AQUA: An Amphibious Autonomous Robot

Gregory Dudek, Philippe Giguere, Chris Prahacs, Shane Saunderson, Junaed Sattar, and Luz-Abril Torres-Mendez, McGill University
Michael Jenkin, Andrew German, Andrew Hogue, Arlene Ripsman, and Jim Zacher, York University
Evangelos Milios, Hui Liu, and Pifu Zhang, Dalhousie University
Martin Buehler, Boston Dynamics
Christina Georgiades, MathWorks

AQUA, an amphibious robot that swims via the motion of its legs rather than using thrusters and control surfaces for propulsion, can walk along the shore, swim along the surface in open water, or walk on the bottom of the ocean. The vehicle uses a variety of sensors to estimate its position with respect to local visual features and provide a global frame of reference.

The aquatic environment is almost ideal for autonomous robot development. First, it provides a range of real tasks for autonomous systems to perform, including ongoing inspection of reef damage and renewal, tasks in the oil and gas industry, and aquaculture. Second, operating in the water requires robust solutions to mobility, sensing, navigation, and communication.

A common theme of many industrial aquatic tasks is *site acquisition and scene reinspection* (SASR). Figure 1 shows AQUA performing a typical SASR task, in which it walks out into the water under operator control and is directed to a particular location where it will make sensor measurements. Once near the site, the robot achieves an appropriate pose from which to undertake extensive sensor readings. After making the measurements, the robot returns home autonomously. Later, the robot autonomously returns to the site to collect additional data.

The SASR task requires solving multiple scientific and engineering problems including pose estimation in an unstructured environment, underwater landmark recognition, robotic navigation, motion control, path planning, vehicle design, environment modeling and scene reconstruction, 3D environment exploration, and autonomous and teleoperated control of a robotic vehicle.

Performing a SASR task is a formidable challenge for terrestrial vehicles. It is even more complex in an aquatic environment. In addition to the increased degrees of freedom (DOF) associated with performing a task under water, working in this domain introduces complications such as station keeping, sensing, and the differences involved in mobility in open water versus shallow water or motion along the surface.

THE VEHICLE

A biologically inspired robot capable of both legged and swimming motions,[1,2] AQUA is based on RHex, a terrestrial six-legged robot developed between 1999 and 2003, in part by the Ambulatory Robotics Lab at McGill University in collaboration with the University of Michigan, the University of California at Berkeley, and Carnegie Mellon University.[3-4] In addition to surface and

Published by the IEEE Computer Society

underwater swimming, AQUA's capabilities include diving to a depth of 30 meters, swimming at up to 1.0 m/s, station keeping, and crawling on the bottom of the sea.

Unlike most underwater vehicles, AQUA does not use thrusters for propulsion; instead, it uses six paddles, which act as control surfaces during swimming and as legs while walking. The paddle configuration gives the robot direct control over five of the 6 DOF that it has: surge (back and forth), heave (up and down), pitch, roll, and yaw. Like a bicycle or an automobile, it lacks the capacity for lateral (side to side or sway) displacement. Its operators use an onboard inclinometer and a compass to control the robot's motion underwater.

The robot is approximately 65 cm long, 45 cm wide (at the fins), and 13 cm high. It has an aluminum waterproof shell and displaces about 16 kg of water. Onboard batteries provide more than three hours of continuous operation. Optionally, a fiber-optic tether can bring signals from cameras mounted within the AQUA vehicle itself, from the sensor systems mounted on the robot, and from the command and control output to a surface-based operator.

Within the robot, two PC/104 stacks support local control, communication, and sensing. One stack runs the QNX real-time operating system and is responsible for real-time control of the vehicle actuators. The second PC/104 runs non-real-time Linux and provides communication and sensing for the vehicle. Each of the robot's fins is controlled by a single degree-of-freedom revolute joint. The onboard computer provides real-time control of the six legs. The legs are compliant, and the spring energy stored in the legs as they bend under load is an integral part of the vehicle's locomotion strategy.

Net vehicle motion is effected through the application of one of several precomputed gaits. Researchers have developed terrestrial walking, surface swimming, and free water swimming gaits for the vehicle. AQUA's unique locomotion strategy provides great flexibility in terms of potential locomotion modes. The walking gait is a basic hexapod motion. The robot uses a rich class of alternative gaits and behaviors to swim in open water with its six 1-DOF actuators (although there is often coupling) performing controlled five degree-of-freedom motion.

