

A Methodology for Consistent Georegistration in Underwater Hyperspectral Imaging

Håvard Snefjellå Løvås¹, Aksel Alstad Mogstad¹, Asgeir J. Sørensen¹, Senior Member, IEEE, and Geir Johnsen¹

Abstract—This article proposes a novel methodology for precise georegistration in underwater hyperspectral imaging (UHI) using a red-green-blue (RGB) camera to build a photogrammetry model that estimates the pose and a 3-D seabed model. The two main scientific contributions are: 1) the development of two methods for geometric calibration of the hyperspectral imager (HSI) with the RGB camera and 2) the development of a methodology for consistent georegistration of the pushbroom hyperspectral imagery on the seabed using the photogrammetry model and the geometrically calibrated HSI. The georegistration uses a calibrated RGB camera, fixed to the HSI, with an overlapping field of view. The RGB images are used to build a photogrammetry model. Through utilizing the geometric HSI parameters from the calibration, the hyperspectral imagery is ray cast onto the 3-D model. This methodology is exemplified for UHI from a remotely operated vehicle on a cold-water coral reef in the Trondheim Fjord, Norway. The precision is unprecedented for georegistration in UHI, demonstrated by the first-ever spatially consistent UHI mosaic to contain multiple transects.

Index Terms—Mapping and monitoring, photogrammetry, underwater hyperspectral imaging (UHI).

I. INTRODUCTION

UNDERWATER hyperspectral imaging (UHI) is a promising field for close-range nonintrusive classification and mapping of objects of interest (OOIs) such as geology [1], biology [2], and archaeology [3]. In contrast to regular red-green-blue (RGB) imaging, UHI captures a high-resolution spectrum for each pixel. This allows for classifying OOIs based on their full spectral fingerprints, rather than coarse RGB color estimates. The spectral resolution also makes UHI suited for inherent/apparent optical property (IOP/AOP) correction since the properties are wavelength specific. Spatially, most hyperspectral imagers (HSIs) capture 1-D images, or slit images,

Manuscript received March 30, 2021; revised July 14, 2021; accepted August 20, 2021. Date of publication November 9, 2021; date of current version April 13, 2022. This work was supported by the Research Council of Norway through the Centre of Excellence funding scheme of the Centre for Autonomous Marine Operations and Systems, Norwegian University of Science and Technology under Grant 223254. (Corresponding author: Håvard Snefjellå Løvås.)

Associate Editor: B. Thornton.

Håvard Snefjellå Løvås and Asgeir J. Sørensen are with the Centre for Autonomous Marine Operations and Systems, the Department of Marine Technology, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: havard.s.lovas@ntnu.no; asgeir.sorensen@ntnu.no).

Aksel Alstad Mogstad is with the Centre for Autonomous Marine Operations and Systems, the Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: aksel.a.mogstad@ntnu.no).

Geir Johnsen is with the Centre for Autonomous Marine Operations and Systems, the Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway, and also with the University Centre in Svalbard, 9171 Longyearbyen, Norway (e-mail: geir.johnsen@ntnu.no).

Digital Object Identifier 10.1109/JOE.2021.3108229

and map an area through pushbroom scanning [4]. The part of mapping concerned with computing earth-fixed coordinates for measurements is called georegistration (also called georeferencing or geometric correction). Building a spatially consistent UHI map, a mosaic, is challenging since it requires a precision in the georegistration of a similar magnitude as the resolution of the map. Achieving this precision underwater requires precise pose (position and orientation) measurements that are consistent with a high-resolution 3-D model of the seabed. A direct consequence of a precise georegistration is that light rays' optical paths can be determined geometrically from light source to seabed to the HSI. This allows for physics-based correction for IOPs/AOPs and generation of consistent reflectance spectra invariant to water body type and optical path.

The georegistration of passive optical sensors, such as an HSI, should have three main components:

- 1) a model of the 3-D surface from which the measurements are taken;
- 2) pose (position and orientation) measurements;
- 3) a calibrated sensor model describing the geometric configuration of spectral measurements relative to the pose.

The method for georegistration is chosen based on the operational considerations, in particular the range and resolution of the mapping. For long-range georegistration, such as remote sensing from satellites, there likely exists a sufficiently accurate 3-D model of the earth, and position measurements are highly accurate compared with the range. However, for close-range georegistration such as for UHI, there is no sufficiently accurate 3-D surface model or pose (especially position) measurements available. In UHI, the spatial resolution is about 1 cm, and the uncertainty in, e.g., acoustic positioning is two orders of magnitude higher. The solutions in close-range georegistration are, therefore, generally characterized by statistical methods that use measurements of the 3-D world to simultaneously estimate the pose and a 3-D model of the scene [5]. These are either direct range measurements from time-of-flight methods using optics or acoustics or methods from computer vision, such as structure from motion (SfM) [6, Ch. 7] that use 2-D images to estimate 3-D shape. More broadly, SfM can be used in building a photogrammetry model. A photogrammetry model built from RGB images simultaneously computes RGB camera pose and a 3-D seabed model in a consistent way. Thus, an RGB camera (or multiple) fixed to an HSI can facilitate consistent georegistration in UHI of near photogrammetry precision. A calibrated geometric sensor model of the HSI enables the transformation from RGB camera pose to the HSI measurement rays. The

georegistration is finalized by finding the intersections of the HSI rays with the 3-D model, a process called ray casting.

The standard georegistration procedure in UHI [7] and [8] uses the measured pose of the remotely operated vehicle (ROV) and combines a point altimeter with a flat seabed assumption to model the 3-D seabed. However, as stated in [1], acoustic positioning alone is not accurate enough for high-quality UHI mosaics. Therefore, they used ROV video and SfM to refine the pose estimates of the vehicle. Similarly, in [9], SfM was used to improve the pose estimates for the ROV and georegistration for UHI. Hence, their method improved the georegistration relative to the standard procedure by improving the second component of georegistration. However, to model the seabed, they simply used one altitude for each slit image and assumed the seabed to be flat for each slit image, as shown in Fig. 6(a). Notably, the HSI was not calibrated with the RGB camera for either of these UHI approaches, and their work did not state how the relative translations and rotations between the two sensors were estimated. Prior work used a factory calibration describing the view angle of each pixel to describe the distribution of measurements along the HSI's field of view (FoV). Hence, former georegistration methods using SfM in UHI did not account for across-track variations in bathymetry and did not calibrate the HSI with the RGB camera. Another related work, although not UHI, [10] showed promising results in the georegistration of single-beam spectrometer measurements. They used a stereo camera with an overlapping FoV with the spectrometer for 3-D georegistration on the seabed. Contrary to the georegistration approaches in UHI, this method had comprehensive models for all three components of georegistration. They used a stereo-camera-based photogrammetry for estimating the pose and 3-D model of the seabed [11]. They also used a calibrated geometric sensor model for the spectrometer, relative to the stereo camera [12]. Their method allowed for state-of-the-art *in situ* IOP/AOP correction.

In this article, we present a methodology for improved georegistration in UHI. The 3-D model and pose estimates are established using an RGB camera with an overlapping FoV to build a photogrammetry model. We improve upon prior georegistration methods in UHI by using the 3-D model from the photogrammetry and by geometrically calibrating the HSI with the RGB camera. A geometric sensor model for the HSI is suggested, and two calibration methods are presented for calibrating the model. To demonstrate the georegistration, we consider a UHI mapping survey of the Tautra cold-water coral reef in the Trondheim Fjord, Norway. In contrast to prior georegistration methods using SfM, we demonstrate a UHI mapping consisting of multiple overlapping transects, allowing us to test and verify the consistency in georegistration.

Mapping of cold-water corals is challenging due to the sites on which they are found. The cold-water corals at Tautra are exposed to high currents. This leads to unwanted translational and rotational motions for an ROV. According to [13, Preface], cold-water coral reefs are also characterized by complex morphology and high ruggedness [see Figs. 2 and 11(a)]. The combined effects of unwanted motion and the complex 3-D structure of the corals make georegistration challenging and

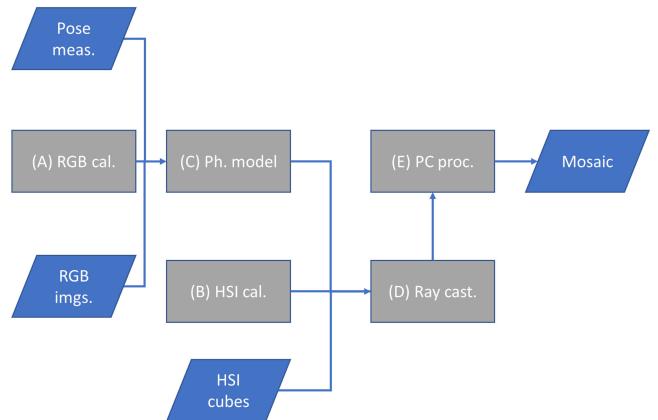


Fig. 1. Georegistration model *in situ*. The gray rectangles denote processes, while the blue parallelograms are data.

a suitable example for demonstrating the performance of the proposed methods.

Two main scientific contributions are proposed: 1) the development of two methods for geometric calibration of the HSI with the RGB camera and 2) the development of a methodology for consistent georegistration of the pushbroom hyperspectral imagery on the seabed with the pose and 3-D model from photogrammetry and the geometrically calibrated HSI.

The rest of this article is organized as follows. Section II presents methods for georegistration in UHI. It presents mathematical models and approaches for georegistration and calibrations in the laboratory and *in situ*. Section III presents results related to the performance of georegistration and calibrations in the laboratory and *in situ*. A discussion of the methods and results are given in Section IV. Finally, Section V concludes this article.

II. METHODS FOR GEOREGISTRATION IN UHI

A. Overview of Methods

In this section, we will present our methods for georegistration and calibration in UHI. We start by defining these two terms in the context of this article.

- 1) *Georegistration* is the process of estimating the 3-D earth-fixed position on the seabed corresponding to a pixel measurement from the HSI. The georegistration model includes the geometric HSI sensor model, pose for the RGB camera, and a 3-D seabed model.
- 2) *Calibration* refers to the process of refining the geometric HSI sensor model to yield the best transformation between the 3-D earth-fixed positions on the seabed and the HSI pixel positions. Hence, the calibration process uses the georegistration model.

