

OCEANS 2021: San Diego – Porto, 2021,
Pages 1-9
Electronic ISBN:978-0-692-93559-0, ISSN: 0197-7385.
September 2021,
<https://doi.org/10.23919/OCEANS44145.2021.9706053>
<https://archimer.ifremer.fr/doc/00759/87116/>

Archimer
<https://archimer.ifremer.fr>

Combined use of a frame and a linear pushbroom camera for deep-sea 3D hyperspectral mapping

Kumar Peeyush ¹, Gracias Nuno ², Istenic Klemen ², Garcia Rafael ², Arnaubec Aurelien ³, Ferrera Maxime ³, Bajjouk Touria ³

¹ Institute of Geodesy and Geoinformation, University of Bonn, Suisse

² Research Centre in Underwater Robotics (CIRS), University of Girona, Spain ³ French Research Institute for Exploitation of the Sea (IFREMER), france

Abstract :

Hyperspectral (HS) imaging produces an image of an object across a large range of the visible spectrum, and not just the primary colors (R, G, B) of conventional cameras. It can provide valuable information for object detection, analysis of materials and processes in environmental science in the deep-sea, especially for the study of benthic environments and pollution monitoring. In this paper, we address the problem of camera calibration towards 3D hyperspectral mapping where GPS is not available, and the platform navigational sensors are not accurate enough to allow direct georeferencing of linear sensors, as is the case with traditional aerial platform methods. Our approach presents a preliminary method for 3D hyperspectral mapping that uses only image processing techniques to reduce reliance on GPS or navigation sensors. The method is based on the use of standard RGB camera coupled with the hyperspectral pushbroom camera. The main contribution is the implementation and preliminary testing of a method to relate the two cameras using image information alone. The experiments presented in this paper analyze the estimation of relative orientation and time synchronization parameters for both cameras through experiments based on epipolar geometry and Monte-Carlo simulation. All methods are designed to work with real world data.

Keywords : Geometry, Sea surface, Three-dimensional displays, Navigation, Nonlinear distortion, Cameras, Sensors

I. INTRODUCTION

Hyperspectral (HS) images convey a large amount of information for environmental research. Recent HS sensors can provide high spectral resolution data with narrow spectral bands for each image pixel, typically of 1 nm width and ranging from 380 to 750 nm, thus generating large quantities of information for terrain interpretation. These data can be used in conjunction with suitable sensors to produce highly accurate spectral maps of large areas [1], from which physical and ecological characteristics can be extracted.

The application of HS imagery for deep-sea studies is becoming increasingly popular. Recent work has focused in topic such as:

- Benthic habitat classification, bathymetry, seawater quality [2].
 - Continental surface properties such as geochemical and geophysical, soil quality [3].
- Monitoring of pollution in the sea [4].
Ecological investigation of the marine environment [5] and deep-sea environmental research dealing with issues related to the critical zone [6].



Fig. 1: Stereo setup of the two cameras on Ariane HROV from IFREMER [7], [8]. The upper is the RGB frame camera. The bottom is the pushbroom HS camera. The Z-axis is in the view direction, X-axis to the right and Y-axis pointing downwards. The cameras are mounted vertically having the y-axis as baseline between them.

Currently, hyperspectral imaging is mostly performed using a satellite or aerial platforms such as aircraft and UAVs [9]. For coastal applications, aerial platforms are limited to depths of a few meters for underwater mapping due to insufficient natural illumination. For shallow water mapping, unmanned surface vessels have been used [10]. All of these platforms have access to GPS and accurate navigational data. To map the deep-sea surface, the hyperspectral sensors must be mounted on an AUV or ROV. However, HS data is only useful when we can establish the geometric relationship between the image pixels and their actual positions on the ground [11].

Some of the factors influencing the direct georeferencing and geometry of raw HS data are:

- Interior projection and distortion parameters of the HS camera.
- Exterior orientation parameters (camera pose with respect to the vehicle).
- Inaccurate navigation sensor data.
- Inaccessibility of GPS in underwater scenario.
- Deep-sea surface topography.
- Insufficient illumination in underwater scenario.

Unlike airborne and satellite platforms, there is no access to GPS to calculate the position of the AUV with high

accuracy, and the use of ground control points (GCPs) is restricted or impossible in many cases. Navigation data from on-board sensors such as Ultra Short Baseline (USBL) are affected by environmental conditions such as distance from the surface and sea waves, and tend to be very inaccurate in the deep-sea.

The problem of accurate positioning is further exacerbated for HS cameras with linear sensors, such as the one used for acquiring the data in this paper. These so-called pushbroom HS cameras, capture one scan line at a time and rely on the vehicle motion to form the images. The HS camera used in this paper captures an image of size 1920×1 pixels, which is only a single line of its view compared to a standard frame camera. This also restricts the use of the algorithm Bundle Adjustment (BA) [12] directly on the pushbroom HS camera data due to insufficient texture and overlap. These combined conditions makes solving this problem a complex one.

As part of its general goal to achieve or maintain marine natural habitats in a favorable conservation status, Marha project (EU LIFE) ¹ experiments underwater hyperspectral imaging capabilities as an emerging tool.

