

3.1 Shallow Water Calibration

The stereo-cameras are pre- or post-calibrated in shallow water, usually in a swimming pool, using the techniques developed by Shortis and Harvey (1998). The standard requirements of a multi-station self-calibration network are required, such as multiple convergent photographs, camera roll at each location and a 3D target array. The 3D target array, usually in the form of a light, easily manoeuvrable calibration fixture, has the size determined by the field of view of the cameras and the likely working distance for the measurements. It is impractical to manoeuvre towed body systems in the same way as a hand-held camera, so instead the calibration fixture is tilted and rotated in the field of view of the camera (see figure 3) to replicate the convergent multi-station network (Harvey and Shortis, 1996). Accurate information for the positions of the targets on the calibration fixture is not required, as co-ordinates of the targets are derived as part of the self-calibration procedure. Hence it is immaterial if the frame distorts or is dis-assembled between calibrations, although the frame must retain its structural integrity during a calibration sequence. Scale is determined by distance constraints on between targets on the rigid arms of the frame.

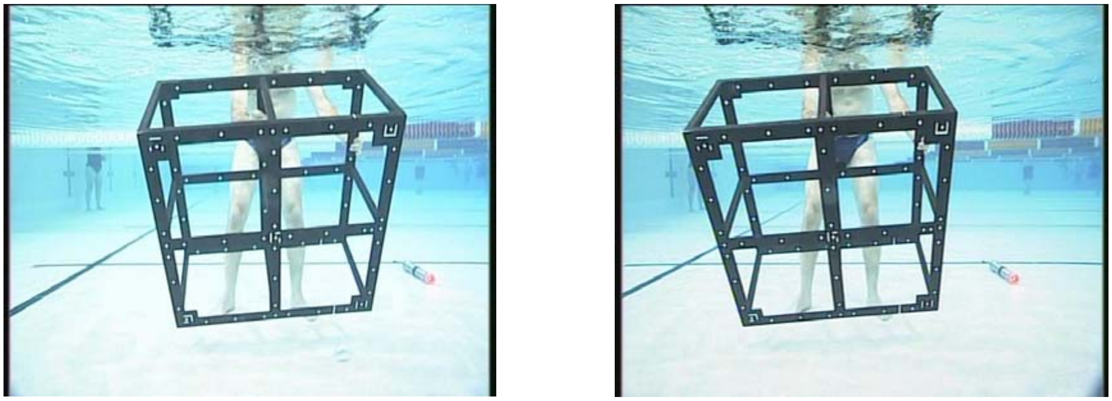


Figure 3. Typical stereo pair of shallow water calibration images (the LED is used for synchronisation checks)

The camera calibration model does not need to contain explicit terms for the refractive effects of the camera ports and the refractive interfaces, as analysis of the effects of the refractive surfaces in the optical path in an ideal camera housing shows that images are displaced radially from the principal point (Li *et al.*, 1996). Whilst the assumptions that the optical components of the housing are symmetrical around the optical axis of the camera and refractive surfaces are in general perpendicular to the optical axis are unlikely to be perfectly fulfilled in practice, it is clear that the primary component of the refractive effect is radial. As a consequence, the approach that has been adopted has been to allow the refractive effects of the optical components and refractive interfaces to be absorbed by the conventional, physical camera calibration parameters (Shortis and Harvey, 1998). The principal component is implicitly taken up by the standard, odd-ordered polynomial model for radial distortion, whilst any residual effects from asymmetric components of the housing are partly or wholly absorbed into other parameters of the camera calibration, such as decentring lens distortion or the affinity term. No assumptions need to be made concerning the refractive indices of the air, glass or water media, and modelling of the optical components of the underwater housing is unnecessary.

Rigid mounting of the camera housings to the frame and a rigid connection between the cameras and the view ports generally ensures the stability of the relative orientation of the cameras (Shortis *et al.*, 2000). Experience has demonstrated that a weakness of the implicit model for refraction is the integrity of the full light path from the first water-port interface through to the image sensor. A consistent spatial relationship between the view port and the camera lens is critical to this stability.

3.2 Deep Water Operations

For deep-water operations there may be measurement inaccuracies resulting from the application of a camera calibration carried out in shallow water to imagery gathered at much greater depths. Stereo camera calibrations are generally carried out at depths of 1-3 m for operational convenience, however the stereo cameras can subsequently be deployed to depths of up to 2000 m. Under these conditions of considerably increased water pressure it is suspected the camera housings and view ports will deform, and the deformation may adversely affect camera calibration and subsequent stereo measurement.