Although AQUA is capable of complex gaits, monitoring complex six-degree-of-freedom trajectories externally can be challenging, so locomotion is usually accomplished by selecting from one of a small number of gaits that permit control of vehicle roll, pitch, yaw, surge, or heave. These behaviors are easy to control and monitor when operating the vehicle in a teleoperational fashion, and they also are the foundation of servo-controlled vehicle motion. Various hydrodynamic vehicle simulators have been developed to aid in tasks as varied as teleoperation rehearsal, leg design and

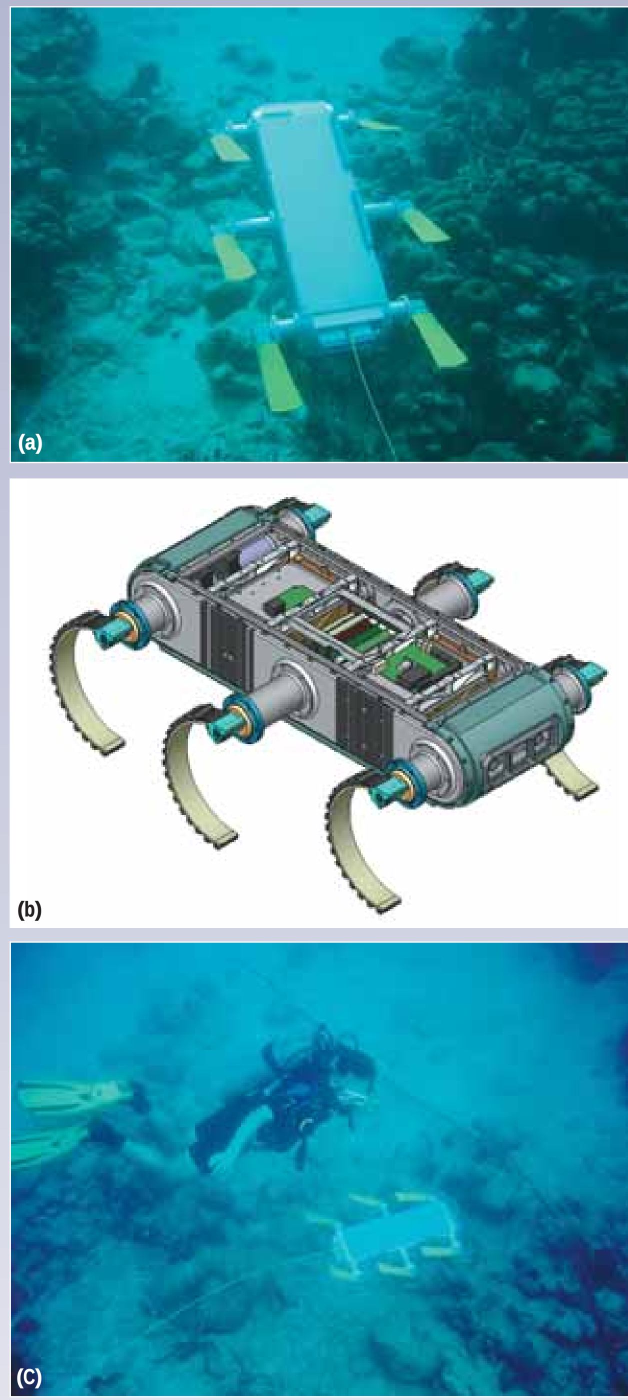

Figure 1. AQUA performing a SASR task. (a) The robot swimming over a coral reef while tethered to an external operator. The vehicle has six fins that can be controlled independently. (b) Arrangement of internal components. The robot has treaded legs for use while walking on the shore or on the bottom of the ocean. (c) AQUA with a diver.

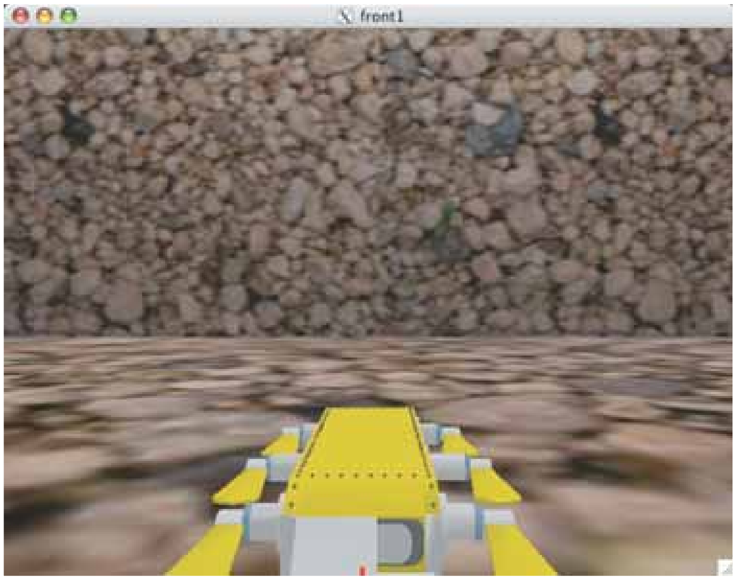

Figure 2. A "co-pilot" view from one of the AQUA simulators. Researchers can use these simulators for task rehearsal as well as hydrodynamic simulation of gait and fin design.

