Terrain-based Navigation for Underwater Vehicles
Using Side Scan Sonar Images

Severin Stalder
Laboratoire de Systèmes Robotiques
EPFL, Switzerland
sstalder@vision.ee.ethz.ch

Hannes Bleuler
Laboratoire de Systèmes Robotiques
EPFL, Switzerland
hannes.bleuler@epfl.ch

Tamaki Ura
Institute of Industrial Science
The University of Tokyo, Japan
ura@iis.u-tokyo.ac.jp

Abstract - Underwater navigation challenges the research community as a reliable navigation system is unavailable. Correctly matched landmarks could compensate the drift of dead reckoning navigation systems. Furthermore, they could be useful in side scan sonar image registration. We propose to integrate both applications to form one landmark detection and matching system. Our approach detects disruptions in the local texture field using level set evolution on Haralick feature maps. During evolution the landmarks are continuously tried to be matched. An energy term which has been used for supervised image registration is used to verify the hypothesized matches. If the energy term based on pixel-wise similarity and hard landmark constraints is minimum, the landmark matches are considered as correct.

I. INTRODUCTION

Autonomous Underwater Vehicles (AUVs) can be used to survey deep ocean ridge systems in order to study earthquakes or underwater volcanos. Side scan sonar images of “r2D4” [1] acquired in about 3000 meters depth were used within the scope of this work.

The low data transmission rate makes underwater navigation difficult. Current AUVs do have expensive inertial navigation systems aboard. Nevertheless, navigation errors accumulate and the estimation of the current position gets less and less accurate. The purpose of this project is to re-calibrate the navigation system with correctly matched landmarks in relatively flat sea bottom.

Side scan sonar images suffer from many kinds of distortions. Recently, synthetic aperture sonar gains popularity as the resolution can be improved by one order of magnitude. The improvement is achieved through sophisticated post-processing of sonar data. However, the required computational power is not available underwater. This project concentrates on side scan sonar images as they can be used to support the underwater navigation.

This paper is organized as follows. In Section 2, related work is briefly presented. Section 3 describes the side scan sonar and some image characteristics. The methods are described in section 4. Implementation details and results are shown in section 5. Finally, section 6 concludes the paper.

II. RELATED WORK

Related work includes side scan sonar image segmentation, landmark matching and image registration. Fig. 1 shows the relationship among these applications.

In image segmentation, the fundamental problem is to find a balance between boundary accuracy and texture characterization as boundaries can be clearly located, but textures are defined over regions. Active contours or deformable templates are widespread because of their ability to smoothly segment region boundaries in noisy images. Karoui et al. [2] use active contours and co-occurrence matrices to describe sea bottom texture. The Kullback-Leibler divergence on co-occurrence distributions was used to measure the similarity between regions.

In image registration there are two different matching techniques. On the one hand, there are landmark matching methods. It is assumed that landmarks are regularly spread over the image. Furthermore, they should be small and their relative position must be conserved from one image to another. A priori knowledge and pre-processing are used to reduce the complexity of the matching algorithm of [3]. However, most of the landmark features do not resist to view point changes.

On the other hand, there are region matching methods. Regions are features which are always present and detectable.

Fig. 1 The framework of the presented algorithm shows the multiple uses of matched landmarks in navigation and image registration. This paper addresses the colored boxes.
from any modalities [4].

This project tries to combine both approaches. Matched landmarks are used to guide image registration. Image registration is supposed to verify the matches at a region level.

III. SIDE SCAN SONAR IMAGE ACQUISITION PROCESS

A side scan sonar is used to efficiently visualize large areas of sea bottom. The side scan sonar emits narrow sound beams equally in about 45 degrees to both sides of the tow fish. The strength of the returned signal is recorded as intensity in the side scan sonar image. The intensity depends on both the surface area and the composition of the reflecting material. As the angle of incidence to the sea bottom varies across the image, uniform sea bottom will not appear uniform in the side scan sonar image. In hilly areas, characteristic shadows appear behind salient objects. The strength of one return signal must be recorded before the next sound pulse is emitted. Thus, one line on the sonar image corresponds to the echo of one beam. Consequently, any deviation from a straight line in the path of the AUV will introduce another source of distortion.

We focus on images of relatively flat areas which can be modelled as textured images. Disruptions in the local texture field are interpreted as landmarks.

IV. METHODS

This algorithm is inspired by manual landmark matching. Most human operators try to match one landmark after another. The most salient and visible landmark is likely to be matched at first. Then, more and more landmarks are taken into account until a maximum number of landmarks are matched without contradiction. Without a hierarchical approach it would be impossible to solve the matching problem as only the information about the location and perhaps the relative size of landmarks can be used.

Unfortunately, the level set evolution proposes possible landmarks in a different manner. At initialization, all landmark candidates are present. Then, landmarks are eliminated respecting their size and saliency in the local texture field. In order to keep as many matches as possible, it is crucial to stop the level set evolution as soon as there is no matching contradiction and the overlaid images are most similar.

