# Robust Vision-Based Semi-Autonomous Underwater Manipulation

**Conference Paper** · January 2008

DOI: 10.3233/978-1-58603-887-8-308



**4 authors:**

Marc Hildebrandt
Deutsches Forschungszentrum für Künstliche Intelligenz


Jochen Kerdels
FernUniversität in Hagen


Jan Christian Albiez
Kraken Robotik GmbH, Bremen, Germany


Frank Kirchner
Universität Bremen

# Robust Vision-Based Semi-Autonomous Underwater Manipulation

Marc HILDEBRANDT [a], Jochen KERDELS [a], Jan ALBIEZ [a] and Frank KIRCHNER [a]
[a] *DFKI Robotics Lab Bremen, 28359 Bremen, Germany*

**Abstract.** The increasing demand for deep sea resources induces an increased amount of robotic technology to work in this harsh environment. In this paper we present a visual servoing approach to release the strain on operators controlling the currently used remotely operated vehicles, by giving them the possibility to hand some of the tasks over to a semi-autonomous control system. The paper deals with the basic control of a commercially available 7 degrees of freedom hydraulic manipulator, the 3D camera system used and the object recognition and visual servoing control. The presented approach focuses heavily on robustness to augment the acceptance with the offshore industry.

**Keywords.** Underwater Robotics, Visual Servoing, Dexterous Manipulation, Control, Kinematics

## Introduction

An increasing trend in the offshore industry is to mine deep sea resources, e.g. oil or mangan. At the moment this is a very expensive endeavor. Especially, cost-efficient 24h systems for the inspection and maintenance of deep sea production facilities are needed.

The current state of the art for deep-sea system-maintenance done at the well heads and pump stations of offshore oil- and gas-rigs are manipulation tasks like opening or closing of valves, setting up structures or guiding equipment lowered down from the surface. All this work is accomplished by remotely operated vehicles (ROVs) relying on a cable connection to a surface vessel for control and energy. The demands on an operator team of such a vehicle are high. They have to control the vehicle in all six dimensions, while trying to interpret two dimensional video feeds as three dimensional data and accomplishing the aforementioned manipulation tasks. The control interface for these systems are classical tele-operation interfaces relying on a master-slave control with limited feedback ability.

Due to the high costs for operating a ROV and its parent ship the offshore industry is extremly conservative in terms of new, untried technology. The environmental conditions for the technical systems are extremly harsh. Up to 600 bar of pressure, low temperatures and a highly corrosive environment while deployed is added to rough handling onboard the support vessels due to sometimes harsh weather conditions resulting on a high strain on every part of the equipment. In this paper we present a visual servoing approach which therefore relies mainly on components well known in the offshore industry like the Orion 7P hydraulic manipulator and extends its capabilites by adding semi-autonomous components. So far there has been some work done in underwater vision for controlling AUV or ROV (e.g. [13]), but in the case of autonomous underwater manipulation closed-loop vision-based control has not been applied to this degree.

## 1. Manipulator Control

Underwater manipulator control is a very current topic. A number of significant differences to traditional manipulator control complicates the transfer of established techniques and algorithms to this specific new environment [3]. Some of the major differences and problems are the following:

- **Instable Base**
  Underwater manipulator systems are usually not ground-based but operate from platforms like heavy-workclass-ROV systems. These systems inherently have 6 DOF and are subject to effects of currents, swaying motions induced by active vehicle movement, and the 'tail wags the dog'-problem [4], placing the manipulator system in a very unstable configuration.
- **Low Control Frequency**
  Since Underwater manipulators are built for manual operator control, their control frequency is rather low (e.g. 60Hz forward/12Hz response for the Orion7P). This poses a problem on e.g. smooth visual servoing.
- **Limited Feedback**
  Due to the difficulties of pressure-resistant sensor systems, most manipulators provide very little sensory feedback. The only two established types of feedback are position and force. Multi-dimensional sensor-arrays for slip-detection in the jaws or similar sensors used in traditional robotics are not available in any form. Additionally, the visual feedback is only available through the on-board camera systems, which usually have mediocre image quality and, due to their high costs, are only used in small numbers.
- **Highly Dynamic Environment**
  The underwater environment is very dynamic. The workspace cannot be cleared of marine life (which is usually surprisingly interested in new light-sources and movement), and movements of the base systems causes changes in size and shape of the workspace.

