VISION SYSTEMS IN THE CONTROL OF AUTONOMOUS UNDERWATER VEHICLES

Jonathan Horgan & Daniel Toal

Mobile & Marine Robotics Research Centre
Department of Electronic and Computer Engineering
University of Limerick, Ireland

Abstract: Over the last decade machine vision has played an important role in the research and development of autonomous underwater vehicles (AUVs) for the exploration and utilization of the deep sea environment. This paper presents a review of research in the application of machine vision for the control of AUVs. The paper focuses on five particular applications: intervention class vehicles, video mosaicking, station keeping, navigation & positioning and cable tracking. Each application is explained and analysed by examining different projects. Based on the authors' observations, an assessment regarding the performance of the strategies is given and possible future trends are identified for each application.

Keywords: mobile robots, autonomous vehicles, marine systems, visual servoing, robot vision, cameras

1. INTRODUCTION

The number of applications for underwater vehicles is ever increasing. The inquisitive nature of humans has lead to the comprehensive exploration and mapping of land masses on planet earth, subsequently scientists are now turning to the oceans to discover new possibilities for telecommunications, biological and geological information and energy sources. Underwater vehicles play an important role in this exploration, as the deep ocean is a harsh and unforgiving environment for human discovery. Autonomous underwater vehicles are rapidly becoming accepted as a viable alternative to remotely operated vehicles (ROV), due to their cost effectiveness and their greater manoeuvrability for use in more hazardous environments (Wernli 2002).

Vision is a very useful sensor in the control of underwater vehicles and has been the subject of increased interest over the last decade as a result of improved processing capabilities of hardware and the need for more flexible and accurate sensor solutions. The increase in research efforts into vision systems is due to its inexpensive nature and its common inclusion on underwater vehicles as a payload sensor. Vision also has a number of advantages over other types of sensors for underwater applications. Vision systems are light weight and do not possess a minimum operating range unlike their acoustic counterparts (Nolan, et al., 2005). Despite these advantages over other sensors, machine vision underwater poses an amount of difficult challenges to be overcome for it to be successfully incorporated into control (Matsumoto and Ito 1995). Marine snow, low contrast, non-uniform illumination and a lack of distinguishable features on the seabed are just some of the inherent difficulties faced when using optics underwater (Cufi, et al., 2002).

This article focuses on investigating five particular applications of vision in AUVs; intervention class AUVs (IAUV), video mosaicking, station keeping, navigation & positioning and cable tracking. Each application is discussed and its evolution into its present state is analysed by looking at a number of projects. An assessment regarding the performance of the strategies is given by discussing the advantages and disadvantages of each solution. Based on the authors' observations, possible future trends are identified for each application.

2. INTERVENTION CLASS AUV (IAUV)

A next step in the evolution of the unmanned underwater vehicle (UUV), after the development of robust autonomous navigation, is to allow interaction with the surrounding environment (intervention class). One of the ways an AUV can be developed into an intervention class vehicle (IAUV) is by equipping the vehicle with manipulators (robotic arms). Manipulators are useful for many undersea applications such as repair of undersea cables and pipes, programming tasks such as valve operation, environment interaction and sample taking. At present, most underwater manipulation tasks are performed by manned submersibles or ROVs. Machine vision can be employed to tackle some of the issues involved with making robot arm control autonomous.

Very little literature exists of research completed in the field of underwater robotic arm control. Smith (Smith, et al., 1994) developed a vision system for controlling a manipulator using a positional based visual servoing approach. A single monochrome CCD camera is used to capture images of the target object. The images are analysed using a combination of edge detection and corner detection to identify simple shapes by comparing them with known reference objects. Once the target is identified, range measurement is performed using a two spot laser triangulation scheme. Using both the acquired location and range measurements the target can be acquired. This method is capable of identifying and interacting with known target objects. Unfortunately, in a real unstructured marine environment, the manipulator may encounter unidentified and unrecognised potential objects of interest in random locations, with unknown position, size and orientation.

