A new approach for Visual Underwater Mapping using Topological Shell Maps

Silvia Silva da Costa Botelho and Paulo Drews Jr and Gabriel Leivas

Abstract— **The use of Autonomous Underwater Vehicles (AUVs) for underwater tasks is a promising robotic field. These robots can carry visual inspection cameras. Besides serving the activities of inspection and mapping, the captured images can also be used to aid navigation and localization of the robots. Visual odometry is the process of determining the position and orientation of a robot by analyzing the associated camera images. It has been used in a wide variety of non-standard locomotion robotic methods. In this context, this paper proposes an approach to visual odometry and mapping of underwater vehicles. Supposing the use of inspection cameras, this proposal is composed of two stages: *i*) the use of computer vision through the algorithm SIFT to visual odometry, extracting landmarks in underwater image sequences and *ii*) the development of topological maps for localization and navigation. The integration of such systems will allow visual odometry, localization and mapping of the environment. A set of tests with real robots was accomplished, regarding online and performance issues. The results reveal an accuracy and robust approach to several underwater conditions, such as illumination and noise, leading to a promissory and original visual odometry and mapping technique.**

I. INTRODUCTION

Autonomous Underwater Vehicles can be applied to many tasks of difficult human exploration [9]. In underwater visual inspection, the vehicles can be equipped with *down-looking cameras*, usually attached to the robot structure [12]. These cameras capture images from the bottom of the ocean. In these images, natural landmarks, also called keypoints in this work, can be detected allowing the AUV visual odometry.

In navigation, classical odometry is the process of determining the position and orientation of a vehicle by measuring the wheel rotations through devices such as rotary encoders. While useful for many wheeled or tracked vehicles, traditional odometry techniques cannot be applied to robots with non-standard locomotion methods, such as AUVs. In addition, odometry universally suffers from precision problems, since wheels tend to slip and slide on the floor, and the error increases even more when the vehicle runs on non-smooth surfaces. As the errors accumulate over time, the odometry readings become increasingly unreliable.

Visual odometry is the process of determining equivalent odometry information using only camera images. Compared to traditional odometry techniques, visual odometry is not restricted to a particular locomotion method, and can be utilized on any robot with a sufficiently high quality camera.

In this paper we propose a new approach to extract and map keypoints between consecutive images in underwater environment. We use Scale Invariant Feature Transform (SIFT), which is a robust invariant method in keypoints detection [17]. Furthermore, these keypoints are used as landmarks in an online topological mapping. We propose the use of self-organizing maps (SOM) based on Kohonen maps [16] and Growing Cell Strutures (GCS) [10] that allow a consistent map construction even in the presence of noisy information.

First the paper presents related works on visual odometry and mapping. Section III presents a detailed view of our approach with SIFT algorithm and self-organizing maps, followed by the implementation, test analysis and results with different undersea features. Finally, the conclusion of the study and future perspectives are presented.

II. RELATED WORKS

Localization, navigation and mapping using vision-based algoritms use visual landmarks to create visual maps of the environment. The extent to which the robot navigates, the map grows in size and complexity, increasing the computational cost and making it difficult to process in real time. Moreover, the efficiency of the data association, an important stage of the system, decreases as the complexity of the map augments. It is therefore important for these systems, to extract a few, but representative, features (points of interest) of the environment.

The problem of extracting points of interest in image sequences has resulted in the development of a variety of keypoint detectors: Shi and Tomasi [26], SIFT [17], Speeded up robust features Descriptor (SURF) [1], affine covariant [20], etc. All these proposals are based on the same approach: extraction points representing regions with high intensity gradient. These regions are highly discriminatory and thus robust to noise and changes in illumination, point of view of the camera, etc [21]

Some approaches using SIFT for visual indoor Simultaneous Localization and Mapping (SLAM) were made by Se and Lowe[24] [25]. They use SIFT in a stereo visual system to detect the visual landmarks, together with odometry, using ego-motion estimation and the Kalman filter. The tests were made in structured environments with new maps.