The consequence of these problems is that the modus operandi of underwater manipulators still is manual operator control. In the CManipulator project, computer-based manipulator control is evaluated. The basic setup of the used system is shown in figure 1.

At this point, various benefits of computer-based manipulator control could already be tested and will be addressed in depth in the following paragraphs.

### *1.1. Improved sensory feedback*

In the original deployment of the used Orion7P manipulator system the only sensory feedback was the joint angles as alphanumeric data in the control panel. In a situation where only parts of the manipulator arm are visible, this makes it very hard to estimate the current manipulator configuration. To improve this shortcoming, a 3D-visualization

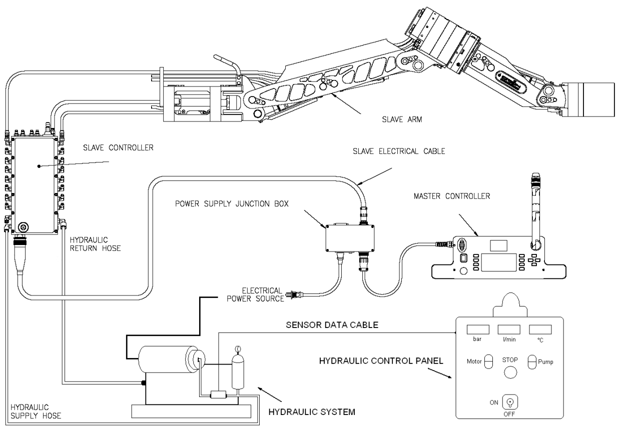


**Figure 1.** Communication connections in a typical manipulator system.

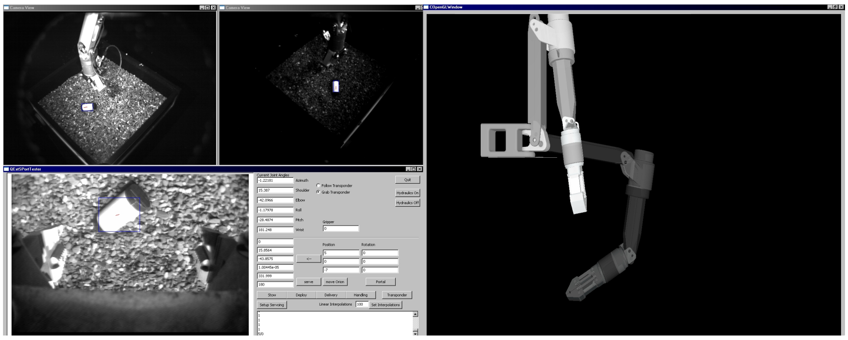

**Figure 2.** View of the Orion7P control center. On the left side the three camera images can be seen, on the right side the 3D representation of both the current (solid) and future (transparent) positions of the manipulator.

of the system's current state was generated. It allows real-time movement and rotation in order to give the operator the best observation position. Additionally, a second 3D visualization of the manipulator arm is projected as semi-transparent overlay. Now it is possible to first visualize a desired manipulator configuration, and allows a much safer movement after validation. This can be seen in figure 2 on the right side.

To improve the ability to precisely manipulate even small objects, a wrist-camera was mounted onto the last joint of the Orion7P. While this restricts its otherwise freely rotable wrist to $\pm 180°$, it provides valuable close-object information and eye-in-hand visual servoing [5]. Currently, different approaches to mount the camera in a way not restricting continuous rotary motion are evaluated.

Another area of interest is jaw-force-feedback. The original Orion7P only provides different rates of jaw movement but does not provide any feedback information about position or force acting on the brackets. This results in always applying its peak force of 4400 N [6]. To prevent damage to manipulated objects and to detect object movement while in the manipulator's jaws, a 2D-sensor matrix is being applied on the jaws. First experiments with a piezo-resistive sensor matrix proved very promising and further research into this area is planned.