Two more recent intervention class vehicle development projects are the SAUVIM (Semi autonomous Underwater Vehicle for Intervention Missions) project at the University of Hawaii (Marani, et al., 2006) and the ALIVE (Autonomous Light Intervention Vehicle) project (Evans, et al., 2003). The ALIVE vehicle was developed to be capable of autonomous docking and intervention of sub sea structures to replace costly tethered ROVs currently used for such tasks. The vision system is used for vehicle pose estimation during the final docking phase. When the docking sequence is initialized, edge detection is performed on incoming images to locate the edges of three circular valves on the docking panel. The centres of these circular valves are then used in combination with a CAD (Computer Aided Design) model, in order to estimate the pose of the vehicle. This system is developed for docking and then manipulating known objects, fusing other sensor information in its operation besides optical. The system is proficient at its intended application but as yet is unable to cope with unmodeled environments.

The main issue with autonomous manipulator arm control underwater is that the environment is unstructured. The two traditional visual servoing control approaches, image based and positional based, are incapable of dealing with unmodeled environments as both techniques require prior knowledge of the target object. To overcome the problems of manipulation in unstructured environments Toal *et al* (2002, 2004) proposed a method of interacting with objects using a technique called 'Pull to Position', a direction based variant on visual servoing. This technique is more capable of working in scenarios where there is greater variability in the robot environment and the target objects. Vision is used to determine the required direction to the object of interest. An incremental advance approach is then implemented in order to acquire the target. Early experiments were performed using simple light dependant resistor array in a honeycomb formation to develop and test the motion control aspect of the technique. This technique has since been tested with camera systems in the laboratory with results highlighting that it is capable of tracking objects in motion (Toal and MacGiollaEáin 2006) (see Fig 1).

Intervention class vehicle development is a young area of research with most marine research institutions concentrating their efforts on the areas of navigation and positioning. However, in more recent years commercial projects have highlighted the potential of IAUVs as a viable alternative to their more expensive ROV counterparts. The challenge of making an intervention class vehicle autonomous proves to be a very difficult task but remains a vital area of research in order to make IAUVs commercially successful and reduce the need for customised solutions.

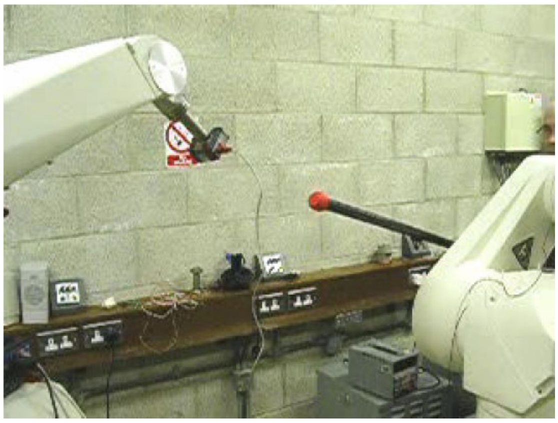

Fig 1. Test setup for 'Pull to position'

3. VIDEO MOSAICKING

Light attenuation and backscatter inhibit the ability of a vision system to capture large area images of the sea floor. Video mosaicking is an attempt to overcome this limitation using a process of aligning short range images of the seabed floor to create one large composite map. Video mosaicking can be used as an aid to other applications such as navigation, wreckage visualisation, station keeping and also to promote a better understanding of the sea floor in areas such as biology and geology. Video mosaicking involves the accurate estimation of vehicle motion in order to accurately position each frame in the composite image (mosaic). The general setup of the vision system remains the same for almost all mosaicking implementations reviewed. A single CCD camera acquires images at right angle to the seabed and at an altitude ranging from 1-3 meters.

Early attempts at video mosaicking were developed by Marks *et al* (1995) who proposed a method of measuring offsets and connecting the images using correlation to create an accurate real-time mosaicking system. The images are filtered using a Laplacian of Gaussian filter in order to highlight zero crossings and pronounce the image textures. The filtering reduces the image noise and also the effect of nonuniform illumination from artificial sources. A stored reference image is correlated with the live incoming images to derive the offset in pixels. The mosaic is created by repeatedly storing images and using the offset calculated to place the image in the map. Correlation methods rely on well contrasted images in order to locate regions of correlation; a lack of texture will inhibit the system from correctly positioning images in the mosaic. Correlations inability to deal with rotations, scale changes and undersea currents (seen from results) may hinder its ability to create multiple column mosaics.