evaluation, and novel gait synthesis. Figure 2 shows an immersive virtual reality robot simulator that can be used for teleoperational task rehearsal.

VISUAL BEHAVIOR CONTROL

One ongoing need for the robot is to estimate its current environmental state. For an amphibious robot like AQUA, this includes having knowledge of whether it is in open water, on the sea bottom, in the surf, or on land. This is particularly difficult in the surf since turbulence, spray, and other artifacts make visual and acoustic sensing difficult. Moreover, using visual or acoustic sensing is computation-intensive, straining the robot's energy budget.

One approach to state estimation uses feedback from the robot's effectors—that is, the legs or flippers. Just as biological organisms can use contact forces to moderate their gait, AQUA can exploit contact forces to estimate surface conditions and incrementally tune its current gait or qualitatively change its behavior.

While walking, the need to make constant adaptive leg placements is, to a large extent, obviated by relying on compliance of the robot's legs. Prior work on the leg dynamics of the RHex vehicle family developed passive adaptability to ground conditions, letting the legs act somewhat like shock absorbers.[3-4] The particular form of this adaptation was strongly motivated by the biological observations of Robert Full,[5] who obtained measurements from cockroaches and made similar morphologies to the RHex and AQUA robots. While this compliance reduces the need for adaptive gait planning to maintain stability on the ground, surface estimation is still important for many other reasons including selecting the optimal gait for speed (as opposed to stability), position estimation, mapping, or behavior selection. A particularly interesting behavior change is the transition from walking to swimming as the robot enters the water.

Our current work estimates environmental properties by measuring drive currents to the robot's six legs as a function of their orientation. We used a statistical classifier to model the difference between the "feeling" of sand, carpet, ice, water, and other terrain types, and we have applied this information to model terrain recognition with accuracies of greater than 80 percent over a single leg cycle, and with higher accuracy if we combine multiple measurements over time.[6]

Another issue with AQUA's gait relates to strategies used as the vehicle transitions from one gait to another. Most terrestrial legged robots achieve gait transition by changing leg motion parameters during the flight phase of the gait where the change has limited indirect effects on the device's trajectory. Due to the constant contact they have with the fluid that surrounds them, underwater robots do not have a flight phase in their gait. This means that there is no way to reposition a leg without applying unwanted forces to the robot. This unwanted motion can be a problem for tasks such as visual servoing, where an unexpected shift in trajectory could cause the vehicle to loose track of its target.

Ongoing work is examining different strategies for gait transition based on ensuring smooth body motion during the transition with the goal of minimizing the energy that the transition consumes.

SENSORS

AQUA relies on vision-based sensing to operate within its environment. Due to the inherent physical properties of the marine environment, vision systems for aquatic robots must cope with a host of geometrical distortions: color, dynamic lighting conditions, and suspended particles known as "marine snow."

The aquatic environment's unique nature invalidates many of the assumptions of classic vision algorithms, and even simple problems—such as stereo surface recovery in the presence of suspended marine particles—remain unsolved.

A fundamental problem with visual sensing in the aquatic robotic domain is that it is not possible to assume that the sensor only moves when commanded. The aquatic medium is in constant and generally unpredictable motion, and this motion complicates already difficult problems in understanding time-varying images. One mechanism to simplify vision processing is to monitor the sensor's true motion, independent of its commanded motion.

Inertial navigation systems (INS) have found applications for the determination of a vehicle's relative pose over time in various autonomous systems. Under normal conditions, INSs measure the physical forces applied to them and provide independent measurements of relative motion. Unfortunately, these systems drift; thus, they typically are employed in concert with some secondary sensing system to counteract this effect.

We use stereo vision as this associated second sensor. Real-time stereo sensors permit the recovery of 3D surfaces. Integrating an inertial navigation system with a trinocular stereo sensor simplifies the registration process by providing a relative motion between frames. With this initial estimate of the camera pose, we require few features to refine the registration to the global coordinate frame.

