Optical Position Sensing Challenges in Autonomously Docking a UUV With A Submerged Submarine

Jared Giesbrecht
DRDC – Suffield Research Centre

George Watt
DRDC – Atlantic Research Centre

MTS/IEEE OCEANS North America 2017
Anchorage, Alaska

Date of Publication from Ext Publisher: September 2017
IMPORTANT INFORMATIVE STATEMENTS

Disclaimer: This document is not published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada, but is to be catalogued in the Canadian Defence Information System (CANDIS), the national repository for Defence S&T documents. Her Majesty the Queen in Right of Canada (Department of National Defence) makes no representations or warranties, expressed or implied, of any kind whatsoever, and assumes no liability for the accuracy, reliability, completeness, currency or usefulness of any information, product, process or material included in this document. Nothing in this document should be interpreted as an endorsement for the specific use of any tool, technique or process examined in it. Any reliance on, or use of, any information, product, process or material included in this document is at the sole risk of the person so using it or relying on it. Canada does not assume any liability in respect of any damages or losses arising out of or in connection with the use of, or reliance on, any information, product, process or material included in this document.

This document was reviewed for Controlled Goods by Defence Research and Development Canada (DRDC) using the Schedule to the Defence Production Act.

© Her Majesty the Queen in Right of Canada (Department of National Defence), 2017
© Sa Majesté la Reine en droit du Canada (Ministère de la Défense nationale), 2017
Optical Position Sensing Challenges in Autonomously Docking a UUV With A Submerged Submarine

Jared Giesbrecht  
Defence R&D Canada  
Suffield Research Centre  
Medicine Hat, Alberta, Canada

George Watt  
Defence R&D Canada  
Atlantic Research Centre  
Dartmouth, Nova Scotia, Canada

Abstract—This paper describes the development of an underwater optical tracking system for the final stage of automated docking of an unmanned underwater vehicle (UUV) with a submerged, slowly moving submarine. During the final docking phase, while the UUV holds course, the docking mechanism will track LED light sources on the UUV and will rapidly adapt to relative motions for final capture. For this phase, the optical system is intended to provide position and pose measurement of the UUV from a camera mounted on the submarine capture mechanism for the final 10 to 15 meters of rendezvous. This requires accurate relative pose tracking of the UUV, real-time operation, and robustness to changing lighting and water turbidity conditions. This report documents the design, development, and test of a preliminary monocular visual optical tracking for this application.

I. INTRODUCTION

An unmanned underwater vehicle deployed from a submarine would be useful for reconnaissance, mine hunting, and other applications. It is relatively straightforward to launch a UUV from the torpedo tubes, but retrieving it while submerged to collect the data is much more difficult. A robotic arm and capture mechanism to retrieve the UUV that accounts for the inherent navigational and relative motion problems is one solution to this problem. In this system, the UUV would use long range acoustic navigation techniques to find the submarine and manoeuver to, and maintain itself within the docking envelope. At this point, the robotic dock mechanism would grab the UUV to stow on-board the submarine (Figure 1).

This project is investigating long range acoustic, medium range electromagnetic, and short range optical position sensing [1]. During the final docking phase, while the UUV holds course, the docking mechanism will rapidly adapt to relative motions of the UUV for final capture. For this to be feasible, a short-range, high frequency sensing method for UUV position and pose is required for feedback control on the dock. Towards this end, we are pursuing an underwater optical system.

The biggest challenges to optical sensing underwater are the variable lighting conditions and water clarity resulting from the need to dock at variable depths, in daytime and nighttime, and in various types of water including open ocean, littoral waters, harbours and estuaries. In these situations water turbidity will be highly variable due to suspended sediment, organic matter, and life forms such as plankton and algae, which reduce visibility and scatter light sources used for tracking.

The optical system presented here uses computer vision algorithms and careful control of camera exposure to accurately pinpoint lights and determine the relative pose of the UUV from an underwater camera. In this design, a monocular camera captures images that are processed by computer vision algorithms to determine the location of the UUV using multiple LEDs mounted on the UUV hull.

The work is focused on determining potential performance of short range optical tracking in a realistic environment. As a proof of concept, lights have been mounted on a surrogate
UUV and tested with the vision system in harbour water at depths to 20m and ranges to 13m.