Garcia *et al* (2001) proposed a method of feature characterisation to improve the correspondences between images in order to create a more accurate mosaic to position an underwater vehicle. Firstly regions of high spatial gradient are selected from the image using a corner detector. Image matching is accomplished by taking the textural parameters of the areas selected and correlating them with the next image in sequence. A colour camera improves the process as the matching is implemented on the hue and saturation components of the image as well as the intensity of the image. A set of displacement vectors for the candidate features from one image to the next is calculated. A transformation matrix can then be constructed to merge the images in the correct location in the final mosaic. This paper takes advantage of the hue and saturation components of a colour camera to improve the image correspondence and also uses textures rather than features to obtain a more robust image match. The mosaic construction was not completed in real-time possibly because of high computational requirements.

Gracias and Neighadaripour (2005) present two methods of creating mosaics using video sequences captured at different altitudes. The first method relies on a rendered mosaic of higher altitude images to act as a map to guide the position of the images in the lower altitude mosaic ('image to mosaic'). The second method does not require rendering of the higher altitude mosaic, just the topology to match each particular image of the lower altitude sequence against the higher altitude images ('image to image'). Ground truth points were used to compare the two methods presented. Both methods obtained good results, while the 'image to image' method showed less distortion, it had the disadvantage of higher computational expense. Unfortunately the method requires a small amount of user input to select correspondences and the flat, static and constant lighting of the environment are assumptions of the technique. Time efficiency is another factor to be considered due to the method requiring runs at different altitudes

.

It is difficult to compare and evaluate the performance of each of the methods described. Each technique has been tested in scenarios where different assumptions are made regarding the environment, vehicle dynamics and processing power available. Negahdaripour and Firoozfam (2001) attempted to compare methods (using a common data set) implemented by different institutions to document the various approaches and performances of different techniques to the marine world. Unfortunately, due to time constraints, only comparative results for feature–based and direct methods are reported. A more comprehensive report would give a better understanding of the strengths and weaknesses of current techniques available. Some recent research efforts in the area have investigated the construction of 3D mosaics, a further step forward in the evolution of mosaicking methods (Nicosevici, et al., 2005). Video mosaicking remains a very complex and challenging application because of the inherent difficulties faced with optics underwater. 3D mocaicking, is a glimpse of what the future could possibly hold for this application and what research institutes will be improving upon with advances in processing capability and vision systems.

4. STATION KEEPING

The ability for submersible vehicles to accurately maintain position and orientation is a necessity. The process of maintaining a vehicle's predefined pose in the presence of disturbances (undersea currents and reaction forces from manipulators attached to vehicle) is known as station keeping. Station keeping can be used for many different underwater applications such as repair of underwater structures and near seabed data collection. Station keeping using a vision system has the advantage of being able to use natural rather than manmade beacons for motion detection. The camera is set up in a similar fashion to that of mosaicking (section 3) and the methods for motion estimation overlap greatly between the two applications. The general method for visual station keeping is to maintain a reference image of the station and compare live incoming frames with this image to estimate and correct for drift of the vehicle.

Stanford/MBARI researchers proposed a method of measuring vehicle drift using a texture tracking strategy (Marks, et al., 1994). The method of motion estimation is the same process as described in video mosaicking (Marks, et al., 1995). As with the mosaicking technique the method relies on having a highly textured image in order to find regions of correlation. Correlation-based methods' inability to deal with changes in the image due to rotations will inhibit accurate motion estimation. Cufi *et al* (2002) makes use of a technique previously developed for a mosiacking application (Garcia, et al., 2001). The mosaicking application is incorporated into the station keeping solution thus allowing for greater interframe motion (mosaic can be referenced) and solving the problem of correlation with rotated images. This method could again suffer from a lack of distinct textures in the sub sea environment.

Negahdaripour *et al* (1999) proposes a method of station keeping by directly measuring motion from spatio-temporal image gradient information. This method allows for the estimation of 3D motion directly using the spatio–temporal derivatives of incremental images captured over a short period of time (Negahdaripour and Horn 1987). A generalized dynamic image motion model was later developed (Negahdaripour 1998) to account for variations in the scene radiance due to lighting and medium conditions underwater. This is of particular importance when using flow–based methods in underwater imagery. A technique for calculating both instantaneous velocity and absolute position is implemented to increase the limit of interframe motion. This method is susceptible to sporadic miscalculations in velocity, which accumulated over time can result in inaccurate position estimations.