Several AUVs localization and mapping methods are based on mosaics [11] [14]. *Delaunoy et al* analyse change detecton techniques available nowadays for on-land applications and propose a method to detect changes in sequences of underwater images [6]. Mahon e Williams [19] propose a visual system for SLAM in underwater environment, using the Lucas-Kanade optical filter and extended Kalman filter

All authors are from Fundação Universidade Federal do Rio Grande (FURG), Km 8, Italia Av., Rio Grande, Brazil silviacb, paulo, gabriel@ee.furg.br

1-4244-2523-5/09/$20.00 ©2009 IEEE

(EKF), with aid of a sonar. *Nicosevici et al* propose an identification of suitable interest points using geometric and photometric cues in motion video for 3D environmental modeling [21].

Booij et al. [3] has the most similar approach to this presented in this work. They do visual odometry with topological maps based on appearance. In this case, the SIFT method is used in omnidimentional images. However, this approach is validated only with mobile robots in terrestrial environment. The use of both *i.* SIFT to extract visual underwater features and *ii.* topological maps for mapping and localization on underwater environment was not found in the literature.

III. A System to Visual Odometry

Figure 1 shows an overview of the approach proposed here. First, the underwater image is captured and pre-processed for the removal of radial distortion and others distortions caused by water diffraction. With the corrected image, keypoints are detected and local descriptors for each one of these points are computed by SIFT. Each keypoint has a n dimensional local descriptors and global pose informations. A matching stage provides a set of correlated keypoints between consecutive images. Considering all correlated points found, *outliers* are removed, using RANSAC [8] and LMedS [23] algorithms.

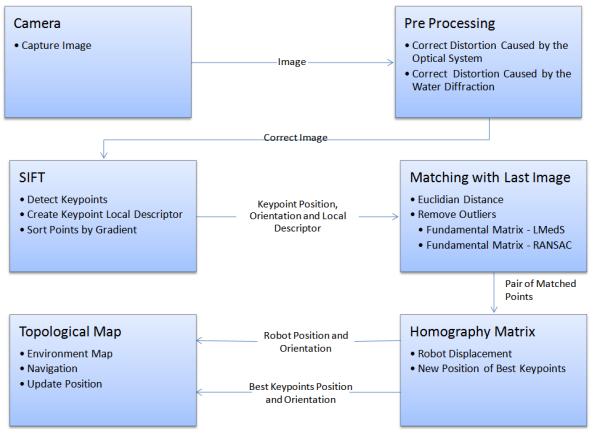

Fig. 1. Overview of the system proposed.

The relative motion between frames is estimated, using the correlated points and the homography matrix.

Moreover, the keypoints are used to create and train the topological maps. A growing cell strutures algorithm is used to create the nodes and edges of the SOM. Each node has a n-dimensional weight. After a trainning stage, the system provides a topological map, where its nodes represent the main keypoints of the environment.

During the navigation, when a new image is captured, the system calculates its local descriptors, correlating them with the nodes of the current trained SOM. To estimate the pose of the robot (center of the image), we use use the correlated points and the homography matrix concept. Thus, the global position and orientation of the center of the image is obtained, providing the localization of the robot.

Each module of the proposed approach is detailed below.

A. Pre Processing

The distortion caused by the camera lenses can be represented by a radial and tangential approximation. As the radial component causes higher distortion, most of the works developed so far only correct this component [13].

In underwater environment, there is an additional distortion caused by water diffraction. Equation 1 shows one method to solve this problem [30], where m is the point without radial distortion with coordinates (m_x, m_y), and M_0 the new point without additional distortion; u_0 and v_0 are the central point coordinates. Also, $R = \sqrt{{m_x}^2 + {m_y}^2}$ and R_0 are defined by 2 with focal distance f.

$$\begin{aligned} m_0 x &= m_x + \frac{R_0}{R}(m_x - u_0) \\ m_0 y &= m_y + \frac{R_0}{R}(m_y - v_0) \end{aligned} \tag{1}$$

$$R_0 = f \tan\left(\sin^{-1}(1.33 * \sin(\tan^{-1}\frac{R}{f}))\right) \tag{2}$$

B. SIFT

The Scale Invariant Feature Transform (SIFT) is an efficient filter to extract and describe keypoints of images [17]. It generates dense sets of image features, allowing matching under a wide range of image transformations (i.e. rotation, scale, perspective) an important aspect when imaging complex scenes at close range as in the case of underwater vision [6]. The image descriptors are highly discriminative providing bases for data association in several tasks like visual odometry, loop closing, SLAM, etc.

