2004 8th International Conference on
Control, Automation, Robotics and Vision
Kunming, China, 6-9th December 2004

Sonar and Vision based Navigation Schemes for Autonomous Underwater Vehicles

Bharath Kalyan*, Arjuna Balasuriya*, Tamaki Ura†, Sardha Wijesoma*

*School of EEE, Nanyang Technological University, Singapore Nanyang Avenue, Singapore 639798
†URA Laboratory, Underwater Technology Research Center, IIS, University of Tokyo, Japan
Email: bharathkalyan@pmail.ntu.edu.sg, earjuna@ntu.edu.sg, eswwijesoma@ntu.edu.sg, ura@iis.u-tokyo.ac.jp

Abstract

In this paper, an autonomous underwater vehicle (AUV) navigation schemes are proposed using Forward Looking Sonar (FLS) and Charged Coupled Device (CCD) camera for near bottom applications. Scans obtained from the onboard FLS are processed using a feature extraction technique discussed in the paper to extract stable point features in the environment. A technique is also proposed for AUVs to track optical features. Through field trials conducted using the test-bed vehicles NTU-UAV and the Twin-Burger2, the performance of the algorithms proposed are analyzed. The paper discusses the above-mentioned individual sensor capabilities and drawbacks in AUV navigation.

1 Introduction

In order to explore the vaguely understood, wide and deep underwater environment, which covers two thirds of our planet, it is necessary to navigate to different locations. The unmanned untethered vehicles called Autonomous Underwater Vehicles (AUVs) have the potential to reach the deep and shallow underwater environments. The development of fully autonomous underwater vehicles (AUVs) is a challenging task due to the highly hostile nature of the environment. It is necessary to navigate to an unknown unstructured underwater environment with limited sensing technologies. Deficiencies of current underwater sensing technologies are well documented, and include lack of resolution and accuracy, high cost, operational complexities etc [1]. AUVs at present are not capable of navigating without any external assistance. Underwater navigation is currently done, using positioning schemes such as long baseline (LBL), ultra-short baseline (USBL) or GPS- buoys. The deployments of these systems are highly cumbersome and expensive.

This paper presents the on going work on sector scan sonar and vision based data processing algorithms to be used in Autonomous Underwater Vehicles (AUVs). These algorithms are developed with a view of building a terrain map and navigating the vehicle concurrently in unstructured underwater environments using onboard sensors. The experiments were performed using the test bed platforms NTU-UAV [2] and Twin- Burger 2 shown in Figure 1.

In section 2 a robust feature extraction algorithm

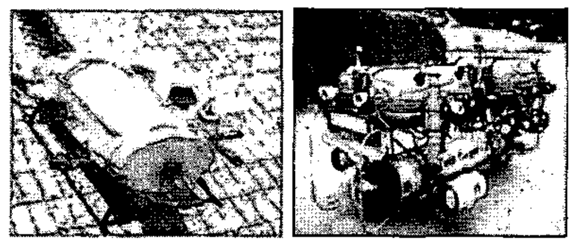

Figure 1: Test bed platforms: NTU-UAV and Twin-Burger2

using clustering techniques is discussed. Results based on experimental data are also presented. Section 3 summarizes the extended Kalman filter (EKF) based stochastic positioning algorithm. Section 4 presents a technique to track the underwater cables in harsh environments using optical vision based systems. Finally Section 5 concludes the paper by summarizing the results along with advantages and drawbacks on the techniques presented in this paper followed by research directions and future work.

2 Sonar Feature Extraction and Navigation

There has been a great deal of research undertaken in developing techniques for automatic feature detection and classification of sonar data, generally with limited success [3] [4].

The development of autonomous feature based navigation scheme relies on the ability of the system to extract appropriate and reliable feature. In this paper, point features are identified from the sonar scans returned by the imaging sonar and are used to build up a feature map of the environment. A technique has been proposed for extracting the point features from the sonar scan returns. A flow chart explaining the algorithm is presented in Figure 3

The data returned by the imaging sonar consists of the complete time history of each sonar ping in a discrete set of bins scaled over the desired range. The first task in extracting reliable feature is to identify the principal return from the ping data. Figure 2 shows a pool environment with targets and its scan from sector scan sonar. The sonar targets produce strong sonar returns that can be characterized as point features for the purpose of mapping.