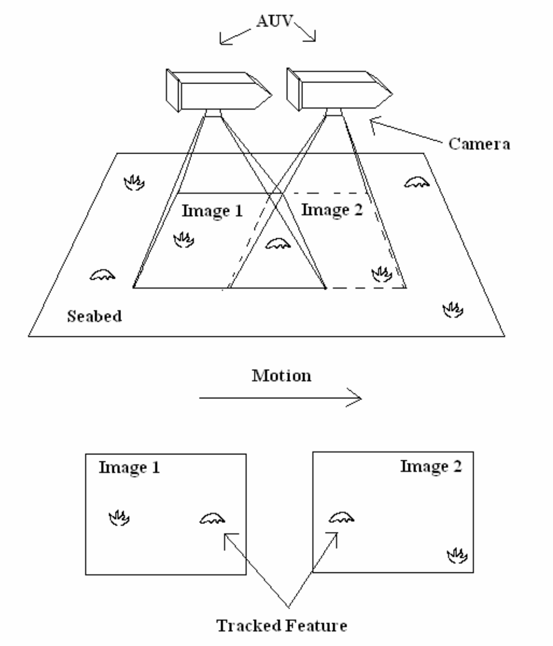

Fig 2. Feature based motion detection

Other methods implement a combination of methods to achieve station keeping. Van Der Zwaan *et al* (2002) uses a technique of integrating optic flow information with template matching in order to estimate motion. A motion model for template matching and an automatic landmark selection scheme were also implemented in this paper. Station keeping, much like mosaicking, has many methods for tracking motion from vision: correlation based, feature based (Fig 2), optical flow based etc. The underwater environment is inherently difficult for the exploitation of optical sensors. Many factors have to be considered to obtain accurate results with the final goal of creating an autonomous real-time station keeping system. The methods discussed are hard to compare due to differing test setups and vehicle dynamics. However, there remains room for improvement in the area, especially with camera and processing hardware improvements.

5. VISION IN NAVIGATION & POSITIONING

Underwater navigation is an extensive area of research and poses a number of difficult challenges. Accurate navigation systems are vital for all underwater missions for correct correspondence between sensor and navigation data and also for final recovery of the vehicle. The main difficulties encountered with navigation and positioning underwater is that GPS is unavailable to the vehicle, so other accurate positional measurement systems must be developed to compensate. In general, navigation underwater is achieved with combinations of; acoustic positioning, inertial navigation systems, Doppler velocity log (DVL), pressure depth and other integrated sensors. Inertial navigation with a Kalman filter fusing position estimates from a complex of different sensors, gives the best possible navigation/position estimates. In this scenario DVL is a common inertial navigation system (INS) aiding sensor. At low velocities DVL performance degrades (Keary, et al., 1999). DVL as with all acoustic instruments has a minimum (blanking) range from the sea bed below which they do not function and it is often a bulky sensor which is an issue for AUVs.

In this paper, the possibilities of using vision systems for navigation are discussed in the case of video mosaicking, station keeping and cable tracking. From the extensive research in the field it is obvious that vision systems are a robust sensor for motion detection and measurement, so it is feasible that vision could be implemented as an aid to INS. It is a standard operating procedure to equip UUVs with camera systems and this approach would thereby use the already present camera in navigation. This would have the advantage over DVL with regard to cost and performance that wouldn't degrade at lower speeds and sub meter altitudes. To our knowledge no system for using vision as an aid to INS has been developed for an underwater application.

However, land based applications of vision aided INS have been developed. Deil *et al* (2005) proposed a method of tracking features to lengthen the period of time which a vehicle can maintain correct positional information in GPS deprived environments. Another example of vision used as a dead-reckoning system is proposed by Tora *et al* (2006) to improve on odometry based solutions. A method very similar to the mosaicking methods explained above is implemented to estimate pose.

Machine vision has been proven as a viable aiding sensor in an unstructured land setting; however, these techniques are not directly transferable to a sub sea scheme. The underwater environment adds the complexity of 3D motion and a lack of distinct features. The author believes that vision is a viable solution for aiding INS in a near seabed environment where acoustic positioning may be prone to inaccuracy e.g. in channels, caves or wrecks.

6. CABLE TRACKING

The necessity for frequent underwater cable/pipe inspection is becoming more apparent with increased construction of undersea piping networks for the oil and gas industry and heavy international telecommunication traffic. Current methods for the surveillance, inspection and repair of undersea cable/pipes utilize remotely operated vehicles (ROV) controlled from the surface. This work can prove to be very tedious and time consuming while also being prone to human error due to loss of concentration and fatigue. An autonomous vision based cable tracking system would prove much less expensive than current ROV based solutions, as the need for a mother ship is removed. Vision systems also possess advantages over magnetometer and sonar based solutions for cable tracking. Vision systems prove less expensive, have the ability to identify faults and require a smaller less powerful vehicle for operation.

