Underwater 3D Mapping and Pose Estimation for ROV Operations

Piotr Jasiobedzki, Stephen Se, Michel Bondy, Roy Jakola
MDA, Space Missions,
9445 Airport Rd, Brampton, ON, L6S 4J3, Canada
piotr.jasiobedzki@mdacorporation.com

Abstract- **Teleoperating Remotely Operated Vehicles (ROV) underwater is a challenging and tedious task as the operators are physically removed from the sites and their situational awareness is low due to restricted camera views, poor visibility under water, and limited three dimensional perception. This paper describes three vision system technologies, developed originally for space applications, which will enable more efficient ROV operations. Automatic detection and tracking of artificial and natural targets will enable autonomous ROV station-keeping (mooring) allowing the operators to focus on their main tasks. The 3D modeling system will create instantly models of underwater worksites and seafloor increasing thus operators' spatial perception and allowing measurements and precise manipulation tasks. Initial results of underwater experiments are presented.**

I. INTRODUCTION

Remotely Operated Vehicles (ROV) are used to provide assistance to divers underwater or extend the human reach to depths where divers cannot operate safely and effectively. ROV operators rely on video and sonar images, monitor data from on-board sensors and subsystems displayed on their console to teleoperate the ROVs and robotic manipulators. ROV's are primarily used to perform two types of tasks: 1) subsea equipment installation, inspection, manipulation and maintenance; 2) sea-floor survey and inspection [1].

Teleoperating an ROV is a challenging and tedious task. The operators are physically removed from the sites and their situational awareness is low due to restricted camera views and visibility under water, coarse sonar imagery and poor three dimensional perception. Operating simultaneously multiple sub-systems of the ROV in the vicinity of the sea floor or underwater structures is particularly demanding. Solutions to these problems will be similar to approaches used in robotic systems for space exploration, servicing satellites on-orbit, and unmanned mining machinery and vehicles. In theses domain the solutions are sought by:

- Increasing autonomy of remote systems and shared control
- Increasing the operator situational awareness by providing him with virtual presence in the remote site

Operating robotic manipulators and tools on board an ROV during underwater construction / recovery operations requires simultaneous control of the robotic devices and maintaining the ROV position. Reactions to contact forces or water currents affect the ROV location requiring regular adjustments.

Shared control of robotic hardware alleviates some of the limitations of teleoperation by autonomous execution of mundane and repeatable operations, and retaining operators' involvement in tasks requiring human skills. Automation of simple ROV tasks such as station-keeping relative to the worksite will allow the operator to focus on the primary task. Such automation can be achieved using sophisticated vision systems that recognize designated structures or visual targets and estimate their relative position. Architecture and results of experiments with of a prototype vision system have been proposed [2]. Simple fiducials (visual targets) were detected by a vision system and the computed bearing and range were fused with gyro measurements providing relative position estimate, which in turn was used to control the ROV thrusters. Similarly, tasks such as visual homing to a target designed by the operator, maintaining safe distance and obstacle avoidance can be performed autonomously.

Ground control of space robotic systems relies on virtual presence by the ground based operators to plan and monitor mission execution. Virtual presence and virtual reality has been postulated for underwater telerobotic use over 20 years ago [3]. The proposed approaches involve presenting, similarly to space operations, a graphic, virtual reality model of the workspace to the operator, who then performs tasks on this representation of the actual work site. Recent real world application of virtual presence included recovery of Ehime Maru [4]. A real-time database was maintained an accurate representation (world model) of all the elements that composed the underwater scene. Location of surface equipment was tracked using GPS and visualized together with 3D underwater scans obtained beforehand. The system was used for task planning and monitoring of their execution. Using similar systems at a smaller scale of a single ROV will require creation a local 3D model of the worksite (using 3D imaging systems) and placing a virtual model of an ROV – similarly to robotic planetary exploration and satellite servicing on-orbit.

