2021 The 21st International Conference on Control, Automation and Systems (ICCAS 2021)
Ramada Plaza Hotel, Jeju, Korea, Oct. 12∼15, 2021

Shallow Depth SIFT Based Approach for Mapping underwater surfaces using AUV's

** RAGHURAM C S and **Sai Anoop Sadineni

Electrical and Electronics Engineering, BITS Pilani Hyderabad Campus,
Hyderabad, India (f20190357@hyderabad.bits-pilani.ac.in)
Electronics and Instrumentation Engineering,BITS Pilani Hyderabad Campus,
Hyderabad, India (f20190506@hyderabad.bits-pilani.ac.in)
(** Authors have contributed Equally to the research work.
** Authors Declare no conflict of interest.)

Abstract: Autonomous Underwater Vehicles(AUVs) have always been used in oceanic exploration. They were used for topographic mapping, studying the various elements of the sea from flora and fauna to synthetic objects and debris alike. This paper aims to produce perceivable maps of the shallow seabed using AUVs, using the Scale-Invariant Feature Transform(SIFT) algorithm for collecting features from the images and using the Brute-force matcher to match the images producing 2-dimensional rectangular maps by applying the proposed algorithm. The algorithm was tested using handheld webcams in a simulated environment and our results were consistent with the expected output, and also account for the uncertainty of noise, distortion due to the reflection of light on the surface of the water at shallow depths. This approach to mapping has a low computational cost and can be deployed to multiple AUVs to map larger areas.

Keywords: **Seafloor Mapping, Feature extraction, Autonomous Underwater vehicles, vision-based mapping.**

1. INTRODUCTION

Underwater robotics research widely involves autonomous vehicles and controlled robots/rovers for undersea exploration for research and entertainment such as photography alike. They have multiple ways of getting SLAM(Simultaneous Localization and Mapping) data using acoustic sensors, laser rangefinders, and visual sensors.

Our paper presents an approach to use on-board optical sensors(cameras) independently of the AUVs on-board navigation computer to map the shallow seabed. Our approach uses feature extraction based on the SIFT algorithm and Matching methods to generate maps in segments and merge them independently to create maps.

The paper goes over the past work and research done in a similar domain; the experiment set up, a brief overview of the SIFT algorithm, the mapping ideology, and our aspired future work.

2. REVIEW OF LITERATURE

Several Underwater Mapping techniques and research projects using AUVs have been performed in the past, but mapping rectangular portions using SIFT and generating visual maps only was not to be found.

In the paper *SLAM in Underwater Environment using SIFT and Topologic Maps* by *Paulo Drews Jr, Silvia Botelho and Sebastião Gomes. October, 2008*.In the Paper they have provided an approach for localization and mapping of underwater terrain, using cameras and the SIFT algorithm. They also have worked on producing topological maps from localization and mapping[1].

There are also some non profit oceanographic research centers and surveying government organisations which have been actively working on seafloor mapping with AUVs. MBARI has two AUVs optimised for seafloor mapping which can chart the seafloor more precisely than hull-mounted or towed sonar systems can in their fleet.[3][4] They are equipped with four SONAR sensors each in which two side scan sonars produce images based on the intensity of the reflected sound. The Swath Multibeam Sonar is responsible for producing high resolution bathymetry. It is the measurement of depth to the ocean floor. The last sensor, Sub-Bottom profiler, detects the layers between sediments and depth to the basement for higher precision.[4]

The SeaFloor Mapping Group (SFMG) supports coastal and marine geographic research at Woods Hole Coastal Marine Science Center (WHCMSC) under the US Geological Survey (USGS).[5] It has been active for 25 Years. It uses Acoustic and Optical techniques for seabed study. SFMG nearly uses the same mapping techniques like MBARI[4] with inclusion of still photographs, seismic reflection systems and sediment collection from the underwater. They also study the sea surface in addition to the Underwater. This group specialises in geophysical and sample data acquisition, analysis, and interpretation in the lacustrine, tidal, and marine ecosystems with a large community of maritime electronics technicians, physicists, geologists, physical scientists, geographers, and visual and photographic experts.[6]

3. METHODOLOGY

Our project on Seafloor mapping aims to map rectangular sections of the seafloor using AUV(Autonomous Underwater Vehicles). The project aims to use only cameras and vision based systems to map the desired region

978-89-93215-21-2/21/$31.00 ©ICROS

on the map. We will choose an area at the selected location to map it, and the AUV will be programmed to follow a path as shown in Fig.1. The cameras present on the AUV underside will capture images continuously with breaks at respective areas, as they go along the route.