Initial trials involved testing a variety of optical sensing methods and investigating the difficulties of using computer vision in turbid water. Further trials involved moving the surrogate UUV through a set of motions to simulate a docking maneuver to explore the feasibility of tracking a representative target underwater. The most recent trials involved using a ground truth measurement apparatus to determine the accuracy of the UUV tracking system over a wide variety of relative positions and poses. We discuss qualitative lessons learned and provide our latest validated position sensing results.

II. BACKGROUND

For this application, optical tracking was selected because of the potential for high data rates using simple, inexpensive, low power hardware. During our preliminary trials we investigated options for underwater optical tracking, including laser ranging [2]–[4], passive tracking using ambient light [5]–[8], and active lighting to illuminate a target. However, with all of these approaches, attenuation of light, as well as backscatter from the light source left the target difficult to image.

Instead, we use camera on the dock/sub side to track LED lights on the UUV to find UUV position and pose. This provides target detection at practical ranges of from 10 to 100 m [9]–[13] depending on the environment and light source. Various geometries and numbers of lights can be used to extract the 3-D pose of a target from a 2D image [10], [12], [14]. Li et. al. [15] use a pair of stereo cameras to further improve the performance of the system. Further discussion on underwater optical imaging can be found in [16].

The method of finding a target pose from a number of control points is often called the Perspective N-Point Problem [17]. Among the many approaches to solving this, it is given that using 3 points produces multiple solutions, while having four or more points produces unique solutions. There are many works in the literature for finding the position of robotic vehicles above the surface [18], [19]. For an underwater system, Bosch et. al. detail an approach that includes the use of camera calibration [14]. However, they do not address tracking LED targets over a wide variety of lighting and turbidity conditions, or a wide variety of ranges from the most distant to immediately in front of the camera.

III. OPTICAL IMAGING OF LEDS UNDERWATER

Placing the LED lights on the UUV, and the camera, tracking equipment, and final docking control on submarine dock allows for larger, higher power, higher quality imaging and computing systems, while reducing complexity and valuable space claim on the UUV. The LEDs are low power devices turned on during the final few minutes of the docking process, and represent an insignificant electrical or physical burden on the UUV. Further, this setup takes advantage of the transverse maneuverability of the dock to capture the UUV despite relative motions of the UUV and submarine, which would be more difficult to accomplish solely with UUV homing. The dock design is discussed in much greater detail in [1].

Some of our earlier experiments involved the general feasibility of tracking lights underwater [20]. There were an number of difficulties encountered, including:

1) Turbid water causes blurring of LED light sources into a diffuse glow, making the resolution of individual LEDs difficult, and making pose estimation impossible.
2) Ambient lighting in turbid water can swamp the intensity of the LED lights, reducing detection range.
3) Directional or narrow angle light sources cause a wide range of light intensities in the image with changes in UUV heading, either overwhelming the image sensor or making it too dim to see.
4) Variation in brightness of LEDs between near and far distances also causes difficulties in imaging. Maximum exposure is desirable to acquire an LED target at long distances, but as the UUV approaches, the lights quickly saturate the camera image. An off-the-shelf camera with auto-exposure that we tested could not control the image brightness adequately for computer vision work.
5) Marine life and other items in the water (such as the UUV hull) can reflect LED or ambient light, and appear brighter than the LEDs themselves (Figure 2).

IV. COMPUTER VISION ALGORITHMS

The computer vision portion of this project involved locating the LED targets in the camera images despite changes in turbidity, ambient lighting, UUV range, UUV heading, etc. The computer vision software was implemented in C++ on a laptop PC, making extensive use of the OpenCV libraries [21]. It runs in real time providing position updates at 10 Hz or faster, with processing speed normally being limited by the camera shutter speed (i.e. at most depths the PC executed the software faster than the amount of time the camera needed to
gather enough light to image the scene). The pose estimation algorithm was implemented in Maple, as described below.

The computer vision system consists of several components described in the following sections.

A. Intensity Thresholding

This algorithm finds LED targets in images by choosing the brightest pixels in the image as being from LED light sources. This threshold is adaptively changed during the docking process to accommodate changes in ambient lighting and exposure. The threshold is chosen to select only a small percentile of the pixels in the image (typically 1% to 0.001%) to choose only LED lights, and not other sources of light such as fish or reflections. Image smoothing, as well as erode and dilate functions are further used to eliminate tiny non-LED points of light.

This method works best for bright light sources in clear water, or when the UUV is near the camera and the light sources are large and distinct in the image. The centre of the LED is chosen as the centre of the brightest, largest point in the image. This may be a poor approximation for lights that are large or have bloomed into odd shapes because of turbidity and odd LED/camera angles (Figure 2). However, with proper exposure control, the LED lights should never become very large in the image, and so the approximation should be a reasonable one.

