Swimming with Robots: Human Robot Communication at Depth

Bart Verzijlenberg and Michael Jenkin
York University, Department of Computer Science and Engineering
4700 Keele St., Toronto, ON, Canada, M3J 1P3
{bartv, jenkin}@cse.yorku.ca

Abstract— **Human-robot communication is a complex problem even in the terrestrial domain. Failure to properly communicate instructions to a robot and receive appropriate feedback can at the very least hamper the ability of the robot to perform its task, and at worst prevent the task from being completed. The problem of providing effective communication between a robot and its operator becomes even more complex underwater. Many communication channels available in the terrestrial domain become unavailable, and communication between team members and task oversight become even more complex. This paper describes initial experiments with the AQUATablet – a robot interaction device designed to be operated by a diver tethered to, or in visual communication with, an underwater robot. The basic requirements of the device are described along with design considerations and results of initial experiments with the device conducted in the pool and in the open ocean.**

I. THE AQUA PROJECT

Except for very simple autonomous systems it is essential that effective communications strategies be developed to enable an operator to communicate with the autonomous device. It is critical to be able to communicate instructions to the device in a clear and effective manner, and it is important to be able to receive easily understood responses from the device. Although important in many domains, the need for effective human-robot communications is particularly important for devices that operate in complex and remote environments, such as underwater and outer space, where human-robot communication is especially difficult.

To illustrate some of the difficulties associated with human-robot communication in harsh environments, consider the problem of communicating with a device such as the AQUA robot[1] shown in Figure 1. Unlike traditional underwater vehicles that rely on thrusters and control surfaces, the AQUA robot is a hexapod robot that utilizes flippers or fins for propulsion. The vehicle relies on internal power and computation and can be accessed via an optical tether in order to communicate with an external operator. The sensing needs of the AQUA robot are met through visual-inertial sensors [11], and a range of sensors that monitor aspects of the vehicle's state, such as 3D orientation and (underwater) operating depth. The robot can operate to over 100' depth and can operate for over two hours on its internal batteries.

Unlike typical, large, underwater robots, the AQUA robot's small size makes it well suited for field deployment, as it can simply be picked up and tossed into the water. The robot is designed to collect data from a target site such as a coral reef, lobster traps, or a ship's hull. The vehicle's on-board stereo cameras and inertial sensors allow visual imagery collected from the robot to be combined into large-scale 3D models, and repeated data collections can be used, for example, to study the long term development of coral reefs. These large-scale 3D models may also be used in situations where human access is limited or potentially dangerous, such as wreck penetration. Although AQUA can be given instructions prior to deployment, additional/updated instructions or sensor feedback may be desired when performing these and similar tasks.

This paper describes the AQUATablet, a device developed in order to provide effective communication between the robot and an operator either operating at depth or on the surface. The AQUATablet is, in essence, a submersible computer that enables communication between a human operator and the robot via either an optical tether or a suite of visual targets.

II. EXISTING UNDERWATER COMMUNICATION TECHNOLOGIES

A variety of technologies have been developed for underwater communication. The primary limitation in many underwater communication strategies is the interference of water on the signal used for transmission. Many signals are quickly absorbed by water: for example, as one descends in the water, colours with longer wavelengths are absorbed more quickly than colours with shorter wavelengths. For visual strategies (i.e., divers using hand signals), the clarity of water can also be a concern. Turbidity, or 'aquatic snow' can reduce visibility in the water column to only a few feet, or even less.