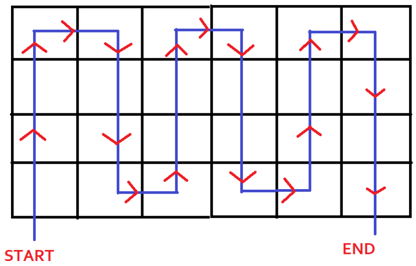

Fig. 1. Path followed by the AUV

The cameras will be capturing images only as the AUV moves in a vertical path and will be off during the turn the AUV makes to realign itself parallel to the direction it transited before. This process continues until the AUV reaches the end of the area it has to cover during the mapping expedition.

The flowchart depicting the mapping algorithm is shown in Fig 2

4. CONSTRAINTS OF THE STUDY

This research was intended to develop an algorithm for Mapping relatively small rectangular areas of the seafloor to produce a geographical map of the seabed. In the study, we simulated the underwater environment of a submersible AUV moving using a handheld digital camera. This mapping algorithm can be deployed to various AUVs which are intended to map similar surfaces.

5. EXPERIMENT SETUP

Using a handheld digital camera, we followed it through the path described in Fig.1 as the trajectory of the AUV; the camera was a 2mp digital webcam. The camera capture 60 frames per second, and the average speed of an AUV is 1.5 to 2m/s. The handled camera roughly maintained the rate of 1.6 to 1.7m/s. The algorithm is aimed at generating clear maps of the seabed and a camera with a focal distance of less than 2 feet (45-50cm).

The AUV(handheld camera) will be at an height of (BLANK) to the seabed(ground/mapping surface). The height will be determined by the camera's focal distance adjusted to parameters of the water surrounding it. in the test case, it will beset to the default focal length

The AUV will be given the predefined Approximate area of the surface to be mapped(area), Approximate Length(L), and Breadth(B). The AUV takes these values and divides the space into segments, as shown in Fig.1(the number of vertical and horizontal segments will vary depending on the mapping area).

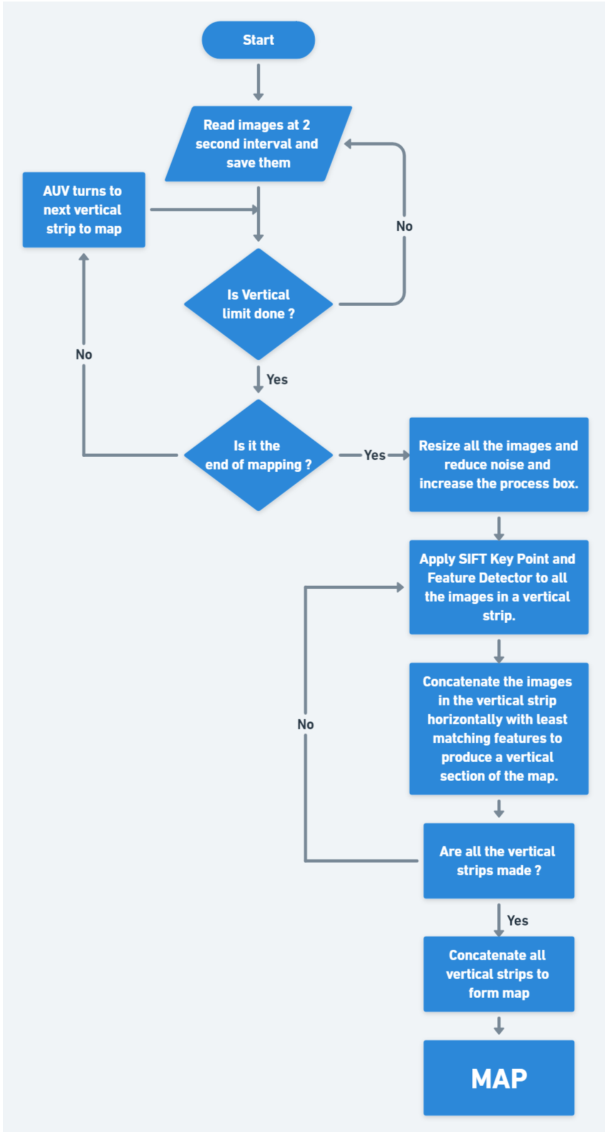

Fig. 2. Mapping algorithm Flowchart

6. ALGORITHM EXPLANATION

6.1 Image Capturing

The cameras will be capturing images only as the AUV moves in a vertical path and will be off during the turn the AUV makes to realign itself parallel to the direction it transited before. This process continues until the AUV reaches the end of the area it has to cover during the mapping expedition.