First, the SIFT algorithm uses the Difference-of-Gaussian filter to detect potential interesting points in a space invariant to scale and rotation. The SIFT algorithm generates a scale space $L(x, y, k\sigma)$ by convolving repeatedly an input image $I(x, y)$ using a variable-scale Gaussian, $G(x, y, \sigma)$, see eq. 3:

$$L(x, y, \sigma) \quad = \quad G(x, y, \sigma) * I(x, y) \tag{3}$$

SIFT analyzes the images at different scales and extracts the keypoints, detecting scale-invariable image locations. The keypoints represent scale-space extrema in the difference-of-Gaussian function $D(x, y, \sigma)$ convolved with the image, see 4:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{4}$$

where k is a constant multiplicative factor.

After the keypoints extraction, each feature is associated with a scale and an orientation vector. This vector represents the major direction of the local image gradient at the scale

where the keypoint was extracted. The keypoint descriptor is obtained after rotating the nearby area of the feature according to the assigned orientation, thus achieving invariance of the descriptor to rotation. The algorithm analyses images gradients in 4×4 windows around each keypoint, providing a 128 elements vector. This vector represents each set of feature descriptors. For each window a local orientation histogram with 8 bins is constructed. Thus, SIFT maps every feature as a point in a 128-dimension descriptor space.

A point to point distance computation between keypoints in the descriptors space provides the matching. To eliminate false matches, an effective method is used to compare the smallest match distance to the second-best distance [17], where through a threshold only close matches are selected.

Furthermore, *outliers* are removed through RANSAC and LMedS, fitting a homography matrix H [1]. In this paper, this matrix can be fitted by both RANSAC and LMedS methods [29]. Both methods are considered only if the number of matching points is bigger than a predefined threshold t_m.

C. Estimating the Homography Matrix and Computing the Camera Pose

We use the homography concept to provide the camera pose. A homography matrix H is obtained from a set of correct matches, transforming homogeneous coordinates into non-homogeneous. The terms are operated in order to obtain a linear system [15], considering the keypoints $(x_1, y_1), .., (x_n, y_n)$ in the image I and $(x_1\prime, y_1\prime), .., (x_n\prime, y_n\prime)$ in the image $I\prime$ obtained by SIFT.

The current global pose of the robot can be estimated using 5, where ${}^1H_{k+1}$ is the homography matrix between image I_1 in the initial time and image $I_k + 1$ in the time $k + 1$. The matrix ${}^1H_1$ is defined by the identity matrix $3x3$ that consider the robot in the beginning position $(0, 0)$.

$$ {}^1H_{k+1} = \prod_{i=1}^{k} {}^iH_{i+1} \tag{5} $$

Thus, the SIFT provides a set of scale invariant keypoints, described by a feature vector. A frame has a m keypoints, and each keypoint, X_i, has 128 features, $f_1, ..., f_128$ and the pose and scale (x, y, s):

$$ X_i = f_1, f_2, ..., f_{128}, x, y, s, i = 1, .., m \tag{6} $$

These m vectors are used to obtain a topological map, detailed in the next section.

D. Topological Maps

In this work, the vectors extracted from SIFT feature are used to compose a topological map. This map is obtained using a self-organizing mapping (SOM) based on Kohonen Neural Networks [16] and the Growing Cell Structures (GCS) method [10]. Like most artificial neural networks, SOMs operate in two modes: training and mapping. Training builds the map using input examples. It is a competitive process, also called vector quantization. A low-dimensional (typically two dimensional) map discretizes the input space of the training samples. The map seeks to preserve the topological properties of the input space. A structure of this map consists of components called nodes or neurons. Associated with each node is a weight vector of the same dimension as the input data vectors and a position in the map space. Nodes are connected by edges, resulting in a (2D) grid.