To address the above limitations, IFREMER devised a stereo configuration consisting of an RGB frame camera and a pushbroom HS camera to acquire the data simultaneously. In the approach of this paper, the frame camera images serve as input to the *Bundle Adjustment* (BA) algorithm, which outputs the poses of the frame camera and a 3D map of the deep-sea surface. Corresponding points between the RGB images and the HS scanline stack are clicked manually. The two cameras are synchronized using a time-shift parameter and the frame rates of both cameras. The frame camera poses are used to compute the pushbroom HS camera poses for the matched points. Then, a nonlinear least squares method is applied to minimize the SED error considering the coplanarity condition and estimate the best values of the unknown parameters. The estimated parameters are then used to project the hyperspectral scan lines over the 3D map created from BA using the frame camera images.

The aim of this paper is to develop an algorithm for creating a 3D hyperspectral map of a deep-sea surface. This is achieved by superimposing a layer of hyperspectral scan lines over a 3D map created from RGB images of the same scene. The contributions in this work to solving the above problem are:

- A new approach to estimate the external calibration parameters that relate a standard frame camera to a HS linear pushbroom camera, including unknown synchronization.
- Development of a method to produce 3D hyperspectral textured underwater maps from imagery captured at close range.
- Use of pure image processing techniques, resulting in a solution that does not rely on GPS or navigation data.
- Analysis of parameter sensitivity using Monte-Carlo.

¹More information on the Marha Project can be found in: <https://www.life-marha.fr/>

II. RELATED WORK

The direct georeferencing of pushbroom HS camera images has been discussed [13] in extensive literature for airborne platforms based on navigation data and the GPS. There is no research work in the literature that directly addresses the problem of 3D hyperspectral mapping for deep-sea scenarios. The focus of this paper is to develop a pure image processing based method on the constraints for the underwater scenario.

Bongiorno *et. al* [14] present a solution for co-registration of hyperspectral and RGB image to create a map of the seafloor by creating a mosaic of images. The system consisted of ship borne sensors to create an air-water model. Also, a pair of upward facing spectrometers are mounted on the ship and on the AUV, to model the light reaching the sea-floor. The AUV also had a stereo configuration of a RGB camera and another spectrometer which was facing downward. The relative pose between the RGB camera and spectrometer was computed in the lab [15]. The result was a mosaic of seafloor RGB images projected with hyperspectral information.

In the solution presented by Barbieux *et. al* [16], the IMU data is used and first the relative pose between the RGB and IMU is computed. These parameters are then used to calculate the trajectory of the pushbroom HS camera by directly using the RGB poses for the pushbroom HS camera due to the very short baseline. Due to the high fps, missing pose for pushbroom is computed by interpolating RGB poses to obtain a pushbroom mosaic. The tie points are identified manually and a least squares problem is formed to minimize the reprojection error and calculate the relative pose between the IMU and the pushbroom HS camera.

In another paper [17], the solution is based on automatic detection of tie points using SURF features [18]. This solution also uses RGB and pushbroom cameras in a stereo setup. First, the trajectory for the RGB camera and a DEM / orthophoto is obtained by bundle adjustment on frame images along with the IMU data fused in the Kalman filter. Using this RGB trajectory for the pushbroom HS camera, the scan lines are projected into a mosaic. Then automatic SURF feature detection is used to detect tie points between the pushbroom mosaic and the orthophoto. Bundle adjustment is used to reduce the errors between the tie points by estimating the relative pose between IMU and the pushbroom HS camera and the intrinsics of the pushbroom HS camera.

The method explained in [19] does not use data from GPS or IMU and is based on correlation of the hyperspectral scan line and the frame image. A kernel is used to correlate the two images by sampling the frame images as line images. This correlation works best for areas that have a good amount of available features. This method was originally developed for the aerial platform and has been shown to work well for urban areas with many buildings with straight features that span multiple images in the hyperspectral image. For coastal areas, the authors mention that the method may not perform well due to the lack of features.

Several other studies [20]–[25], for pushbroom mapping used direct georeferencing method based on GPS and navigation data either for a satellite or an aircraft.

In our approach, we use a rigidly coupled setup with two unsynchronized cameras, an RGB frame sensor and a hyperspectral pushbroom HS camera. The intrinsics of the two cameras are precomputed and are not part of the estimation process. The time-shift between the two cameras and the relative orientation are estimated by minimizing the SED geometric distance error for each observed point pair. Finally, the hyperspectral information is projected onto the 3D map obtained from BA using the estimated relative pose and time-shift.

Our approach formulates the problem of 3D hyperspectral mapping of underwater surfaces as a bundle adjustment approach that eliminates the dependence on GPS and navigation sensors and uses only image processing techniques.

III. METHODOLOGY

The proposed 3D hyperspectral mapping algorithm for deep-sea surfaces works in several steps, as shown in Fig. 2. In a preprocessing step, additional data required for the algorithm is obtained. The RGB images are used to build a 3D model and estimate the trajectory of the RGB frame camera. The 3D point cloud is further refined to create a textured mesh of the scene. The starting point for the algorithm is to manually select corresponding points between the RGB images and the pushbroom scan line stack. Second, using the initial guess of the time-shift between the two cameras and the relative orientation values for each observed point, the pose of the pushbroom HS camera is calculated based on the pose of the frame camera. Third, the poses of both cameras are used to calculate the Symmetric Epipolar Distance (SED) which serves as the error criteria for each observed matching point. Fourth, the error is minimized using a nonlinear least squares method to estimate the optimized parameters of relative orientation and time-shift between the two cameras. Fifth, the complete trajectory for the push-broom HS camera is calculated using these estimated parameters. Finally, the hyperspectral 3D map is created by projecting the pushbroom scan lines onto the 3D map from frame images.