The AUV captures images in the vertical paths called "strips" consecutively. Each upright portion will be stored and processed separately before merging them to generate an image. To make the vertical portions "strips," we will be applying the SIFT algorithm[2] to find each image's key points and descriptors.

We start by capturing 200 images frame by frame as the AUV moves through the vertical segments. It will store the images in a folder temporarily.

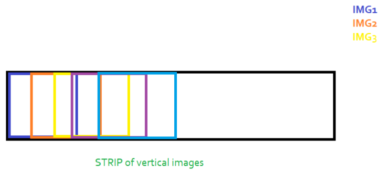

Fig. 3. The capturing and comparison of images in the vertical segments

6.2 Loading the images into a useable array

The images will be updated into a list, simultaneously applying the SIFT algorithm using OpenCV.[2] The images are updated into the array and will be resized to (256, 256, 3), i.e., 256 pixels wide, 256 pixels long, and with the three channel.The camera captures RGB images which is converted to BGR images.(ref. Fig.4)

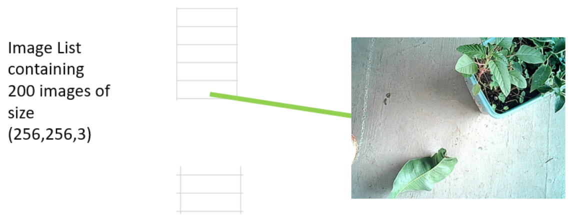

Fig. 4. The image array with images stored as the order captured in

6.3 SIFT Application

6.3.1 **Understanding SIFT**

SIFT, or **Scale-Invariant Feature Transform**, is a feature detection algorithm in Computer Vision. SIFT helps locate the key points, which are local features in an image. These key points are scale rotation invariant that can be used for various computer vision applications, like image matching, object detection, scene detection, etc.

The SIFT algorithm being the backbone of our project, it is used to find the key points and the feature descriptors of each of the images. The SIFT method uses a Difference of Gaussian (DoG) to see the difference between the different Gaussian blurring of the image with different sigma (σ) values. This same process will be repeated for different octaves in the Gaussian pyramid.[2] The scale space function represented by L(x, y , σ), is found by convolving the The standard Gaussian function G(x, y ,σ) with the Input Image I(x, y). The Difference of Gaussian function(DoG), D(x, y, σ) can be computing the difference of two nearby scales wiht a constant multiplicative factor k.[2]

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{1}$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

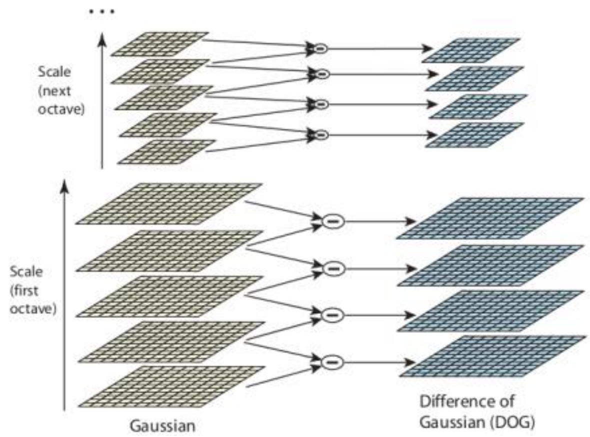

Fig. 5. The Difference in Gaussian for a specific pixel of the image

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{2}$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

After finding the Difference of Gaussian , maxima and minima of these images is found by comparing the neighbours of individual pixels with the previous and further images. The paper summary gives us empirical data, no