The georegistration model is illustrated by the simple flow diagram shown in Fig. 1. The inputs in Fig. 1 are the timestamped pose measurements, RGB images, and HSI slit images denoted by *Pose meas.*, *RGB imgs.*, and *HSI cubes*, respectively. The

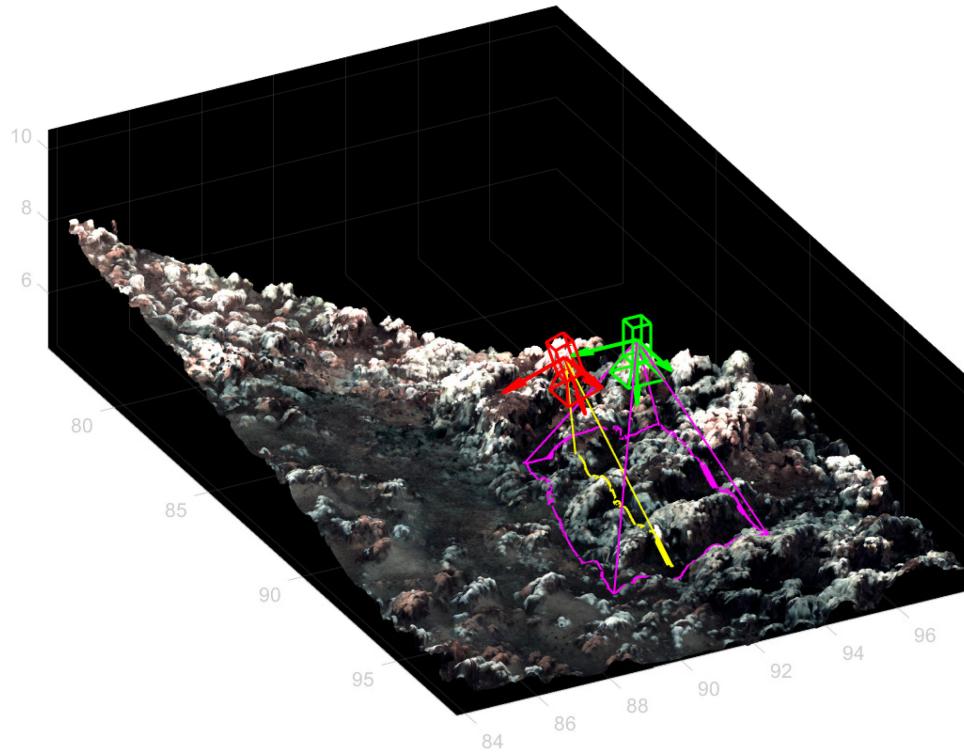


Fig. 2. Georegistration concept. The red camera is the HSI, while the green camera is the RGB camera, and these are fixed to each other. The arrows show their corresponding reference frames denoted by the HSI frame and the RGB frame. The straight yellow lines show the FoV, and the curly yellow line shows ray intersections of the HSI with the seabed, or the georegistration of a slit image. The purple outline shows the FoV of the RGB camera corresponding to an image. Note that rotation and translation between the two sensors are highly exaggerated for the purpose of visualization. For the actual configuration, see Fig. 7(a). Also, in reality, the altitude is about half of this example, meaning that the ruggedness of the seabed is around twice as high in the perspective of the HSI [see Fig. 11(a)]. The earth-fixed coordinates are expressed in meters.

output is a hyperspectral map, a mosaic of the seabed referred to as *Mosaic*. The processes indicated by the gray rectangles in Fig. 1 correspond to the subsections in this section, which are the following.

- 1) *RGB calibration (RGB cal. in Fig. 1)*: Section II-C gives a brief description of how the RGB camera is calibrated in the laboratory. The calibration yields RGB intrinsics and RGB laboratory extrinsics. The intrinsic parameters are used for building the photogrammetry model, while the RGB laboratory extrinsics are used for the HSI laboratory calibrations.
- 2) *HSI calibrations (HSI cal.)*: Section II-D describes the two HSI calibration methods for estimating the HSI sensor model, including the HSI intrinsics and the HSI-RGB extrinsics. HSI calibrations are divided in the following parts.
 - a) *HSI model*: In Section II-D1, the geometric model of the HSI is described. The model describes the HSI rays geometrically.
 - b) *Method 1*: In Section II-D2, a conventional laboratory calibration method is proposed.
 - c) *Method 2*: In Section II-D3, a method for both laboratory and *in situ* calibration is proposed.
- 3) *Photogrammetry model (Ph. model)*: Section II-E describes the establishment of the photogrammetry model.

The main photogrammetry process uses the photogrammetry software, *Agisoft Metashape Professional*.

- 4) *Ray casting (Ray cast.)*: Section II-F describes how the photogrammetry model and HSI calibration parameters are used for finding the intersections between HSI rays and the 3-D seabed. This yields 3-D georegistration for each spectral measurement, creating a point cloud. The ray casting is also used for *in situ* HSI calibration with Method 2.
- 5) *Point cloud processing (PC proc.)*: Section II-G explains how the slit images are mosaicked using the point cloud. It also explains a simple IOP/AOP correction that is applied.

In the laboratory, the georegistration model is slightly simpler than in Fig. 1. The photogrammetry step 3 is removed. Instead, the RGB extrinsics are computed in the RGB calibration in step 1, and the 3-D model of the seabed is exchanged with the plane defined by the calibration target.

B. Key Definitions

Fig. 2 visualizes the geometry of the georegistration. The RGB camera is symbolized by the green camera, while the HSI is shown with the red camera. The seabed is modeled by a colored 3-D triangular mesh from the photogrammetry model. The following key definitions are described from Fig. 2.

- 1) *RGB intrinsics*: The intrinsic parameters define the projection from the RGB camera frame (indicated by green arrows) onto the image plane. For example, how the straight purple lines in Fig. 2 are projected from the RGB frame to the corner pixels of the RGB image. In this article, we include the distortion parameters in the intrinsics.
- 2) *RGB extrinsics*: The extrinsic parameters define the translations and rotations of the RGB camera relative to the earth-fixed frame, as indicated by the green camera in Fig. 2. In the context of laboratory calibration, the earth-fixed frame is defined by the planar calibration target.
- 3) *HSI-RGB extrinsics*: The HSI-RGB extrinsic parameters define the fixed pose of the HSI relative to the RGB camera. Combining them with RGB extrinsic parameters yields the HSI extrinsic parameters or HSI pose, as indicated by the red camera in Fig. 2.
- 4) *HSI intrinsics*: The HSI intrinsic parameters define the mapping from the slit image plane to the HSI frame indicated by red arrows.
- 5) *Photogrammetry model*: The photogrammetry model simultaneously computes the 3-D model, here shown as the colored 3-D triangular mesh model of the seabed, and the RGB extrinsics for each image taken, here exemplified by the green camera.
- 6) *HSI ray*: By HSI ray, we are referring to the parametric description of a line in the earth-fixed frame that corresponds to a pixel measurement. The HSI ray is defined by the RGB extrinsics, the HSI-RGB extrinsics, and the HSI intrinsics. The straight yellow lines in Fig. 2 are HSI rays corresponding to the outermost pixels in the slit image. The ray casting procedure computes the intersection of rays with the mesh model as shown for all pixels by the curly yellow line.

C. RGB Calibration

We start by calibrating the RGB camera model in the laboratory to find the RGB intrinsics and RGB extrinsics. Note that vectors and matrices are denoted by bold font. The RGB camera is calibrated using MATLAB's *Camera Calibrator application* based on [14]. It uses the pinhole camera model as shown in the following equation with radial and tangential distortion as described in (3):

$$w[u', v', 1] = [X, Y, Z, 1] \begin{bmatrix} \mathbf{R}_1(\Theta) \\ \mathbf{t} \end{bmatrix} \mathbf{K} \quad (1)$$

where w is a scalar, and $[u', v']$ are the distorted pixel positions in the image plane. $[X, Y, Z]$ are earth-fixed points. The rotation matrix $\mathbf{R}_1(\Theta)$ and the translation vector \mathbf{t} transform points between the RGB frame and the earth-fixed frame. The RGB extrinsics are the rotations Θ and translations \mathbf{t} of the earth-fixed frame relative to the RGB frame. This definition is common in camera calibration, but for *in situ* calibration, Θ refers to the rotation of the RGB frame relative to the earth-fixed frame, meaning that the rotation matrix in (1) should be transposed.

Moreover, the intrinsic camera matrix \mathbf{K} in (1) is expressed as

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ c_x & c_y & 1 \end{bmatrix} \quad (2)$$

where the intrinsic parameters include the focal lengths f_x and f_y and the principal points c_x and c_y . Skew is not included in the matrix since we assume the pixels to be orthogonal. These are used to project normalized camera coordinates, $[x, y]$, to distorted pixel positions, $[u', v']$. To account for radial and tangential lens distortion and refraction effects, the following equations transform the distorted pixel positions $[u', v']$ to undistorted pixel positions in the image plane $[u, v]$:

$$x = (u' - c_x)/f_x \quad (3a)$$

$$y = (v' - c_y)/f_y \quad (3b)$$

$$r = \sqrt{x^2 + y^2} \quad (3c)$$

$$x' = x(1 + k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + x^2) \quad (3d)$$

$$y' = y(1 + k_1 r^2 + k_2 r^4) + 2p_1(r^2 + y^2) + 2p_2xy \quad (3e)$$

$$u = f_x x' + c_x \quad (3f)$$

$$v = f_y y' + c_y \quad (3g)$$

where x' and y' are normalized image plane coordinates. The RGB intrinsics also include the radial distortion coefficients k_1 and k_2 and the tangential distortion coefficients p_1 and p_2 . In total, (1)–(3) allow for reprojecting an earth-fixed 3-D point $[X, Y, Z]$ to a reprojected pixel position in an RGB image $[u, v]$.

The camera calibration uses the position of known calibration points and corresponding detected pixel positions of calibration points in images to find the RGB extrinsics and intrinsics that yield the smallest error between the detected pixel positions and the reprojected pixel positions. The calibrated intrinsics are used for the photogrammetry and for HSI calibration method 2 in Section II-D3, while the RGB extrinsics are used for both HSI calibrations in Section II-D.

D. HSI Calibrations

This section presents the HSI calibration including the geometric model of the HSI and the two methods for calibrating the parameters of that model: the HSI intrinsics and the HSI-RGB extrinsics. Prior georegistration methods used a factory calibration from the manufacturer, a table describing the view angle of each pixel measurement, rather than an HSI intrinsic parameter model. The HSI-RGB extrinsic parameters were not calibrated for prior methods in UHI, but rather used a best guess. In our case, the factory calibration was not accurate enough, and we wanted a calibrated parametric model for both the HSI intrinsics and the HSI-RGB extrinsics.