B. Gradient Detection

We also detected LEDs in the images using the Sobel operator to find areas in the image with the highest rate of change in brightness [22]. This operator approximates the first derivative or rate of change in intensity of the pixels over a specified area. Once again smoothing and erode and dilate functions are used to remove small sources of light.

This method works best for dim points of light in the image that are not much brighter than the surrounding background (i.e. at long ranges or high turbidity). LEDs under these conditions created areas that may still be of relatively low intensity compared to the overall image, but for which the rate of change is higher in the local area of the image. These light sources would typically be missed by the intensity threshold detector.

The gradient detector also works well to reject areas of higher ambient light that create a bright arc of high intensity pixels in the top half of the image (Figure 5). In this case, the local rate of change is not very high in the bright, ambient light area, and is not picked up by the gradient detector.

The gradient detector does not work well to find the centre of the light source for large, bright, LED images. The rate of change in intensity is higher at the edges of the light source than in the middle, which causes this algorithm to pick the edges of large light exposures as a potential LED source. For this reason, the gradient detector is used mostly at the start of docking sequence not the end.

C. LED Selection and Tracking

Because of the complementary nature of the threshold and gradient trackers, the results from both are combined to maximize tracking performance. The first step is to assign a relative certainty value to the candidate LED locations selected by the algorithm above. This is based on its intensity and size, as larger, brighter targets are more likely to be LED sources than noise or reflections off of the UUV or marine life.

The threshold method typically gives us more reliable results when the LED light source is obvious, with a better estimate of the LED position in the light part of the image. In situations where there are few obvious LED sources in the image, then the gradient method will provide greater sensitivity at the cost of the potential for false positive results, and less accurate estimation of LED position.

The last step is to combine the results of the threshold and gradient methods. The candidates are first sorted to remove any that are too close to each other (and therefore probably the same LED). Then, they are sorted from highest to lowest certainty value, so that the best candidates for light sources are used to calculate the UUV pose.

D. Tracking and Exposure Control

The docking scenario will have the UUV moving over the full range of the camera system, from beyond the visible range, to immediately in front of the camera. The LED lights appearance in the image varies drastically as well. It was found that exposure control of the camera is critical to provide any level of functionality over the full range of docking, and that commercial auto-exposure systems were not appropriate for the image processing tasks. Exposure control also provides robustness to variable ambient light conditions due to depth, turbidity, etc.

In order to provide dynamic range to the computer vision system, two tactics were used:

- Controlling the camera exposure during the docking process using “states” of LED tracking: Initialing, Acquiring, Tracking, and Lost. This allows us to have maximum sensitivity when trying to first find the UUV lights at long ranges, while still accurately finding the LED positions in the image at shorter range.
- Increasing how selective we are about accepting results from the computer vision algorithm based on what “state” we are in. This also allows us to have maximize sensitivity at long ranges, and reduces false positives and errors in LED positions at shorter ranges.

During initialization, the UUV is assumed to be absent, and the exposure is set to be as high as possible without incurring false positives. During “Acquiring”, our tolerance for accepting candidate points of lights as LEDs is low, to allow the system to be as sensitive as possible. During this phase it is most likely that the gradient tracker will first acquire the LED lights. During acquisition, the system waits to see a point of light in the same spot for several seconds before moving into “Tracking” mode. In tracking mode, the system
begins to report UUV position and pose. As the LED lights become larger and brighter in the image, our tolerance for accepting potential LED lights becomes higher to improve noise rejection. Points that move too far or too fast in the image are also rejected.

During tracking, exposure control is also used to maintain the LED sources at the appropriate size in the image to pinpoint their location. The exposure control tries to maintain the brightest LED in the image between 40 and 70 pixels. If it is outside this range, the exposure is either increased or decreased by a step percentage (typically 20%). Although this method is fairly simplistic, it proved robust for limited underwater dynamics conditions, and did not result in continuous changes are made to the system. The first is that we must increase the minimum certainty value and adjusting the minimum amount of pixel separation between the LEDs should increase, and adjusting the minimum amount of pixel separation between candidate LED sources allows improved noise rejection. Points that move too far or too fast in the image are also rejected.

Finally, as the UUV nears, our separation between the LEDs should become larger and brighter in the image, our tolerance for accepting potential LED lights becomes higher to improve noise rejection. Points that move too far or too fast in the image are also rejected.