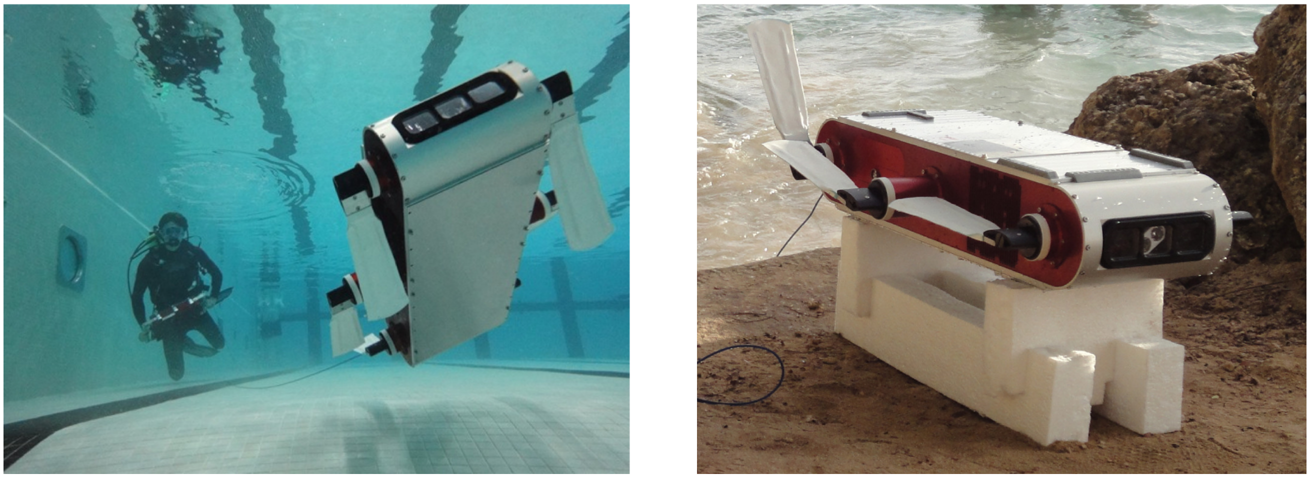
In this section we outline some existing technologies for diver-robot communication, including tethered, wireless, and visual strategies.

A. Surface tethered control

A very common approach to the control of submerged devices is via a simple communications tether (see [16], [14], [3], and Figure 1(b) for examples). While this can provide for excellent communication (as long as the tether remains intact), it also presents several problems. Perhaps

---

[1] The robot used for this paper is actually the KROY version of the AQUA family of robots [9], [5].



(a) The AQUA robot during pool trials. The AQUAtablet is visible in the background, being used for control of the robot.

(b) The AQUA robot ready for deployment during field trials.

Fig. 1. The AQUA robot is a 6 legged amphibious robot, utilizing fins for underwater movement, as shown in (a), which facilitates movement in any direction except sideways. The robot can be tethered using a fiber-optic cable attached at the rear of the robot.

the most critical of these is the inability of the operator to communicate directly with any robot handlers or other users present in the water, either at the surface or at depth. In the field this is complicated by the distance between the operator station (on a beach or boat) and robot handlers. Larger systems may require the operator to be positioned within a separate room onboard ship at considerable distance from operations. The resulting communications delay and lack of situational awareness can be hazardous; for this reason, the Canadian military never deploys (surface controlled) robots and divers simultaneously.

The cable itself also presents a number of problems. When using small robots, such as AQUA, the tether can be fragile and require care in handling: as such, AQUA typically needs two to four cable wranglers when deployed off the beach or from a boat. When working in the field, the operator's controlling computer, being in close proximity to water, may be unintentionally exposed to environmental contaminants such as water from ocean spray or rain. Reducing this risk requires that the operator be placed at a significant distance from operation. This implies longer cables between the robot and semi-dry operator locations, increasing cable management issues and handler communication concerns. The problems with tethered surface operations become even more acute at depth. Once submerged, an accompanying diver may be required to relay commands to the surface operator in control of the robot. Divers are limited in their ability to assess the state of the robot, relying instead on confirmation from the operator, cable wranglers and the limited communication abilities of the robot.