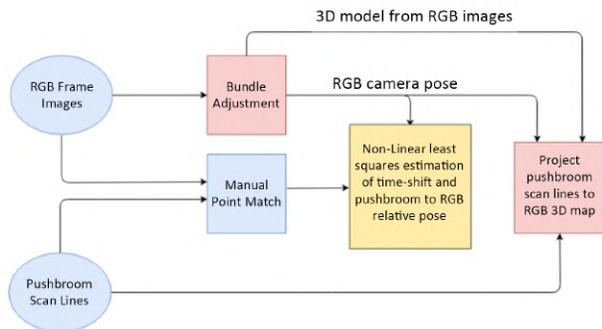


Fig. 2: Schematic representation of the algorithm steps

IV. DATA PRE-PROCESSING AND INPUT

A. Bundle Adjustment of Frame Images

In the absence of accurate navigation data and GPS to determine the platform pose, the information from the frame camera images can be used to estimate the camera pose, which is used as input to the algorithm.

The state-of-the-art bundle adjustment algorithm can be used to estimate the trajectory for the frame camera. Using the open source software openMVG [26] for Structure from Motion [27], the single frame camera images are processed in the bundle adjustment. Each pose is associated with one image. The openMVG works by automatic correspondence matching and then stitching the RGB images to create the 3D map and estimate the pose of the frame camera in the world coordinate system over the whole scene. Fig. 3 shows the point cloud of the 3D scene reconstructed from the RGB images along with the trajectory.

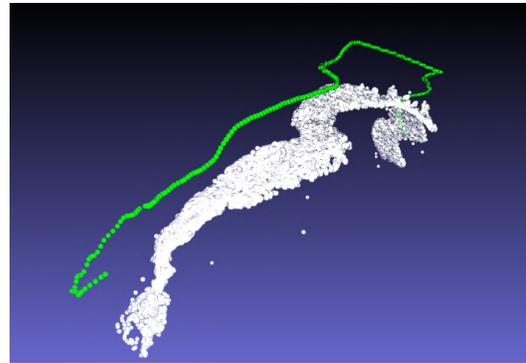


Fig. 3: Point cloud and RGB frame camera pose generated from openMVG



Fig. 4: Textured mesh generated from openMVS

The point cloud obtained from the openMVG software is further processed using another open source software openMVS [28] to create a textured 3D mesh representing the detailed 3D map of the deep-sea surface. The final solution is then to project the hyperspectral pushbroom scan lines over this textured 3D mesh of the surface using the estimated relative pose and time-shift parameters. Fig. 4 shows a section of the textured mesh of the surface.

B. Corresponding points selection

The corresponding point matching depends on finding the key points in the image, e.g., an edge or a corner. For con-



Fig. 5: Input RGB image from the frame camera and selected points

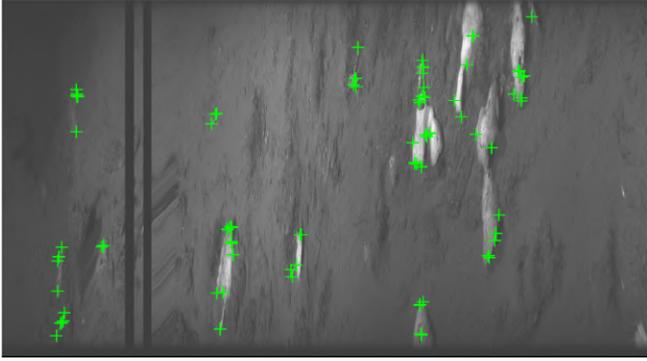


Fig. 6: Input pushbroom HS camera image and selected points. x-axis represents pushbroom poses (Total poses: 3130) and y-axis represents pixels on one frame (1920 pixels).

ventional images, automatic feature detection methods can be used. These methods work by extracting the neighbourhood of a keypoint in the image that describe the keypoint. SURF and SIFT [18], [29] are commonly used feature detection algorithms that are invariant with respect to the scale and rotation of the image. These methods cannot be used to detect features on an image from a pushbroom HS camera that is $N \times 1$. Therefore, for each pixel in the scanned line, there is no neighbourhood texture around it to detect a keypoint and extract the texture around it to uniquely represent it as a feature. To solve this problem, the corresponding keypoints in the frame camera image and in the pushbroom scan lines stack must be manually clicked using an app developed in MatLab. Fig. 5 shows an image of the RGB frame camera and manually clicked points. Fig. 6 shows a complete HS image stack and manually clicked points for HS images stack.

V. OPTIMIZATION

The matched points are used as observations to perform the nonlinear least squares adjustment [30] and estimate the time-shift synchronization and relative orientation of the two cameras. The estimation is performed using the Marquardt-Levenburg algorithm [31], starting with an adequate initial

guess for the unknown parameters.

A. Unknown parameter vector

The total number of unknown parameters for least squares estimation is seven, where one parameter is used for inter-camera synchronization (referred to as time-shift), 3 are used for translation, and the last 3 for orientation. The vector of unknown parameters can be represented as follows:

$$u = [\tau \quad \alpha \quad \beta \quad \gamma \quad t_x \quad t_y \quad t_z] \quad (1)$$

where τ is the time-shift in unit frames, α is roll, β is pitch, γ is yaw rotation angle, unit in radian and t_x, t_y, t_z are the translation in x, y, z direction respectively, unit in meters.