Color correction

For many inspection and observation tasks, obtaining high-quality image data is desirable. We have developed a technique for image enhancement based on training from examples. This allows the system to adapt the image restoration algorithm to the current environmental conditions and also to the task requirements.

Image restoration involves the removal of some known degradation in an image. Traditionally, the most common sources of degradation are imperfections in the sensors or in analog signal transmission and storage. For underwater images, additional factors include poor visibility (even in the cleanest water), ambient light, and frequency-dependent scattering and absorption both between the camera and the environment and also between the light source (the sun) and the local environment (this varies with both depth and local water conditions). The result is an image that appears bluish, blurry, and out of focus.

Most prior work used idealized mathematical models to approximate the deblurring and noise processes. Such approaches are often elegant, but they might not be well suited to the particular phenomena in any specific real environment. Image restoration is difficult since it is an ill-posed problem: There is not enough information in the degraded image alone to determine the original image without ambiguity.

Our approach is based on learning the statistical relationships between image pairs as proposed in the work of B. Singh and colleagues.[7] In our case, these pairs are both the images we actually observe and corresponding color-corrected and deblurred images. We use a Markov random field model to learn the statistics from the training pairs. This model uses multiscale representations of the corrected (enhanced) and original images to construct a probabilistic enhancement algorithm that improves the observed video. This improvement is based on a combination of color matching, correspondence with training data, and local context via belief propagation, all embodied in the Markov random field. Training images are small patches of regions of interest that capture the maximum intensity variations from the image to be restored. The corresponding pairs—that is, the ground truth data containing the restored information from the same regions—are captured when lights mounted on the robot are turned on.

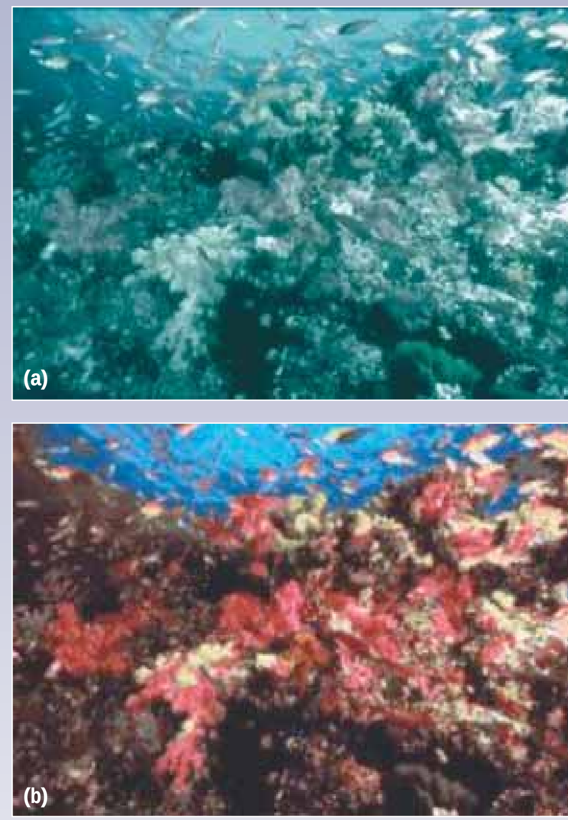

Figure 3. Image restoration. (a) Uncorrected and (b) corrected images. Applying a learning-based Markov random field model accomplishes color correction and deblurring.

Figure 3 shows some experimental results. Several factors influence the quality of the results, including having an adequate amount of reliable information as an input and the statistical consistency of the images in the training sets.

Sensing for environmental recovery and pose maintenance

AQUA combines inertial sensors with a stereo camera rig to construct local environmental models and to aid in pose maintenance. To estimate camera motion, we use both 2D image motion and 3D data from the extracted disparities. First, we use the Kanade-Lucas-Tomasi feature-tracking algorithm[8-9] to extract good features from the left camera at time t and then track these features into the subsequent image at time $t + 1$. Using the disparity map previously extracted for both time steps, we eliminate tracked points that do not have a corresponding disparity at both time t and $t + 1$. We tri-