The principal returns are extracted based on a pre-

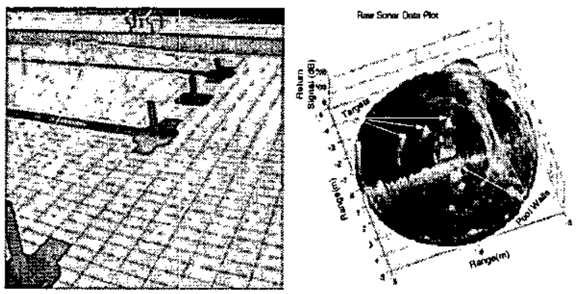

Figure 2: Experimental site and its Sonar Scan

set threshold under the assumption that, the amplitude of the returns is not higher than that of the multi-path principal return. Adaptive thresholding techniques using CFAR based techniques can also be employed for the extraction of the principle returns [5]. The cross talk and the multi-path effects

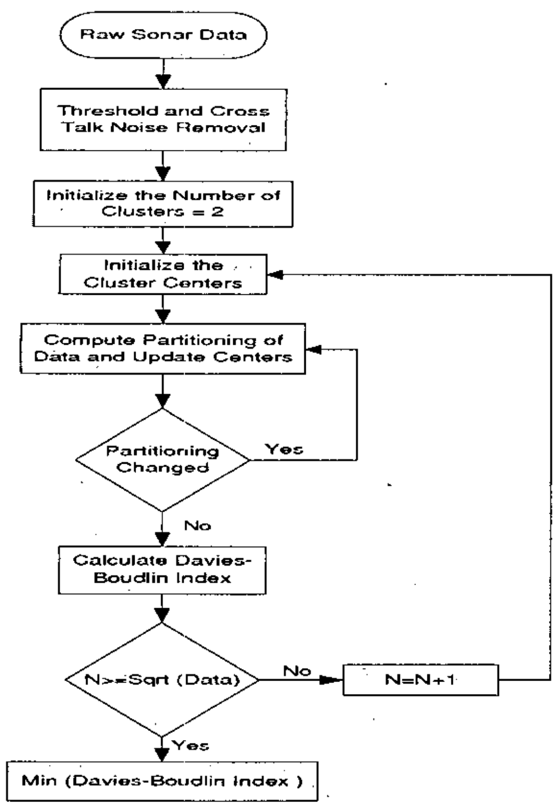

Figure 3: Flowchart indicating various stages of the Feature Extraction process

are clearly visible in Figure 2, which is removed based on a known preset threshold. Following this is the cluster formation, where the data is clustered based on a distance criterion explained in detail in the following section.

2.1 Cluster formation

There are several techniques available for cluster formation; and the partitive clustering techniques are employed here. It divides a data set into a number of clusters, typically trying to minimize an error function. The number of clusters is usually predefined, but it can also be a part of error function. The algorithm consists of the following steps [6].

Step1: *Determine the number of clusters*
Step2: *Initialize the cluster centers*
Step3: *Compute partitioning for data*
Step4: *Compute (update) cluster centers*
Step5: *Partitioning changed(or the algorithm has not converged), return to step3 else stop.*

If the number of clusters is unknown, the partitive algorithm can be repeated for a set of different number of clusters, typically from two to $\sqrt{N}$, where N is the number of samples in the data set. An example of a commonly used partitive algorithm is the k-means, which minimizes the error function given in Equation (1)

$$E = \sum_{k=1}^{C} \sum_{x \epsilon Q_k} \|x - z_k\|^2 \tag{1}$$

Where C is the number of clusters, z_k is the center of cluster k. K-means clustering technique is employed here for clustering the returns of the sonar scan data. The k-means method aims to minimize the sum of squared distances between all points and the cluster center. It is a known fact that, K-means requires the number of clusters to be produced as an input. This is of utmost concern, as this requires some prior knowledge of the number of clusters present in the data, i.e., number of targets present in one sonar scan, which, in practice, is highly unlikely. In order to determine appropriate number of clusters present in the data, the clusters have to be validated based on certain indices. In this work Davies-Bouldin validity index is employed for cluster validation.