a) Building the map: Our proposal operates in Scale Invariant Feature vectors Space, SIFT space, instead of image space, in other words, our space has $n = 131$ values (128 by the SIFT's descriptor vector and 3 by the feature's pose). A Kohonen map must be created and trainned to represent the space of descriptors. To build the map, feature vectors are presented to the SOM. The learning algorithm is based on the concept of nearest-neighbor learning using KD-Tree algorithm [17]. When a new input arrives, the topological map determines the feature vector of the reference node that best matches the input vector. As our system uses several feature vectors associated with each captured image, the nearest-neighbor algorithm is applied to each feature vector separately. The results of the nearest-neighbor algorithms are combined with a simple scheme based on unanimous voting.

The Growing Cell Structures method allows the creation and removal of the nodes during the learning process. The algoritm constrains the network topology to k-dimensional simplices whereby k is some positive integer chosen in advance. In this work, the basic building block and also the initial configuration of each network is a $k = 2$-dimensional simplex. For a given network configuration a number of adaptation steps are used to update the reference vectors of the nodes and to gather local error information at each node. This error information is used to decide where to insert new nodes. A new node is always inserted by splitting the longest edge emanating from the node q with maximum accumulated error. In doing this, additional edges are inserted such that the resulting structure consists exclusively of k-dimensional simplices again.

After a set of training steps, the kohonen map represents the descriptors space. This SOM can be used to locate the robot during the navigation.

b) Location of the robot on the Map: New frames are captured during the navigation. For each new frame F, SIFT calculates a set of m keypoints X_i, see equation 6. A $n = 131$ dimensional descriptor vector is associated to each keypoint. We use the trainned SOM to map/locate the robot in the environment. A mapping stage is runned m times. For each step i there will be one single winning neuron, N_i: the neuron whose weight vector lies closest to the input descriptor vector, X_i. This can be simply determined by calculating the Euclidean distance between input vector and weight vectors. After the m steps we have a set of m winner nodes, N_i, associated with each feature descriptor, X_i. With m pairs (X_i, N_i), we can use the homography

[1]We suppose a plannar motion - the altitude of the vehicle makes the relief differences of the scene neglectable. On the other hand if the scene were 3D, we could use a fundamental matrix F for outliers methods and structure-from-motion to compute the pose of the camera.

TABLE I

UNDERSEA FEATURES FOR EACH DISTORTION USING IN THE TESTS.

Distortion	1	2	3	4	5
Light Source Distance (m)	0.2	0.22	0.25	0.25	0.3
Attenuation Value (%)	0.05	0.05	0.06	0.05	0.05
Gaussian Noise (σ)	2	2	2	4	4
Gray Level Minimum	20	30	20	20	20
Number of Flakes of Snow Marine	30	30	30	30	30

concept to obtain a linear matrix transformation, H_{SOM}. Equation 7 gives the map localization of the center of the frame, $X_C\prime = (x_c\prime, y_c\prime)$:

$$X_C\prime = H_{SOM} * X_C, \tag{7}$$

where X_C is the position of the center of the frame. Moreover the final topological map allows the navigation in two ways: through target positions or visual goals. From the current position, graph search algorithms like Dijkstra [7] or A^* algorithm [5] can be used to search a path to the goal.

IV. SYSTEM IMPLEMENTATION, TESTS AND RESULTS

In this work, the robot presented in figure 2 was developed. This robot is equipped with a Tritech Typhoon Colour Underwater Video Camera with zoom, a MiniKing sonar and a set of sensors (altimeters and accelerometers) [4].

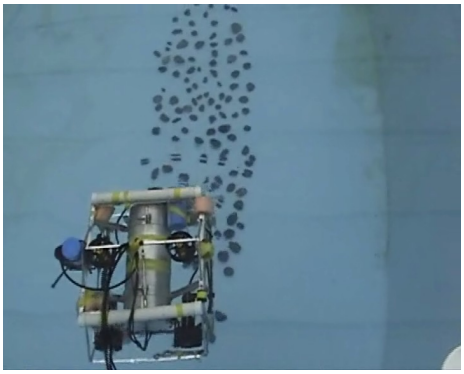

Fig. 2. ROVFURGII in test field.