B. Time-shift

The start time of data recording is different for the two cameras. To compute the corresponding frame camera image j for a particular scan line k , given initial value of time-shift (τ), the equation is [32]:

$$j(k) = \tau + \rho k \quad (2)$$

where $\rho = fr'/fr$ is the frequency ratio of the acquisition frequency of the pushbroom HS camera (fr') and the frequency of the frame camera (fr). While the frame rate can differ between surveys, it has to be known and remain constant within a single survey.

C. Pushbroom HS camera pose

The pose of the pushbroom HS camera at each instance (${}^W T_{HS_k}$) can be calculated based on the estimated trajectory of the frame camera (${}^W T_{RGB_j}$) and the relative orientation between the frame camera and the pushbroom HS camera (${}^{RGB} T_{HS}$):

$${}^W T_{HS_k} = {}^W T_{RGB_k} * {}^{RGB} T_{HS} \quad (3)$$

Due to the difference in the acquisition frequencies and an unknown relative pose between the two sensors, the pose of the pushbroom HS camera cannot be extracted directly from the pose of the frame camera.

As described in Sec. V-B, the image of the frame camera and the pushbroom scan line can be synchronized using a parameter τ , which allows us to estimate the pose of the frame camera at the moment at which the pushbroom scan line was acquired (${}^W T_{RGB_k}$). Each pose is computed by interpolating between two boundary poses of the frame camera, with the interpolation of the position being calculated using linear interpolation [33], while the SLERP method [34] is used for the orientation.

The unknown relative pose ${}^{RGB} T_{HS}$ is finally estimated by solving nonlinear least squares problem based on the previously described corresponding matches between the RGB images $x_{i,j}$ and the HS stack $x'_{i,k}$. For each pair of matched points, the error is calculated as *Symmetric Epipolar Distance* [35] where the distance of a point from its projected epipolar line is computed in each of the images. The fundamental matrix F is constructed using the estimated pose of

pushbroom HS camera and known intrinsic parameters of the cameras, while the homogeneous coordinates of the epipolar lines in both images are $l = Fx'_{i,k}$ and $l' = F^T x_{i,j}$. The error is then computed as:

$$\begin{aligned} \text{SED}_i^2 &= d^2(x_{i,j}, l) + d^2(x'_{i,k}, l') \\ &= \left(\frac{1}{l_1^2 + l_2^2} + \frac{1}{l_1'^2 + l_2'^2} \right) R_i^2 \end{aligned} \quad (4)$$

To estimate the unknown parameters Eq. (1), the cost function is minimized over the complete set of matches:

$$\operatorname{argmin} \sum_{i=1}^m \text{SED}_i^2 \quad (5)$$

where m is the total number of corresponding matches over the RGB image set and the HS stack.

VI. HYPERSPECTRAL DRAPING

Once the complete trajectory of the pushbroom HS camera is computed using the estimated unknown parameters, the 3D model, reconstructed using RGB frame camera images, can be draped with the hyperspectral information by projecting data utilizing the pushbroom trajectory. An RGB image of the hyperspectral is created for visual analysis by extracting data centered on the R,G,B filters with wavelengths of 700 nm, 530 nm, 470 nm. The hyperspectral image stack is draped over the 3D model as follows:

- For each 3D point in the textured map of 3D model, \mathbf{X}_i , the point is projected onto the image plane of the pushbroom HS camera at all pushbroom HS camera poses:

$$\mathbf{x}_a = P_a \mathbf{X}_i \quad (6)$$

where $a = 1 \dots \text{numPose}$ is the HS frame and x_a is the projected image point for all pushbroom poses. The projected point is further undistorted using known distortion parameters for the pushbroom HS camera.

- Given that the pushbroom image is a single scanline represented by $y = 0.5$ with N pixels on the x-axis, the hyperspectral frame (\mathbf{a}) of the projection for the point \mathbf{X}_i is selected based as the one with the minimum distance from y component of the projected image point:

$$a = \operatorname{argmin} \mathbf{y}_a - 0.5 \quad (7)$$

- Once the best hyperspectral frame (\mathbf{a}) is selected, the projected point is additionally filtered based on the distance from the scan line and that the projection is within the sensor size of the pushbroom HS camera in the x-axis. If the distance of the projected point from step 2 to the closest actual pixel in pushbroom image is greater than a threshold (i.e., 5 pixels), the point is discarded because it is considered to not be seen by the pushbroom camera. While ideally, the threshold would be 1 pixel, small increase helps compensate for slight misalignment.

- Finally, for the computed image pixel in the hyperspectral image, the hyperspectral information is extracted and assigned to the current 3D point on the 3D model in Fig. 4.

VII. EXPERIMENTAL EVALUATION

The method described in this paper was developed for offline processing of data to create the 3D hyperpsectral map. The pre-processing of RGB images to create 3D map using openMVG/openMVS took 2 hrs to complete. The optimization run finished in 2 mins while the final draping of hyperspectral information took 2-3 hrs.

A. Equipment and setup

The deep-sea data was acquired with the Ariane HROV from IFREMER [36]. Two different cameras were used, a frame camera and a pushbroom HS camera as shown in the setup in Fig. 1. The specifications of the cameras are listed in the table below:

	Resolution	Frame rate	Channels
Frame camera	1920 × 1080	25 fps	(R,G,B)
Pushbroom HS camera	1920 × 1	33 fps	320-750 nm (1 nm each)

TABLE I: Camera specifications.

B. Initial values for the unknown parameters

The translation between the two cameras was physically measured with a tape measure. The relative orientation of the hyperspectral camera to the RGB camera follows a typical parallel stereo pair configuration, where the angles are approximately zero, and these values were used for the initialization.