II. VISION SYSTEMS FOR UNDERWATER APPLICATIONS

MDA is currently adopting some of the sensing and robotic technologies, originally developed for space applications, for use underwater. Operating space robotic systems is similar to operating ROVs and AUVs: remote and inaccessible locations, low situational awareness of the operator, lack of accurate maps or models of the environment, low communication bandwidth and latency (AUVs) and significant cost of losing the vehicle or retrieval. There are obvious differences: design for high vs. low ambient pressure, low communication bandwidth and long latency for space vs. high bandwidth and real time links for ROVs. Currently at MDA we are focusing on two aspects

- Vision systems for shared ROV control to offload operators from routine tasks during robotic operations
- Creation of a photorealistic and calibrated three-dimensional models of sea floor or underwater worksites

III. VISION SYSTEMS TO SUPPORT SHARED CONTROL

One of the simplest functions and useful to the ROV operator will be an ability to moor (station-keep) the vehicle at a fixed position with respect to a worksite without any physical restraints. This position may be determined by on-board vision systems that detect and estimate pose of a special target embedded or temporarily attached within the worksite. Alternatively, the operator may designate natural features already existing in the worksite as visual targets. The relative pose computed by vision systems can be directly used to control the ROV or to provide warnings to operator about unexpected motion. The following sections describe in detail vision systems that perform such functions.

A. Marker Based Pose Estimation

MDA vision systems can automatically recognise and estimate relative position and orientation (pose) of visual markers and natural features. Space Vision Marker System (SVMS) [5], originally developed for autonomous docking and servicing on-orbit, uses coded markers. Unique features of this system that make it particularly suitable for use in underwater ROV operations include

- Design of the markers enabling their reliable detection from a wide range of distances and viewing angles
- Encoding information in the markers allowing unique identification of objects and their position and orientation
- Redundant design and encoding allowing correct operation even when partial data is lost due to poor visibility and occlusion
- Three dimensional structure of the markers that allows computing their pose accurately
- Marker detection in every frame without inter-frame tracking, which enables instant recovery from occlusion or image loss
- Ability to use multiple markers in the same workspace for different objects or to increase the pose estimation accuracy

The basic SVMS marker consists of a planar base with a central black square surrounded by a white stripe. Attached to this stripe are data bits, which encode the marker identity. A black post with a white top is mounted in the centre of the black square – location of the post tip in images allows accurate estimation of the marker orientation. Fig.1 shows three markers of different sizes (note different encoding); the image on the left shows markers on a flat surface and the image on the right shows markers detected in the scene. Overlays of different colours indicate identified markers.

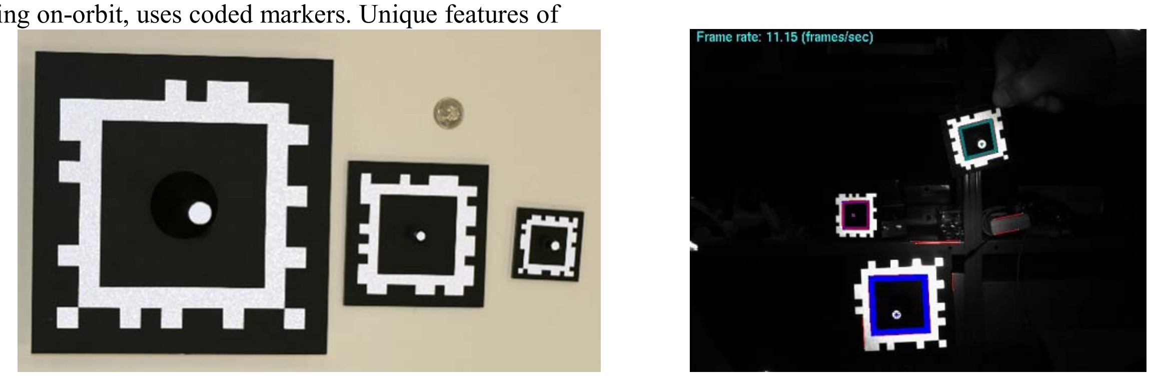

Figure 1. SVMS markers on a planar surface (left) and detected in a scene (right)