2.2 Cluster Validity measures

Many criteria have been developed for determining the cluster validity [6], all of which have a common goal to find the clustering which results in compact clusters which are well separated. In this work, the Davies - Bouldin index [7] has been employed. According to this index, the best clustering minimizes [7]

$$\frac{1}{C} \sum_{k=1}^{C} \frac{S_c(Q_k) + S_c(Q_j)}{d_{ce}(Q_k, Q_j)} \tag{2}$$

Where, C is the number of clusters, $d_{ce}(Q_k, Q_j) = \|z_k - z_j\|$ is the Euclidean distance between cluster Q_k with center z_k and Q_j with center z_j and S_c is intra cluster distance given by

$$S_c(Q_k) = \sum_{x \epsilon Q_k} \frac{\|x - z_k\|}{N_k} \quad (3)$$

Where N_k is the number of samples in cluster Q_k. Here the aim is to minimize the intra-cluster distance and to maximize the inter cluster distance measures. As it can be seen from Equation (2), the intra-cluster distance measure term is present in the numerator, and the inter-cluster distance measure term in the denominator. Thus maximizing the denominator and minimizing the numerator will lead to an index measure term which is commonly referred to as the Davis-Bouldin index [7] as given by Equation (2). Therefore, the clustering which gives the minimum value for the Davies - Bouldin validity index will indicate the ideal value of the number of clusters (C) present in the k-means procedure. The computational complexity is proportional to

$$\sum_{k=2}^{C_{max}} Nk \quad (4)$$

Where $C_{max} = \sqrt{N}$ is the maximum number of clusters. Thus from the Davies-Bouldin validity index

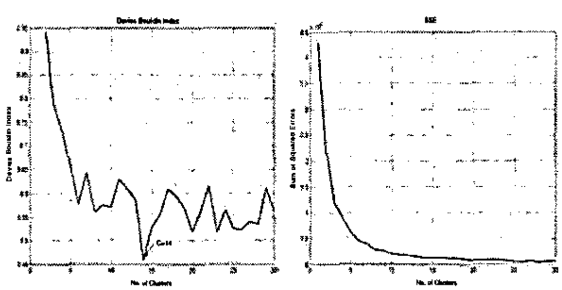

Figure 4: Davies Bouldin index and The Error function

plot in Figure 4, it can be seen that the index is minimum at $C = 14$ given an indication of possible compact clusters. The error function plot shown in Figure 4, gives an estimate of the sum of squared errors (SSE) as explained in Equation (1) and can be observed that the error function is minimized when the number of clusters are optimal. The output of the clustering algorithm returns clusters with the cluster centroids, which would be as point features for further tracking. The output of the clustering algorithm indicating different clusters are superimposed on the original data as shown in Figure 5. The centroids can be used as point features, and the problem can be formulated as a feature-tracking problem. A study showing the utilization of these point features

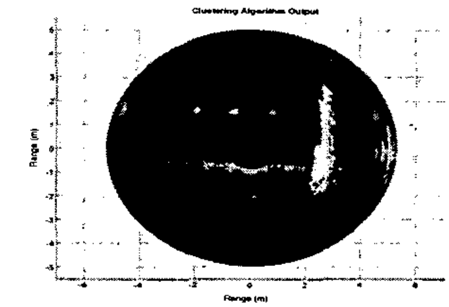

Figure 5: Clustering Algorithm output superimposed on threshold Sonar data

for AUV localization is discussed in the following section.

3 EKF Based Stochastic Localization And Map Building

A series of seminal papers by [8] introduced a powerful statistical framework for concurrently solving the mapping problem and the induced problem of localizing the vehicle relative to its growing map. One family of probabilistic approaches employ Kalman filters to estimate the map and the vehicle location [8] [9]. It consists of a recursive three stage process comprising prediction, observation and update.

In the current implementation, the vehicle pose is made up of two dimensional positions (x_v, y_v) and orientation ψ_v of the vehicle. An estimate of vehicle ground speed V_v, slip angle γ_v, and the gyros rate bias $\dot{\psi}_{bias}$ is also generated by the algorithm. The slip angle γ_v is the angle between the vehicle axis and the direction of the velocity vector. Although the thrusters that drive the vehicle are oriented in the direction of the vehicle axis, the slip angle is often non zero due to disturbances caused by the ocean currents, and wave effects. The positioning filter estimates the vehicle position (x_v, y_v), orientation ψ_v, velocity V_v. The landmarks tracked in this implementation of the algorithm are assumed to be point features. Thus $X_v(k)$ consists of six states that describe the state of the vehicle

$$X_v(k) = [x_v(k), y_v(k), \psi_v(k), V_v(k), \gamma_v(k), \dot{\psi}_{bias}(k)] \quad (5)$$

The augmented state matrix the vehicle state along with N states that describe the position of the N observed landmarks $x_i(k), i = 1, 2...N$.

$$X(k) = [X_v(k)X_i(k)]^T \quad (6)$$

In this implementation, depth information is kept separate from the position and orientation of the vehicle. Though it is not an entirely accurate reflection of the motion of the vehicle in the underwater domain but serves to allow the algorithms to be developed and tested. For many of the missions, this model is acceptable given that the vehicle will generally be operating in close proximity to the sea floor.