*1.2. Cartesian manipulator control*

A number of tasks requires manipulator control in the Cartesian coordinate system rather than the axial coordinate system intrinsic to 6DOF manipulators. The design of a fast algorithm for computing this transformation was given a high priority at the beginning of the CManipulator project. The inherent flaws of numerical algorithms (namely speed and computational stability) made them unsuitable for this application, and a closed-form geometric/analytic solution for the inverse kinematics problem of the Orion 7P was developed by solving (1) for $\theta_1$ to $\theta_6$ for a given $\mathbf{R}_6^0$ (the transformation from base coordinate system to tip coordinate system) containing the rotational ($r_{1n}$, $r_{2n}$ and $r_{3n}$) and translatory ($p_x, p_y$ and $p_z$) components of the desired end-effector position. $\mathbf{T}_{i+1}^i$ are the transformations along the kinematic chain of the manipulator according to Denavit-Hartenberg (see [15]).

$$\mathbf{R}_6^0 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p_x \\ r_{21} & r_{22} & r_{23} & p_y \\ r_{31} & r_{32} & r_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{T}_1^0(\theta_1) \cdot \mathbf{T}_2^1(\theta_2) \cdot \mathbf{T}_3^2(\theta_3) \cdot \mathbf{T}_4^3(\theta_4) \cdot \mathbf{T}_5^4(\theta_5) \cdot \mathbf{T}_6^5(\theta_6) \quad (1)$$

With highly optimized implementation, the time for computation of one random inverse configuration could be reduced to $64\mu$s on a desktop computer, enabling it to be calculated in real-time without consuming too much computational power. This is vital since resulting from the dynamic nature of the system a precalculation of position is not possible, so it needs to be done on-line. At the same time it has to be computationally stable to prevent undesirable movement, which is covered by using a closed-form solution.

One of the benefits of Cartesian manipulator control is the possibility to compensate much more intuitively for vehicle movement, which can be seen as rotation and translation of the manipulator's base. If this movement is known it could directly be compensated by a simple inverse movement of the end-effector position prior to the inverse-kinematics calculation. So in world coordinates this would lead to equation (2) with $\mathbf{P}_{TIP}$ (the absolute transformation to the tip coordinate system from world origin, including the ROV-position $\mathbf{P}_{ROV}$) staying constant at both times $t_0$ and $t_1$ as long as the ROV's movement between these two time-frames $\Delta\mathbf{P}_{ROV}$ is known.

$$\begin{aligned} &\mathbf{P}_{ROV}(t_0) \cdot \mathbf{R}_6^0 = \mathbf{P}_{TIP} \\ &\mathbf{P}_{ROV}(t_1) \cdot (\Delta\mathbf{P}_{ROV}(t_1 - t_0))^{-1} \cdot \mathbf{R}_6^0 = \mathbf{P}_{TIP} \end{aligned} \quad (2)$$

This idealized equation will not be directly applicable in practice. The exact measurement of vehicle movement is the subject of intensive ongoing research, but to date the results are of rather poor long-term stability. The consequence is that $\Delta\mathbf{P}_{ROV}$ is not exact enough to allow such direct compensation of vehicle movement. The approach used in the CManipulator project to meet this problem is visual servoing.

*1.3. Visual Servoing*

The basic idea of visual servoing is to use the visual information from a camera system to control a manipulator's movement. A good overview of basic visual servoing methods can be found in [7]. Usually, the goal is to change the end-effector's position relative to the object of interest (e.g. moving closer in order to manipulate it). Visual servoing assumes that the object of interest has already been detected in the camera image and will stay recognizable during the whole visual servoing process. This already poses an aggravated problem in the underwater environment, as described in section 2. Having surpassed this challenge the second problem is the dynamics of the system. Usually the object will not only move because of manipulator movement, but also because of vehicle or environment movement. Fortunately for a eye-in-hand-system this poses only minor problems, as the necessary transformation usually is not calculated for execution in one single step, but iteratively optimized in a closed-loop control fashion. This way small vehicle/object movements are compensated directly by the visual servoing algorithm.

A close analysis of a common underwater manipulation task, i.e. gripping of an object in the field of view, yields another problem: the object might be out of manipulator reach, or become out of manipulator reach due to vehicle movement during the servoing process. To prevent this from happening, not only the manipulator has to be controlled, but the vehicle as well. This would involve using a second visual servoing control loop which moves/keeps the vehicle in a desired position relative to the object, while the object is gripped by the manipulator. First steps into this direction were already made in [2]. Up to this point the two systems were operated separately, but for the future a coupled system is planned. The necessity for this coupling arises from the problem, that a system with two actively controlling components, which are connected in a kinematic chain as in the vehicle-manipulator system, tend to over-compensate: the manipulator-visual servoing loop may issue a movement command along e.g. the x-axis. At the same time the vehicle loop may issue a similar command along the same axis, resulting in a situation where the wrist-camera of the manipulator looses the object due to much larger movement than estimated. This can either be addressed by very small control steps, simply minimizing the probability of the occurrence of such a situation, or by the aforementioned coupling of the two systems. Since smaller control steps would need higher control rates as available in the Orion7P system, the second approach will be investigated.