The visual system was tested in a desktop Intel Core 2 Quad Q6600 computer with 2Gb of DDR2-667 RAM. The camera is NTSC standard using 320x240 pixels at a maximum rate of 29.97 frames per second.

Different undersea features were applied in the images, like turbidity, sea snow, non-linear illumination, and others, simulating different underwater conditions. Table I shows the applied features.

A. The method in different underwater features

The visual system was tested in five different underwater environments, corresponding the image without distortion and first four filters presented in table I (the effects were artificially added to the images) . Figure 3 enumerates the detected and matching keypoints obtained in a sequence of visual navigation. Even though the number of points and correlations has diminished with the quality loss because of underwater conditions, it is still possible to localize the robot, according to figure 4. In this figure, the motion referential is represented in blue, executed by a robotic arm composed by an *harmonic drive* actuator with a coupled encoder supplying angular readings in each 0.651 ms, with a camera coupled to this. It is possible to see that the approach proposed is robust to underwater environment changes. All graphics shown in this paper use centimeter as metric unit, including figure 4.

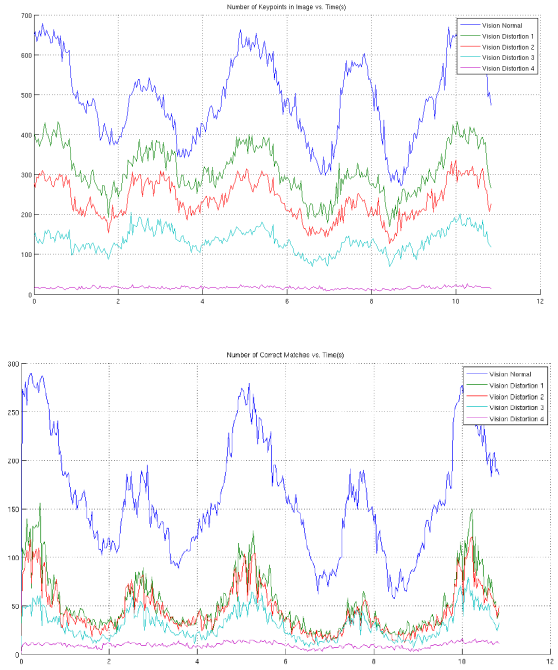

Fig. 3. Number of keypoints detected and true correlation during the robotic arm movement.

B. Online Robotic Localization

Some tests were performed to evaluate the SIFT algorithm performance considering a comparison with another algorithm for robotic localization in underwater environment: KLT [22] [28]. This one was proposed by Lucas and Kanade [18] and was later improved by [27] [26]. The KLT tests was performed over the Birchfield implementation [2] with some modifications, such as search of new points after each 5 images processed.

Figure 5 shows the performance results using SIFT and KLT methods. SIFT has obtained an average rate of 4.4 fps over original images, without distortion, and a rate of 10.5 fps with the use of **filter 5**, the worst distortion applied. KLT presented higher averages, 13.2 fps and 13.08 fps, respectively. Note that SIFT has its worst performance in high quality images because of the large amount of detected points and, consequently, because of the higher number of descriptors to be processed. The KLT, on the other hand, keeps an almost constant performance. However, due to the

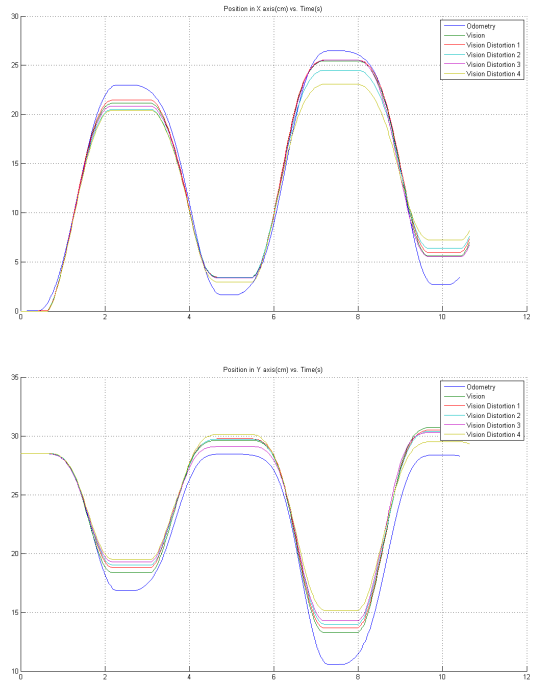

Fig. 4. Position determinated by the robotic arm odometry and a visual system, without and with distortion.