The first calibration method suggested is an in-laboratory pushbroom calibration method using a calibration target, while the second calibration method is an *in situ* calibration method that uses correlation between light measured by the HSI with light measured by a calibrated RGB camera. The main practical

benefits of the *in situ* calibration method are that it allows for setups where laboratory calibration is infeasible, it can be used for old surveys where a laboratory calibration would no longer be valid, and it can be used in cases where the RGB camera is calibrated *in situ* to account for differences in refraction from a laboratory environment. The second calibration method also has an advantage in the laboratory because it does not depend on pushbroom motion. Thus, the acquisition of calibration data can be done in conjunction with a regular RGB camera calibration with no additional effort. In our case, we studied a field dataset acquired in 2017, prior to some maintenance work on the HSI. Thus, small changes in HSI-RGB rotations and HSI intrinsics were expected between the calibration parameters in the field and the 2020 laboratory trials. Therefore, an *in situ* calibration method was especially useful for our case.

1) *HSI Model*: In Section II-C, the RGB extrinsics, including the translations \mathbf{t} and rotations Θ , were found. This reduces the number of extrinsic parameters to the HSI-RGB extrinsics consisting of three relative rotations and three relative translations. The relative rotations $\delta\Theta = [\delta\phi, \delta\theta, \delta\psi]$ are expressed about the RGB frame's x -, y -, and z -axes, respectively. The fixed rotation matrix $\mathbf{R}_2(\delta\Theta)$ defines the rotations between the HSI frame and the RGB frame. The three relative translations $\delta\mathbf{t} = [\delta x, \delta y, \delta z]$ are expressed along the RGB frame's x -, y -, and z -axes, respectively. $\delta\Theta$ are expressed as the rotations from the RGB frame to the HSI frame and $\delta\mathbf{t}$ is the vector from the origin of the HSI frame to the origin of the RGB frame, expressed in the RGB frame.

To avoid confusion in notation, HSI-specific parameters are denoted with the superscript h . The camera model for the HSI is adapted from the line-scan camera model, as described in [15]. They describe a calibration method for a line-scan camera using a calibrated 2-D camera (equivalent to our RGB camera) to find the RGB extrinsics. Our HSI camera model is described in the set of following equations:

$$u^h = c_x^h + \Delta u^h + f^h x^h \quad (4a)$$

$$\Delta u^h = k_1^h (u^h - c_x^h)^5 + k_2^h (u^h - c_x^h)^3 + k_3^h (u^h - c_x^h)^2 \quad (4b)$$

$$x^h = (u^h - c_x^h - \Delta u^h) / f^h \quad (4c)$$

$$y^h = 0 \quad (4d)$$

$$\mathbf{a} = [x^h, y^h, 1] \mathbf{R}_2(\delta\Theta) \mathbf{R}_1(\Theta)^\top \quad (4e)$$

$$\mathbf{b} = -(\mathbf{t} + \delta\mathbf{t}) \mathbf{R}_1(\Theta)^\top \quad (4f)$$

$$as + \mathbf{b} = [X, Y, Z] \quad (4g)$$

where u^h is the HSI pixel position starting from the leftmost pixel. The HSI intrinsics include the principal point c_x^h , the focal length f^h , the radial distortion coefficients k_1^h and k_2^h , and the tangential distortion coefficient k_3^h . The distortion coefficients compensate for the lens and refraction effects through the distortion term Δu^h . The normalized HSI frame is denoted by $[x^h, y^h, 1]$. The direction vector \mathbf{a} and the origin of the HSI frame, \mathbf{b} , describe an HSI ray corresponding with u^h . Finally, s denotes the intersection parameter so that the HSI ray intersects

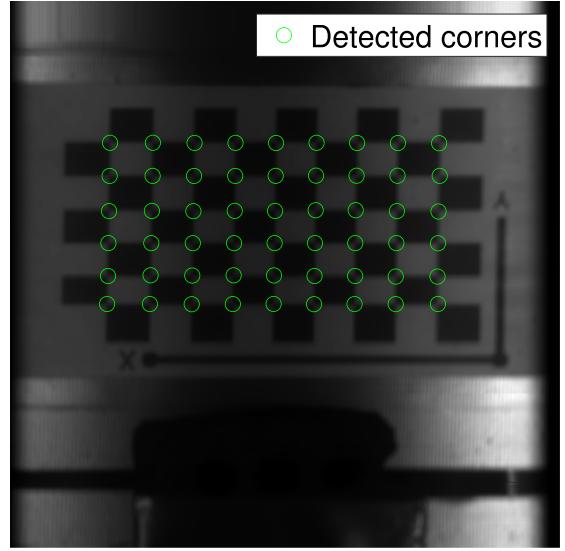


Fig. 3. Grayscale data cube. The x - and y -axes denote the pixel number u^h and the slit image number j_i , respectively. The green circles indicate corner detection in the data cube $[u_i^h, j_i]$.

with the first surface in the viewing direction at some earth-fixed position $[X, Y, Z]$. The computation of s in (4g) is trivial in the laboratory since the surface is defined by $Z = 0$, which allows for solving for s directly. However, *in situ*, the 3-D model consists of a vast number of surfaces. This ray casting problem is handled in Section II-F.

2) *Method 1*: The calibration points have known earth-fixed coordinates $[X_i, Y_i, 0]$. Each calibration point has a corresponding ground truth pixel positions u_i^h and slit image number j_i . These are found in Fig. 3 using MATLAB's checkerboard detection tool to find the corners. Fig. 3 shows a stacked sequence of slit images and a data cube, converted to grayscale. Each slit image number j_i has a time stamp corresponding with RGB extrinsics, \mathbf{t}_i and Θ_i , found through interpolation.

To evaluate the quality of a set of HSI intrinsics and HSI-RGB extrinsics, a distance invariant metric is used. According to [6, Ch. 7], an alternative formulation of the quality of a camera point projection is to minimize the normalized residuals in the measurement frame $[x^h, y^h]$. In this trial, the residuals were multiplied by the focal length to get the equivalent pixel residuals. $[x^h, y^h]$ are estimated in two ways. The first estimate $[\hat{x}_i^h, \hat{y}_i^h]$ uses $[X_i, Y_i, 0]$ with \mathbf{t}_i and Θ_i in

$$\hat{\mathbf{a}}_i \hat{\mathbf{s}}_i = [X_i, Y_i, 0] - \hat{\mathbf{b}}_i \quad (5a)$$

$$w_i[\hat{x}_i^h, \hat{y}_i^h, 1] = \hat{\mathbf{a}}_i \hat{\mathbf{s}}_i \mathbf{R}_1(\Theta_i) \mathbf{R}_2(\delta\Theta). \quad (5b)$$

The second estimate $[\tilde{x}_i^h, \tilde{y}_i^h]$ uses u_i^h in

$$[\tilde{x}_i^h, \tilde{y}_i^h] = [(u_i^h - c_x^h - \Delta u_i^h) / f^h, 0] \quad (5c)$$

with $\tilde{y}_i^h = 0$. The objective function J_1 in the following equation is the sum of squared residuals between the two estimates:

$$J_1 = \sum_{i=1}^n ((f^h \tilde{x}_i^h - f^h \hat{x}_i^h)^2 + (f^h \hat{y}_i^h)^2) \quad (5d)$$

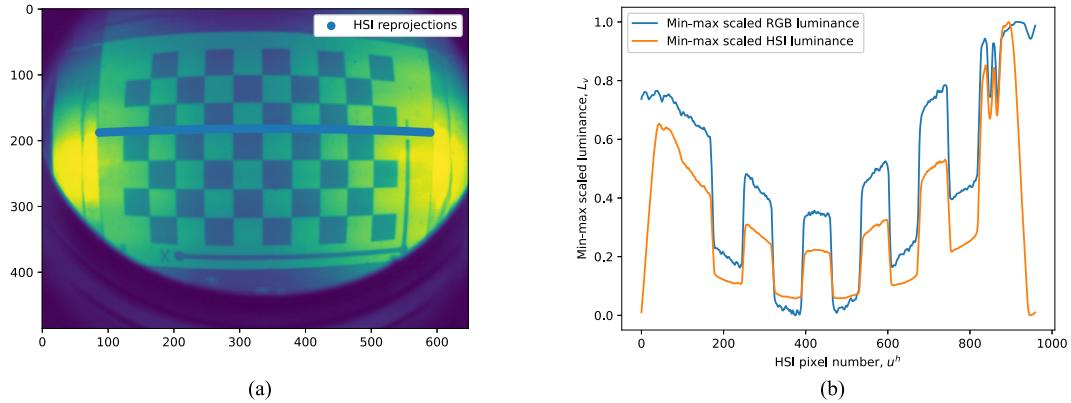


Fig. 4. Luminance-based calibration. (a) RGB image converted to luminance, $L_{v,\text{rgb}}(u, v)$, with HSI pixel reprojections, $[u_{\text{rgb}}, v_{\text{rgb}}]$, shown as blue dots. The RGB luminance of reprojected positions, $L_{v,\text{rgb}}(u_{\text{rgb}}, v_{\text{rgb}})$, are extracted and plotted in (b) with the measured HSI luminance. (b) Min–max scaled HSI luminance, $L_{v,\text{hsi}}(u_{\text{hsd}}^h)$, and min–max scaled RGB luminance, $L_{v,\text{rgb}}(u_{\text{rgb}}, v_{\text{rgb}})$. Note that correlation is invariant to min–max scaling. The extremities of the HSI measurements suffer from a sort of vignetting.

where n is the number of calibration points.

3) *Method 2*: This section suggests a new method for HSI calibration, which compares spectral measurements from the HSI with color measurements from the RGB camera. It is hypothesized that an RGB image acquired within a small time interval contains the same light pattern as the HSI given that the FoV has sufficient overlap. The calibration methodology is based on the assumption that HSI measurements and spatially corresponding RGB measurements correlate most for the best HSI-RGB extrinsics and HSI intrinsics. As a common metric of light intensity for the two sensors, luminance approximations, L_v , are used.

The following assumptions are made for the luminance-based calibration.

- 1) The luminance L_v in the underwater scene is a function of location, direction, and time.
- 2) Assume that time t_1 corresponds to an HSI slit image and t_2 corresponds to an RGB image. L_v is assumed unchanged with respect to time given that $|t_1 - t_2| < \epsilon$. That is, we select the HSI slit image that is closest in time to an RGB image to create an RGB/HSI pair, given that the time difference is small.
- 3) L_v is assumed invariant to the difference in position and view angle of the HSI and RGB given that $\delta\Theta$ and δt are small.