Finally another underwater-specific problem has to be addressed for visual servoing. The used system was characterized with the term eye-in-hand system, which computationally is absolutely correct. Unfortunately the design of the Orion7P does not allow the mounting of the wrist-camera in a place where the object of interest will be visible through the whole visual servoing process. The best way to compensate for this would be the use of a second wrist-camera, which is not applicable due to a number of reasons. The only currently available solution is a visual servoing up to a certain point, and a subsequent teached-in final movement. This last movement should be as small as possible to minimize changes in the system's state. In our hitherto experiments this last correctional movement has not posed major problems.

## 2. Underwater Computer Vision

The underwater environment poses some challenges with respect to computer vision. For example, the optical properties of water are dependent on various parameters like salinity, temperature, and pressure (depth). Changes in these parameters cause a change in the density of the water which modifies the projection of the image in the optical system, e.g., at a depth of several thousand meters in cold saltwater, objects appear to be closer to the camera.

Another optical effect is caused by the different absorption rates of the particular wavelengths of light in water. In the visible spectrum the absorption rate rises with the wavelength, i.e. the red spectrum is absorbed first, then the green spectrum and after that the blue spectrum [11]. This effect causes the perceived color of an object to change depending on the distance to that object.

A third challenging property of the underwater environment is water turbidity. The turbidity reduces the visibility range and the contrast of the images. In addition, the lights used to illuminate the field of vision of an underwater vehicle are usually located near the cameras due to mechanical restrictions. Due to the small angle between the light and camera axis, most of the light is directly reflected back by small particles in the water (i.e. marine snow, etc.), which further degrades the visibility conditions.

Further details on these challenges in underwater computer vision are given in e.g. [13] and [14].

### *2.1. Stereo Camera System*

In addition to the wrist-mounted camera a stereo camera system is used in the CManipulator project. The two high-quality digital cameras are capable of 200 FPS noise-free image transmission at their full resolution. This enables very good image qualities even in low lighting conditions often encountered in underwater environments. This stereo camera system allows 3D-reconstruction of object positions in observed scenes. In order to make these calculations the system has to be calibrated both intrinsically and extrinsically. Intrinsic calibration includes the determination of the camera-lens parameters, e.g. distortion, for each individual camera. Extrinsic parameters describe the camera's position and orientation. Since the optical properties of water change with its chemical composition and pressure as described above, the intrinsic calibration has to be conducted just before the usage of the system, on site. The same applies to the extrinsic calibration, on the one hand since it relies on the intrinsic calibration, on the other hand since its accuracy depends on an unchanged system after calibration, which would not be certain if conducted prior to submersion.

The conclusion is, that a method had to be developed which calibrates the camera system at any specified time without external operator actions necessary. Our solution involves a tool-change, where the Orion7P grabs a 2D calibration rig and moves it into the FOV of the cameras in a pre-defined fashion. The images acquired in this fashion are then calibrated using the algorithm described in [12]. The result is a fully calibrated camera system with all the benefits described above.

### *2.2. Object Recognition*

In one of the application scenarios of the CManipulator project the operator of the system selects the object which should then be picked up. The operator selects the object by clicking on it in one of the images coming from the stereo vision system. The object then is searched in the second image by scanning along the corresponding epipolar line based on the previously described camera calibration. With these object positions in the images, a position in 3D space is calculated and the robotic arm is moved to an intermediate position above the object, so that the object is visible in the image coming from the wrist-camera. The object is then tracked during the visual servoing process (s. section 1.3).

**Figure 3.** The Orion 7P in the Underwater Testbed of the DFKI-Lab Bremen.

In the current stage of development, the object which is picked up is a transponder mock-up. The experimental setup is shown in fig. 3. The transponder is of cylindrical shape, 75mm in diameter, and 200mm in length. It has a yellow-colored body with black caps at both ends. Although this object seems to be easily identifiable in an image, the different perspectives, varying lighting conditions, and the usage of different cameras make it quite hard to find a stable detection method. Our current approach scans the area with a scan grid to detect the borders of the transponder. To detect the borders, an unsharp mask filter is applied to the scan line prior to the use of a standard edge detection filter. The collected border points are then checked against the shape properties of the transponder, e.g., parallel edges and aspect ratio. From the remaining set of points the orientation and size of the transponder is extracted. This straightforward approach turns out to be very stable under various lighting conditions and on different camera systems.