E. Position and Pose Estimation

The UUV location and pose relative to the camera are estimated using the setup shown in Figure 3. A right handed Cartesian coordinate system \( \xi, \mu, \nu \) is fixed to the camera with its origin at the middle of the camera charge-coupled device (CCD). The \( \xi \) axis is normal to the CCD and points in the direction the camera faces. The \( \mu \) and \( \nu \) axes are aligned with the horizontal and vertical pixels respectively in the camera image. Sometimes it is useful to talk in terms of range \( \rho \), bearing \( \beta \), and elevation \( \delta \) to the UUV:

\[
\begin{align*}
\xi &= \rho \cos \beta \cos \delta, \quad \mu = \rho \sin \beta \cos \delta, \quad \nu = \rho \sin \delta
\end{align*}
\]  

(1)

The \( x, y, z \) axes in Figure 3 are standard body-fixed axes for underwater vehicles [23, 24]. Their origin is typically on the UUV hull centerline at or close to the vehicle center of buoyancy; the \( x \) axis points forward along the hull centerline, the \( y \) axis points to starboard, and the \( z \) axis points down. The UUV orientation (pose), relative to the camera, is defined using a standard underwater vehicle Euler transformation. If the camera and UUV axes are initially aligned, UUV pose is achieved by first yawing an angle \( \psi \) about the \( z \) axis, then pitching an angle \( \theta \) about the \( y \) axis, and finally rolling an angle \( \phi \) about the \( x \) axis. Thus, UUV body axes vectors \( \mathbf{X} \) are reoriented to camera axes using the Euler transformation \( \mathbf{A} \cdot \mathbf{X} \) where:

\[
\mathbf{A} = \begin{bmatrix}
\cos \theta \cos \psi & \sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi & -\sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi \\
\cos \theta \sin \psi & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi & \sin \phi \cos \theta \\
-\sin \theta & \cos \phi \sin \theta & \cos \phi \cos \theta
\end{bmatrix}
\]

(2)

Let \( p, q \) be the pixel coordinates giving location in a camera image. Their common origin is the top left corner of the image and they increase in the same directions as \( \mu, \nu \). If \( p_0, q_0 \) locate the center of the image, \( p_j, q_j \) locate LED \( j \) in the image, and \( F \) is the focal length of the camera in seawater in pixels, then:

\[
\begin{align*}
\xi_j P_j &= \frac{p_j - p_0}{F} = \frac{p_j}{\xi_j} \quad \Rightarrow \quad \xi_j P_j = \mu_j \\
\xi_j Q_j &= \frac{q_j - q_0}{F} = \frac{q_j}{\xi_j} \quad \Rightarrow \quad \xi_j Q_j = \nu_j
\end{align*}
\]

(3)

where:

\[
\begin{bmatrix}
\xi_j \\
\mu_j \\
\nu_j
\end{bmatrix} = \begin{bmatrix}
\xi \\
\mu \\
\nu
\end{bmatrix} + \mathbf{A} \begin{bmatrix}
x_j \\
y_j \\
z_j
\end{bmatrix}.
\]

(4)

The fixed coordinates \( x_j, y_j, z_j \) of each LED in UUV axes must be known. For \( j = 1, 2, 3 \), these are the 6 equations required to solve for the 6 unknowns \( \xi, \mu, \nu, \phi, \theta, \psi \) giving the position and pose of the UUV.

The UUV position coordinates \( \xi, \mu, \nu \) appear linearly in (3) and can be eliminated using 3 of the equations. This leaves three equations that are nonlinear in the Euler angles \( \phi, \theta, \psi \) and which must be solved iteratively. Newton’s method is ideal for this because the formulation (2), (3), and (4) is simple enough that analytical expressions for the necessary derivatives can be calculated analytically and rapidly evaluated. CPU time is not an issue. The problem, however, is that there can be several unique solutions and in some situations these can be very close together. A robust pose prediction requires more than three LEDs and these should not be coplanar.

With \( n \) LEDs in an image, there are \( N = n!/(3!(n - 3)!) \) groups of three independent equations to solve for only three unknowns. There will not be a unique solution because of inherent error in the pixel coordinates, so a least squares approach is used. For large \( n \), this is more efficient than
solving all the equation groups separately and averaging the results.