The next step is to define the conversion to luminance for the two sensors. For illustrations on the implementation, see Fig. 4. First, the luminance measured by the HSI, $L_{v,\text{hsi}}$ ($\text{Im} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$), can be computed as a weighted integral of spectral radiance $L(\lambda)$ ($\text{W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \text{nm}^{-1}$), as shown in

$$L_{v,\text{hsi}}(u_{\text{hsd}}^h) = K_m \int_0^\infty L(\lambda) \hat{y}(\lambda) d\lambda. \quad (6)$$

The conversion formula, values of the photopic luminosity function $\hat{y}(\lambda)$, and the maximum luminous efficacy K_m are adopted from [16, Ch. 2]. However, since the correlation is invariant to scale, using K_m has no practical purpose. Second, the RGB luminance, $L_{v,\text{rgb}}(-)$, is approximated as the gamma-compressed luma channel. The luma channel is computed as a weighted sum

of the 8-bit gamma-compressed $R'(u, v), G'(u, v), B'(u, v)$ channels, as suggested by Szeliski [6, Ch. 2], in the following expression:

$$L_{v,\text{rgb}}(u, v) = 0.2125R'(u, v) + 0.7154G'(u, v) + 0.0721B'(u, v). \quad (7)$$

Szeliski [6, Ch. 2] suggests using gamma decompression for linearization, but it is not applied to the RGB images. It was tried and improved correlation for well-lit images, but decreased correlation for low-light conditions.

So far, we have defined a geometric mapping from an HSI pixel, u_{hsd}^h , to an earth-fixed point, $[X_{\text{hsd}}, Y_{\text{hsd}}, Z_{\text{hsd}}]$, in (4). Furthermore, this point is reprojected onto the RGB image at pixel $[u_{\text{rgb}}, v_{\text{rgb}}]$, using (1)–(3). Spectrally, a mapping from HSI spectral radiance to luminance was established in (6) and a mapping from RGB images to luminance was established in (7). We can now compare the luminance converted slit image, $L_{v,\text{hsi}}(u_{\text{hsd}}^h)$, with corresponding RGB luminance, $L_{v,\text{rgb}}(u_{\text{rgb}}, v_{\text{rgb}})$. These two signals are shown in Fig. 4(b) with the corresponding luminance converted RGB image, $L_{v,\text{rgb}}(u, v)$, in Fig. 4(a). Note that the example uses calibrated parameters for the mapping from u_{hsd}^h to $[u_{\text{rgb}}, v_{\text{rgb}}]$.

Assume that u_{hsd} denotes the vector of HSI pixels in a slit image and that $[u_{\text{rgb}}, v_{\text{rgb}}]$ denotes the corresponding reprojected RGB image pixel positions. We denote the vector of the HSI luminance as $\mathbf{L}_{v,\text{hsi}} = L_{v,\text{hsi}}(u_{\text{hsd}}^h)$ and the RGB luminance vector as $\mathbf{L}_{v,\text{rgb}} = L_{v,\text{rgb}}(u_{\text{rgb}}, v_{\text{rgb}})$. These equate to the blue and orange plots in Fig. 4(b). The objective function in the following equation calculates the average luminance correlation for a set of n RGB/HSI pairs:

$$J_2 = \frac{1}{n} \sum_{i=1}^n \rho_{\mathbf{L}_{v,\text{hsi}}^i, \mathbf{L}_{v,\text{rgb}}^i} \quad (8)$$

where the subscript i refers to the i th RGB/HSI pair. By $\rho_{\mathbf{L}_{v,\text{hsi}}^i, \mathbf{L}_{v,\text{rgb}}^i}$, we refer to Pearson's correlation coefficient between the two luminance estimates. The correlation can also be computed pixelwise rather than samplewise, but this was not investigated. Notably, this method is not dependent on pure

translational motion such as Method 1 and can be used for any sort of motion, in the laboratory or *in situ*.

E. Photogrammetry Model

This section will briefly explain the procedure behind building the underwater photogrammetry model supplying the RGB extrinsics and the 3-D mesh model for the georegistration. Photogrammetry uses features in observed terrain that are identified and matched between images, and the differing positions of features across the images imply pose relations between the images. The pose relations between images in conjunction with position and orientation measurements enable us to compute the RGB extrinsics. Given the multiple views of the same features on the seabed, we can compute the 3-D shape of the seabed.

For self-consistent underwater photogrammetry, the mapping should facilitate nonsequential image matching, or loop closure, as seen in [11]. To enable this, the ROV mapping method follows a lawnmower pattern using automatic trajectory tracking and altitude control [17]. Additionally, the altitude and line spacing in the lawnmower pattern are selected to ensure that the FoV of the RGB camera (and HSI) overlaps sufficiently between adjacent lawnmower lines. This ensures multiple views of objects and increased consistency in the model. The mapping methodology results in several straight parallel survey lines called transects.

The mapping operation yields timestamped pose measurements, RGB images, and HSI slit images (or data cubes). The photogrammetry model uses the measured pose, the RGB images, and the RGB intrinsics. The procedure for building the underwater photogrammetry model is adapted from [18] as follows.

- 1) *Establish camera pose estimates:* The software uses pose estimates to find image matching candidates, and pose estimates are key to nonsequential image matching. Pose estimates are also used to compute the scale of objects when using a single RGB camera for photogrammetry. A robotic platform usually has the following sensors for underwater navigation.
 - a) *Acoustic positioning:* Typically, an ultrashort baseline system (USBL) is used for positioning relative to a topside unit with GPS. This yields an earth-fixed position for east, $X = E$, and north, $Y = N$.
 - b) *Pressure sensor:* The pressure sensor is used to calculate the global Z -position, or the depth below the sea surface. The tidal contribution is then added to achieve the depth below mean surface level. Using the east–north–up (ENU) convention, this measurement is denoted by $Z = U$.
 - c) *Inertial measurement unit (IMU):* An IMU yields orientations $\Theta_{\text{imu}} = [\phi, \theta, \psi]$, or roll, pitch, and yaw, respectively. The IMU also measures translational accelerations. Notably, the rotations have opposite definitions as Θ so that $R_1(\Theta) = R_1(\Theta_{\text{imu}})^\top$.
 - d) *Doppler velocity log (DVL):* The DVL measures the altitude with four beams and the velocity over ground. These raw measurements are further processed in an observer, as described in [19], to generate improved pose

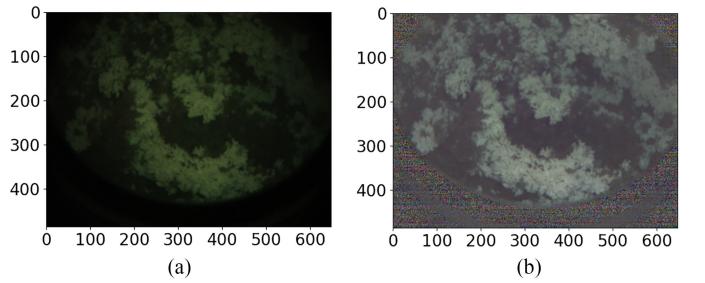


Fig. 5. Gray world correction. The rim of the underwater housing is shown in the images and was masked out during photogrammetry. (a) Before correction. (b) After correction.

estimates for the control system of the ROV. We consider the RGB camera to be the center of origin and need to translate and rotate pose measurements to the RGB camera's reference frame. The translation vector is expressed as $\mathbf{t} = -[E, N, U]\mathbf{R}_1(\Theta)$.

- 2) *Preprocessing of RGB images:* The RGB images are affected by several effects ranging from IOPs/AOPs to variations in the sensor response [20]. A simple underwater correction scheme is the gray-world correction [21]. The correction is formulated in the following equation and describes the relation between a raw image I_x and the corrected image I_y :

$$\frac{I_y(u, v, \lambda_c) - \mu_y(\lambda_c)}{\sigma_y(\lambda_c)} = \frac{I_x(u, v, \lambda_c) - \mu_x(u, v, \lambda_c)}{\sigma_x(u, v, \lambda_c)} \quad (9)$$

where $\mu_x(u, v, \lambda_c)$ and $\sigma_x(u, v, \lambda_c)$ are the empirical average intensity and standard deviation in intensity at pixel position $[u, v]$ for RGB color channel λ_c , respectively. Conversely, $\mu_y(\lambda_c)$ and $\sigma_y(\lambda_c)$ are the desired average intensity and standard deviation for the image in color channel λ_c , respectively. The operation is shown in Fig. 5.

- 3) *Image alignment:* The mapping procedure runs in a lawnmower pattern and has a higher density of images in the along-track direction. Although high-density along-track is advised for high-frequency pose estimates for the HSI, it complicates the matching procedure. Thus, it is suggested to begin with an initial alignment with a subset of images, still with a reasonable overlap. Afterward, the remaining images can be aligned. This yields a sparse point cloud of tie points (features matched between multiple images) and RGB extrinsics for each image.
- 4) *Adjust RGB camera parameters (optional):* After alignment using RGB intrinsics from the laboratory, it is possible to adjust the intrinsics with respect to the reprojection error, if there is reason to believe that they differ from the laboratory parameters. This optimized alignment yields the refined RGB extrinsics for each RGB image. It also alters the sparse point cloud.
- 5) *Photogrammetry pipeline:* The multiple views of features are used to build depth maps for each RGB image and a dense point cloud. This point cloud is further used

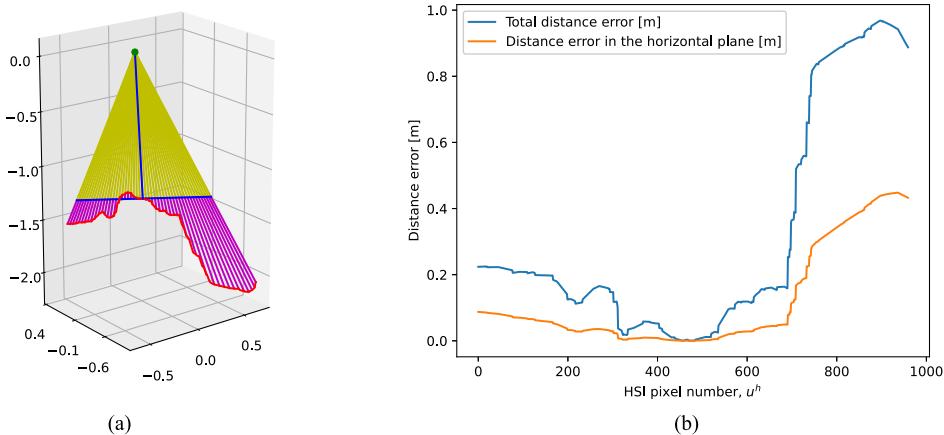


Fig. 6. Georegistration of a slit image. The red line in (a) shows an actual georegistration of a slit image (HSI-relative) in a transect using ray casting. All axes are in meters and are in the ENU coordinate system. The horizontal blue line shows the georegistration assuming the seabed to be flat, such as assumed by prior methods in UHI. The vertical blue line is an altitude measurement along the middle of the slit. (b) Length of the purple lines in (a), or the error in georegistration due to the flat seabed assumption.

to create a 3-D triangular mesh model, digital elevation model (DEM), and photomosaic.