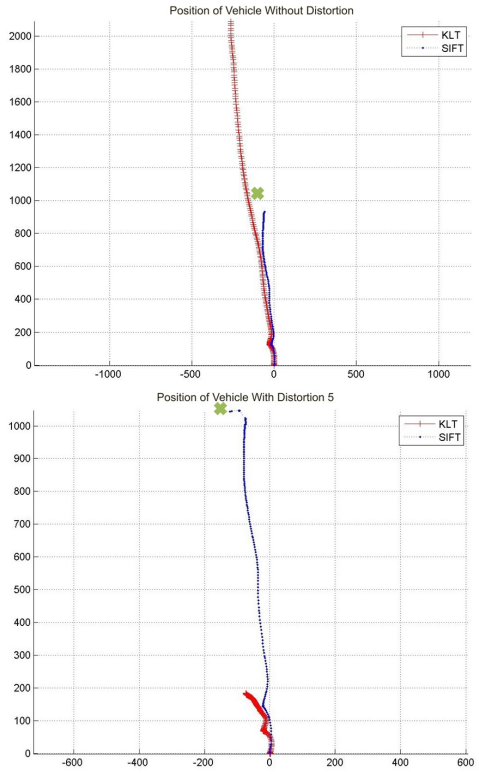

Fig. 5. Real Robot Localization in online system, without and with artificial distortion.

slow dynamic associated with undersea vehicle motion, both methods can be applied to online visual odometry for AUV. The green cross represent the real final position and the metric unit is centimeter.

The SIFT results related to the robot localization were considered satisfactory, even with extreme environment distortions (**filter 5**). On the other hand, KLT gives unsatisfying results for both cases, once it is too much susceptible to the robot's depth variation, or image scale, that occurs constantly in the AUV motion, despite the depth control.

C. Robustness to Scale

Tests were performed to estimate the robustness of the proposed system to the sudden scale variation. In this case, a translation motion with height variation was performed with the camera to simulate a deeper movement of the robot in critical conditions.

The figure 6 shows the SIFT results, considered satisfactory, even in critical water conditions. Considering the use of some filters in extreme conditions, SIFT is superior to KLT although it shows an inexistent movement in Y axis. Over the tests, SIFT has shown an average rate of 6,22 fps over original images captured by the camera and a rate of 7.31 fps using **filter 1** and 10.24 fps using **filter 5**. The KLT have shown 12.5, 10.2 and 11.84 fps, respectively. The green cross represent the real final position, it is the same to all graphics in figure 6, the metric unit is centimeter.

D. Topological Maps

Tests to validate the mapping system proposed were performed. Figure 7 show the final map, using images acquired during the underwater vehicle navigation. This map was used to localize the vehicle, and to aid the navigation to visual targets, validating the visual odometry, mapping and online issues associated to AUVs inspection tasks. The position is updated when the robot recognizes one local known.

V. CONCLUSION

This paper proposes a new approach to visual odometry and mapping, using a SIFT space as reference for topological maps. This system can be used either in autonomous inspection tasks or in control assistance of robot closed-loop, in case of a human remote operator.

A set of tests were performed under different underwater conditions. The effectiveness of our proposal was evaluated inside a set of real scenarios, with different levels of turbidity, snow marine, non-uniform illumination and noise, among others conditions. The results have shown the SIFT advantages in relation to others methods, as KLT, in reason of its invariance to illumination conditions and perspective transformations. The estimated localization is robust, comparing with the vehicle real pose.

Considering time performance, our proposal can be used in online AUV visual odometry and mapping, even in very extreme sea conditions.