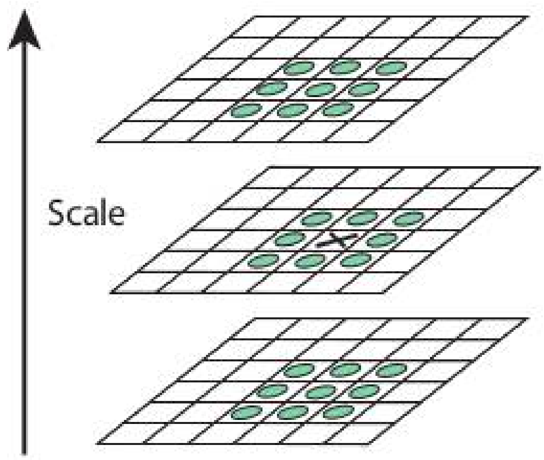

Fig. 6. Maxima and minima of the difference-of-Gaussian images are detected by comparing a pixel (marked with X) to its 26 neighbors in 3x3 regions at the current and adjacent scales (marked with circles).

octaves = 4, number of scale levels = 5, initial , etc. DoG has higher response for edges, so edges also need to be removed. For this, a concept similar to Harris corner detector is used. They used a 2x2 Hessian matrix (H) to compute the principal curvature. We know from Harris corner detector that for edges, one eigen value is larger than the other. So here they used a simple function,If this ratio is greater than a threshold, called edgeThreshold in OpenCV, that keypoint is discarded. It is given as 10 in paper.So it eliminates any low-contrast keypoints and edge keypoints and what remains is strong interest points. For further understanding of SIFT refer to paper.[2]

6.3.2 **Application of SIFT**

The images loaded into their specific arrays of each vertical segment are passed one image array at a time through a function that applies Scale Invariant Feature Transform(SIFT) on each of the images. The SIFT function in OpenCV allows us to detect the key points and feature descriptors of the images. The key points give us information about the image location, scale, and orientation. [2]

All these parameters embedded within the keypoint descriptor estimate a 2D coordinate system that constantly describes the local image region throughout and different for each image. The function also computes the descriptor of the images by adding each image's gradient magnitude and orientation around a keypoint location a Gaussian window, then weights the values to define the final descriptor.[2]

These values also account for the change in 3D orientation and illumination to remove false-positive key points and descriptors. We still need to retain the values of the key points and descriptors for each image to use them further, which will append into corresponding arrays.

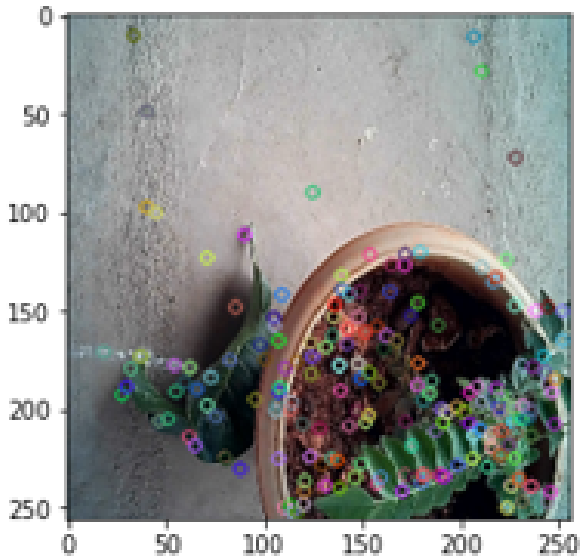

Fig. 7. A captured image with key points detected

6.4 Forming Each vertical Segment

. In the test case we have considered an are of 4×6 segments, with each segment of $(30cm \times 30cm)$.The image arrays of each of the vertical segments have the images with the detected key points drawn onto them, with their respective feature descriptors also stored in respective arrays. To match images to form the vertical images we have utilized the Brute Force Matcher(BFmatcher), toolkit available in Open-CV.

- The first image in each segment which is the first image of each vertical segment is matched with all the images in the array, the no of matches with each of the images is appended into an array simultaneously.
- The mean of the sorted matches is found and the values higher than that are discarded, the repeating values less than the mean are also discarded, we will be left with only different frames from the same vertical segment.
- The images corresponding to the remaining matches are arranged in the order they were captured.
- The first image of this new list which represents the second frame of the vertical segment is considered and the process is repeated with the first 10 images in the sorted macthes list, now we have found the third frame of the vertical segment.
- The first image in the new list which represents the third frame of the vertical segment is considered and the process is repeated with the first 5 or less images of the new sorted macthes list, the fourth frame of the segment is obtained.