An early attempt at a cable tracking system using machine vision was developed by Matsumoto (Matsumoto and Ito 1995). The method takes advantage of the lack of straight line edges found in the underwater environment. An edge image of the sea floor is acquired using a Laplacian of Gaussian filter. The Hough transform is then applied to the edge pixel image in order to find the most likely pipe edge candidates. A method of candidate evaluation is implemented by examining the length and width of each edge pixel line candidate. The direction of the cable in the present image and the previous image are used to predict the angle of the Hough transform to be applied to the subsequent image to reduce computation time. This paper is a very early implementation and uses different techniques to address the problems of sediment covered pipes and nonuniform illumination. While achieving reasonable results in a controlled environment, factors such as spurious edge detection from other pipes or elements, abrupt pipe direction changes and a search algorithm (when cable is undetected) have not been accounted for and could have reduced performance.

Balisuriya developed on previous work (Balasuriya, et al., 1997) by adding an *a priori* map of the cable location to his technique (Balasuriya and Ura 2002). The main objectives of the method are to follow the cable when it is not visible to the vision system and selection of the correct cable in the image (in the case of multiple cables). These objectives are addressed by assuming that an *a priori* map of the cable is available. The *a priori* map serves three purposes; to predict the region of interest (ROI), to avoid misinterpretations with other cables in the image and to be used as a navigation map in the case where the cable disappears from view. A similar method to Matsumoto (1995) is implemented to locate the cable in the image and again the Hough transform is utilized. This paper fuses other inputs besides optical to track the cable and has attempted to overcome the issues of tracking a cable when it invisible (sediment or algae covered) and correct cable selection. The method demonstrates that the extra information, in the form of a map, fused with optical sensing can greatly improve performance. Unfortunately, having an *a priori* map of the cable location is not always a feasible assumption.

Ortiz *et al* (2002) developed a method for real-time cable tracking using only visual information and again takes advantage of the cables shape to locate strong alignment features along its side. The contour pixels are examined to locate pixel alignments that display strong pipe characteristics (long pixel alignments, parallel alignments and alignments in a y direction on the image). Once the cable has been located in the image a Kalman filter is implemented to reduce the ROI for the proceeding image to reduce computation time. When anomalies occur in the prediction phase special actions are taken in order to correct the algorithm; either the frame causing the anomaly is discarded or if a number of frames are incorrect the Kalman filter is reset. This method achieved a 90 percent success rate for trials at 25 frames/sec for old installations of cable. The technique dealt reasonable well with partially covered cables however, a minimal presence of the cable is required in the image at all times. No backup system exists in the scenario where the cable becomes completely invisible to the camera.

Many different methods for cable tracking systems exist each with there own advantages and disadvantages. The work reviewed uses similar techniques for cable detection (looking for straight line edges) but differ in their approaches to cable direction prediction to save on computational expense. The need for a robust system for tracking a cable that has become invisible for a short segment remains a priority. Sensor fusion is a good approach to robust cable following when in the cable is in view. The future will focus on refining the tracking methods and working towards the development of vision systems for inspection, fault identification and localisation to fully automate the process and reduce human input.

7. CONCLUSIONS

This work has highlighted the importance of vision systems for the control of autonomous underwater vehicles. Each application has shown the distinct advantages of vision systems as a sensor and the different ways in which optical information can be manipulated to produce successful applications. Machine vision research appears to be growing strongly, with publications in the field increasing dramatically in recent years. From the review it can be seen that using optics underwater has a number of inherent difficulties to be overcome and that selection of the correct approach to any application is not trivial. However, despite these challenges it appears that the advantages of vision based solutions outweigh the disadvantages and vision has an important role to play in the future of AUV development.

REFERENCES

Balasuriya, A. and T. Ura (2002). Vision-based underwater cable detection and following using AUVs. Oceans '02 MTS/IEEE.

Balasuriya, B. A. A. P., M. Takai, et al. (1997). Vision based autonomous underwater vehicle navigation: underwater cable tracking. OCEANS '97. MTS/IEEE Conference Proceedings.

Cufi, X., R. Garcia, et al. (2002). An approach to vision-based station keeping for an unmanned underwater vehicle. IEEE/RSJ International Conference on Intelligent Robots and System, 2002.