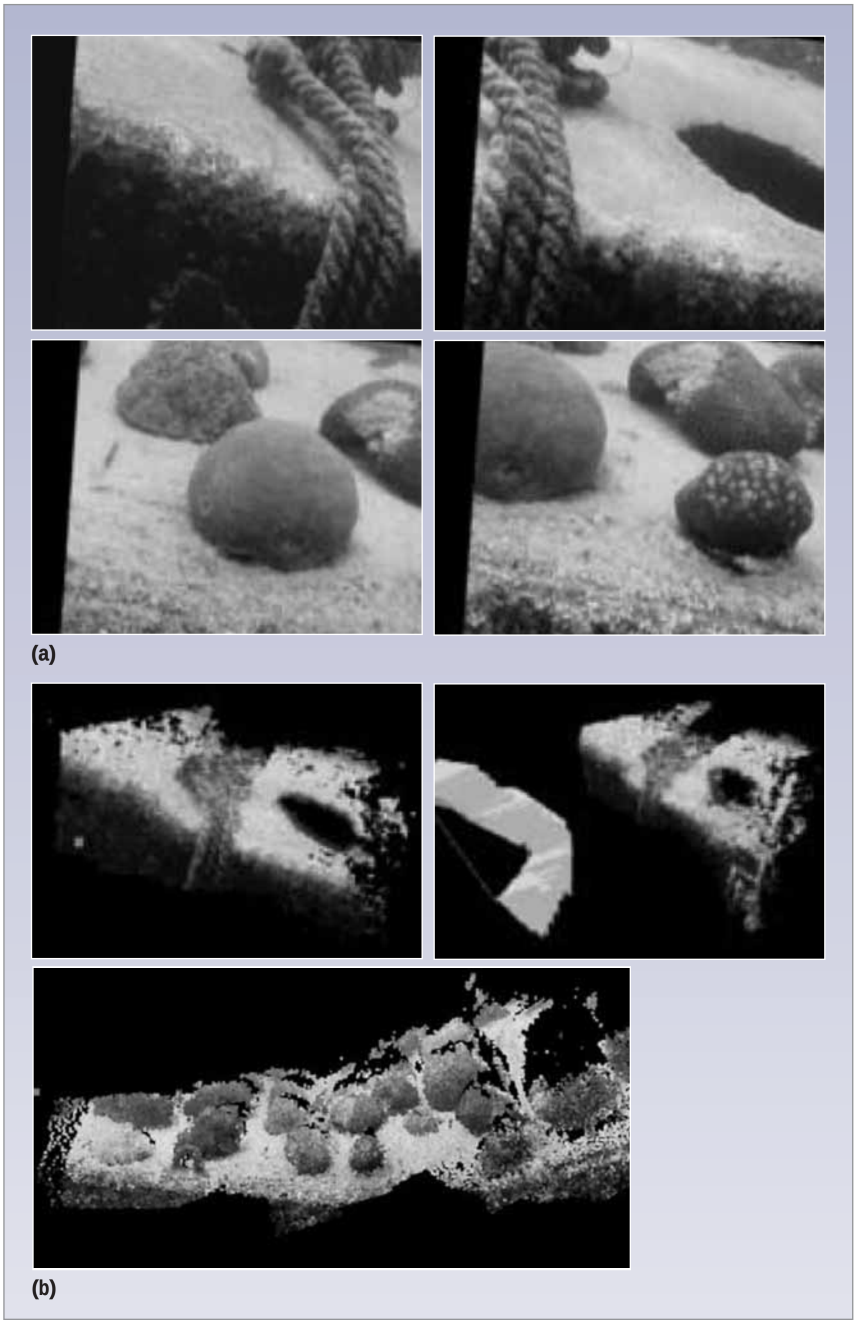

Figure 4. Fig10). Underwater stereo imagery. (a) Reference images from underwater sequences. (b) Recovered 3D underwater structure.

angulate the surviving points to determine the metric 3D points associated with each disparity.

Because many objects and points are visually similar in underwater scenes, many of the feature tracks will be incorrect. Dynamic illumination effects and moving objects—fish, for example—increase the number of incorrect points tracked from frame to frame. To overcome these problems, we employ robust statistical estimation techniques to label the feature tracks as either static or nonstatic. We achieve this by creating a rotation and translation model with the assumption that the scene is stationary. We associate the resulting 3D temporal correspondences with stable scene points for later processing.

We use a volumetric approach to visualize the resulting 3D model. We register the 3D point clouds into a global frame using the previously computed camera pose, and we add each point to an octree. We average the points added to the octree to maintain a constant number of points per node. We then prune the octree to remove isolated points, which produces a result that is less noisy in appearance and can be manipulated in real-time for visualization. The octree can be viewed at any level to produce a coarse or fine representation of the underwater data. Subsequently, we can use standard algorithms such as the constrained elastic surface net algorithm[10] to extract a mesh.

Figure 4 shows some sample reconstructions from underwater stereo imagery.