In order to quantify depth effects on camera calibration, a purpose-built laser projection system has been developed (see figure 4). The system uses 16 individual lasers arranged in an approximate 4 by 4 grid pattern. Each laser is 25mw at a wavelength of 658nm. Red lasers were chosen because, although they have poor transmission properties through water, they are inexpensive and provide a unique red feature that can be automatically extracted from imagery. A small servomotor is used to rotate the entire laser projection system.

Figure 4. Detail of the 16 laser array mounted between the stereo-video cameras.

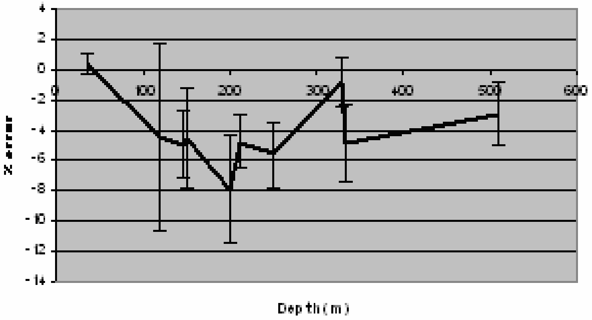


Figure 5. % error in distance measurement caused by depth effects.

The laser projection system is designed to introduce a pattern of dots into the view of the stereo cameras. The laser dots are used as well-defined targets, and their locations are measured in a series of images as the towed body flies over the ocean floor. Variations in the ocean floor topography and the flying height of the towed body are used in conjunction with rotations of the laser projection system to provide wide variety of projection geometries. The measured laser dot locations are used in a least squares estimation process to simultaneously determine the camera calibration parameters as well as the physical properties of the laser projector. The simultaneous determination of camera and laser projector parameters effectively removes any depth effect. The process gives a measure of the effect of depth, or the effect of using shallow water calibrations at depth, and also provides a method for calibrating the cameras at depth.

Over the various laser system deployments it was found that depth did have an effect on the stereo camera calibration. Figure 5 shows the percentage error for measured distances caused by depth. The maximum error observed is around 8% at a depth of 200 m. Interestingly, there is little correlation between depth and error. It is assumed that the observed depth effect is caused by somewhat unpredictable settling of the camera housing view port. The view port component is made of a lens and several housing components, and involves multiple o-ring seals. It appears that the view port exhibits a precessional motion as it settles, and does so in a slightly different manner each time the camera system is deployed. These experiments have led to some design changes in the camera housings in order to try to eliminate the effects of view port and o-ring settling.

Although deployments of the laser projection system have suggested some significant depth effects, the experiment has been slightly deficient in that it has lacked a reliable scale reference. A pair of parallel lasers was used to provide a scale reference, and although these lasers had a very small housing and view port, stability with depth could not be validated. Deployments of the laser projection system are being repeated using an accurate scale reference to confirm the results.

4. DATA ANALYSIS

The stereo-video images enable accurate 3D measurements of marine fauna (see figure 6), or seafloor dimensions such as boulder size or shelf height. The system was originally developed for length measurement of fish as a tool for population distribution sampling within MPAs (Harvey and Shortis, 1996) and is based on operator identification of points of interest in the stereo-images. Manual measurement and analysis of large volumes of video sequences is time consuming, labour intensive and therefore costly. The efficiency of the measurement process is particularly important as CSIRO researchers collect hundreds of hours of video recordings annually during biodiversity and fishery habitat surveys.

Collaborative research between CSIRO and the universities is driving the development of automated techniques for analysing the video sequences. Motion analysis, image segmentation against a decaying background, and colour matching are being used to identify the presence and percentage cover of benthic fauna and substratum types in the video sequences (see figure 7). Stereo-measurement can then provide the sizes of individual animals or substratum features within selected image pairs to estimate population characteristics.

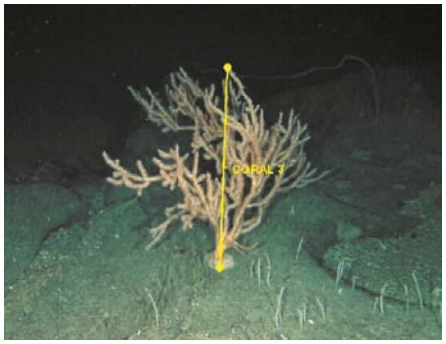


Figure 6. Example of an operator measurement (left image of a pair) of the height of a deep-water coral.