3.1 Prediction

The prediction stage uses a model of the motion of the vehicle to update the predicted vehicle position. A discrete form of the vehicle model is used to predict the state of the vehicle $x_v(k)$ given the previous state $x_v(k-1)$.This defines the discrete, non-linear vehicle prediction equation,

$$X_v(k) = f_v(X_v(k-1), u(k)) \tag{7}$$

3.2 Observation

Point features are currently tracked, which could be extracted from the sonar scans using the feature extraction technique discussed above. Given the current vehicle position and the position of the observed feature $x_i(k)$, the observation, consisting of range $z_R(k)$, and bearing $z_\theta(k)$, can be modeled as

$$\begin{aligned} z(k) &= \begin{bmatrix} z_R(k) \\ z_\theta(k) \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{(x_v(k)-x_i(k))^2 + (y_v(k)-y_i(k))^2} \\ \tan^{-1}\left(\frac{y_v(k)-y_i(k)}{x_v(k)-x_i(k)} - \psi_v(k)\right) \end{bmatrix} \\ &\quad + \begin{bmatrix} w_r(k) \\ w_\theta(k) \end{bmatrix} \end{aligned} \tag{8}$$

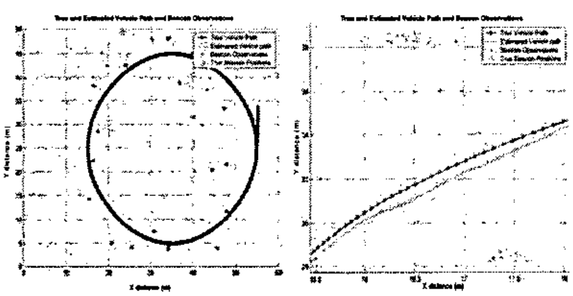

Figure 6: Estimated vs. True vehicle trajectory and its zoom in on a fragment

3.3 Update

For the work reported here, the Extended Kalman Filter (EKF) is employed to estimate the pose of the vehicle $\hat{x}_v(k|k)$ [9]. The innovation function consists of the difference between the predicted and the actual observation. The features that are observed must be associated with the features in the map. The data association technique that is currently employed here is the gated nearest neighbor approach [9]. A plot showing the estimated and true trajectory is presented in the Figure 6.

The disadvantage of sonar is the fact that it will generally have rather poor angular accuracy. An underwater (monocular) camera has a better resolution than sonar and can give much more details of the observed objects.

4 Optical Vision Based Cable Tracking

This section discusses the implementation of the navigation system for underwater cable tracking using optical vision system. The application of underwater vision is greatly limited by the poor image quality due to forward scattering and backscattering and marine snow [10]. To facilitate the application of available feature extraction techniques, the initial image should be restored to get rid of the unwanted components.

4.1 Edge map with spatial band-pass Filters

It has been determined that scatter components have the characteristic of a low-pass filter, which spatially attenuates the high frequencies in the interested object. Low spatial frequencies account for the slowly varying gray levels in an image such as the variation of intensity over the continuous surface (scattering effect). High frequency components are associated with"quickly varying" information such as edges due to non-uniformity of lighting, marine snow etc. [10]. The unwanted features mentioned above fall into the low and high spatial frequency domains. Therefore a band pass filter has to be designed to filter out the unwanted components [10] [11]. The image is convolved with a Laplacian of Gaussian (LoG) mask, which serves as a band pass spatial filter. LoG filter in two dimensions is given by

$$\nabla^2 G(x,y) = \frac{1}{2\Pi\sigma^4}\left[2 - \frac{x^2+y^2}{\sigma^2}\right] e^{-\left(\frac{x^2+y^2}{\sigma^2}\right)} \tag{9}$$

where σ is the space constant of the Gaussian and (x, y) is the pixel position in the image. From the theoretical analysis [10], it was derived that the lower cutoff frequency introduced by the LoG filter is $W = 2\sqrt{2}\sigma$ and the higher cutoff frequency is the size of the kernel $3W$. The LoG filter extracts edges with size in the range of $W \sim 3W$. A sample of a raw image and its convolution with a LoG kernel is shown in Figure 8.