For the initial value of the time-shift, there is a dark frame from the HS scan lines stack at the time when the lighting on the platform was turned off. This frame is taken as the reference for the frame ID of the images in both cameras. The frame rate for both cameras is known. Using Eq. (2) for the referenced dark frame, the initial value of time-shift in number of frames is computed.

C. Optimization

The optimization was performed in two steps. In the first step, the initial guess (Sec. VII-B) for the unknown parameters was used. In the second optimization step, the estimated parameter values from the first step were input as the initial parameter values.

The first optimization uses all the observation points. The outliers are discarded when the SED error is greater than 25 pixels and the remaining observation points are used for the second step. The maximum change occurs in the time-shift and a small change in t_y and t_z , while all others have negligible changes.

Fig. 7 shows plot of the SED error. The red plot shows the residues for the initial guess of the parameter vector before optimization. The first optimization step reduces the error, which is shown in the blue plot. The plot before optimization and the plot of the first optimization step uses

all observation points. The green plot is after the second optimization which uses only the observations points after removing outliers from the first optimization result. The error was further reduced. For some points, the error increased by a small value in the second step, which is due to the shift of the optimization towards the region with more data points.

Fig. 8 shows the distribution plot of the SED error. The red points are marked as outliers with SED errors larger than 25 pixels after the first optimization. The yellow points show the points with error larger than 10 pixels and the green points show the points with error smaller than 10 pixels computed after the second optimization. The distribution of residues shows many green points on the right side of the scene, which has more points concentrated in a small region. The left side of the scene is less distinct due to insufficient illumination, which also affects the geometry. The middle and the lower part again have less green points due to the poor geometry of the scene. There could be two reasons for the high error, first that the points do not match well and second that the optimization shifts towards the region with more concentrated points and better geometry, so the points in the other region produce a higher error.

The blue plot in Fig. 7 shows the error after the first optimization and depicts that the SED error was higher for many points which were discarded before the second step. Since, the points were manually clicked and matched, in the next section we discuss the quality of outliers.

D. Outlier analysis

In parameter estimation, the outliers, i.e., the points with SED error larger than 25 pixels, are selected. We have visually analysed 20 outliers with highest error values in the optimization described in above section. From the visual analysis, it can be deduced whether the points are poorly matched or not. Moreover, Fig. 8 shows that the region with more outliers has a lower number of matched points.

However, discarding outliers for the second optimization

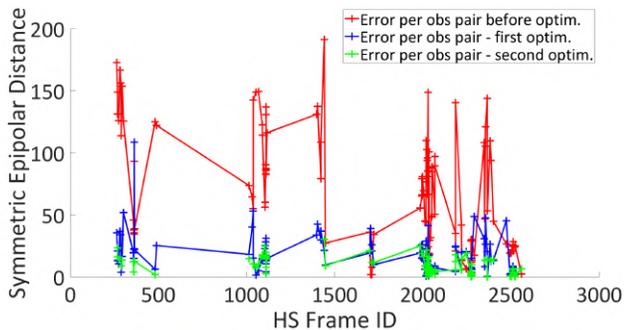


Fig. 7: SED error per obs. pair. Red: Before optimization using initial guess for params value. Blue: first estimation step using the initial guess for params values and all the observation points. Green: second estimation step using the estimate params of first step as initial values and excluding the outliers selected after first optim. Total points: 115 ; Inliers: 88.

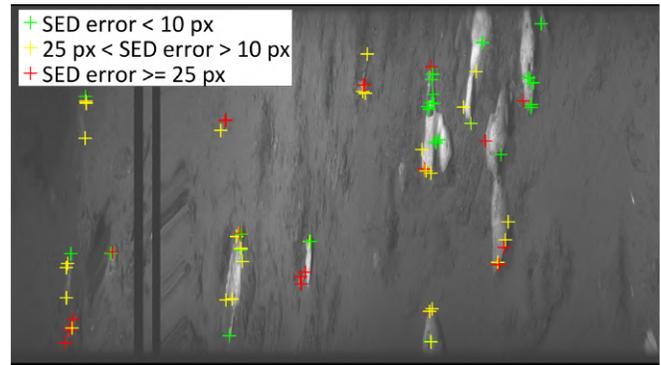


Fig. 8: Plot of the observed points on the pushbroom image with SED error distribution. Green: error less than 10 pixels, Yellow: error between 10 and 25 pixels, Red: error larger than or equal to 25 pixels.

step did not greatly change the value of the estimated parameters. Visual analysis shows that these points are not a poor correspondence match. Fig. 9 shows the point with highest SED error after first optimization and it can be seen that the point is approximately in the same location in both the hyperspectral image (left) and the RGB image (right). Therefore, it can be said that the points with the highest error are not outliers because they represent a good match. Out of 27 outliers, 20 were visually analysed and they were a good match. The reason for the high error values could be that due to less number of points in these regions, the optimization shifts towards the region with more points, resulting in high error values for the rest. The reason for fewer matching points in these regions is the poor visibility of the objects and also the distorted geometry which makes it difficult to click the matching points manually. This can be improved by an automated feature detection and matching method, which can detect a larger number of points in regions with fewer objects or poor visibility.