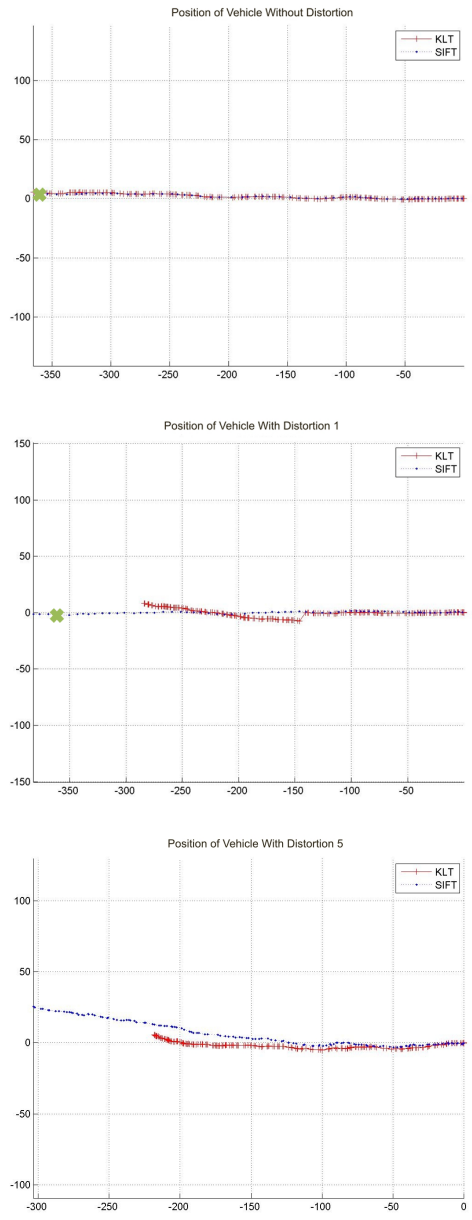

Fig. 6. Localization with translation and scale movement without and with distortion.

The correlations of interest points provided by SIFT were satisfying, even with the presence of many *outliers*, i.e., false correlations. The proposal of using homography matrix estimated in robust ways in order to remove *outliers* through RANSAC and LMedS algorithms shows good results.

The original integration of SIFT and topological maps for AUV navigation is a promissing field. The topological mapping based on Kohonen Nets and GCS showed potential in underwater applications using visual information due to its robustness to sensory impreciseness and low computational cost. Although preliminary, this method presents promising results which validate the approach.

As future work, we propose to detail the analysis of our topological mapping system, executing a set of tests with different scenarios and parameters. We are also proposing the use of scale information provided by SIFT in conjunction with the altimeter information, allowing the estimation of the depth motion of the vehicle. The utilization of stereoscopic vision is also a possibility in order to provide more accuracy to the system.

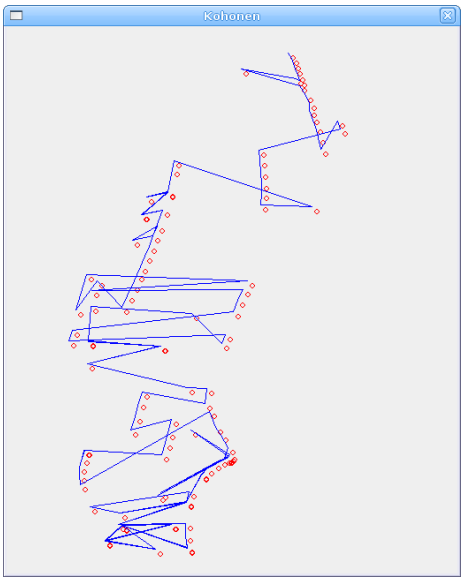

Fig. 7. Topological Map generated by ROVFURGII in movement.

REFERENCES

[1] H. Bay, T. Tuytelaars, and L. J. Van GoolRafael. Surf: Speeded up robust features. In *European Conference on Computer Vision*, 2006.

[2] Stan Birchfield. Klt 1.3.4 : A c implementation of klt tracker. http://www.ces.clemson.edu/ stb/klt/, 2007. [Last Acessed in February, 2008].