The SVMS vision system uses calibrated cameras and provides complete information on the relative target location (position and orientation – 6 degrees of freedom). Fig.2 shows images of the SVMS target detected in a hand held camera sequence, captured underwater in a swimming pool: blue overlays on the target indicate correct detection (marker is black and white) and the numbers on the top of the screen indicate the computed pose. The operating range (minimum and maximum distance, viewing angles) depends on the marker size, camera field of view and image resolution. Several versions of the markers have been developed and tested for space (spacecraft docking, robotic satellite capture and tool grasping) and terrestrial applications (autonomous convoy driving).

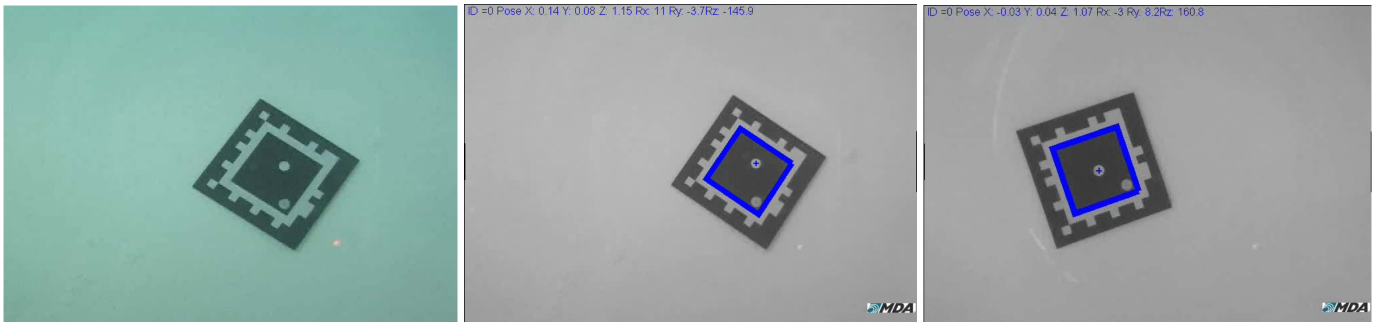

Figure 2 An input image from an underwater image sequence (left) and the SVMS target detected with computed pose (middle and right)

B. Pose Estimation with Natural Features

A similar concept for estimating relative pose has been developed for recognizing and tracking natural objects. The operator designates an object of interest in an image and, provided that the object contains enough distinct features, the vision system automatically recognizes the object and estimates relative position in subsequent images.

Description

The SIFT Object Recognition and Tracking (SORT) system uses a set of high level natural visual features called Scale Invariant Feature Transform (SIFT) for tracking. SIFT was developed [6] for image feature generation in object recognition applications. The SIFT features are invariant to image translation, scaling, rotation, and partially invariant to illumination changes and to affine or 3D projections. The SIFT features are highly distinctive for tracking as each feature contains a 128-element local image descriptor.

Visual tracking is performed by first matching each keypoint independently to the database of keypoints extracted from the training image. Many of these initial matches may be incorrect due to ambiguous features or features from background clutter.

Therefore, a cluster of features are identified that agree on the tracked object, as these clusters have a much higher probability of being correct than individual feature matches.

Based on the initial distance of the training image, we estimate the distance for the tracked object. The X and Y angles are estimated based on where the tracked object is found in the image. A confidence measure is computed based on the number of SIFT features that match the training image.

Initial results

For testing, we have used an underwater image sequence from Folkestone Marine Reserve in Holetown, Barbados. Selected images from the sequence are shown in Fig. 3.