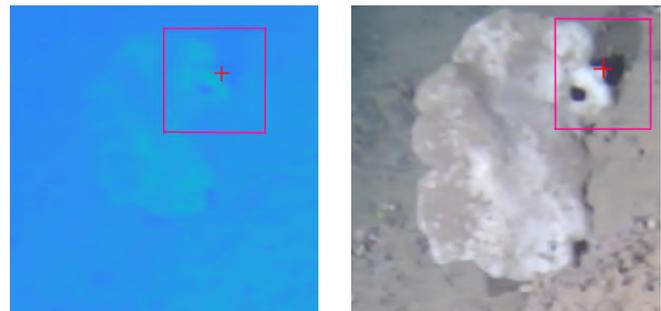


Fig. 9: SED error = 117.08 pixels; Left: hyperspectral image; Right: RGB camera image

E. Monte Carlo simulation

Monte Carlo simulation [37] is a stochastic algorithm based on a large number of identical experiments with random inputs. It is used to understand the effects of noise and uncertainty. It is performed by randomly selecting samples

and adding noise to the observations which are given as input to the algorithm to estimate the model parameters for each sample. This is repeated for a very large number of times to study the properties of the estimator and also the variance in the estimated parameters.

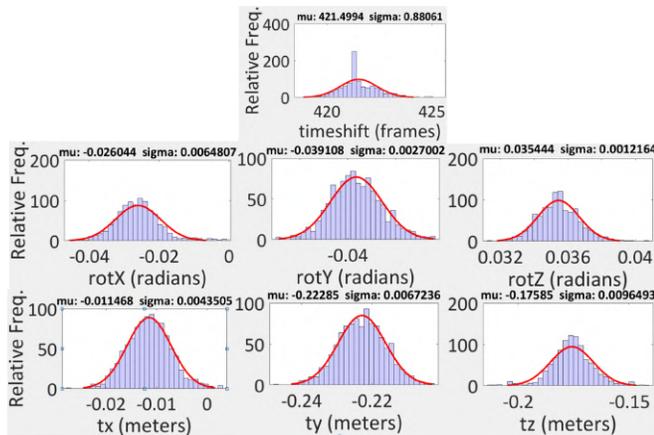


Fig. 10: Monte Carlo plot for all the seven estimated parameters.

The plots in Fig. 10 show that the estimator is consistent because the estimated values of the parameters for a large set of samples, which in this case is 1000, are close to the expected values.

F. Hyperspectral mapping

The final goal of the paper is to use the estimated parameters and create a 3D map of the deep-sea surface on which the hyperspectral information from the pushbroom scan lines is draped. The mapping process is described in Sec. VI.

Fig. 4 is the 3D mesh created from the RGB images that serves as input to the hyperspectral mapping function. For each 3D point on this mesh, the hyperspectral data shown in Fig. 11 is projected as explained in Sec. VI.

Fig. 12 shows the top view of the two 3D maps, where the right image is the 3D mesh created from frame images, and the left is the 3D mesh projected with hyperspectral information using the estimated parameters. The whole area of the 3D mesh is well fitted with the hyperspectral information and the draping quality is good. However, there are local distortions in the hyperspectral map. The object under the pink rectangle has half of the area occupied by ground color information. The reason for this distorted information draping can be deduced from Fig. 11. The view angle of the pushbroom HS camera was limited and it never captured the information from the side of the object in question. The frame camera has a wide angle of view and thus the frame images capture greater detail, which is better represented when a mesh is created. The single line captured by the pushbroom HS camera restricts such detailed information.

The object under the green rectangle in Fig. 12 has a good fit with the hyperspectral information. There are some shifts, but they are negligible. Again, these shifts can be attributed



Fig. 11: Input RGB image extracted from hyperspectral image to be used as a visual template to assess the quality of hyperspectral draping. The marked regions are analysed for the quality of hyperspectral draping.

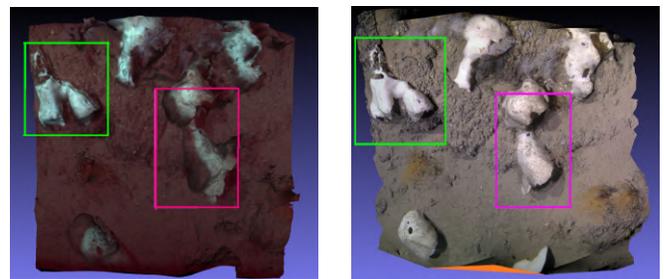


Fig. 12: (Left) 3D map projected with hyperspectral information. (Right) 3D map generated from frame camera images. Marked areas show local distortions and are analyzed in the text

the larger field of view of the frame camera, which captures more detail than the pushbroom HS camera.

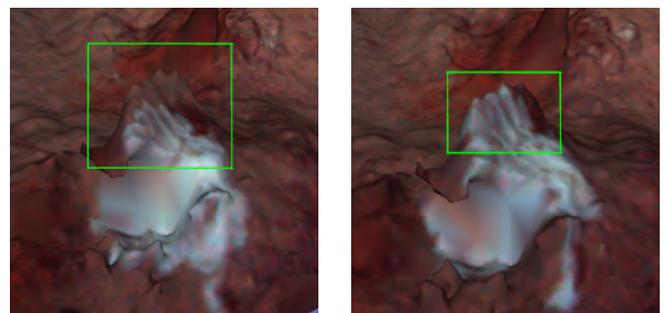


Fig. 13: (Left) Object on HS map with initial parameter values. (Right) Object on HS map with estimated parameter values

Fig. 13 shows an object on the hyperspectral map. The left image is with the parameter values before optimization, the hyperspectral information does not fit completely on the object structure and a shift can be observed. The right image is the hyperspectral map using the estimated parameter values. It shows that the hyperspectral information covers the full object and the shift has been corrected.