[3] O. Booij, B. Terwijn, Z. Zivkovic, and B. Krose. Navigation using an appearance based topological map. In *IEEE International Conference on Robotics and Automation*, pages 3927–3932, April 2007.

[4] Mario Centeno. Rovfurg-ii: Projeto e construção de um veículo subaquático não tripulado de baixo custo. Master's thesis, Engenharia Oceânica - FURG, 2007.

[5] Rina Dechter and Judea Pearl. Generalized best-first search strategies and the optimality af a*. *Journal of the Association for Computing Machinery*, 32(3):505–536, July 1985.

[6] O. Delaunoy, N. Gracias, and Garcia R. Towards detecting changes in underwater image sequences. In *MTS/IEEE Oceans08*, April 2008.

[7] Edsger W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.

[8] Martin Fischler and Robert Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[9] Stephen D. Fleischer. *Bounded-Error Vision-Based Navigation of Autonomous Underwater Vehicles*. PhD thesis, Stanford University, 2000.

[10] Bernd Fritzke. Growing cell structures - a self-organizing network for unsupervised and supervised learning. Technical report, University of California - Berkeley, International Computer Science Institute, May 1993.

[11] Rafael Garcia, Xavier Cufi, and Marc Carreras. Estimating the motion of an underwater robot from a monocular image sequence. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 1682–1687, 2001.

[12] Rafael Garcia, V. Lla, and F. Charot. Vlsi architecture for an underwater robot vision system. In *IEEE Oceans Conference*, volume 1, pages 674–679, 2005.
[13] N. Gracias, S. Van der Zwaan, A. Bernardino, and J. Santos-Vitor. Results on underwater mosaic-based navigation. In *IEEE Oceans Conference*, volume 3, pages 1588–1594, 10 2002.
[14] Nuno Gracias and Jose Santos-Victor. Underwater video mosaics as visual navigation maps. *Computer Vision and Image Understanding*, 79(1):66–91, July 2000.
[15] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
[16] Teuvo Kohonen. *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001.
[17] David Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
[18] Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conferences on Artificial Intelligence*, pages 674–679, 1981.
[19] I. Mahon and S. Williams. Slam using natural features in an underwater environment. In *Control, Automation, Robotics and Vision Conference*, volume 3, pages 2076–2081, December 2004.
[20] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *International Journal on Computer Vision*, 65:43–72, 2005.
[21] T. Nicosevici, R. Garcia, S. Negahdaripour, M. Kudzinava, and J. Ferrer. Identification of suitable interest points using geometric and photometric cues in motion video for efficient 3-d environmental modeling. In *International Conference on Robotics and Automation*, April 2007.
[22] K. Plakas and E. Trucco. Developing a real-time, robust, video tracker. In *MTS/IEEE OCEANS Conference and Exhibition*, volume 2, pages 1345–1352, 2000.
[23] Peter Rousseeuw. Least median of squares regression. *Journal of the American Statistics Association*, 79(388):871–880, December 1984.
[24] Stephen Se, David Lowe, and James Little. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *The International Journal of Robotics Research*, 21(8):735–758, 2002.
[25] Stephen Se, David Lowe, and James Little. Vision-based global localization and mapping for mobile robots. *IEEE Transactions on Robotics*, 21(3):364–375, June 2005.
[26] Jianbo Shi and Carlo Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
[27] Carlos Tomasi and Takeo Kanade. Detection and tracking of point features. Technical report, Carnegie Mellon University, April 1991.
[28] T. Tommasini, A. Fusiello, V. Roberto, and E. Trucco. Robust feature tracking in underwater video sequences. In *IEEE OCEANS Conference and Exhibition*, volume 1, pages 46–50, 1998.
[29] P. H. S. Torr and D. W. Murray. The development and comparison of robust methodsfor estimating the fundamental matrix. *International Journal of Computer Vision*, 24(3):271–300, 1997.
[30] Xun Xu and Shahriar Negahdaripour. Vision-based motion sensing for underwater navigation and mosaicing of ocean floor images. In *MTS/IEEE OCEANS Conference and Exhibition*, volume 2, pages 1412–1417, October 1997.