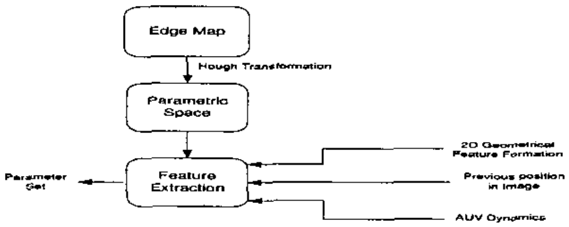

Figure 7: Feature Extraction

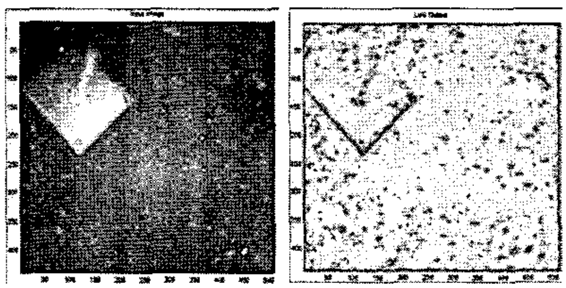

Figure 8: Input image $I(x, y)$ and image after convolution with LoG mask indicating the edge map

4.2 Grouping and Feature Extraction

The feature extraction process is highly dependent on the kind of application for which the vision is employed. In this paper, it is assumed that the objects of interest in the underwater images are the man-made objects, which consist of well defined geometric features. The edge map generated as explained in section 4.1, represents all the geometrical features in the vicinity, which means that there is too much data to be processed. Since the desirable features are definable, their geometrical parameters are known. Therefore the grouping is done by transforming the edge map into a parametric space which defines the geometric features of the interested object in the vicinity as shown in Figure 7. The Hough transformation technique is used to its simplicity in transformation and also due to the interest in underwater objects such as pipelines, cables, structures which consist of straight line geometrical edges. Therefore, the main concern is to extract straight line geometric features from the visual image and to group them in connected sets to find the most suitable set representing the interested object being observed, for navigation. With transformation into the parametric space, feature extraction is possible by looking into the special geometrical features of the interested object as shown in Figure 7. For real time performance, the entire parametric space is not searched, but a rough estimation is done by looking at the dynamics of the AUV and previous parametric data to know the whereabouts of the object in the frame and its corresponding parameters as represented in Figure 9. If a peak representing the cable cannot be found in the region predicted, the features of the model line are used for the navigation of the AUV. This avoids the misinterpretation of similar features in the image. If the cable is not detected in this region then, the navigation is carried out by following the model line [12]. The model

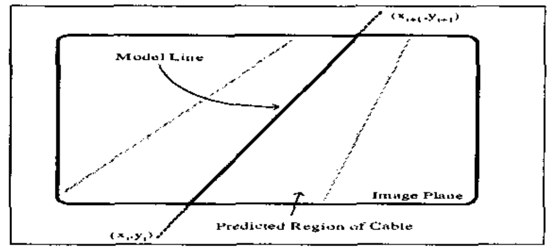

Figure 9: Predicted Region in the Image

line features $(\rho, \theta)$are used in the Hough plane too as shown in Figure 10. The cable image introduces a high concentration of pixels in a particular direction forming a line in the image plane. This line feature introduces a peak in the Hough space and its existing region can be predicted using the uncertainty of the model line as shown in Figure 10. This technique avoids the extraction of other peaks in the Hough plane. In other words, it is possible to distinguish the cable of interest even when there are similar cables appearing in the image. Once the instantaneous

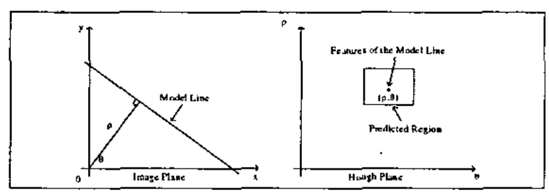

Figure 10: Region Estimation in the Hough Plane

features of the cable are determined, these features are verified to check whether they represent the cable. If the features extracted by image processing represents the properties of the cable then those 2D features are transformed into the vehicle coordinate system for determining the navigational references. Else the model line features are used for determining the navigational references. The generation of navigational references is discussed in [10].