In the first test, a region at on the sea floor is selected as the object of interest in the first image as shown in Fig. 4. SORT can subsequently track this object automatically and consistently throughout the image sequence, estimating the distance, angles and confidence. The estimated distance is relative to the initial distance; therefore even if the precise initial distance to the object is not known, the estimated distance would still be useful for ROV station-keeping.

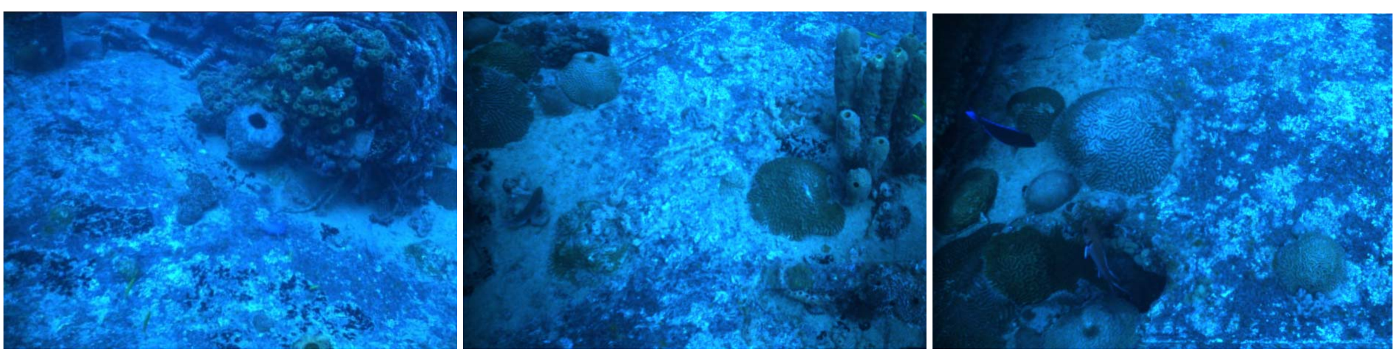

Figure 3 Selected images from an underwater image sequence

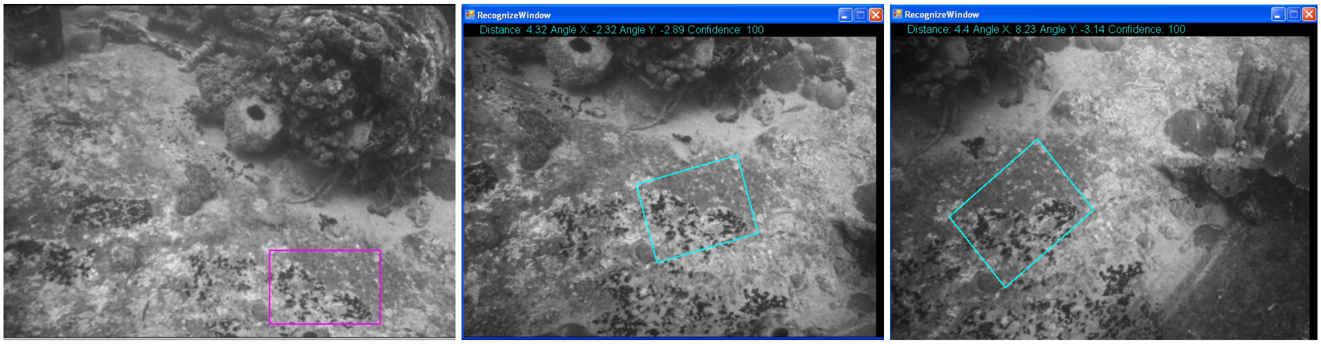

Figure 4 An object of interest is selected in the first image (purple rectangle in the left image) and automatically tracked in the subsequent images (middle and right)

In the second test, the system has been trained with an image of an object beforehand, in this case the coral shown in Fig. 5 (left). SORT detects the coral when it appears in images for the first time and then tracks it robustly. Partial occlusion does not affect the processing - notice the swimming fish shown in Fig. 5 (right). By The pose information can be used for ROV station-keeping, to maintain a certain relative distance and bearing angles between the ROV and the object.