F. Ray Casting

Fig. 6 quantitatively shows one of the key methodical differences from our georegistration method and prior. Our method actually computes the intersection between the HSI rays and the 3-D model from the photogrammetry, while prior methods use a single altitude and assume the seabed to be flat along the slit image. The maximal total and horizontal distance error in Fig. 6(b) is 0.97 and 0.45 m, respectively. Meanwhile, the altitude (vertical blue line) in Fig. 6(a) is 1.30 m.

In the context of this article, we consider ray casting to be the process of intersecting the HSI rays described in (4) with the 3-D triangular mesh model from the photogrammetry model. For a visual illustration, the reader is advised to Figs. 2 and 6(a). For clarity, assume that the rays are expressed according to

$$\mathbf{r}_{i,j}(s) = \mathbf{a}_{i,j}s + \mathbf{b}_j \quad (10)$$

where $\mathbf{r}_{i,j}$ denotes a ray with pixel number i of slit image j . The ray direction vector $\mathbf{a}_{i,j}$ is dependent on both pixel number and HSI pose, while the origin of the HSI, \mathbf{b}_j , is only dependent on the HSI pose. Note that the HSI pose is computed from interpolating the RGB pose. Each mesh element has a face with three vertices at known locations. The intersection criterion is that $\mathbf{r}_{i,j}$ intersects with the face. The intersection tests are computed quickly as in [22], but computation time becomes infeasible when there are millions of rays to be tested with millions of triangles. Therefore, it is efficient to use a spatial data structure that enables quick searches for triangle candidates for ray intersection. It is common to use hierarchical tree structures for these operations. One memory-efficient, fast, tree structure is the Oriented Bounding Box Tree (OBTree) [23]. The structure recursively builds smaller and smaller bounding boxes that encompass the triangles. The intersection tests have complexity $O(n \log(k))$, where n is the number of rays and k is the number

of triangles. The intersections of the HSI rays in Fig. 2 were found using an OBTree of the triangle mesh.

The implementation of the OBTree and intersection tests in Python are based on the methodology in [23]. The implementation is based on the *vtk* library in Python and adapted from [24]. The module takes $\mathbf{a}_{i,j}$, \mathbf{b}_j , and a max ray length as input and outputs the intersection of each ray, $\mathbf{r}_{i,j}(s_{i,j}) = [X_{i,j}, Y_{i,j}, Z_{i,j}]$, solved for s . Per definition, the georegistration of radiance spectrum $L_{i,j}^0(\lambda)$ has been done.

The next step is to perform the *in situ* calibration using method 2 from Section II-D. The ray casting described using OBTree is the equivalent to the line/plane intersection in (4g) used for georegistration in the laboratory. By inserting the intersection, $[X_{i,j}, Y_{i,j}, Z_{i,j}]$, along with RGB extrinsics from photogrammetry into the RGB camera's reprojection equations in (1)–(3), we obtain the corresponding pixel position $[u_{rgb}, v_{rgb}]$. Aside from the use of OBTree for solving (10) and photogrammetry for computing RGB extrinsics, the *in situ* calibration method is the same as in Section II-D3.

When the calibration is finished, the ray casting procedure is executed for all pixels in all slit images in each transect. This yields a point cloud for each transect corresponding with each data cube. In other words, each spectral radiance measurement $L_{i,j}^0(\lambda)$ has a corresponding earth-fixed position $[X_{i,j}, Y_{i,j}, Z_{i,j}]$.

G. Point Cloud Processing

The ray casting concludes the final step in the georegistration of each measurement. However, it seems relevant to describe the processing of the point clouds and corresponding data cubes into a hyperspectral mosaic of spectral reflectance. The following three steps are suggested.

- 1) *IOP/AOP correction*: This step explains how each spectrum in the data cube is corrected for the effects of the IOPs/AOPs using the optical path of each measurement. This yields spectral reflectance for each measurement.

- 2) *Gridding*: Each data cube corresponds to irregularly spaced data in the point cloud, which needs to be converted to an orthographic mosaic. This step produces a mosaic of spectral reflectance for each transect data cube.
- 3) *Blending overlapping mosaics*: Overlapping mosaics need to be merged into a single mosaic. Notably, the overlapping mosaics are used to validate the georegistration precision.

In this work, it was chosen to do the steps in the order shown above. Steps 1) and 2) are steps applied to the individual transects. Considering the pushbroom motion, it is fair to believe that spectra that fall within a small grid cell are subjected to highly similar light conditions and optical paths. Because spectra in different transects within the same grid cell are exposed to varying light conditions and optical paths, 3) is a separate step.

1) *IOP/AOP Correction*: Previously in UHI, the effects of the ambient light field and the water column, the AOPs and IOPs, were corrected for through statistical interpretations of the raw data cube. In [9], Mogstad *et al.* divided the data cube by the average spectrum of the whole data cube and applied polynomial corrections along-track and across-track, while Dumke *et al.* [1] divided the data cube by the median spectra along-track. Some studies, like [25], used a reflectance standard on the seabed in the transect and divided the data cube by this reference spectrum. These approaches have in common that they did not use any spatial information from the georegistration in IOP/AOP correction for the data cube. However, in [10], although using a point spectrometer and not an HSI, they compensated for attenuation along the light's optical path, both from the sun and from two light sources on the vehicle using Beer–Lambert's law.

In our work, it was desired to implement a simple IOP/AOP correction scheme to facilitate classification of the spectra. We model the exponential decay in radiance with along the optical path using Beer–Lambert's law to compensate for both the change in spectral shape and the wavelength-independent changes in intensity. In order to convert the radiance metric into reflectance, we propose to use reference objects on the seafloor with known reflectance spectra.

We had two identical halogen lamps around 35 cm to the starboard/port side of the HSI. We assumed these offsets to be negligible and computed the optical path as $2d_{i,j}$, where $d_{i,j} = |\mathbf{b}_j - \mathbf{r}_{i,j}(s_{i,j})|$, essentially modeling it as one light source that coincided with the imager. Then, we corrected the raw radiance spectra using Beer–Lambert's law for nonscattering media adapted from [16, Ch. 5]

$$L_{i,j}(\lambda) = L_{i,j}^0(\lambda) e^{K(\lambda) \cdot 2d_{i,j}} \quad (11)$$

where $L_{i,j}$ and $L_{i,j}^0$ represent the corrected and measured spectral radiance, respectively. The diffuse attenuation coefficient, $K(\lambda)$ (1/m), describes the exponential decay of the radiance. We assume that the spectral directional reflectance $R_{i,j}(\lambda)$ (–) can be approximated as

$$R_{i,j}(\lambda) = C(\lambda) \cdot L_{i,j}(\lambda) \quad (12)$$

where $C(\lambda)$ is a constant spectrum inversely proportional to the light source spectrum. The IOP/AOP spectra $C(\lambda)$ and

$K(\lambda)$ are estimated using *in situ* spectra of the scleractinian coral (*Desmophyllum pertusum*), as described in Appendix A. This approach was chosen since we had several ground-truth reflectance spectra of this species from the laboratory, and it was highly abundant in the transects.

2) *Gridding*: The next step is to use each 3-D transect point cloud to convert each reflectance data cube to a regular spatial grid. For this case, an orthographic projection of the point clouds is used to create hyperspectral mosaics of the reflectance data cubes. This problem is similar to the problem of compositing in computer vision [6, Ch. 9]. A general definition in this context is that the grid cell spectrum is a weighted sum of the measurements within the cell. Our two contradicting objectives of this gridding are to minimize the degree of spectral mixing and to maximize the signal-to-noise ratio (SNR). The typical approach in remote sensing is to select the measurement closest to the grid cell center, the so-called nearest neighbor. This approach has no degree of spectral mixing between measurements. However, in our case, SNR was the priority, and it was, therefore, chosen to use the mean of all samples within the grid cell. The nearest neighbor method and the mean method are actually the two extremes of the inverse distance weighting interpolation with exponents ∞ and 0, respectively. The gridding of measurements is not a topic previously discussed in UHI, but could be an interesting topic in the future to develop weighting functions that are more optimal.

A reflectance measurement is put in a spatial cell at X_n and Y_m if $X_{i,j}$ and $Y_{i,j}$ are within that cell. For each transect, we find the average of all the measurements within each cell, $R_{n,m}(\lambda)$. The same averaging is done for the distances $d_{i,j} = |\mathbf{b}_j - \mathbf{r}_{i,j}(s_{i,j})|$ into $d_{n,m}$. This yields a raster of distances and a mosaic of reflectance for each transect. This is based on the pushbroom assumption that all measurements in a cell, X_n and Y_m , have the same optical path, and the distance rasters are used for blending overlapping mosaics. The gridding process is repeated for each transect. This yields a set of overlapping hyperspectral mosaics and corresponding distance rasters for each data cube.

3) *Blending Overlapping Mosaics*: The final step is to create one mosaic of all the transect mosaics. This is highly similar to gridding, and there exist multiple solutions to selecting and weighting overlapping pixels. However, since this is the first time in UHI where multiple overlapping IOP/AOP-corrected mosaics are blended, we suggest a new method for handling this problem. We believe that weighting or selection of pixels should be based on the optical distance $d_{n,m}$ since reflectance measurements with high optical distances have a poorer quality. In our implementation, overlapping transects are handled by simply selecting the reflectance spectrum $R_{n,m}$ with the shortest $d_{n,m}$. This avoids mixing of reflectance spectra across transects. A weighting approach is likely more optimal, but this requires further research into the topic.

Each distance raster and hyperspectral reflectance mosaic have the same corners and same cells containing data. Therefore, we can rapidly check if two rasters overlap using their corners. If they overlap, we check if there are overlapping data points. For the overlapping cells, we simply remove the spectrum in a

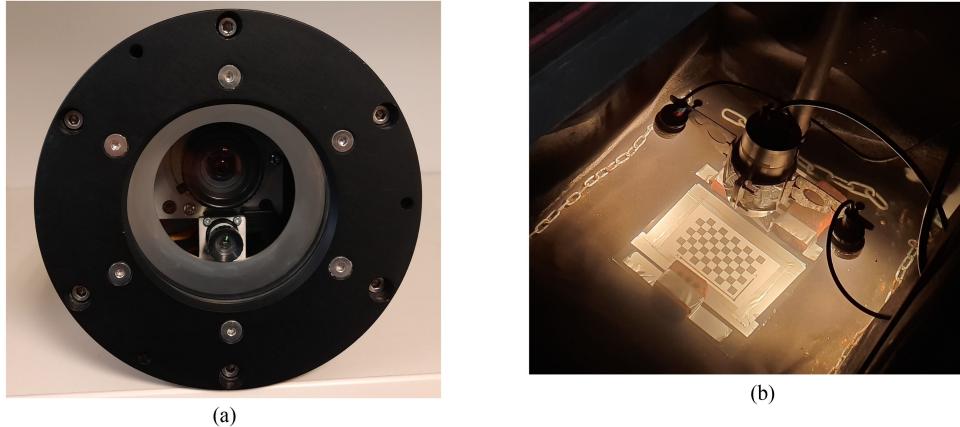


Fig. 7. UHI-4 by Ecotone. (a) Image of the UHI-4 showing the RGB camera (small lens) and HSI (big lens). They correspond to the green and red camera in Fig. 2. The vertical distance between the lenses is around 30 mm. (b) Setup in the tank during fixed translation. The UHI-4 is in the center of the bracket. To get varying orientations, the sensor was dismounted from the bracket and moved around the checkerboard.