However, the final object recognition approach will use a more sophisticated method which allows the detection and tracking of arbitrary objects. The particular algorithms were developed in a parallel project which implements a vision-based station-keeping on a ROV [2]. The approach will search for interesting features in the region surrounding the point where the operator clicked on the target object. The interesting features are detected by an improved version of the Harris corner detector [8]. The features are then described with the SIFT descriptor [10], and subsequently the corresponding set of features is searched along the corresponding epipolar line in the second image. After the robotic arm has reached the intermediate position above the object, the set union of features from both stereo camera images is used to identify the object in the wrist-camera image. Once the object is detected, the features of the object are tracked using a pyramidal implementation of the Lucas Kanade feature tracker [9].

## 3. Conclusion and Outlook

In this paper we presented an underwater manipulation system which is, at the current stage of development, capable of autonomously grasping a transponder-like object in a static underwater environment where neither the base of the robotic arm nor the object are moving. In the next development stage we will enhance the system to be capable of grasping various objects in a dynamic underwater environment. In this *dynamic* case, the base of the robotic arm as well as the target object will move. Furthermore we will introduce the concept of skills to represent other common underwater manipulation tasks, e.g., opening/closing of valves or coupling of underwater connectors.


## Acknowledgements

This work is funded in the CManipulator project by the German Ministry of Economics (BMWI), grant number 03SX231


## References


[1] D. Spenneberg and J. Albiez and F. Kirchner and J. Kerdels and S. Fechner, *CManipulator: An Autonomous Dual Manipulator Project for Underwater Inspection and Maintenance*, Proceedings of OMAE 2007 ASME 2007 International Conference on Offshore&Mechanics and Arctic Engineering June 10-15, 2007, San Diego, USA.

[2] J. Kerdels and J. Albiez and F. Kirchner, *A Robust Vision-Based Hover Control for ROV*, Proceedings of Oceans 2008 April 8-11, 2008, Kobe, Japan.

[3] M. Hildebrandt and J. Albiez and F. Kirchner, *Computer-Based Control of Deep-Sea Manipulators*, Proceedings of Oceans 2008 April 8-11, 2008, Kobe, Japan.

[4] A.J. Healey and A.M. Pascoal and F. Lobo Pereira, *Autonomous underwater vehicles for survey operations: theory and practice*, Proceedings of the American Control Conference 1995 June 21-23, 1995, Seattle, USA.

[5] G. Chesi and K. Hashimoto, *Static-eye against hand-eye visual servoing*, Nippon Robotto Gakkai Gakujutsu Koenkai Yokoshu Vol. 19, 2001, Japan.

[6] Schilling Robotics, *Orion7P Technical Manual*, 1st ed. Davis, CA: Schilling Robotics 2007.

[7] F. Chaumette and S. Hutchinson, *Visual Servo Control*, IEEE Robotics and Automation Magazine, vol. 13, no. 4, 2007.

[8] C. Harris and M. Stephens, *A combined corner and edge detector*, Proceedings of The Fourth Alvey Vision Conference, 1988, S. 147-151

[9] J.Y. Bouguet, *Pyramidal Implementation of the Lucas Kanade Feature Tracker*, Intel Corporation Microprocessor Research Labs, 2000

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

[11] R. M. Pope and E. S. Fry, *Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements* Applied Optics, vol. 36, pp. 8710-8723, 1997.

[12] Z. Zhang, *A flexible new technique for camera calibration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.

[13] G.L. Foresti, *Visual inspection of sea bottom structures by an autonomous underwater vehicle*, IEEE Trans. Syst. Man and Cyber, Part B 31, 691-705, 2001.

[14] Y.Y. Schechner, N. Karpel, *Clear Underater Vision*, Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004.

[15] J. Denavit and R. S. Hartenberg,*A Kinematic Notation for Lower- Pair Mechanisms Based on Matrices*,Journal of Applied Mechanics, Trans. Of the ASME, June 1955.