In the least squares formulation, we take advantage of the fact that the $a_{ij}$ from (2) appear linearly in (3). If:

$$V^T = [a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}]$$

(5)

then the equations to solve can be put in the form $B \cdot V = 0$ where $B$ is a $3N \times 9$ coefficient matrix independent of the unknown Euler angles. The sum of the squares $S$ of these equations is:

$$S = (B \cdot V)^T \cdot B \cdot V = V^T \cdot B^T \cdot B \cdot V.$$  

(6)

The nice result here is that the inner product $B^T \cdot B$, which has constant coefficients, is reduced to a $9 \times 9$ symmetric matrix that only needs to be calculated once; using iteration to solve for the pose uses $B^T \cdot B$, not $B$.

Newton’s method is again used to minimize $S$. This requires the gradient $G$, the iteration correction $\Delta$, and the Hessian $H$:

$$G = \begin{bmatrix} \frac{\partial S}{\partial \phi} \\ \frac{\partial S}{\partial \theta} \\ \frac{\partial S}{\partial \psi} \end{bmatrix} \quad \Delta = \begin{bmatrix} \Delta \phi \\ \Delta \theta \\ \Delta \psi \end{bmatrix},$$

(7)

$$H = \begin{bmatrix} \frac{\partial^2 S}{\partial \phi^2} & \frac{\partial^2 S}{\partial \phi \partial \theta} & \frac{\partial^2 S}{\partial \phi \partial \psi} \\ \frac{\partial^2 S}{\partial \theta \partial \phi} & \frac{\partial^2 S}{\partial \theta^2} & \frac{\partial^2 S}{\partial \theta \partial \psi} \\ \frac{\partial^2 S}{\partial \psi \partial \phi} & \frac{\partial^2 S}{\partial \psi \partial \theta} & \frac{\partial^2 S}{\partial \psi^2} \end{bmatrix}.$$  

The gradient $G$ is a coefficient matrix independent of the unknown Euler angles. The sum of the squares $S$ of these equations is minimized when $G = 0$ (3 equations in 3 unknowns) but to ensure the solution is not a saddlepoint or maximum, $H$ must be positive definite. A pure Newton approach makes an initial guess for the pose and then repeatedly solves:

$$G + H \cdot \Delta = 0$$  

(8)

for a series of corrections $\Delta$. This method was unreliable for processing the 3 LED images. Therefore a modified Newton method called a ‘restricted step’ method by Fletcher [25] and a ‘trust-region’ method by Nocedal and Wright [26] was adopted. The method limits $|\Delta|$ to some value $\Delta_{max}$. The pure Newton method is allowed to proceed normally unless $|\Delta|$ exceeds $\Delta_{max}$. $S$ increases, or the Hessian is not positive definite. $\Delta_{max}$ is adjusted up or down depending on the algorithm’s ability to decrease $S$. If necessary, $H$ is made positive definite by increasing the magnitude of its diagonal elements by $\lambda \geq 0$ following:

$$(H + \lambda I) \cdot \Delta = -G \quad \text{where} \quad \lambda(\Delta_{max} - |\Delta|) = 0.$$  

(9)

The method has been customized for a $3 \times 3$ Hessian with $G$ and $H$ calculated analytically. The images discussed below were processed by running Maple scripts in hardware floating point mode; the CPU time to process each image averaged 0.03 s. The initial condition was the solution from the previously processed image. Of the 908 validation images processed, 682 contained 4 or 5 LEDs and the remainder only 3 LEDs. All of the former group and 122 of the latter could be processed without requiring a modified Hessian and these averaged 3.3 iterations per image to reduce $|\Delta|$ below $10^{-4}$ radians. The remaining 104 images, all with only 3 LEDs, required a modified Hessian for at least one of the iterations but usually only for the first few; they averaged about 12 iterations per image, with one aborted solution when the iteration count exceeded 100.

V. EXPERIMENTAL SETUP

In order to capture images for the computer vision system, we used a SubC uLux (Stargazer) camera, which uses a Retiga 1350B (QIClick) scientific video camera from QImaging. It has 1392 x 1040 resolution1, software-controllable gain and exposure settings, and a fixed-focus lens with a horizontal field of view of about 40 degrees in water. Images were relayed over a IEEE-1394 digital video from the camera using a pair of FireNEX-COAX-S800 IEEE-1394 repeaters and through 25 meters of coaxial underwater cable. The images were captured from the IEEE-1394 interface using a laptop computer.

The lights used in the experiments were Cree XLamp XP-E LEDs, which output 250 lumens each at 1A of current, powered by a cable from the surface. For most of the tests, LED current of 100mA was sufficient, except for the longest ranges at shallow depth. Importantly, the LEDs have a viewing angle of 115 to 130 degrees, which allows the UUV to be aimed almost to 90 degrees away from the camera and still visible. This also meant that light diffusers were not required, and less exposure control was needed during UUV heading changes.