TABLE I
RGB AND HSI SETTINGS

	n_λ [-]	$n_{u,v}$ [-]	t_{exp} [ms]	[FPS]	vFoV [deg]	hFoV [deg]
RGB lab		648 × 486	3/30	5	61/0.13	77/0.12
RGB i.s.		-	10	-	-	-
HSI lab	211	960 × 1	20	50	0.4	70/0.07
HSI i.s.	224	960 × 1 (480 × 1)	-	-	-	70/0.07(0.15)

cell if there exists a lower value of $d_{n,m}$ in the same cell of the other raster.

III. RESULTS

Section III-A presents the laboratory setup from September 2020 and how the RGB and HSI imagery was acquired. Section III-B presents the results from the laboratory RGB calibration, including the intrinsics, the extrinsics, and validating metrics for the calibration. The optimized intrinsics from the field are presented for comparison. Section III-C presents the HSI calibration results from the two calibration methods, including metrics for validation. The *in situ* calibration is also presented. Section III-D presents the Tautra coral reef mapping from March 2017 and the results from the *in situ* georegistration.

A. Laboratory Setup

The RGB camera and HSI were encapsulated into an imaging device called the underwater hyperspectral imager 4 (UHI-4) from the manufacturer Ecotone AS (Trondheim, Norway). The UHI-4 is a cylindrical device with the HSI and RGB camera configured, as shown in Fig. 7(a). It was mounted as shown in Fig. 7(b) during translational laboratory experiments with light sources on each side.

A summary of the parameters of the RGB and HSI in the laboratory and *in situ* (denoted i.s. in Table I) are shown in Table I. The UHI-4 had a spectral calibration of the HSI during the summer of 2017. Therefore, the number of wavelengths n_λ is different from laboratory to *in situ* in Table I. Geometrically,

the sensors are configured in the same brackets, and only minor changes in HSI and RGB parameters are expected.

n_λ and $n_{u,v}$ in Table I are the number of wavelengths and the spatial dimensions, respectively. t_{exp} is the exposure time in milliseconds, which was increased to 30 ms for the RGB during the manual acquisition due to low-light conditions. Frames per second (FPS) denotes images, or slit images, per second. vFoV describes the FoV along with average per-pixel FoV along the vertical axis of the image, v . hFoV is the same along the horizontal axis, u . These are described in air. The spatial dimensions, $n_{u,v}$, was changed to 480 × 1 pixel (pix) during parts of field trials, so that the hFOV increased to 0.15° per pixel. The dashed lines in Table I indicate no change between laboratory and *in situ*.

The laboratory trials were conducted in September 2020 at Trondheim Biologiske Stasjon, Norwegian University of Science and Technology (NTNU), in a testing tank with seawater. The tank is custom-made for hyperspectral imaging, as shown in Fig. 7(b). That is, there was an engine moving the UHI-4 forward at constant speed along a straight track. Between trials, the track could be elevated and moved transversely. A calibration pattern, a checkerboard, with square size 21 mm, was laid on the bottom and weighed down to avoid unwanted motion. For each track, the altitude from the UHI-4 to the checkerboard was measured.

At first, we did a 20-cm-altitude track straight over the checkerboard, as shown in the upper left image in Fig. 8. The track was then moved transversely to get calibration points in the outermost pixels, as shown in the upper right image in Fig. 8. Then, the altitude of the track relative to the checkerboard was changed to 42.5 cm and run again, as shown in the lower left

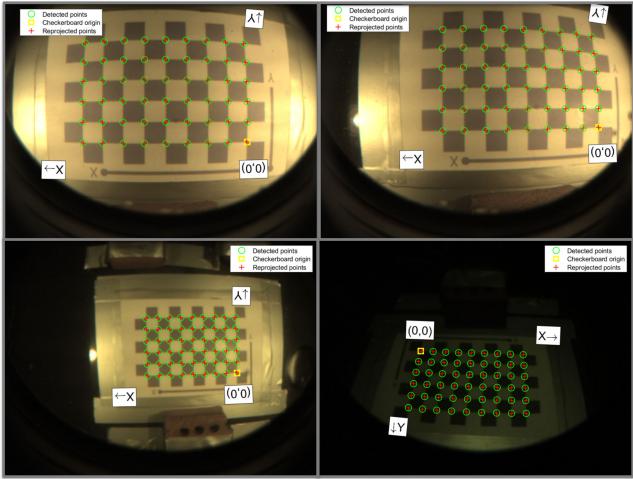


Fig. 8. Both upper images and the lower left image are from the three translational tracks. The lower right image is from the manual acquisition. Green circles indicate located checkerboard points in each image and red crosses show the reprojected points.

TABLE II
RGB INTRINSICS

	RGB Lab	RGB In-Situ	Difference
f_x	536.6 pix	534.8 pix	-1.8 pix
f_y	533.2 pix	531.1 pix	-2.1 pix
c_x	319.7 pix	321.2 pix	+1.5 pix
c_y	236.8 pix	239.2 pix	+2.4 pix
k_1	-0.4249	-0.4175	+0.0074
k_2	0.1455	0.1351	-0.0104
p_1	0.001	-0.001	0.002
p_2	-0.002	-0.002	0.000
RMSE	0.26 pix*	0.44 pix*	

*Laboratory trials consider reprojection errors for calibration points, while *in situ* consider errors for tie points.

image of Fig. 8. Using the straight track allowed for variation in translation, but not rotation. Therefore, the rig was removed, and the UHI-4 was moved around manually in the tank to vary the rotations, as shown in the lower right image of Fig. 8. In total, there were 431 images containing the full checkerboard, which were used for the RGB calibration.

B. RGB Calibration

From 431 images, a root-mean-square error (RMSE) of 0.26 pixels was achieved for the RGB camera. Notably, reprojection errors were higher for manual acquisition with 30-ms exposure time. The RGB intrinsics were estimated and are shown in Table II in comparison with the adjusted RGB intrinsics from the photogrammetry model. The corresponding laboratory RGB extrinsics are illustrated by Fig. 9.

The three tracks in Fig. 9(a) show the estimated 20-cm-altitude tracks and the estimated 42.5-cm-altitude track. The two

TABLE III
TWO INITIAL CALIBRATION PARAMETER SETS

Name	x_0	x_1	Factory calibration*
$\delta\phi$	0 deg	-0.53 deg	0 deg
$\delta\theta$	0 deg	2.44 deg	0 deg
$\delta\psi$	0 deg	-0.39 deg	0 deg
δx	0 mm	-3.1 mm	0.9 mm
δy	30 mm	20.0 mm	29.1 mm
δz	30 mm	34.6 mm	35 mm
f^h	909.8 pix	975.2 pix	859.2 pix
c_x^h	480 pix	484.84 pix	480.74 pix
k_1^h	0	12.74e-13	-3.03e-13
k_2^h	0	0.08e-07	5.11e-13
k_3^h	0	5.10e-05	0.00e-05
J_1 (MAE)	39.27 pix	8.43 pix	37.13 pix
J_2	0.4655	0.8927	0.5261

x_0 and x_1 are initial guesses for parameters for the HSI model. x_0 was chosen arbitrarily, while x_1 was found from trial and error. *The factory calibration is a parameterization of the tabular calibration from the manufacturer fitted with a 0.01 pix error. Notably, the factory calibration was given the same translation parameters as those calibrated. An illustration comparing the difference between the factory calibration and the final calibration is shown in Fig. 11.

20-cm-altitude tracks were estimated with means of 25.32 and 25.26 cm, and with standard deviations of 0.08 and 0.06 cm, respectively. The 42.5-cm track was estimated to 47.00 cm with a standard deviation of 0.13 mm.

C. HSI Calibrations

The HSI slit images were acquired during the same trials as described in Section III-B. HSI calibration method 1 used the three translation tracks to create three data cubes. Two of the grayscale data cubes are shown in Fig. 10. Note that Fig. 10 displays six detected rows of calibration points for each track, but the uppermost and lowermost rows of points were discarded since the corresponding RGB images did not capture the full checkerboard. Hence, from the three grayscale data cubes, 9 by 4 calibration points per track, a total of 108 points were used for calibration. Although the metric used in the optimization method was the summed squared error, a more intuitive related metric is the mean absolute error (MAE) in pixels. This is found by summing the square root rather than the squared residuals in (5d). Hence, J_1 refers to the MAE in Tables III and IV.

Method 2 used all the 431 RGB images corresponding with the RGB extrinsics in Fig. 9(a) and (b). For every RGB image, it selected the closest HSI slit image in time, given that the time difference was smaller than a threshold, $\epsilon < 30$ ms. However, the expected time difference was $|\delta|t = 10$ ms and varied between images. The quality of the calibration was defined by the average luminance correlation, J_2 . The HSI in the laboratory had black spots in the extremities causing the dips for $u^h < 30$ pix and

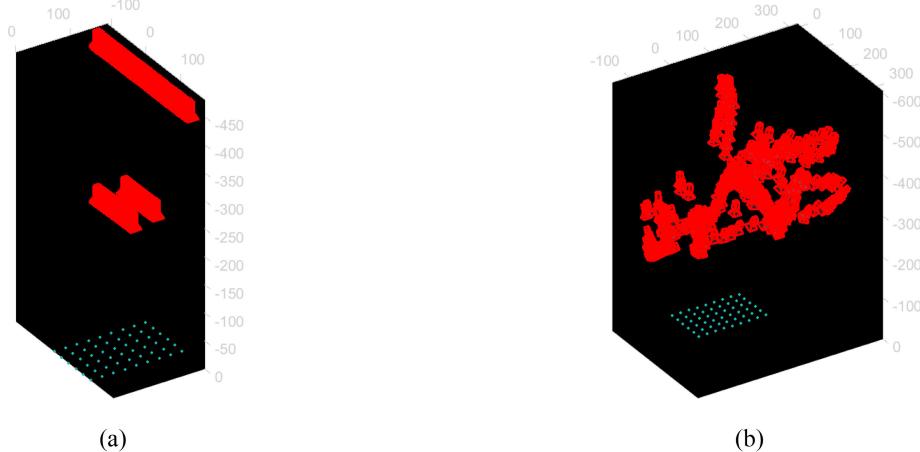


Fig. 9. RGB extrinsics in the laboratory. The blue points show the calibration points, which are the checkerboard corners. The red shapes indicate RGB camera pose. The earth-fixed coordinates are given in millimeters. (a) Three translational tracks (135 images in total). (b) Manual motion of the UHI-4 (296 images in total).