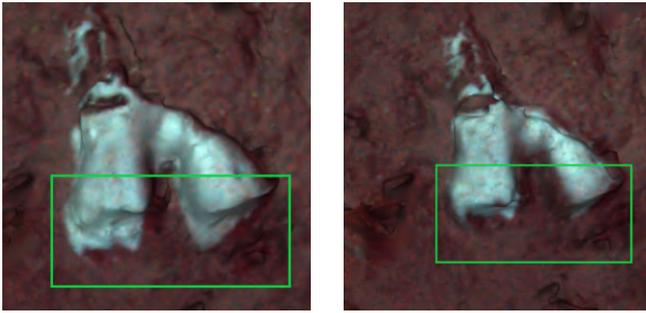


Fig. 14: (Left) Object on HS map with initial parameter values. (Right) Object on HS map with estimated parameter values

Fig. 14 shows the shift correction for another object. The estimated parameters fits the hyperspectral information perfectly which covers the correct geometry of the object.

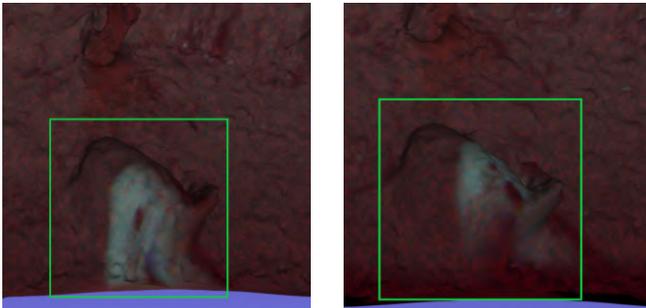


Fig. 15: (Left) Object on HS map with initial parameter values. (Right) Object on HS map with estimated parameter values

Fig. 15 shows a third example. However there is a local distortion for this object as it lies at the edge of the hyperspectral. Fig. 11 shows this object marked with blue rectangle at the bottom. As compared with the 3D RGB map in Fig. 4, the geometry of the objects near the edge is distorted which could be a possible reason for a bad fit of the hyperspectral information on this object. Therefore, from the observation of three objects in different regions of the map, it can be concluded that the estimated parameters drapes the hyperspectral information on the 3D map matching the geometry of the object.

VIII. CONCLUSION

In this paper, we present an approach to 3D hyperspectral mapping using only image processing methods. The imaging setup consists of an RGB frame camera and a pushbroom HS camera in a parallel stereo configuration. Data were acquired for the seafloor by mounting the setup on an HROV. The points are manually matched between the RGB images and the pushbroom scan lines. Then, the symmetric epipolar distance (SED), error is calculated for all matched points and minimized using the nonlinear least squares method to estimate the relative pose and time-shift between the two cameras.

There is no research work that addresses the problem of deep-sea 3D hyperspectral mapping yet. All available approaches are for terrestrial or shallow water scenarios using aerial platforms or surface vessels. The main contributions of this work are:

- Development of a 3D hyperspectral mapping method for deep-sea surfaces by coupled use of a frame camera and a linear pushbroom HS camera.
- Eliminate the use of GPS and navigation data and only image processing based solution.

Our approach gives good results of the final hyperspectral mapping of the deep-sea surface. The optimization problem works well to estimate the unknown parameters. As future work, the method is being extended with automatic feature detection and matching which can eliminate the problem of few points in certain regions with poor visibility and distorted geometry which is the drawback in manual search. Also, there are local distortions in the final result of the generated hyperspectral map, which can be improved with automatic feature detection.

ACKNOWLEDGMENTS

This study is part of the joint Marha “Marine Habitats” (LIFE16 IPE FR001_MARHA) and “EUMarineRobots” projects funded by the European Union. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 712949 (TECNIOspring PLUS).

REFERENCES

- [1] M. Moroni, E. Lupo, E. Marra, and A. Cenedese, “Hyperspectral image analysis in environmental monitoring: setup of a new tunable filter platform,” *Procedia Environmental Sciences*, vol. 19, no. 0, p. 885, 2013.
- [2] W. M. Klonowski, P. R. Fearn, and M. J. Lynch, “Retrieving key benthic cover types and bathymetry from hyperspectral imagery,” *Journal of Applied Remote Sensing*, vol. 1, no. 1, p. 011505, 2007.
- [3] S. Asadzadeh and C. R. de Souza Filho, “A review on spectral processing methods for geological remote sensing,” *International journal of applied earth observation and geoinformation*, vol. 47, pp. 69–90, 2016.
- [4] I. Leifer, W. J. Lehr, D. Simecek-Beatty, E. Bradley, R. Clark, P. Dennison, Y. Hu, S. Matheson, C. E. Jones, B. Holt, *et al.*, “State of the art satellite and airborne marine oil spill remote sensing: Application to the bp deepwater horizon oil spill,” *Remote Sensing of Environment*, vol. 124, pp. 185–209, 2012.
- [5] P. J. Mumby, W. Skirving, A. E. Strong, J. T. Hardy, E. F. LeDrew, E. J. Hochberg, R. P. Stumpf, and L. T. David, “Remote sensing of coral reefs and their physical environment,” *Marine pollution bulletin*, vol. 48, no. 3–4, pp. 219–228, 2004.
- [6] M. Richardson and P. Kumar, “Critical zone services as environmental assessment criteria in intensively managed landscapes,” *Earth’s Future*, vol. 5, no. 6, pp. 617–632, 2017.
- [7] E. Raugel, J. Opderbecke, M. C. Fabri, L. Brignone, and V. Rigaud, “Operational and scientific capabilities of ariane, Ifremer’s hybrid roV,” in *OCEANS 2019 - Marseille*, 2019.
- [8] A. Arnaubec, J. Opderbecke, A.-G. Allais, and L. Brignone, “Optical mapping with the ariane hrov at ifremer: The matisse processing tool,” in *OCEANS 2015 - Genova*, 2015.
- [9] R. Muller, M. Lehner, R. Muller, P. Reinartz, M. Schroeder, and B. Vollmer, “A program for direct georeferencing of airborne and spaceborne line scanner images,” *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 34, no. 1, pp. 148–153, 2002.