Figure 5 A coral reef is selected as target a(left) and automatically detected and tracked in the image sequence (middle and right).

IV. 3D Mapping

The 3D mapping systems developed by MDA process sequences of images from mobile stereo cameras. The images may be acquired while ROV is moving above the seabed, along an underwater structure or the camera is panned and tilted to observe the worksite. The 3D mapping software computes the camera motion from the images alone or it may additionally use data from inertial sensors or telemetry. The 3D data from individual stereo pairs is combined using estimated motion to create continuous representations of the scene. The resulting 3D models can be viewed in 3D from any viewpoint, used for interactive or automated measurements of objects' sizes and distances, and in virtual environments for task planning and monitoring their execution.

A. Description

The MDA instant Scene Modeler (iSM) quickly generates calibrated photo-realistic colour 3D models of unknown environments from a mobile stereo camera [7]. It computes the 3D data, estimates the camera motion in 6 degrees of freedom and registers successive frames together. Fig. 6 shows the architecture of the iSM system. Images are captured by the stereo camera and dense stereo disparity is computed using a correlation based algorithm for each pair to obtain 3D data.

The system does not require any external sensors for computing the camera motion as it automatically extracts and tracks natural tie points in the images using the earlier described SIFT algorithm. The processing is based on a Simultaneous Localization And Mapping (SLAM) approach that uses the SIFT 3D features to localize and simultaneously build a database map [8]. Instead of frame to frame matching, we match the SIFT features at each frame with the database to reduce error accumulation. As the SIFT features are highly distinctive, they can be matched with very few false matches. A weighted least squares procedure is carried out taking into account the feature uncertainty.

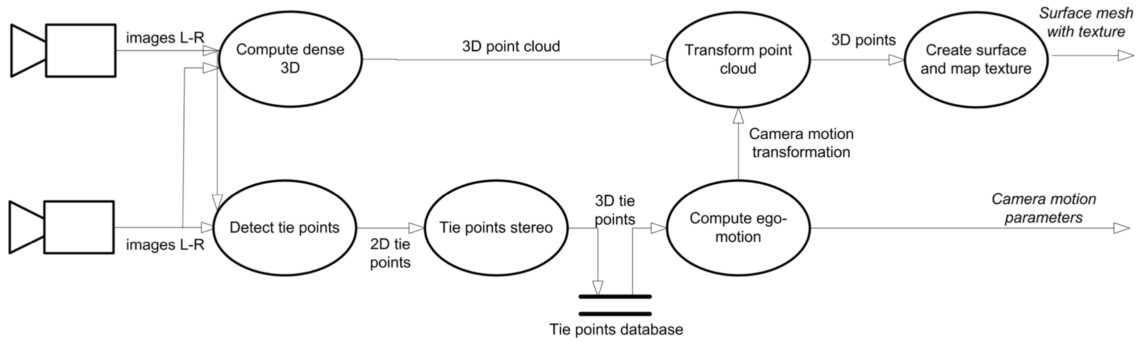

Figure 6 iSM system architecture

3D data is computed in the camera reference frame and is transformed using the camera pose estimated for this frame. Typically, the initial camera pose is used as the reference. Using all 3D points obtained from the stereo processing is not efficient as there are a lot of redundant measurements, and the data may contain noise and missing regions. Therefore representing 3D data as a triangular mesh reduces the amount of data when multiple sets of 3D points are combined. The photo-realistic appearance of the reconstructed scene is created by mapping camera images as texture. Such surfaces are more visually appealing and easier to interpret as they provide additional surface details. Colour images from the stereo camera are used for texture mapping.