Acoustic vehicle localization

A critical problem in a SASR task is relating scene structure recovered at different times to a common (global) reference frame. To localize the robot within a global frame, the AQUA project has developed a global acoustic localization sensor.

The acoustic localization component consists of arrays of commercially available omnidirectional hydrophones attached under a surface-floating buoy, the absolute position of which can be measured via a combination of GPS, compass, inclinometers, and inertial sensors.

Suppose that the vehicle is augmented with an acoustic source. Using time-delay estimation on a planar hydrophone array receiving sounds the vehicle emits, we can estimate the direction line in a 3D space emanating from the array's reference point and pointing toward the vehicle. If multiple arrays are available, we can estimate the sound source's position as the intersection of their respective direction lines.

Computationally, the optimal estimate of the source position is the point that has minimal overall distance from these lines. The overall distance to the unknown source position P(x, y, z) is a quadratic function leading to a linear system of equations in x, y, and z that can be solved using standard techniques.

To calculate reliable time delays between the arrival of the sound signals, we correlate two channels of audio data from two different hydrophones and identify peaks of the correlation function. The peak's location corresponds to the time-delay estimate. Before correlation, we filter the sound signalsto reduce noise and then perform a signal variance test to detect the presence of a sound source.[11] The audio frequency region of interest is 200-4,000 Hz to eliminate high-frequency noise as well as the common 60 Hz electric interference and its second harmonic at 120 Hz, extracted using a band pass digital finite impulse response filter.

Valid time delays from a hydrophone pair must be no greater than the maximum time delay, equal to the length of the baseline divided by the speed of sound in water. This reduces the likelihood of false peaks.

The final step for the time-delay estimation is to cluster the time delays estimated from a number of consecutive, nonoverlapping signal time windows. We discard outliers and compute the mean value over the remaining windows as the final time delay estimate.

Experimental results include software simulations, pool tests using hydrophones, and in the air using microphones with a geometry similar to the pool (properly scaled to account for the different sound propagation speeds in the two media). The listening apparatus consists of four DolphinEar/PRO omnidirectional hydrophones, which are attached at the corners of a 1 m × 1 m square buoy, shown in Figure 5.

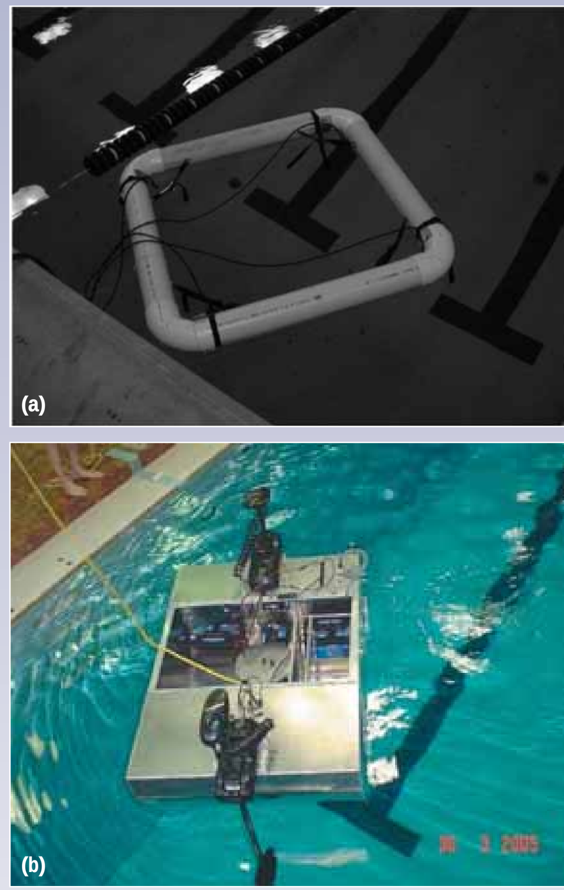

Figure 5. Surface-based sensing. (a) The passive acoustic raft with four omnidirectional hydrophones attached to a 1 m × 1 m square buoy. (b) Self-propelled robotic surface buoy that locates itself in a surface-coordinate system and tracks the AQUA robot via a hydrophone array.

ROBOT LOCALIZATION AND MAPPING

AQUA's vision, inertial, and acoustic sensors provide a foundation for constructing large-scale metric representations of the robot's environment. Such representations support performing a SASR task and presenting task-related sensor readings to a human operator. Indeed, we can envision the construction of a globally consistent metric map that contains the positions of the landmarks in a world coordinate system, thus permitting performing a SASR task over multiple locations.