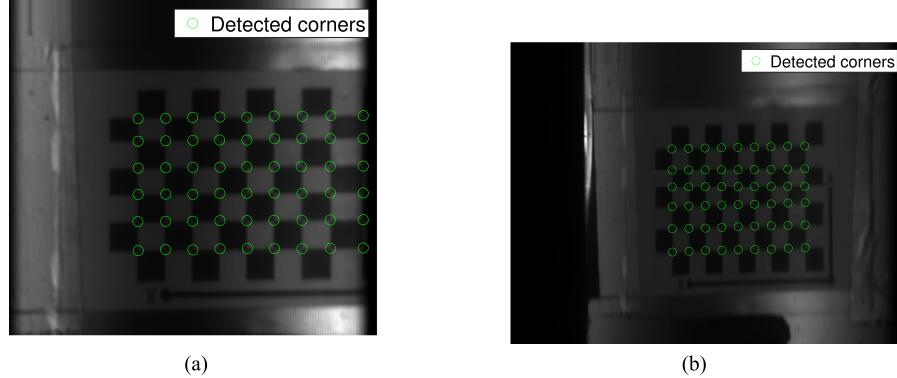


Fig. 10. Grayscale converted data cubes as defined in Fig. 3. The green circles indicate corner detections. (a) Second 20-cm-altitude track. Rightmost six calibration points had to be annotated manually. (b) 42.5-cm-altitude track.

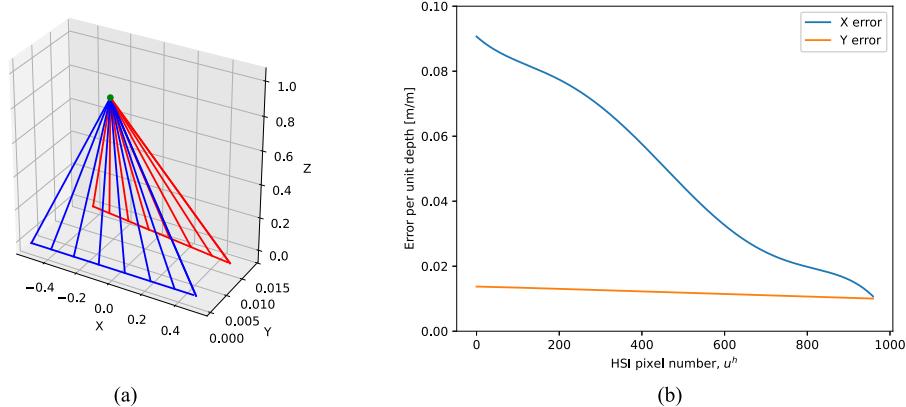


Fig. 11. Factory calibration. (a) Comparison of the calibrated model (first column in Table IV) in red and the factory calibration in blue. The 3-D plot is shown in the RGB camera's reference frame. (b) Errors per unit depth in the across-track X and along-track Y directions. The largest error is 9.1 cm/m, across-track. Almost all the errors are across-track, caused by the HSI intrinsics. Note that (a) makes it look like the largest error is along-track, but this is because of the aspect ratio of the axes.

TABLE IV
HSI LABORATORY CALIBRATIONS

Method	Method 1		Method 2			
Algorithm	<i>lsqcurvefit</i>		<i>least_squares</i>		Powell's	
Initial parameters	x_0	x_1	x_0	x_1	x_0	x_1
$\delta\phi$	-0.69 deg	-0.69 deg	-0.70 deg	-0.59 deg	-0.68 deg	-0.70 deg
$\delta\theta$	1.89 deg	3.12 deg	1.27 deg	2.72 deg	2.41 deg	3.00 deg
$\delta\psi$	-0.25 deg	-0.27 deg	-0.26 deg	-0.26 deg	-0.28 deg	-0.32 deg
δx	0.9 mm	0.9 mm	1.6 mm	2.3 mm	0.9 mm*	1.6 mm*
δy	29.1 mm	29.1 mm	29.3 mm	28.5 mm	29.1 mm*	29.3 mm*
δz	35.1 mm	34.6 mm	40.0 mm	41.4 mm	35.1 mm*	40.0 mm*
f^h	970.1 pix	970.0 pix	977.5 pix	984.9 pix	966.6 pix	989.2 pix
c_x^h	467.0 pix	487.9 pix	455.2 pix	478.6 pix	475.9 pix	484.7 pix
k_1^h	8.62e-13	10.57e-13	0.04e-13	4.38e-13	0.09e-13	12.603e-13
k_2^h	0.95e-07	0.76e-07	2.77e-07	1.91e-07	2.64e-07	0.15e-07
k_3^h	-0.3e-07	3.38e-05	-1.57e-05	4.14e-05	2.18e-05	3.81e-05
J_1 (MAE)	1.04 pix	1.04 pix	1.39 pix	1.50 pix	1.22 pix	1.32 pix
J_2	0.9198	0.9194	0.9214	0.9214	0.9209	0.9209

*Powell's algorithm did not adjust relative translations. Note that results from method 1 had best MAE and results from method 2 had best J_2 , as expected.

$u^h > 930$ pix. Therefore, only pixels between 30 and 930 were used for estimating the correlation.

The two initial calibration parameters for optimization were chosen to be x_0 and x_1 in Table III. The factory calibration is also shown. These initial parameters in Table III were calibrated with methods 1 and 2. Method 1 used MATLAB's *lsqcurvefit* [26], [27], which utilized the default trust region reflective (TRR) algorithm with default settings. The algorithm takes as input the array of all detected calibration points, an objective function computing the reprojection error, and an array of zeros, which is the ideal value of the objective function.

For the optimization in method 2, two different optimization algorithms were tried using Python's *SciPy* library. The first algorithm, implemented as *least_squares* in Python, used the default TRR algorithm [28] with the nondefault setting of an iteratively updated characteristic scale of each variable. The second Python algorithm was the derivative-free Powell's algorithm [29] with default settings. Notably, these algorithms minimize the scalar objective function, $1 - J_2$, as input, where J_2 is from (8). The results from the two calibration methods in the lab are shown in Table IV.

The quality of the georegistration in the laboratory is based on the MAE of the 108 calibration points. It is also converted to projection error in millimeters. The best calibration from method 1 yielded an MAE of 1.04 pix, where the y^h -MAE was 0.88 pix, while the x^h -MAE was 0.37 pix. The corresponding MAEs in projection on the checkerboard were 0.132 and 0.343 mm for X and Y , respectively. Moreover, the best calibration from method 2 with the highest J_2 had corresponding MAE of 1.39 pix and x^h -MAE of 0.78 pix and y^h -MAE of 0.95 pix. The corresponding MAEs in projection on the checkerboard were 0.269 and

TABLE V
IN SITU CALIBRATION PARAMETERS

	HSI Lab	HSI in-situ	Difference
$\delta\theta$	-0.70 deg	-0.07 deg	+0.63 deg
$\delta\phi$	1.27 deg	0.80 deg	-0.47 deg
$\delta\psi$	-0.26 deg	-0.43 deg	-0.17 deg
δx	1.6 mm	-	-
δy	29.3 mm	-	-
δz	40.0 mm	-	-
f^h	977.5 pix	972.4 pix	-5.1 pix
c_x^h	455.2 pix	455.4 pix	+0.2 pix
k_1^h	0.04e-13	2.24e-13	+2.20e-13
k_2^h	2.77e-07	2.74e-07	-0.03e-07
k_3^h	-1.57e-05	-3.47e-05	-1.90e-05
$J_{2,cal}$	0.8415	0.9538	+0.1123
$J_{2,val}$	0.8502	0.9533	+0.1031

0.362 mm for X and Y , respectively. Notably, there was a higher uncertainty in the corner detection in the y^h -direction since the FoV was wider and translation speed increased motion blur.

In Table V, the initial calibration was the calibration from Table IV with the highest J_2 (fourth column from left). Powell's algorithm was used for *in situ* calibration of parameters due to TRR not converging well for *in situ* samples. Note

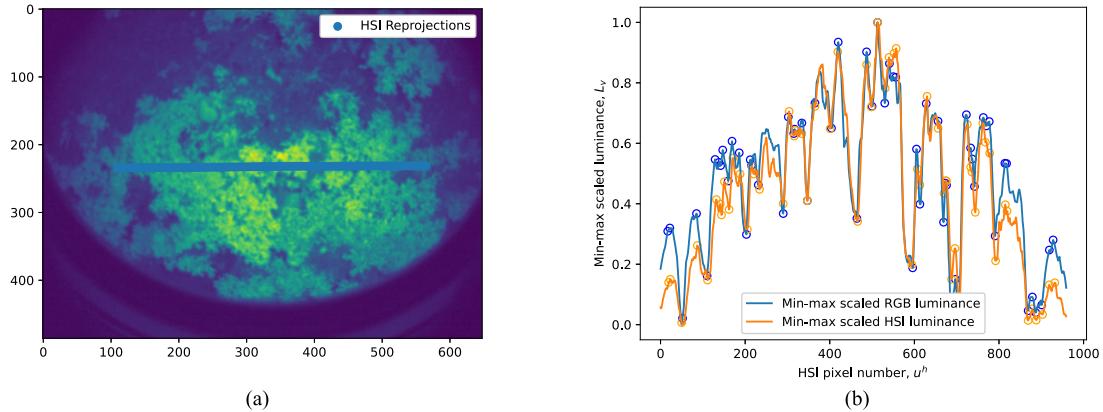


Fig. 12. *In situ* calibration using the calibrated parameters from Table V. To illustrate the quality of the calibration geometrically, we have selected some peaks and troughs and considered the across-track error in placement of these features. Circles in (b) show the selected features. The MAE of the 60 features is 1.47 pix equivalent to ≈ 0.3 cm for a 2-m altitude. This indicates that the *in situ* calibration performed well. (a) Luminance converted RGB image of corals corresponding with (b) with HSI pixel reprojections $[u_{\text{rgb}}, v_{\text{rgb}}]$ shown as blue dots. (b) Min–max scaled HSI luminance $L_{v,\text{hsr}}(u_{\text{hsr}}^h)$ and min–max scaled RGB luminance $L_{v,\text{rgb}}(u_{\text{rgb}}, v_{\text{rgb}})$.

that the translations, δx , δy , and δz , were not optimized in Table V, since the potential changes in these parameters were neglectable. For the *in situ* calibration in Table V, a calibration set of $512 + 373 = 885$ RGB/HSI pairs were selected from two transects in opposite directions to limit overfitting. The results from these are indicated by $J_{2,\text{cal}}$. For validation, $372 + 371 = 743$ RGB/HSI pairs from two other transects were used. The validation results are denoted as $J_{2,\text{val}}$. Contrary to the laboratory trials, the final calibration used all 960 pixels. Note that for efficiency, only a subset of RGB/HSI pairs and pixels were used for initial calibrations. *In situ*, a mesh model with 400 000 triangles was used to represent the seabed model.