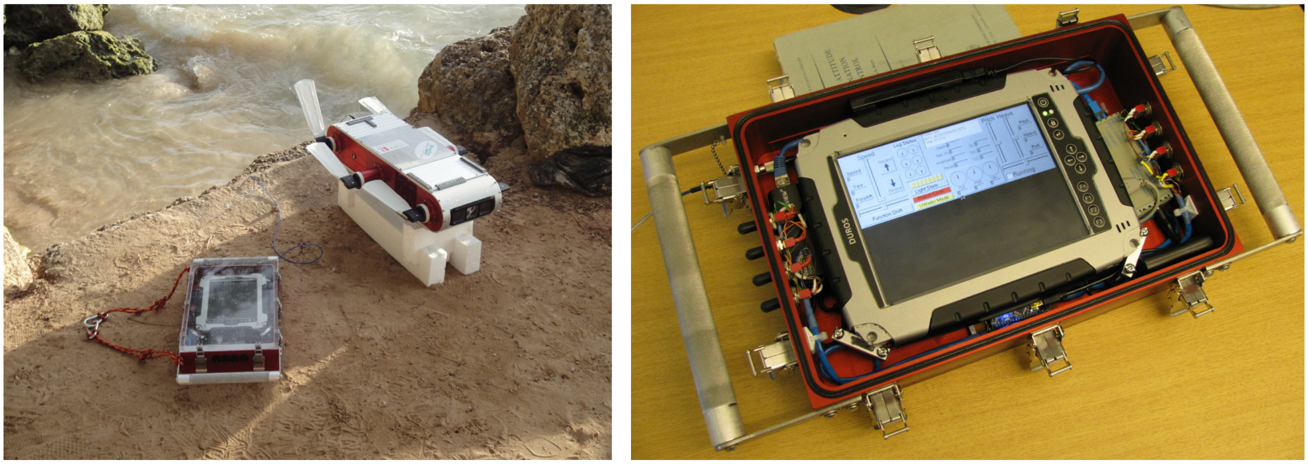
An alternative to tethered operation is to untether the vehicle and utilize some type of wireless communication between the operator and the vehicle. This is discussed in the next section.

B. *Wireless*

While standard wireless systems are ineffective underwater due to the interference of water, low frequency modems and light modems, able to work underwater, have been developed. Low bandwidth ($\sim$19kb/s) acoustic modems are available for underwater applications [10], and high bandwidth ($\sim$20mb/s) optical modems, such as [4] and [2], have been shown to be effective for short range communication with underwater robots. However, use of either of these relatively large devices on small robots is problematic: mounted externally they impact the dynamics of the robot, while internally there is typically very little space. For example, the acoustic UWM1000 modem [10] has a data rate of 19,200 baud, measures 24cm long x 12cm diameter and weights 4.2kg, while the 1013C1 high-bandwidth underwater optical transceiver [2] has a data rate of 10Mbps, measures 27cm long x 10cm diameter and weights 2.7kg. Either of these systems would account for roughly 1/4 of AQUA's size and weight (16kg). Direct electrical communication in saline environments is also possible, although clearly not applicable in fresh water: [13] has shown 1Mbps transmission over short range (1-2m) using 6V signals. However, this has the additional risk of introducing electricity into the environment, which may not be desirable in sensitive aquatic environments such as near coral reefs.

C. *Visual communication*

Many robots are equipped with visual sensors, which can be used for direct communication. [6] has shown the effectiveness of using visual markers (2D bar codes, based on ARTag [7] markers) to communicate commands to an underwater robot. Since such tags carry limited data (for example, an ARTag is limited to 10 bits per tag), [6] makes use of individual tags as tokens in a BNF grammar. While this has been shown effective in communicating basic commands to the robot, this method is both slow and does



(a) Robot and tablet prior to deployment

(b) Closeup of underwater tablet

Fig. 2. (a) AQUA and tablet ready for deployment at Bellairs Research Institute, (b) The tablet housing, showing switches (left and right), the fiber optic converter (far right) and the tablet PC.

not provide for a sophisticated response opportunity from the robot. [6] also explores the natural use of hand signals, already used for diver↔diver communication, for diver→robot communication. [20] makes use of visual servoing to control the pose of the robot using a yellow ball as a diver held target, while [18] has developed a diver tracking system using an onboard camera to track the up-down gait of a diver's fins.