To simulate the proportions of a UUV, the LEDs were mounted on UUV hull made from PVC and aluminum. One LED was mounted on the nose, and one on the end of each tail fin. This configuration provides a decent tradeoff between having several LEDs visible at all poses, with having maximum separation between the lights, and without having lights overlapping and interfering with each other. Having a single LED on the nose also provides the ability to do pure “homing”, where the dock could ignore all of the pose information of the UUV, and simply keep the nose LED in the centre of the image for terminal docking.

Tests were conducted at the DRDC acoustic calibration barge facility in Halifax’s Bedford Basin, shown in Figure 4. The interior has a 18m by 9m “well” that is open to the harbour water to conduct experiments up to 42 metres deep. It has two movable bridges that span across the well, allowing the horizontal movement, as well lowering and raising of camera and UUV equipment in the water.

In order to “ground truth” our experiments, a mechanical linkage was designed to position the UUV and cameras in established relative poses, so that the measurements from the computer vision system could be verified for accuracy and precision. Previously, a short-baseline acoustical system was used in an attempt to provide this functionality, but was found to be insufficiently accurate. As such, a mechanical linkage

1Higher resolution may result in slightly higher accuracy at the price of execution time. In these tests, the camera was used at a resolution of 696x520.
was designed to “fix” the relative positions of the equipment in the water. This “Pose Validation Rig” consisted of the camera and UUV (with LEDs) hanging at the end of variable length vertical shafts in the water from the barge. These shafts were then held at a fixed distance apart by a third “range pole”. The range pole attached to the vertical shafts by a weighted tray that stabilized the vertical shafts, while allowing both the camera and UUV to be rotated in the horizontal plane. Anchor cables to the surface further stabilized the position of the UUV and camera in the water. The Pose Validation Rig will be documented in a subsequent DRDC technical report.

Using this setup, the UUV could be varied in yaw and distance from the camera. Further, the vertical and horizontal position in the camera field of view could be changed by altering depth and horizontal position, resulting in a wide variety of poses and viewing angles of the UUV in the images. Only UUV pitch and roll could not be varied in this setup.

VI. RESULTS

A. Computer Vision

In general, the computer vision methods of identifying LED locations in the images were effective, with the caveat that the parameters need to be adapted over the process of docking as described earlier. At 20m depth, we were able to detect the LEDs over the full range of our test facility (13m to 2m). If tuned correctly, the gradient method could reliably detect LEDs before a human observer looking at the video feed. Meanwhile, at close ranges the intensity threshold method was effective at rejecting reflections off the UUV hull, etc.

The detection range was reduced when the UUV was orientated at angles away from the camera, although this was mostly mitigated by having LED sources with a wide viewing angle. In practice, the effective range was not significantly reduced until relative UUV orientations of 75 degrees or more. Some images of threshold and gradient detection are shown in Figure 10.

Of course, in shallow and turbid water, the detection range is greatly reduced by both ambient light and particles in the water. In the same experiments, reducing the depth from 20m to 5m reduced the reliable detection range from at least 13m to only 8m. In turbid water, the exponential increase in absorption and scattering with range is such that increasing LED power provides only a modest increase in detection range. The conclusion is that automated docking works better at depth away from ambient light.

The effect of ambient lighting is particularly noticeable in Figure 5, with the UUV above the camera. In this image, the large blob at upper left is the sun shining into the water, while the UUV is in the upper right. In this case, the gradient detector still found the LED lights, but if the sunlight had been directly behind the UUV, it would have been totally invisible.

B. Position and Pose Results

Using the pose validation rig discussed earlier, a set of UUV images were taken at a variety of UUV positions and poses. These images were processed as discussed in Section III.E above. Of these, a representative set 782 predictions were selected in which duplication is minimized, UUV range varies from 2 to 13m, UUV yaw varies between 0 and 85 degrees, and there are over 10 different horizontal positions and two different vertical positions in the images.

Our goal was to build the ground-truth pose validation rig system to an accuracy of ±3cm, but there is no independent means of validation. As shown in the results below, the computer vision system and the pose validation rig are in agreement to within several centimetres of range and a few degrees of angle. This seems acceptable given the ranges involved and unknown water currents or other sources of error.
For this set of data, we found a suitable focal length for predicting range by optimizing our results for range over our test data using the ground truth information. This focal length could have also been found through a camera calibration procedure or perhaps through manufacturer data.