B. Space/terrestrial applications

iSM technologies are currently used in various space and terrestrial applications, such as planetary rovers, mining, forensics, and security. For planetary exploration, creation of photo-realistic 3D models of unknown environment is used for obstacle avoidance and path planning, as well as for mission planning on Earth. In underground mining, photo-realistic 3D models are currently used for both survey and geology. The 3D model allows the mine map to be updated daily to minimize any deviation from the plan, and allow the mining companies to monitor how much ore is taken at each blast. Grade assessment is performed on the 3D model so that the blasting can follow the ore body better. Past and recent models are linked together and used for planning for exploration and tunnel advancement.

iSM can create a 3D model of the crime scene quickly without much disturbance to the crime scene and measurements can be performed on the 3D model afterwards [9,10]. The models can help with the police investigation, and also help officers who have not been to the crime scene to understand the crime scene better. Equipped with iSM, Unmanned Ground Vehicles (UGVs) can create 3D models to send back to the base station, while they explore unknown environments [11]. Rapid access to 3D models will increase the operator situational awareness and allow better mission planning and execution.

C. Underwater results

Underwater stereo image sequences from Folkestone Marine Reserve, Barbados have been processed with iSM, to reconstruct a coral bed and a sunken barge. Fig. 7 shows images selected from the underwater sequence; screenshots of the reconstructed 3D model are shown in Fig. 8. The camera trajectory recovered by iSM is overlaid on the 3D model in Fig. 9. It can be seen that the trajectory and the model are locally quite consistent and represent 3D shape correctly.