D. *Systems facilitating robot response*

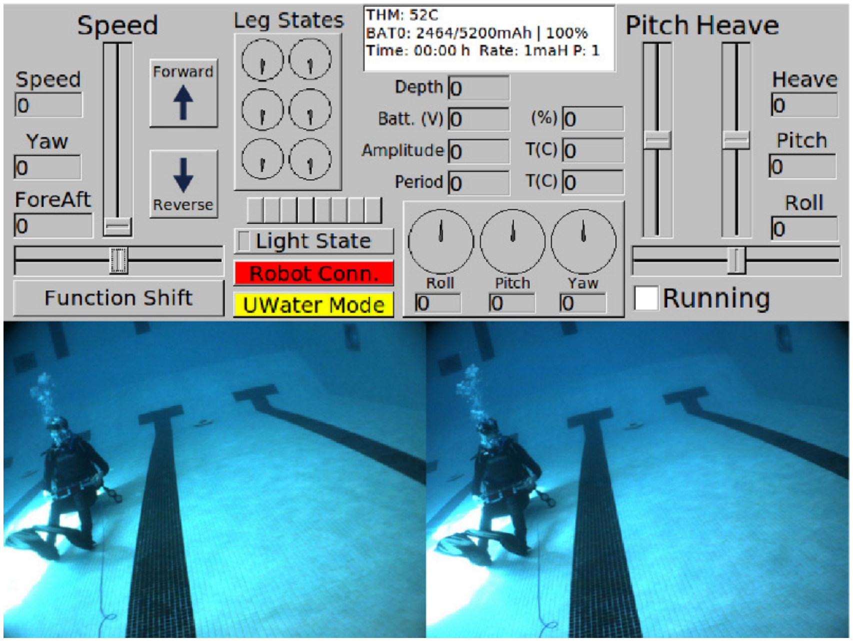
As discussed earlier, there are a wide variety of diver→robot communication technologies, utilizing both standard digital communication as well as more human oriented visual cues. While various, but often limited, technologies exist to allow terrestrial robots to communicate with human handlers, such as verbal communication, hand gestures etc., the ability for an underwater robot to respond to a diver is even more limited. An early example of robot→diver communication is seen in the TwinBurger [8] underwater robot, which makes use of a panel containing five large LED panels (bits) for robot→handler communication. Four of the bits are used to display a coded signal, while the fifth is used as a status indicator. The use of individual status lights is a common strategy for robot→human communication. For example, the AQUA robot acknowledges instructions by flashing its onboard light and the RWI-B12 used Morse code to communicate error states through a single onboard LED. Other options are also possible; for example, AQUA's error states may be communicated by placing its fins in predetermined configurations. Although these communications strategies work, after a fashion, they are often difficult to interpret, especially for a novice user or one not fluent in Morse code.

A sufficiently large robot can incorporate a dedicated communications device for robot→human communications, such as a computer screen mounted within the robot frame, common in museum robots such as [21]. While this facilitates easy access to various information, it requires close proximity to the robot, which may be undesirable underwater. The use of a screen is in essence an extension of the TwinBurger approach, with a (vast) increase in pixels. Essentially, operator control units function in a similar manner, moving the screen location away from the robot.

III. UNDERWATER CONTROL UNIT

Even though surface operators have direct control of the robot, communication between support divers in the water, the operator, and the robot is complex and difficult. It is hard to overestimate the difficulty in having an operator try to catch the attention of a diver at depth or for a diver at depth to signal the surface operator that the vehicle is operating improperly. Eliminating the surface operator would significantly reduce this problem. Unfortunately, operator control units (OCU) are generally meant to be used above water, able to handle, at best, brief rinses in water. In order to address this shortcoming we have developed an OCU capable of being operated at depth. Essentially a watertight enclosure for an off-the-shelf tablet PC, the housing allows the OCU to be brought underwater with the diver. This provides the diver with a direct link to the robot while in close proximity to it; actions taken by the robot, along with robot status, remote video and experiment progress are directly accessible by the diver. The close proximity between the operator and the robot also enables straightforward tele-operation, including the use of orientation sensors integrated in the housing; a mode where the diver need only use simple "wii-mote"-like motions to control the robot's position and orientation is available. Alternatively, auto-pilot modes can be activated, allowing the robot to take a more active role in maintaining its pose.

The underwater control unit (UCU) is shown in Figure 2(b); it consists of a watertight housing that has been tested to 60'. The housing protects a tablet PC, ether-



```
THM: 52C
BAT0: 2464/5200mAh | 100%
Time: 00:00 h  Rate: 1maH P: 1
```

Speed 0
Yaw 0
ForeAft 0
Depth 0
Batt. (V) 0 (%) 0
Amplitude 0 T(C) 0
Period 0 T(C) 0
Roll 0
Pitch 0
Yaw 0
Heave 0
Pitch 0
Roll 0

Fig. 3. Operator GUI for underwater mode, operating in a pool. The upper section of the display shows switch functions as well as robot, tablet and link status. The lower half is used to display the stereo image data from the robot, in which the operator is visible. With the ability to view a live video feed, the accompanying diver can aid the robot in determining what should be done next, without requiring the robot to return to the surface. Unlike terrestrial robots which are limited to a single plane, a wrong decision by the robot can easily result in the loss of the device. Besides operating in an effectively unbounded 3D environment, the robot can easily exceed diver depth restrictions, out-swim a diver into the open ocean or enter parts of overhead environments not safely accessible by the diver. Many of these actions could prevent even the recovery of robot hardware, emphasizing the need for effective communication between diver and robot.