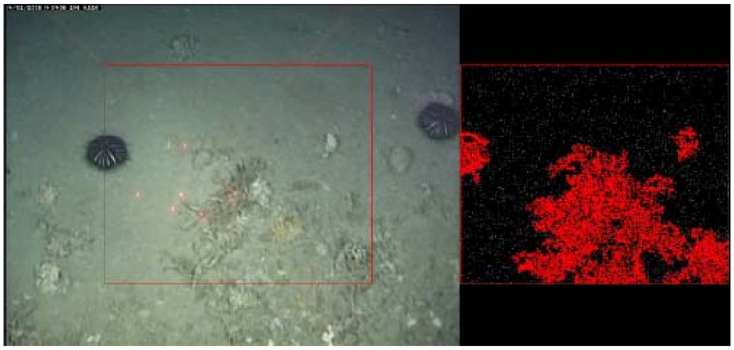


Figure 7. Example of candidate regions detected from motion analysis.

The motion analysis techniques were initially developed to identify candidates for counting and sizing of fish in aquaculture (Harvey *et al.*, 2004). Motion analysis is first being used to identify sections of the image sequences that contain features of interest, effectively eliminating portions of the video that are devoid of features and not of direct interest in terms of habitat mapping. This processing is effectively an image compression technique that dramatically reduces the amount of video sequences requiring inspection, and reduces AVI file sizes. The motion analysis is then used to estimate the percentage cover of interest regions within the video transects. The motion detector can be tuned to detect featureless versus feature-rich regions, or specific marine fauna or flora. Current research is focussing on using stereo camera geometry, colour-based affine image matching and projected laser dots to accurately determine the true geometric area of regions extracted by motion analysis. Ultimately the research will lead to automation of the process of identifying and estimating the true percentage cover of benthic features.

Future development of the algorithms will focus on stereo matching techniques to automatically determine 3D volumes and surface areas. A critical factor in the effectiveness and robustness of the algorithms will be the improvement of image quality and resolution to be provided by the digital progressive scan cameras and direct-to-disk system. As can be seen from figure 7, the image quality from the standard video system and the general reduction in image contrast caused by attenuation through the multiple interfaces and water medium is a limiting factor.

5. HABITAT MAPPING APPLICATIONS

The primary contribution of video data to multi-scale surveys of the seabed is the verification of habitat. Video transects add fine scale detail to areas mapped and differentiated by multi-beam acoustics at intermediate scales. Video transects typically target contrasts in acoustic maps to validate changes between habitats (see figure 8). Information on the biological associations with physical components of habitats enable mapped acoustic data, which has large coverage and is relatively inexpensive to collect, to be used as a proxy for the distribution of biodiversity. Based on analysis of the video sequences, abundance measures such as numerical, biomass, density or cover of several components of benthic megafauna can be made at a variety of scales of taxonomic resolution, and can be related to habitat types at a variety of spatial scales (Williams *et al.*, 2007). These are the building blocks for community biodiversity metrics such as richness and diversity.

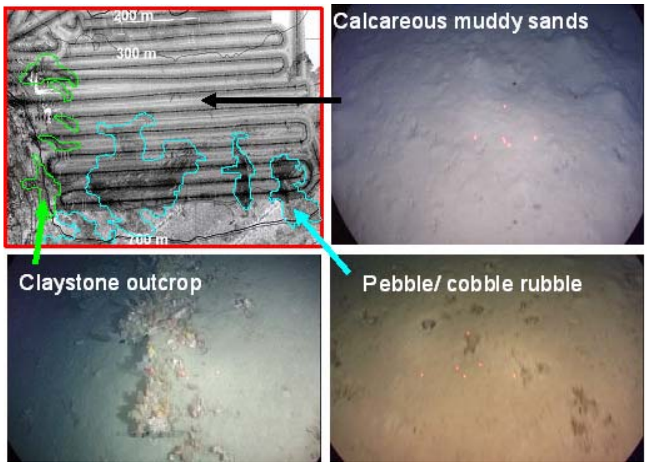


Figure 8. Verification of acoustic terrains by video imagery at canyon edge

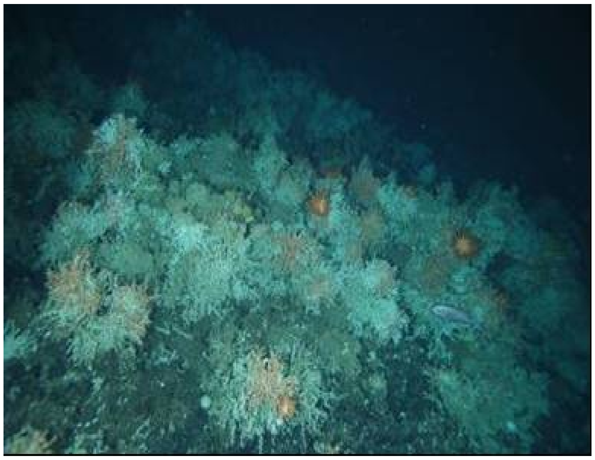


Figure 9. Fragile stony coral and rich biodiversity on a seamount at 1100 m depth.