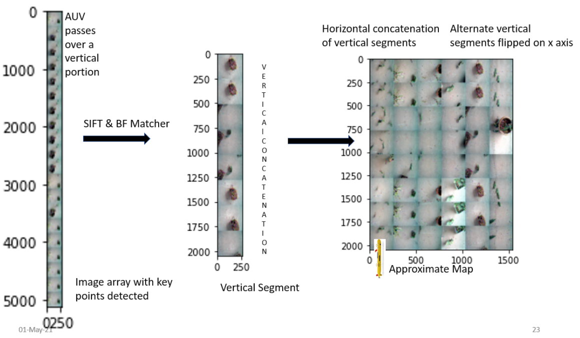

Fig. 8. The complete mapping algorithm

- The four images stored in a new list are concatenated vertically to obtain the vertical segment.
- Each of the image in the image arrays correspoding vertical segments are processed as follows to form each of the vertical segments,

6.5 Forming the Map

- The alternate vertical segments are mirrored along the x-axis to account for the 180 degrees change in direction of the AUV during the mapping.
- All of the vertical segment images are concatenated horizontally to produce the map.

7. CONCLUSIONS

The mapping algorithm produces an accuracy of around 55 to 60%.Is is efficient as it requires less compute power and can also be run on basic cameras.The algorithm can be deployed to AU Vs with various configurations indented to map similar areas with very few changes. The parameters needed to be altered would be the number of images per vertical segments , number of vertical segments, the match parameters such as the number of images in the first pool, camera parameters such as brightness, the amount of noise reduction based on depth and visibility in water.

Fig. 9. Map preview

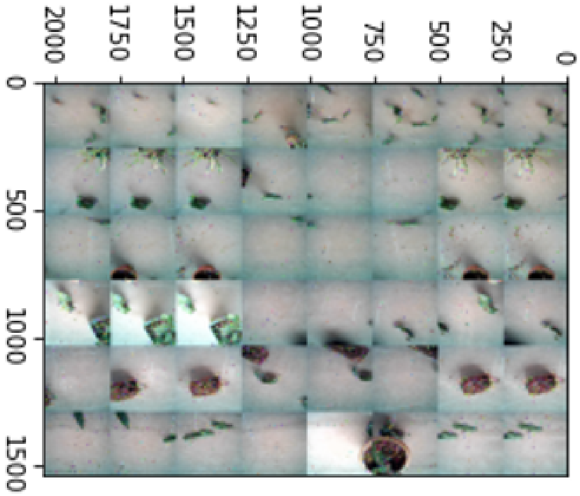

Fig. 10. Generated map

FUTURE WORK

The paper continues it work, by exploring multiple passes of the AUV over the same area to improve in noise reduction in the images, improve accuracy by reducing the false positive matches, and produce further detailed maps[1]. We hope to test our algorithm with a swarm of AUVs to achieve larger mapping surfaces, and reduce the load of a single robot. Using the SONAR data and Vision data simultaneously to achieve 3-dimensional Map of the seabed.

REFERENCES

[1] P. Drews Jr, S. Botelho and S. Gomes, ”SLAM in Underwater Environment Using SIFT and Topologic Maps,” 2008 IEEE Latin American Robotic Symposium, 2008, pp. 91-96, doi: 10.1109/LARS.2008.32.

[2] LLowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004). https://doi.org/10.1023/B:VISI.0000029664.99615.94

[3] Clague D.A., Paduan, J.B., Caress, D.W., Chadwick Jr, W.W., Le Saout, M., Dreyer, B.M., Portner, R.A. (2017). High-resolution AUV mapping and targeted ROV observations of three historical lava flows at Axial Seamount, Oceanography, 30(4), 82-99, doi: 10.5670/oceanog.2017.426.

[4] Caress, D. W., H. Thomas, W. J. Kirkwood, R. McEwen, R. Henthorn, D. A. Clague, C. K. Paull, J. Paduan, and K. L. Maier (2008), High-resolution multibeam, sidescan, and subbottom surveys using the MBARI AUV D. Allan B, Marine habitat mapping technology for Alaska, 47–69. PDF

[5] Schwab, W.C., Denny, J.F., and Baldwin, W.E., 2014, Maps showing bathymetry and modern sediment thickness on the inner continental shelf offshore of Fire Island, New York, pre-Hurricane Sandy: U.S. Geological Survey Open-File Report 2014–1203, http://dx.doi.org/10.3133/ofr20141203. ISSN 2331–1258

[6] High-resolution swath interferometric data collected within Muskeget Channel, Massachusetts; 2014; OFR; 2012-1258; Pendleton, Elizabeth A.; Denny, Jane F.; Danforth, William W.; Baldwin, Wayne E.; Irwin, Barry J.