The algorithm can be decomposed in three stages. First, the texture is analyzed and a binary segmentation is carried out to initialize the level set evolution. Second, the level set evolution proposes landmarks. Third, the landmark matches are verified using an image registration technique. Stage two and three are alternating to assure that landmark detection and landmark matching are concurrent.

A. Texture Descriptors: Haralick Features

Following the work of [2], textures are characterized using co-occurrence matrices. The matrices can be described as local second-order spatial averages. Parzen windows are applied to estimate the unknown probability density function in a non-parametric manner (see Fig. 2). Traditionally, a set of five Haralick features is used to describe any kind of texture. But two Haralick features (correlation and the local homogeneity) turned out to be sufficient for level set evolution in flat sea bottom.

B. Landmark Detection: Level Set Evolution

Active contours are suited to obtain smooth boundaries in noisy side scan sonar images. The contours evolve according to a partial differential equation called energy functional. This equation contains typically an internal energy term $P$ and an external energy term $E$ (1). The $\phi$ function is called level set function on which the curve is implicitly defined to allow for topology changes.

$$E(\phi) = \mu P(\phi) + E_m(\phi) \quad \mu > 0$$  \hspace{1cm} (1)$$

The idea for this algorithm is to use the Correlation feature map to calculate the length of the contour $L_g$ and the LocalH feature map to calculate the area term $A_g$ in (2) (see Fig. 3). Therefore, the energy term is minimized considering two Haralick texture features simultaneously. Originally, the term $A_g$ was only used to speed up the curve evolution.

$$E_m(\phi) = \lambda L_g(\phi) + \mu A_g(\phi)$$  \hspace{1cm} (2)$$

Fig. 2 The co-occurrence matrices are used to characterize the sea bottom. Haralick feature maps are calculated on these distributions.

Fig. 3 The two selected Haralick feature maps. Left: Correlation (indicates best the boundaries between different regions). Right: LocalH (indicates the areas).
C. Landmark Matching: Image Registration

The convergence of the level set evolution is critical. There are very few salient obstacles in relatively flat areas. However, there are always some disruptions in the texture field. Instead of waiting for the eventual convergence of the level set evolution on salient obstacles, the disruptions in the texture field are continuously tried to be matched. In doing so, the number of landmarks to be matched is increased.

The image registration is formulated in an energy minimization framework which can be minimized using graph cuts. The goal is to find a discrete label field respecting a smoothing and a data term

\[ E(f) = E_{smooth}(f) + E_{data}(f) \]  

For each pixel \( p \in P \) in the image there is a label \( f \in L \) which corresponds to a displacement. The transformation between two images can be non-rigid.

The data term has the form of

\[ E_{data}(f) = \sum_{p \in P} D_p(f_p) \]  

\[ D_p(f) = d_p(f) + \mu \sum_{k \in K, f(k) \neq f} \frac{r^2}{|p[k]-p|} \]

The data term was proposed in [5] for manual landmark matching. As the landmark matches are supposed to be correct they are introduced as hard constraints in the deformation field. The influence of a landmark matches decreases with the distance to the landmark site \( k \in K \). If there was only the pixel-wise similarity term \( d_p(f) \), the algorithm would prefer to register the background rather than the important landmark sites.

V. IMPLEMENTATION AND EXPERIMENTAL RESULTS

The size of the co-occurrence matrices is proportional to the square number of gray levels. Therefore, the image was quantized to 10 intensity levels using k-means.

The set of co-occurrence distributions was calculated for a displacement of one pixel in the directions (0°, 45°, 45°, 90°) and for a displacement of two pixels in the direction of 0°. As the texture of flat sea bottom is relatively fine, a window size of 12 by 12 pixels is sufficient to analyze the texture.

The initial level set is built from thresholding the LocalH feature map, a binary segmentation. The landmarks are supposed to be inside the initial contour, e.g. \( \nu < 0 \) in (2).

Only the landmark constraints were implemented to calculate the energy in the image registration framework. The motion constraints of big landmarks was preferred over small landmarks in a circular neighborhood. If there were two landmarks in the same neighborhood, their motion had to be locally coherent. However, it is strongly suggested to use the pixel-wise similarity term \( d_p(f) \) in (5) to obtain a complete label field allowing for true image registration. For example, [4] proposed a useful similarity measure based on mutual information on two Haralick features (energy and contrast). In addition, any features can be used.

Experimentally, the level set evolution was stopped after 260 iterations as the motion field was consistent.

So it was possible to match 7 landmarks correctly in about 250 meters of sea bottom. For comparison, the Concurrent Mapping and Localization (CML) algorithm [6] assumes about 1-3 perfect matches in 200 meters of a multi-pass mission.

VI. CONCLUSION

There were two main difficulties to match landmarks correctly. First, landmarks in side scan sonar images can not be described geometrically as landmarks are seen from opposite view points. Second, the transformation between two side scan sonar images is generally non-rigid due to the irregular path of the AUV.

Therefore landmark detection and landmark matching are proposed to be concurrent. In fact, correctly matched landmarks can guide image registration and image registration can verify the matches at a global scale.

This leads to a framework that integrates the multiple uses of landmarks to form one elegant system based on side scan sonar images.

REFERENCES