Diel, D. D., P. DeBitetto, et al. (2005). "Epipolar Constraints for Vision-Aided Inertial Navigation." IEEE Workshop on Motion and Video Computing (WACV/MOTION'05) **2**: 221-228.

Evans, J., P. Redmond, et al. (2003). Autonomous docking for Intervention-AUVs using sonar and video-based real-time 3D pose estimation. OCEANS 2003. Proceedings.

Garcia, R., J. Batlle, et al. (2001). Positioning an underwater vehicle through image mosaicking. IEEE International Conference on Robotics and Automation

Gracias, N. and S. Negahdaripour (2005). Underwater Mosaic Creation using Video sequences from Different Altitudes. Proceedings of MTS/IEEE OCEANS

Keary, A., M. Hill, et al. (1999). "Simulation of the correlation velocity log using a computer based acoustic model." Proceedings of the 11th International Symposium on Unmanned Untethered Submersible Technology: 446-454.

Marani, G., J. Yuh, et al. (2006). "Autonomous manipulation for intervention AUV." Advances in *Unmanned Marine Vehicles IEE Control Series* Volume **68**: 217-237.

Marks, R. L., S. M. Rock, et al. (1995). "Real-time video mosaicking of the ocean floor." IEEE *Journal of Oceanic Engineering* **20**(3): 229-241.

Marks, R. L., H. H. Wang, et al. (1994). "Automatic visual station keeping of an underwater robot." OCEANS '94. 'Oceans Engineering for Today's Technology and Tomorrow's Preservation.' Proceedings **2**: 137-142.

Matsumoto, S. and Y. Ito (1995). Real-time vision-based tracking of submarine-cables for AUV/ROV. OCEANS '95. MTS/IEEE. 'Challenges of Our Changing Global Environment'. Conference Proceedings. .

Negahdaripour, S. (1998). "Revised definition of optical flow: integration of radiometric and geometric cues for dynamic scene analysis." IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(9): 961-979.

Negahdaripour, S. and P. Firoozfam (2001). Positioning and photo-mosaicking with long image sequences; comparison of selected methods. OCEANS, 2001. MTS/IEEE Conference and Exhibition.

Negahdaripour, S. and B. Horn (1987). "Direct Passive Navigation." IEEE Transactions on Pattern Analysis and Machine Intelligence **9**(1).

Negahdaripour, S., X. Xu, et al. (1999). "Direct estimation of motion from sea floor images for automatic station-keeping of submersible platforms." I*EEE Journal of Oceanic Engineering* **24**(3): 370-382.

Nicosevici, T., S. Negahdaripour, et al. (2005). Monocular-Based 3-D Seafloor Reconstruction and Ortho-Mosaicing by Piecewise Planar Representation. Proceedings of MTS/IEEE OCEANS

Nolan, S. P., D. Toal, et al. (2005). A novel wideangle ultrasonic sensor utilizing a curved radiator developed for use in an AUV obstacle avoidance system. Oceans 2005 - Europe

Ortiz, A., M. Simo, et al. (2002). "A vision system for an underwater cable tracker." *International Journal of Machine Vision and Applications* **13**(3): 129-140.

Smith, J. S., R. Yu, et al. (1994). Computer vision control of an underwater manipulator. OCEANS '94. 'Oceans Engineering for Today's Technology and Tomorrow's Preservation.' Proceedings.

Toal, D., C. Flanagan, et al. (2004). "Equipping autonomous underwater vehicles with manipulators." 3rd IFAC Symposium on Mechatronic Systems.

Toal, D. and C. MacGiollaEáin (2006). "Direction Based Visual Servo – A New Visual Servoing Technique." The 16th International Conference on Flexible automation and Intelligent Manufacturing.

Toal, D. J. F. and C. Flanagan (2002). "'Pull to position', a different approach to the control of robot arms for mobile robots." Journal of Materials Processing Technology **123**(3): 393-8.

Tora, H., B. Koku, et al. (2006). "A vision based optical dead-reckoning system." The 10th Mechatronics Forum Biennial International Conference.

van der Zwaan, S., A. Bernardino, et al. (2002). "Visual station keeping for floating robots in unstructured environments." Robotics and Autonomous Systems **39**(3-4): 145-155.

Wernli, R. L. (2002). AUVs-a technology whose time has come. Proceedings of the 2002 International Symposium on Underwater Technology.