The non-destructive nature of video sampling gives it a significant advantage over conventional physical sampling with an epibenthic sled or trawl, particularly for monitoring. Biodiversity mapping relies on initial physical collection to provide an inventory of fauna, but sensitive environments such as seamount coral communities (see figure 9), especially in conservation areas, benefit greatly from subsequent monitoring that is non-destructive. While video surveys will never replicate the species-level resolution possible from collections of benthic fauna, there is substantial scope to take data for species that are distinctive in imagery, for species where area (density or cover) measures are most relevant, for example encrusting species or aggregated communities, where species cannot be sampled by traditional equipment, such as on very steep or craggy bottom, or where classes of similar species types are sufficient for the needs of an indicator of community change. Where species have strong habitat associations and habitats have high spatial heterogeneity at scales of tens to hundreds of metres, video sampling will also provide more robust measures of abundance because the data are continuous and do not integrate across habitats. In contrast, samples from mobile collecting devices such as sleds or trawls integrate across habitats, mixing the fauna and adding considerable uncertainty to abundance estimates.

A combined measure of the heights of many individual animals and the plan area of their distributions provide measures of habitat heterogeneity, habitat value for other structural habitat-associated fauna such as fishes, and importantly for changes over time. Size-related metrics provide the basis for tracking the slope and intercepts of size spectra which have been identified as a reliable indicator for the health of fish populations (Rice, 2000).

Recognition of direct impacts on benthic habitat resulting from human activities, notably bottom contact fishing techniques, is a key element of defining biodiversity conservation needs, and in managing fishery areas. Image data provide both a means to evaluate and predict habitat vulnerability and to record direct impacts. In deep upper and mid-slope depths in the SE region of Australia, this includes the overturning of friable rocks, gouging of soft mudstones and sediments, and loss of fishing gear (Williams *et al.*, 2006). Where impacts are visible, photographic data permit the condition to be expressed qualitatively as present/absent, relatively as high, medium or low, or more ideally quantitatively as occurrence rates per swept area of video transect.

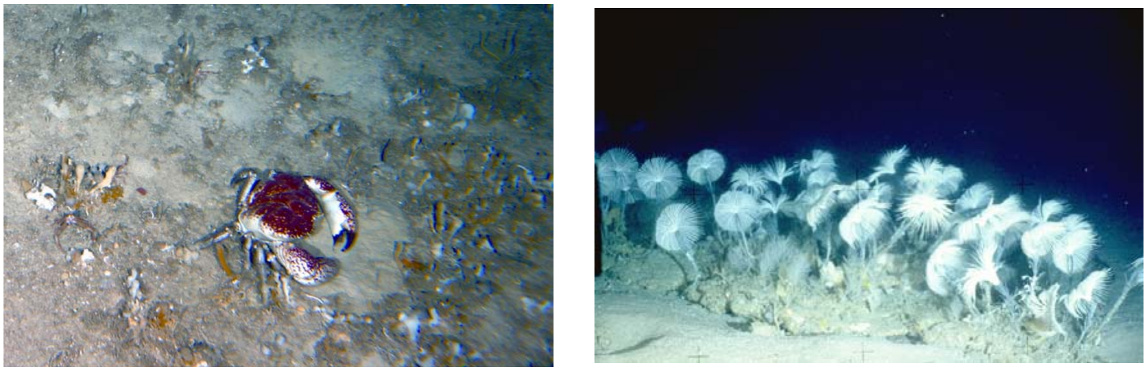


Figure 10. Digital still images showing a giant crab *Pseudocarcinus gigas* amongst sponges and bryozoans at 340 m and the rare stalked crinoid *Metacrinus cyaneus* at 200 m water depth.

In Australia, image data have been used as the basis for risk assessment approaches both for MPA zoning (regulating a range of different activities within specific sub-areas of MPAs), and for spatial management of the ecosystem effects of fishing (Hobday *et al.*, 2006). Established risk assessment methods that draw on catalogues of benthic habitats (Williams *et al.*, 2005) provide a means of translating a variety of fishing impacts into maps of vulnerability for habitat (physical and biological components), species and communities. Both fishery and MPA risk assessment methods rely on the data produced by the towed body system described here. One example of applying video data in this way was to understand biodiversity distribution and impacts on the upper continental slope (150 to 450 m depths) where bottom fish-trawling and giant crab trap fisheries overlap (Williams *et al.*, 2007). Mapping shows that the distribution and cover of a dominant component of habitat, low-relief, bryozoan-based ‘thickets’ (see figure 10) is both an important foraging area for giant crabs, and vulnerable to being degraded by trawling.