- [10] A. A. Mogstad, G. Johnsen, and M. Ludvigsen, "Shallow-water habitat mapping using underwater hyperspectral imaging from an unmanned surface vehicle: A pilot study," *Remote Sensing*, vol. 11, no. 6, p. 685, 2019.
- [11] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "Hydice system: Implementation and performance," in *Imaging Spectrometry*, vol. 2480, pp. 258–267, International Society for Optics and Photonics, 1995.
- [12] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—a modern synthesis," in *International workshop on vision algorithms*, pp. 298–372, Springer, 1999.
- [13] J. P. Kerekes and J. E. Baum, "Hyperspectral imaging system modeling," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 117–130, 2003.
- [14] D. L. Bongiorno, M. Bryson, T. C. Bridge, D. G. Dansereau, and S. B. Williams, "Coregistered hyperspectral and stereo image seafloor mapping from an autonomous underwater vehicle," *Journal of Field Robotics*, vol. 35, no. 3, pp. 312–329, 2018.
- [15] D. L. Bongiorno, A. J. Fairley, M. Bryson, and S. B. Williams, "Automatic spectrometer/rgb camera spatial calibration," in *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pp. 4431–4434, IEEE, 2013.
- [16] K. Barbieux, D. Constantin, and B. Merminod, "Correction of airborne pushbroom images orientation using bundle adjustment of frame images," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 41, 2016.
- [17] K. Barbieux, "Pushbroom hyperspectral data orientation by combining feature-based and area-based co-registration techniques," *Remote Sensing*, vol. 10, no. 4, p. 645, 2018.
- [18] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [19] Y. Rzhanov and S. Pe'eri, "Pushbroom-frame imagery co-registration," *Marine Geodesy*, vol. 35, no. 2, pp. 141–157, 2012.
- [20] D. Schläpfer and R. Richter, "Geo-atmospheric processing of airborne imaging spectrometry data. part 1: parametric orthorectification," *International Journal of Remote Sensing*, vol. 23, no. 13, pp. 2609–2630, 2002.
- [21] H. Tuo and Y. Liu, "A new coarse-to-fine rectification algorithm for airborne push-broom hyperspectral images," *Pattern recognition letters*, vol. 26, no. 11, pp. 1782–1791, 2005.
- [22] J. Wang, Y. Ge, G. B. Heuvelink, C. Zhou, and D. Brus, "Effect of the sampling design of ground control points on the geometric correction of remotely sensed imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 18, pp. 91–100, 2012.
- [23] D. Poli, L. Zhang, and A. Gruen, "Orientation of satellite and airborne imagery from multi-line pushbroom sensors with a rigorous sensor model," *International Archives of Photogrammetry and Remote Sensing*, vol. 35, no. B1, pp. 130–135, 2004.
- [24] C.-K. Yeh and V. J. Tsai, "Direct georeferencing of airborne pushbroom images," *Journal of the Chinese Institute of Engineers*, vol. 38, no. 5, pp. 653–664, 2015.
- [25] A. Habib, W. Xiong, F. He, H. L. Yang, and M. Crawford, "Improving orthorectification of uav-based push-broom scanner imagery using derived orthophotos from frame cameras," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 1, pp. 262–276, 2016.
- [26] P. Moulon, P. Monasse, R. Marlet, and Others, "Openmvg." <https://github.com/openMVG/openMVG>.
- [27] K. Istemic, V. Ila, L. Polok, N. Gracias, and R. García, "Mission-time 3d reconstruction with quality estimation," in *OCEANS 2017-Aberdeen*, pp. 1–9, IEEE, 2017.
- [28] D. Cernea, "Openmvs: Open multiple view stereovision," 2015.
- [29] T. Lindeberg, "Scale invariant feature transform," 2012.
- [30] D. D. Morrison, "Optimization by least squares," *SIAM Journal on Numerical Analysis*, vol. 5, no. 1, pp. 83–88, 1968.
- [31] J. J. Moré, "The levenberg-marquardt algorithm: implementation and theory," in *Numerical analysis*, pp. 105–116, Springer, 1978.
- [32] C. Albl, Z. Kukulova, A. Fitzgibbon, J. Heller, M. Smid, and T. Pajdla, "On the two-view geometry of unsynchronized cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4847–4856, 2017.
- [33] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 710–719, 2004.
- [34] V. E. Kremer, "Quaternions and slerp," *Embots. dfki.de/doc/seminar_ca/Kremer-Quaternions.pdf*, 2008.
- [35] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [36] L. Brignone, E. Raugel, J. Operderbecke, V. Rigaud, R. Piasco, and S. Ragot, "First sea trials of hrov the new hybrid vehicle developed by ifremer," in *OCEANS 2015 - Genova*, pp. 1–7, 2015.
- [37] C. Z. Mooney, *Monte carlo simulation*, vol. 116. Sage publications, 1997.