Figure 7 Selected images from an underwater stereo sequence

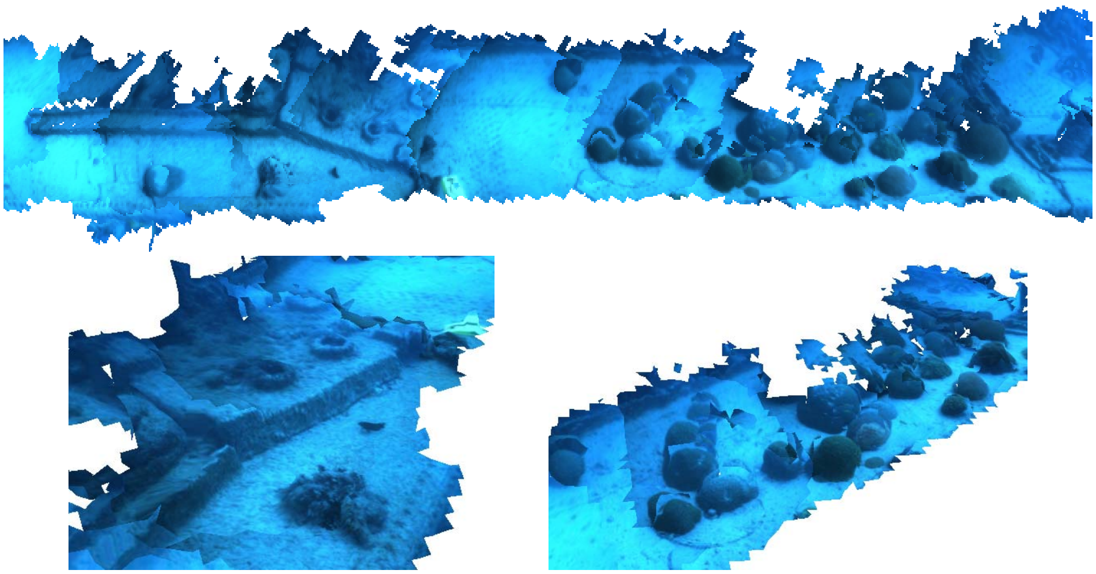

Figure 8 Screenshots of the reconstructed 3D model

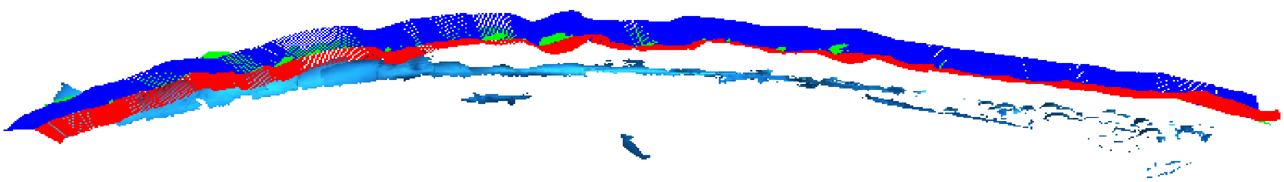

Figure 9 Sideview of the 3D model showing the recovered camera trajectory

However, the side view of the reconstructed path and the model shows a slight bend introduced by imprecise camera calibration underwater and accumulation of camera pose error in the reconstruction, see Fig. 9. The solutions are sought by development of special underwater camera calibration procedures and. incorporating data from external sensors such as gyros [12], inertial measurements units or inclinometers in the camera path estimation. We are already using such sensors and approaches in our space and terrestrial projects.

V. CONCLUSIONS

This paper describes current research and development at MDA on two underwater applications: 3D mapping of sea floor and underwater worksites, and vision guided station-keeping for ROVs. Both of these applications rely on sophisticated computer vision technologies originally developed for space applications: terrain mapping using mobile cameras, and object recognition / pose estimation. Presented results of initial underwater tests of are very promising. Current work focuses on specific issues introduced by the underwater environment: accurate camera calibration for underwater, characterization of the systems' performance (accuracy, reliability and robustness) in representative conditions underwater, packaging the systems and integration with ROVs (hardware, software and operation concepts).

The barge and coral reef stereo image sequences have been kindly provided by the AQUA project team [12].

VI. REFERENCES

[1] Whitcomb L: Underwater Robotics: Out of the Research Lab and out into the Field. IEEE Conference on Robotics and Automation 2000

[2] Plotnik A and Rock S, Visual Servoing of an ROV for Servicing of Tethered Ocean Moorings, Proceedings of the MTS/IEEE OCEANS Conference, Boston, September 2006

[3] Steve Murray and Doug Murphy Underwater Telerobotics and Virtual Reality: A New Technology Partnership. Report, Naval Command Control and Ocean Surveillance, 1986

[4] Bosch, A Op den, Santamaria, J C: Monitoring underwater jobs using virtual environments. Sea Technology, Apr 2002

[5] M. Bondy, R. Krishnasamy, D. Crymble, and P. Jasiobedzki: Space Vision Marker System (SVMS). AIAA SPACE 2007 Conference & Exposition, Long Beach, September 2005

[6] D.G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60(2):91-110, 2004.

[7] S. Se and P. Jasiobedzki, Photo-realistic 3D model reconstruction, IEEE International Conference on Robotics and Automation (ICRA), pages 3076-3082, Orlando, Florida, May 2006.

[8] S. Se, D. Lowe, and J. Little, Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks, International Journal of Robotics Research, vol. 21, no. 8, pages 735–758, August 2002.

[9] S. Se and P. Jasiobedzki, Instant scene modeler for crime scene reconstruction, IEEE Workshop on Advanced 3D Imaging for Safety and Security (A3DISS), San Diego, USA, June 2005.

[10] P. Jasiobedzki, H.-K. Ng, M. Bondy, and C. McDiarmid. Three-dimensional modeling of environments contaminated with chemical, biological radiological and nuclear agents. Proc. SPIE Technologies for Homeland Security and Homeland Defense VII. Proceedings of SPIE, vol. 6943, pages 17–20, 2008.

[11] S. Se and P. Jasiobedzki, Stereo-Vision Based 3D Modeling for Unmanned Ground Vehicles, SPIE Vol.6561: Unmanned Systems Technology IX, Orlando, Florida, USA, April, 2007.

[12] Hogue, A., German, A. and Jenkin, M., Underwater environment reconstruction using stereo and inertial data. Proc. IEEE SMC, Montreal, 2007