In order to demonstrate the performance of the proposed feature extraction algorithm an experiment was carried out using a test bed AUV, Twin-Burger 2. In this experiment, a yellow colored hose pipe was used as the underwater cable to be tracked as shown in Figure 11. The dark dotted lines superimposed on the images as shown in Figure 12, indicate the position of the cable detected by the vision system.

Figure 11: Underwater Cable at Lake Biwa

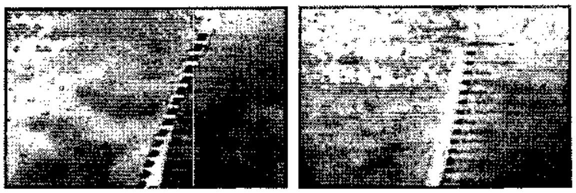

Figure 12: Underwater Cable Detection

5 Conclusions

Techniques have been presented for extracting features from the sonar scan returns and optical images. A simulation of underwater navigation using point features extracted from sonar data returns has also been demonstrated.

A close look at the properties of sonar and a monocular camera clearly shows the complementary characteristics of the sensors. Sonar has good range properties, both in terms of actually measuring distances and (dependent on the frequency) good range resolution, which is indeed important in determining slow relative motion as described in section 2. Further the sonar is well suited for coarse mapping of larger areas and objects. The disadvantage of sonar is that it has rather poor angular accuracy. On the other hand an underwater (monocular) camera has a better resolution than sonar and can give much more details of the observed objects. The angular accuracy is better and translational motion can be relatively easily detected. A camera is also able to detect geometric properties as well as optical properties such as texture, and color, which is totally invisible to sonar. But, on the other hand, the range of underwater vision is limited due to lack of visibility in water. Having looked at individual sensor performances based on experimental results, this paper concludes that, the information from multiple sensors needs to be combined for more reliable position estimation and robust underwater navigation.

Acknowledgements

The authors would like to thank Ship Research Institute, and Biwako Research Institute, in Japan for allowing us to use their facilities for these experiments.

References

[1] Leonard.J.J, Bennett.A, Smith.C.M, and Feder.H.J.S. Autonomous underwater vehicle navigation. In *MIT Marine Robotics Laboratory Technical Memorandum*, 1998.

[2] Balasuriya.A, Wijesoma.S, BharathKalyan, and ThomasLim. Development of a test-bed auv: The ntu-uav. In *Second International Conference on Computational Intelligence Robotics and Autonomous Systems(CIRAS)*, Singapore, 2003.

[3] Brutzman.D.P, Compton.M.A, and Kanayama.Y. Autonomous sonar classification using expert systems. In *IEEE Oceans*, Newport, RI, 1992.

[4] Lane.D.M and Stoner.J.P. Automatic interpretation of sonar imagery using qualitative feature matching. In *IEEE Journal of Oceanic Engineering*, volume 19(3), pages 391–405, 1994.

[5] BharathKalyan and Balasuriya.A. Sonar based automatic target detection scheme for underwater environments using cfar techniques: A comparative study. In *IEEE Underwater technology*, Taiwan, 2004.

[6] Juha.V and Alhoniemi.E. Clustering of the self-organizing map. In *IEEE Transactions on Neural Networks*, volume 11, May 2000.

[7] Davies.D.L and Bouldin.D.W. A cluster separation measure. In *IEEE Transactions Pattern Analysis and Machine Intelligence*, volume Vol. PAMI-1, pages 224–227, April 1979.

[8] Smith.R, Self.M, and Cheeseman.P. Estimating uncertain spatial relationships in robotics. In I.J. Cox and G.T. Wilfong, editors, *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag, 1990.

[9] Barshalom.Y and Fortman.T.E. *Tracking and Data Association.* Academic Press, 1988.

[10] Balasuriya.B.A.A.P. *Computer Vision for Autonomous Underwater Vehicle Navigation.* PhD thesis, Dept. of Naval Architecture and Oceanic Enginerring, University of Tokyo,Japan, 1998.

[11] Marks.R.L, Rock.S.M, and Lee.M.J. Real-time video mosaicking of the ocean floor. In *IEEE Journal of Oceanic Engineering*, volume 20, pages 229–241, 1995.

[12] Balasuriya.B.A.A.P and Ura.T. Vision based autonomous underwater vehicle navigation: Underwater cable tracking. In *OCEANS IEEE/MTS*, Nova Scotia, Canada, October 1997.