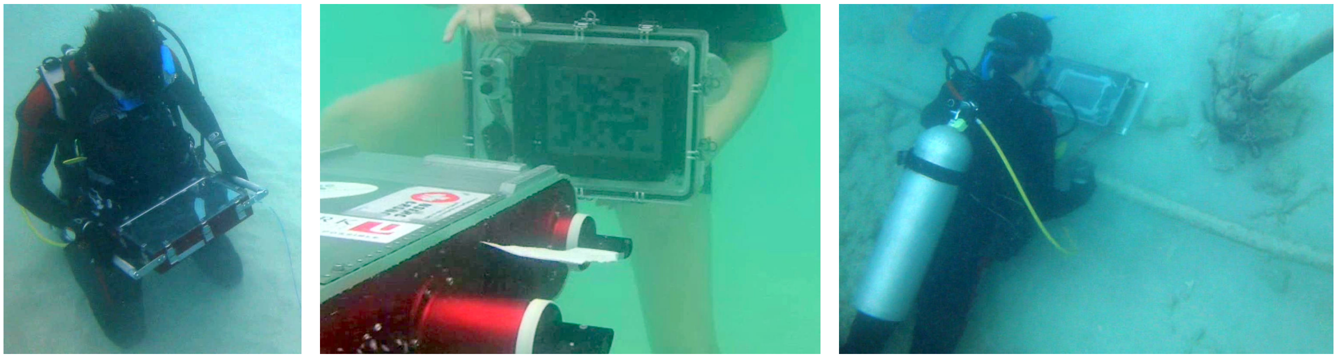
optical fiber converter, an IMU and electronics to interface with a number of waterproof switches. The operator can view the tablet display and interact with it using eight momentary double-throw switches mounted on the housing. Alternatively, the onboard IMU allows the tablet to be used as a giant joystick. An optical connector penetrates the tablet housing providing a waterproof connection for the exterior fiber-cable.

The tablet can be used to communicate with the robot in a variety of different ways. The first of these is as an underwater version of the OCU. As the AQUATablet can be operated at depth this allows a tethered operator to interact directly with the robot. This enables a number of different operational modes not possible with a ship-based operator. For example, a diver operating at a relatively safe depth (say 80’) can teleoperate a vehicle operating 30’-40’ deeper, without exposing the diver to the increased dangers associated with the lower dive profile. An underwater tether can also be used to enable a diver to remain outside of potentially dangerous environments while the robot operates within them. For example, a robot could be sent to investigate the inside of a wreck, while allowing the diver to remain outside.

The OCU also operates in a tether-less mode in which the robot and the OCU communicate using visual communication strategies. The system operates by having the operator command the OCU to display fiducial markers on the tablet screen and then showing these to the robot. The robot in turn responds using it’s internal light as a low bandwidth modem, which the tablet detects and interprets. This method is discussed in further detail in section IV.

A. *Hardware*

The availability of submersible computer systems is extremely limited. Hardened tablets, typically able to be submersed to 1m for 30 minutes, are available but are only intended to simply survive accidental immersion. Existing solutions such as the SeaSlate [17] and SharkMarine’s Navigator [15] are able to go to greater depth (30m), but the SeaSlate is a research unit and is not being manufactured while the Navigator is cost prohibitive. The development of the AQUATablet was in part motivated by this lack of viable off the shelf units. The enclosure described in this paper is milled out of solid aluminum and contains off-the-shelf hardware which allows for easy addition of sensors and hardware upgrades. This also reduces the cost associated with parts replacement should water enter the housing. The housing is designed to contain a tablet PC; it is currently outfitted with a 1.2Ghz processor, 1Gb RAM and a 60Gb hard drive. In addition to acting as a control unit the OCU can act as secondary storage for off-robot data-collection. The use of the housing in warm environments and the added lack of internal air circulation necessitates alternative cooling. An earlier prototype of the housing built from plastic, see Figure 4(b), had heat dissipation issues, which could lead to heat-related hardware failures during operation.