The 6-degree-of-freedom results for one of these sets of poses is shown in Figure 6. This test was conducted at 13 metres range (the limit of our experimental setup), and consisted of moving the UUV horizontally across the images, while also changing its yaw heading. In the graph, blue represents the theoretical pose based on our mechanical ground truth system, while red indicates the result as reported by the computer vision system. Those samples where only three LEDs were visible in the image are marked with an “x”. As can be seen, the system works well except for a reduction in accuracy in range, roll, pitch and yaw estimation when the UUV has yawed away from the camera by 75 degrees and only 3 LEDs were visible.

As expected, as range increases, the accuracy in predicted position decreases. At longer range, there are fewer pixels between each of the LEDs to use as a measurement baseline for range estimation. A summary of the range accuracy over the entire set of 908 poses is shown in Figure 7.

Higher relative UUV yaw also decreases accuracy in range, although less dramatically (Figure 8). With UUV yaw, the distance between nose and tail LEDs increases, which should improve range accuracy. However, with higher yaw, the LED positions from the computer vision software are less accurate, fewer LEDs are visible, and the spacing between tail LEDs is smaller, resulting in poorer estimation of UUV pose angles (Figure 9). As such, roll is poorly predicted when we can’t see the LEDs on the back side of the UUV hull. Because all of the angles are interrelated in the nonlinear S function, if roll is miscalculated, all the other angles will have a poorer result as well.

Statistical results from all 908 poses are shown in Table I. Accuracy improves with more LEDs, although this is somewhat overstated because the poses where more LEDs are visible are less eccentric (i.e. the UUV is located more in the
center of the image without significant yaw when 5 LEDs are visible).

| TABLE I
<table>
<thead>
<tr>
<th>STANDARD DEVIATION OF 6-DOF ERROR OVER ALL POSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 or more LEDs</td>
</tr>
<tr>
<td>Bearing (deg)</td>
</tr>
<tr>
<td>Elevation (deg)</td>
</tr>
<tr>
<td>Roll (deg)</td>
</tr>
<tr>
<td>Pitch (deg)</td>
</tr>
<tr>
<td>Yaw (deg)</td>
</tr>
<tr>
<td>Range (% of measurement)</td>
</tr>
</tbody>
</table>

C. Simulated Docking

In order to try to understand more about the performance of the pose estimation system, several “simulated rendezvous” were performed, by continuously moving the UUV surrogate from the farthest range of the facility to the nearest range (without the pose validation rig). While doing this, the surrogate UUV was also yawed and translated side to side to simulate UUV motion. This was subject to the limits of the barge facility, with UUV pitch and roll remaining fixed during this experiment.

A sequence of images is shown in Figure 10 illustrating the appearance of the UUV with exposure control. Note that due to exposure control, the LEDs appear at a reasonable size for detection despite the reduction in exposure time by two orders of magnitude. Detection locations by the threshold and gradient computer vision algorithms is also shown.

Results from this particular docking simulation are shown in Figure 11. The system begins to provide reasonably reliable pose estimates as soon as multiple LEDs are visible, at a range of about 10.7m. Full pose results are shown in Figure 12.

This test was conducted at a depth of only 3m, in somewhat turbid water, with the LEDs operated at the brighter end of their range (750mA operating current). If conducted at depth, the range achievable is expected to increase. However, even under these conditions, reliable pose estimates were achievable almost immediately when multiple LEDs were detected. As expected, at longer ranges, and when only 3 LEDs were detected, the accuracy suffered.

There is also some outlier data that typically occurs when the 5th LED that can be obscured by the UUV body either appears or disappears (blue dots, Figure 11). This is because when the LED first starts to appear, the center of the light is not visible, but the projection of the light in the water is. At this point, the computer vision system inaccurately finds the LED center, causing inaccurate pose measurements. Pose filtering using a Kalman filter or other method would easily eliminate these outliers.

VII. DISCUSSION

Despite the encouraging results obtained, there remain improvements to be made.

Fig. 10. Sample images from a UUV docking sequence with active exposure control. Gradient detections in yellow with intensity threshold detections in green.
ranges [15]. It might also be possible to extend the stereo baseline (and range) by placing a second camera further away on the dock. In the extreme, an extra camera system could be used on another location on the sub, providing triangulation of UUV position, and greatly improved accuracy. It is imagined that extra outward facing LEDs on the side of the UUV would aid in this.