To solve the mapping problem, the robot needs to estimate its position in relation to the environment at all times, leading to the formulation of the 3D simultaneous localization and mapping problem. The SLAM problem is particularly difficult under water because of issues such as the scarcity of solid objects with distinct features, poor visibility, lack of odometry information, and the inherent 6-DOF limitations.

We use parallel approaches to address the 6-DOF SLAM problem. To overcome low sensor precision, we are investigating two extensions to standard SLAM techniques. The first establishes sophisticated dynamic models that consider earth self-rotation, measurement bias, and system noise. The second uses a sigma-point (unscented) Kalman filter for system-state estimation. We have evaluated this approach through experiments on a land vehicle equipped with an inertial measurement unit, GPS, and a digital compass.[12]

We have explored monocular image-based SLAM in the context of consistent image mosaicing. Here, we address the problem of constructing a globally consistent map using a two-step optimization process. The first step is local optimization: relating the robot's current environmental measurements to its previous measurements based on their overlap—for example, the overlap between the current and previous image. The second step is global optimization, which is carried out as soon as a loop is detected in the robot's path—that is, a sequence of measurements in which the first and the last measurement have substantial overlap. This second step generates or updates a globally consistent map.

For the underwater environment, we have developed a method for estimating the robot's position using a single calibrated image with at least three visual features, the

position of which is known in a world-centered coordinate system. If the feature set in the working environment is sufficiently rich, we use a binocular stereo system to estimate the robot's position and its related uncertainty.

We also have used the stereo-inertial AquaSensor to perform SLAM based on entropy minimization.[13] This algorithm uses the interframe 6-DOF camera egomotion and applies a global rectification strategy to the dense disparity information to estimate an accurate environmental map. The global rectification step reduces accumulated errors from the egomotion estimation that occur due to improper feature localization and dynamic object tracking.

The solution we envision for underwater SLAM is to couple the robot with a self-propelled robotic surface buoy equipped with a sensor suite including GPS and a hydrophone array, such as the one shown in Figure 5. The underwater AQUA robot will be augmented with a transponder that emits a periodic chirp pulse that the hydrophones can detect and the surface buoy can use for localization and tracking. In this manner, the human operators can estimate the underwater robot's absolute position in a world coordinate system and incorporate it into the 3D map.

We have tested AQUA in both terrestrial and aquatic modes, and also in the complex environment the robot encounters as it enters and exits the water. In recent trials, we tested the physical robot, trinocular vision system, and other components at depths up to 40 feet in the Caribbean Sea and the Atlantic Ocean. We also have conducted sea trials near Chester, Nova Scotia, demonstrating the effectiveness of the robot and its sensors in the less clear waters of the North Atlantic.

Experimental testing of the robot to date has concentrated on the independent evaluation of individual components. Over the next few years, we anticipate integrating the various sensor components within the robot itself and performing long-term evaluation of the SASR protocol on reef structures near Holetown, Barbados.

Acknowledgments

We gratefully acknowledge the funding provided by NSERC, the Candian Space Agency, MDRobotics, Waterline Sports, and IRIS NCE for the AQUA project, and DARPA for funding the development of RHex and Rugged RHex, upon which the AQUA robot is based. We also thank the McGill AQUA team for engineering support and the McGill Bellairs Research Institute for providing a positive atmosphere for the field research trials. The research team also thanks Groots for their support during the field trials.