(a) Diver working with the AQUATablet (b) Custom tag being shown to the robot using a tablet in the original housing (c) Aquatablet in use

Fig. 4. The underwater housing in operation in (a,c) tethered mode, or (b) untethered visual communication mode.

The current housing has a raised platform for the tablet PC, providing a large heat-sink directly into the ocean, effectively managing heat dissipation for components within the housing. Below the tablet, and around the heat-sink, there is space for interface electronics and sensors. A fiber-optic converter provides a gigabit link between the tablet and the robot platform.

The internal components are visible through a clear acrylic cover. During surface operation (such as inspecting the underside of a ship), this cover may be removed and normal tablet interactions can be used for control. When operated near the water or submerged, the cover is sealed and the tablet is controlled using two rows of four double throw switches (located on either side by the handles), allowing easy access to robot functions through on-screen prompts. These switches are interfaced through an Arduino Nano [1] micro-controller board. To simplify control of the robot, the housing is outfitted with an orientation sensor which allows the tablet to be used as a haptic device, similar to Wii-mote functionality. When the robot is operated in this manner, tilting the tablet left-right or up-down instructs the robot to alter it's orientation to match the tablet. This makes visual remote control, such as when the robot is swimming some distance below the diver, particularly straightforward.

The tablet runs a standard Linux OS, supporting the use of existing experiment software. For example, the control software uses the standard RHEX communication libraries to maintain connectivity with the robot. The GUI provides labels on each side of the screen to guide the diver which switch to use for any needed function. The tablet screen is used to both prompt the diver for interaction and to display telemetry such as stereo video feeds from the robot (see Figure 3). Robot operating parameters can be accessed from a settings menu, where data collection and other functionality can also be (de)activated.

IV. DIGITAL TAG LIBRARY

When operating in a tether-less mode, the AQUATablet is used as a dynamic generator of visual tags that are communicated to the robot. Following the approach described in [19], a grammar based on ARTag primitives [7] is used to communicate with the vehicle. Fundamental to this communication strategy is the use of ARTags to define tokens in the communications language. The token library for the language described in [19] has grown to a 3" stack of markers, which must be sorted through while underwater to access the correct tag, and without accidentally showing the robot an incorrect one. Through the use of a tablet computer, these markers are available digitally and instead of searching through many tags, the tablet can be used to generate and display 2D tokens with considerably more flexibility. The operator can choose a tag using software operating onboard the tablet and then turn the tablet over and show the digital tag to the robot. Moreover, the tags can be generated on the fly, and the number of bits per token varied depending on the clarity of the water. Using this approach, the diver can select a sequence of commands/tokens and when satisfied with the selected commands, generate appropriate tags, compressing the commands into a compact sequence of tags which can be communicated quickly. Extending this work, [12] uses the tags as a low bandwidth communication channel, where the bit stream is broken into packets (instead of pre-defined tokens), and encoded into successive tags.

The robot, having observed a tag, needs to signal the diver's computer to proceed to the next packet using a coded light stream. To facilitate this, the robot's onboard light can be flashed, which the tablet is able to detect. A basic optical modem may be constructed from existing robot hardware in this manner, where the robot's light is flashed many times a second (instead of flashing the light at operator detectable speeds), creating a low bandwidth upload channel. This channel can be used to acknowledge tag packets and upload telemetry, potentially including (in future work) efficient transfer of images.

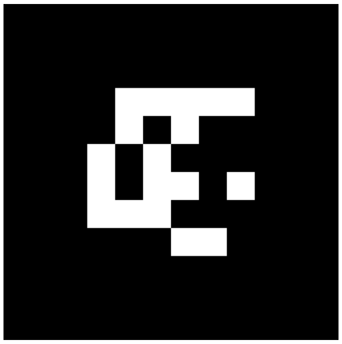


Fig. 5. A fiducial tag [7] using 36 bits to redundantly store 10 bits; the tag shown encodes the number 33. ARTag applies a CRC-16 convolution to the binary string to be encoded and attaches Forward Error Correction (FEC) data, resulting in a 36 bit code. These bits are arranged in a 2D, 6x6 grid-pattern, which can be detected in video streams. For each quad detected in a video frame, ARTag attempts to decode it; each of the 4 possible orientations of the tag is attempted individually. When decoding, FEC is used to detect and correct corrupt bits and the marker's correctness tested using CRC. If the CRC test is passed, this then yields the original 10 bit binary string.