A camera calibration procedure would help remove distortion from the images and allow the calculation of a focal length for position/pose calculations. However, this is not as straightforward for underwater cameras due to refraction in the glass housing [27], [28]. In the end, our results were reasonably accurate without this step.

Number and positions of the LEDs on the UUV could also be optimized. It is helpful to have as much separation between the LEDs for maximum accuracy of pose estimation, and prevent the computer vision system from identifying them as one single light at long ranges and in turbid water. It is also important to position the LEDs where they can be seen from the widest variety of angles and not be obscured by the UUV body (i.e. right at the nose, and on the tail planes as far as possible away from UUV). However, it is also beneficial to have as many lights on the UUV as possible, as pose detection performance improves with more data points, and more LEDs also means that more will be visible at any given time despite some being obscured by the UUV body. This tradeoff would be easier on large UUVs as compared to the small one used for these experiments.

Finally, during the previous sections, there was no mention of the method of determining which LED is which from the images (i.e. nose vs. tail, etc.). This is an important point, as the pose solution relies completely on this information to avoid ambiguous pose solutions. It is especially critical given that there are many UUV poses which may obscure some of the LEDs from view. So far, we have made the assumption that this problem would be solvable with existing technology, such as using pairs of LEDs in key spots, using coloured LEDs, or using synchronized flashing LEDs.

VIII. Conclusion

We have demonstrated an effective means of determining the pose of a submerged UUV using LED lights and a monocular camera. This process is made more difficult by changing ambient light conditions and turbidity in the water. Our system uses a combination of computer vision algorithms and adapts their parameters during the course of UUV approach to maintain maximum sensitivity at long ranges and noise rejection and accuracy at close range. Pose estimation of the UUV was found using at least three LEDs, but is improved with four or five LEDs visible. The software system was able to run at data rates high enough for feedback control.

There are a number of potential improvements to the methods, particularly in the area of LED identification, but on the whole the system was very effective across a wide variety of UUV positions and orientations. It remains for us to test this system against a live UUV manoeuvering realistically, and as part of an active dock to understand how the movement and dynamics of a dynamic scenario would affect performance.

References


<table>
<thead>
<tr>
<th>DOCUMENT CONTROL DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)</td>
</tr>
</tbody>
</table>

| 1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.) |
| DRDC – Suffield Research Centre  |
| Defence Research and Development Canada  |
| P.O. Box 4000, Station Main  |
| Medicine Hat, Alberta T1A 8K6  |
| Canada  |

| 2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) |
| CAN UNCLASSIFIED  |

| 2b. CONTROLLED GOODS |
| NON-CONTROLLED GOODS  |
| DMC A  |

| 3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) |
| Optical Position Sensing Challenges in Autonomously Docking a UUV With A Submerged Submarine  |

| 4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used) |
| Giesbrecht, J.; Watt, G.  |

| 5. DATE OF PUBLICATION (Month and year of publication of document.)  |
| September 2017  |

| 6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) |
| 10  |

| 6b. NO. OF REFS (Total cited in document.) |
| 28  |

| 7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) |
| External Literature (N)  |

| 8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) |
| DRDC – Suffield Research Centre  |
| Defence Research and Development Canada  |
| P.O. Box 4000, Station Main  |
| Medicine Hat, Alberta T1A 8K6  |
| Canada  |

| 9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) |
|  |

| 9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.) |
|  |

| 10a. ORIGINATOR’S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) |
| DRDC-RDDC-2017-N030  |

| 10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.) |
|  |

| 11a. FUTURE DISTRIBUTION (Any limitations on further dissemination of the document, other than those imposed by security classification.) |
| Public release  |

| 11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Any limitations on further dissemination of the document, other than those imposed by security classification.) |
|  |
This paper describes the development of an underwater optical tracking system for the final stage of automated docking of an unmanned underwater vehicle (UUV) with a submerged, slowly moving submarine. During the final docking phase, while the UUV holds course, the docking mechanism will track LED light sources on the UUV and will rapidly adapt to relative motions for final capture. For this phase, the optical system is intended to provide position and pose measurement of the UUV from a camera mounted on the submarine capture mechanism for the final 10 to 15 meters of rendezvous. This requires accurate relative pose tracking of the UUV, real-time operation, and robustness to changing lighting and water turbidity conditions. This report documents the design, development, and test of a preliminary monocular visual optical tracking for this application.

unmanned underwater vehicle, computer vision, autonomy