References

1. C. Prahacs et al., "Towards Legged Amphibious Mobile Robotics," *J. Eng. Design and Innovation* (online), vol. 1, part. 01P3, 2005www.cden.ca/JEDI/index.html**.**
2. M. Théberge and G. Dudek, "Gone Swimmin'," *IEEE Spectrum*, June 2006, pp. //inclusive page numbers??//.
3. R. Altendorfer et al., "RHex: A Biologically Inspired Hexapod Runner," *Autonomous Robots*, **//volume and issue numbers??//**, 2001, pp. 207-213.
4. U. Saranli, M. Buehler, and D. E. Koditschek, "RHex: A Simple and Highly Mobile Hexapod Robot," *Int'l J. Robotics Research*, vol. 20, no. **7**, 2001, pp. 616-631.
5. R.J. Full and C.T. Farley, "Musculoskeletal Dynamics in Rhythmic Systems: A Comparative Approach to Legged Locomotion," *Biomechanics and Neural Control of Posture and Movement*, C. Winter, ed., Springer-Verlag, 2000, pp. 192-205.
6. G. Dudek, P. Giguere, and J. Sattar, "Sensor-Based Behavior Control for an Autonomous Underwater Vehicle, *Proc. Int'l Symp. Experimental Robotics*, **Springer-Verlag,** 2006pp. **//inclusive page numbers??//**.
7. B. Singh, W.T. Freeman, and D. Brainard, "Exploiting Spatial and Spectral Image Regularities for Colour Constancy," *Proc. Workshop Statistical and Computational Theories of Vision*, **<<publisher??>>,** 2003, pp. **//inclusive page numbers??//**.
8. J. Shia and C. Tomasi, "Good Features to Track," *Proc. Symp. Computer Vision and Pattern Recognition* (CVPR), IEEE CS Press, 1994, pp. 593-600.
9. B. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision, *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI), **Morgan Kaufmann,** 1981, pp. 674-679.
10. S. Frisken, "Constrained Elastic Surface Nets: Generating Smooth Models from Binary Segmented Data," *Proc. Int'l Conf. Medical Image Computing and Computer-Assisted Intervention*, Springer, 1998, pp. 888-898.
11. H. Liu and E. Milios, "Acoustic Positioning Using Multiple Microphone Arrays," *J. Acoustical Soc. Am.*, vol. 117, no. 5, 2005, pp. 2,772-2,782.
12. P. Zhang, E. Milios, and J. Gu, "Vision Data Registration For Robot Self-Localization in 3D," Proc. *Int'l Conf. Intelligent Robots and Systems (*IROS), IEEE Press, 2005, pp. 2315-2320.
13. J.M. Saez et al., "Underwater 3D SLAM with Entropy Minimization," *Proc. Int'l Conf. Intelligent Robots and Systems* (IROS), IEEE Press, 2006**, //inclusive page numbers??//.**

Gregory Dudek is an associate professor in the Department of Computer Science and the Centre for Intelligent Machines at McGill University. He received a PhD in computer science from the University of Toronto. Contact him at dudek@cim.mcgill.ca.

Philippe Giguere is a PhD candidate at McGill University. He received an MS in computer science from Northeastern University. Contact him at philg@cim.mcgill.ca.

Chris Prahacs is a research associate with the Center for Intelligent Machines at McGill University. He received a BS in mechanical engineering from McGill University. Contact him at cprahacs@cim.mcgill.ca.

Shane Saunderson is a research associate with the Center for Intelligent Machines at McGill University. He received a BS in mechanical engineering from McGill University. Contact him at shane@cim.mcgill.ca.

Junaed Sattar is a PhD candidate in computer science at McGill University. He received an MSc in computer science from McGill University. Contact him at junnaed@cim.mcgill.ca.

Luz-Abril Torres-Mendez is a researcher-professor in the Robotics and Advanced Manufacturing Section at CINVESTAV (Research Centre of Advanced Studies), Coahulla, Mexico. She received a PhD in computer science from McGill University. Contact her at abril.torres@cinvestav.edu.mx.

Michael Jenkin is a professor of computer science and engineering at York University. He received a PhD in computer scienc/ from the University of Toronto. Contact him at jenkin@cse.yorku.ca.

Andrew German is a PhD candidate in computer science at York University. He received a BSc in computer science from the University of Western Ontario. Contact him at german@cse.yorku.ca.

Andrew Hogue is a PhD candidate in the Department of Computer Science at York University. He received an MSc in computer science from //degree-granting institution??//. Contact him at hogue@cse.yorku.ca.

Arlene Ripsman is a PhD candidate in the Department of Computer Science at York University. She received an MSc in computer science from York University. Contact her at arlene@cse.yorku.ca.

Jim Zacher is a research associate in the Centre for Vision Research at York University. He received a BA and BEd in experimental psychology/education from the University of Windsor/University of Toronto. Contact him at zacher@cvr.yorku.ca.

Evangelos Milios is a professor on the Faculty of Computer Science at Dalhousie University. He received a PhD in electrical engineering and computer science from the Massachusetts Institute of Technology. Contact him at eem@cs.dal.ca.

Hui Liu is a graduate student on the Faculty of Computer Science at Dalhousie University. She received an MSc in computer science from Dalhousie University. Contact her at hiliu@cs.dal.ca.

Pifu Zhang is a PhD candidate on the Faculty of Computer Science at Dalhousie University. He received a DE in mechanical engineering from Hunan University. Contact him at pifu@cs.dal.ca.

Martin Buehler is the director of robotics at Boston Dynamics. He received a PhD in electrical engineering from Yale University. Contact him at buehler@BostonDynamics.com.

Christina Georgiades is //business title?// at MathWorks. She received an MEng in mechanical engineering from McGill University. Contact her at cgeorg@cim.mcgill.ca.