V. FUTURE WORK

We are working to incorporate auto-pilot modes for the AQUA robot, such that the robot can swim with less guidance from the diver. Wreck penetration tests are also planned. Work is underway to use an eigenspace decomposition of images for compact image transmission over alternative, non-tethered, low-bandwidth communication channels. Extension of the digital tag library to enable fast streaming of data is also envisioned.

VI. ACKNOWLEDGMENTS

The technical assistance of Jeff Laurence and Jim Zacher is gratefully acknowledged. This project would not have been possible without the generous financial support of NSERC Canada.

REFERENCES

[1] Arduinno Nano Microcontroller board. www.arduino.cc.

[2] Ambalux. 1013C1 High-Bandwidth Underwater Transceiver. www.ambalux.com.

[3] T. Aoki, T. Murashima, Y. Asao, T. Nakae, and M. Yamaguchi. Development of high-speed data transmission equipment for the full-depth remotely operated vehicle-KAIKO. In *OCEANS '97. MTS/IEEE Conference Proceedings*, volume 1, pages 87–92 vol.1, October 1997.

[4] G. Baiden, Y. Bissiri, and A. Masoti. Paving the way for a future underwater omni-directional wireless optical communication systems. *Ocean Engineering*, 36(9-10):633–640, 2009.

[5] G. Dudek, M. Jenkin, C. Prahacs, A. Hogue, J. Sattar, A. German, H. Liu, S. Saunderson, A. Ripsman, S. Simhon, L. Torres, E. Milios, P. Zhang, and I. Rekletis. A visually guided swimming robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1749–1754, 2005.

[6] G. Dudek, J.Sattar, and A. Xu. A visual language for robot control and programming: A human-interface study. In *IEEE International Conference on Robotics and Automation*, pages 2507–2513, Rome, Italy, 2007.

[7] M. Fiala. ARTag, a fiducial marker system using digital techniques. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*, pages 590–596, Washington, DC, 2005.

[8] T. Fujii and T. Ura. Development of an autonomous underwater robot twin-burger for testing intelligent behaviors in realistic environments. In *Autonomous Robots*, volume 3, pages 285–296, 06 1996.

[9] A. German, A. Hogue, H. Liu, C. Prahacs, A. Ripsman, R. Sim, L. Torres, P. Zhang, M. Buehler, G. Dudek, M. Jenkin, C. Georgiades, and E. Milios. Aqua: An aquatic walking robot. In *Proceedings of the IEEEE/RSJ/GI International Conference on Intelligent Robots and Systems (IROS)*, pages 3525–3531. IEEE Press, 2004.

[10] LinkQuest Inc. Underwater acoustic modem.

[11] M. Jenkin, A. Hogue, A. German, S. Gill, A. Topol, and S. Wilson. Modeling underwater structures. *International Journal of Cognitive Informatics and Natural Intelligence*, 2(4):1–14, 2008.

[12] M. Jenkin and B. Verzijlenberg. Visual underwater communication. In *Intelligent Systems Poster Session*, 2008.

[13] J. Joe and S.H. Toh. Digital underwater communication using electric current method. In *OCEANS 2007 - Europe*, pages 1–4, June 2007.

[14] P. Lee, B. Jeon, S. Hong, Y. Lim, C. Lee, J. Park, and C. Lee. System design of an rov with manipulators and adaptive control of it. In *Proceedings of the 2000 International Symposium on Underwater Technology*, pages 431–436, 2000.

[15] Shark Marine Technologies. Navigator. http://www.sharkmarine.com/navigator.htm.

[16] M. Nokin. Rov 6000-objectives and description. In *OCEANS '94*, volume 2, pages 505–509, September 1994.

[17] Australian Institute of Marine Science. Seaslate. wetpc.com.au/html/technology/underwater.htm.

[18] J. Sattar and G. Dudek. Underwater human-robot interaction via biological motion identification. In *Proceedings of Robotics: Science and Systems*, Seattle, USA, 2009.

[19] J. Sattar, P. Giguère, and G. Dudek. Sensor-based behavior control for an autonomous underwater vehicle. *International Journal of Robotics Research*, 28(6):701–713, 2009.

[20] J. Sattar, P. Giguere, G. Dudek, and C. Prahacs. A visual servoing system for an aquatic swimming robot. In *in IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2071–2077, 2005.

[21] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A.B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva, 2000.