‘New’ discoveries of fauna are a valuable outcome from video photography during ‘exploratory’, as opposed to monitoring, surveys for mapping benthic habitats. For example, recent surveys identified dense aggregations of a stalked crinoid at 200 m depth in the shelf-edge head of a submarine canyon. Crinoids have been in the fossil record of the Earth for millions of years but are now rare off SE Australia, with aggregations known only from this single location (see figure 10).

6. CONCLUSIONS

This paper describes the development and application of a deep water towed body system to map sea bed habitat in deep water. The system provides the ability to extract quantitative data, such as abundance, size and area measurements, from stereo video with known estimates of error. It has important applications for conservation and fishery management, particularly by providing fine-scale, continuous, non-integrated, non-destructive data on animal and habitat distributions. On-going enhancements of the system, namely progressive-scan high resolution video imagery and laser array self-calibration, will substantially improve the accuracy, resolution and utility of the system in the future.

7. ACKNOWLEDGMENTS

The authors gratefully acknowledge the skilled team in CSIRO Marine and Atmospheric Research Engineering and Workshop Services, especially Matthew Horsham, for design and manufacture of camera system components, Jeff Cordell for electronics design and, together with Mark Lewis, for operational support both at sea and in the laboratory. The depth calibrations were made possible due to the generosity of the University of Western Australia for the loan of the laser array system. Thanks are due also to Julian Finn of the Museum of Victoria for permission to publish the photographs used in figures 2 and 4, and Geosciences Australia for contributing part of figure 1.

REFERENCES

Department of the Environment and Water Resources, 2007. New Proposed Network of Commonwealth Marine Reserves for Australia's South-east Marine Region, www.environment.gov.au/coasts/mpa/southeast/index.html, accessed April 29, 2007.

Harvey, E. S. and Shortis, M. R., 1996. A system for stereo-video measurement of sub-tidal organisms. *Marine Technology Society Journal*, 29(4): 10-22.

Harvey, E. S., Shortis, M. R., Seager, J. W. and Hall, N., 2004. Refining non intrusive stereo video techniques and protocols for SBT transfers. Report R03/1428 for the Australian Fisheries Management Authority, 37 pages.

Hobday, A. J., A. Smith, H. Webb, R. Daley, S. Wayte, C. Bulman, J. Dowdney, A. Williams, M. Sporcic, J. Dambacher, M. Fuller and T. Walker, 2006. Ecological Risk Assessment for the Effects of Fishing: Methodology. Report R04/1072 for the Australian Fisheries Management Authority.

Li, R., Li, H., Zou, W., Smith, R. G., and Curran, T. A. 1996. An underwater digital photogrammetric system for fishery geomatics. *International Archives of Photogrammetry and Remote Sensing*, 31(B5): 524-529.

Rice, J. C. 2000. Evaluating fishery impacts using metrics of community structure. *ICES Journal of Marine Science*, 57(3): 682-688.

Shortis, M. R. and Harvey, E. S., 1998. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. *International Archives Photogrammetry and Remote Sensing*, 32(5): 792-799.

Shortis, M. R., Miller, S., Harvey, E. S. and Robson, S. 2000. An analysis of the calibration stability and measurement accuracy of an underwater stereo-video system used for shellfish surveys. *Geomatics Research Australasia*, 73: 1-24.

Williams, A., R. Kloser, B. Barker, N. Bax and A. Butler. 2005. A seascape perspective for managing deep sea habitats. pp 89-97, in Proceedings of ‘Deep Sea’: *Conference on the Governance and Management of Deep-sea Fisheries*. Part 1: Conference reports. Queenstown, New Zealand, 1-5 December, 2003. FAO Fisheries Proceedings. No. 3/1. Rome, FAO. 2005. 718 pp.

Williams, A., N. Bax and B. Barker. 2006. Integrating fishing industry knowledge of fishing grounds with scientific data on seabed habitats for informed spatial management and ESD evaluation in the South East Fishery. Final report to the FRDC, project 2000/153.

Williams, A., F. Althaus, B. Barker, C. Gardener and D. Mills. 2007. Understanding shelf-break habitat for sustainable management of fisheries with spatial overlap. Draft final report to the FRDC, project 2004/066.