A visual representation of the *in situ* calibration results is shown in Fig. 12. The luminance converted image in Fig. 12(a) shows the reprojections of the HSI pixels as blue dots. Luminance from the blue dot locations is shown in Fig. 12(b) as the blue line. The orange line shows the luminance from the HSI pixels.

D. Georegistration

A hyperspectral survey of corals was conducted on March 7, 2017 on the Tautra cold-water coral reef in the Trondheimsfjord, Norway. The survey was conducted by the NTNU Applied Underwater Robotics Laboratory (AUR-Lab) using the ROV SUB-fighter 30k operated from NTNU's research vessel *R/V Gunnerus*. The background was that AUR-lab had previously conducted a survey of the site with synthetic aperture sonar. The mapping lasted for approximately 1.5 h and covered an area of $\approx 15 \times 60 = 900$ m². The ROV automatically followed 13 transects in a lawnmower pattern in the North–East direction, as shown in Fig. 14. The transect spacing was 1 m and altitude 2 m to ensure overlap in the mapping between transects. With an FoV of $\approx 52^\circ$ for the HSI, 2-m altitude yields a swathe width of about 1.8 m on flat seabed. Hence, transects overlapped significantly.

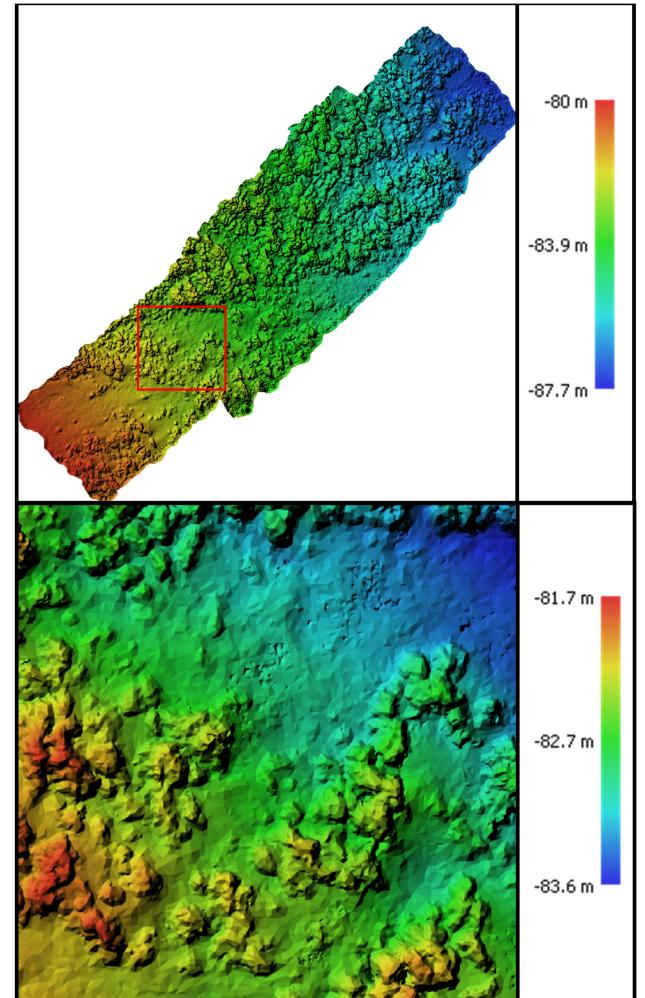


Fig. 13. DEM of the 3-D seabed model used for ray casting. The DEM corresponds with the RGB and hyperspectral mosaics shown in Fig. 14.

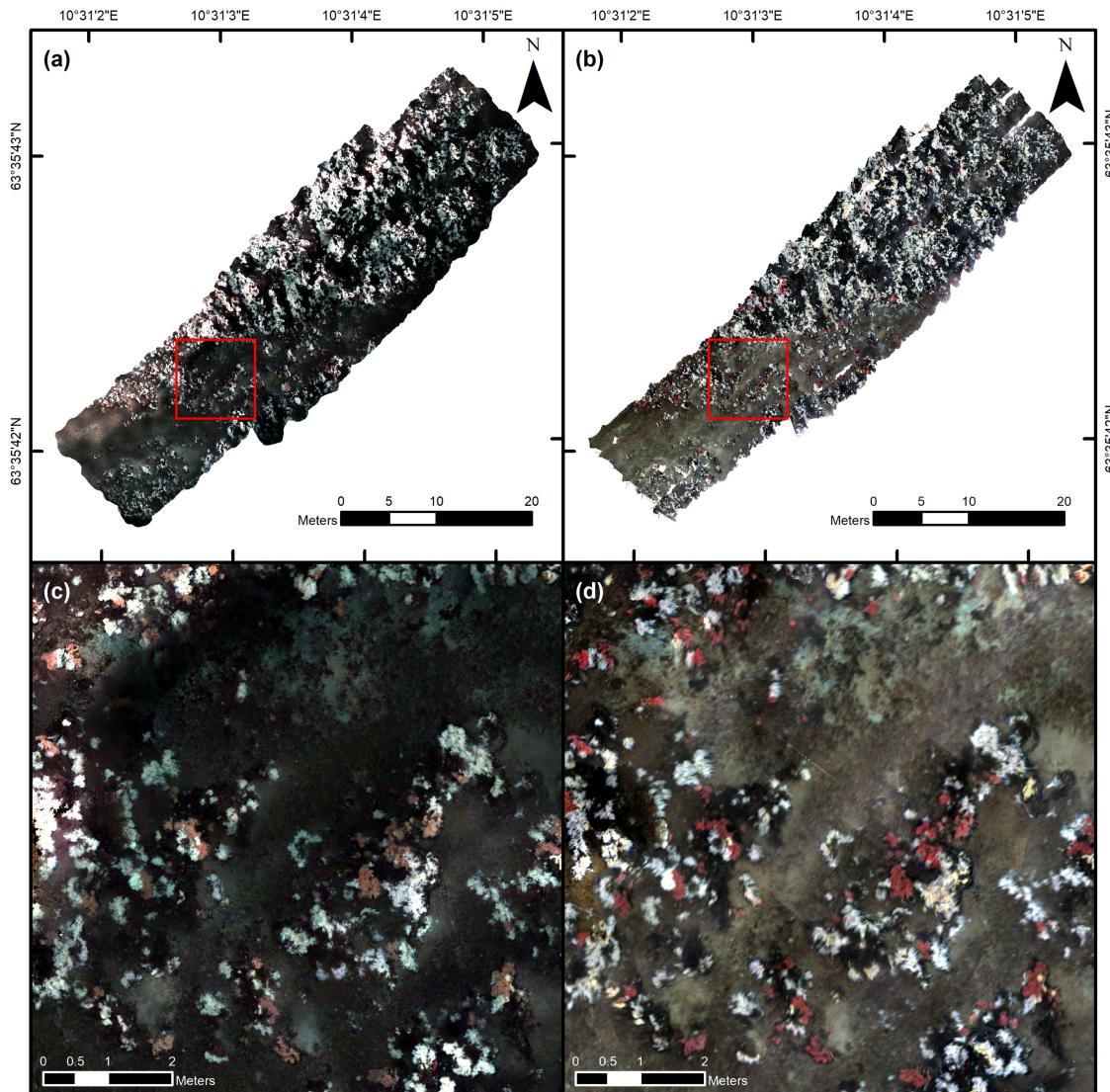


Fig. 14. Final georegistration of the cold-water coral reef. The figure shows the RGB photomosaic from the photogrammetry model in (a) and (c) with the IOP/AOP corrected hyperspectral mosaic in (b) and (d). The RGB visualization of the hyperspectral mosaic uses 590-, 530-, 460-nm wavebands for red, green, and blue channels. The spatial resolutions of the hyperspectral mosaic and photomosaic are 1 cm and \approx 4 mm, respectively. The mosaics in (c) and (d) are made up from 12 transects in the North–East direction. In particular, note the similarities between (c) and (d), the seamless nature of (d) and the consistency in colors of (d).

During the mapping, approximately 22 000 images (at 5 FPS) and more than 200 000 hyperspectral slit images (at 50 FPS) were acquired, resulting in 55 data cubes along the 13 transects. The ROV was positioned in longitude/latitude with a USBL system, Kongsberg HiPaP 500, with high precision. For depth, an independent pressure sensor was used in conjunction with tide cycle data. The ROV was also instrumented with a four-beam DVL that it used for robust altitude control of the vehicle on the rugged seabed. Moreover, the ROV was instrumented with an IMU for estimating orientations. All navigation sensor configurations on the ROV had been measured in relation to the RGB camera frame, so that measurements could be converted to the RGB frame.

In total, 21 702 RGB images were included in the photogrammetry model. The resulting point cloud had an RMSE of \approx 0.44 pix. Moreover, the model resulted in a mesh model with 6.6

million triangles and RGB extrinsics for each image. Since each slit image was recorded with 960 rays (a few segments with 480 rays) per slit image, around 200 million rays had to go through the ray casting procedure. To speed up ray casting, the mesh model was downsampled to 400 000 triangles. This corresponded to an average triangle size of 22.5 cm^2 versus 1.36 cm^2 for the original mesh. The DEM of the downsampled mesh model is shown in Fig. 13.

The HSI rays were ray casted onto the mesh model to create 55 point clouds corresponding to the data cubes. The data cubes were then IOP/AOP corrected and converted into 55 mosaics. The final step was to transform them into one hyperspectral mosaic. The final hyperspectral mosaic is shown and compared to the photomosaic from the photogrammetry model in Fig. 14. The wavelengths used for the RGB representation of the hyperspectral mosaic corresponded to the peak quantum

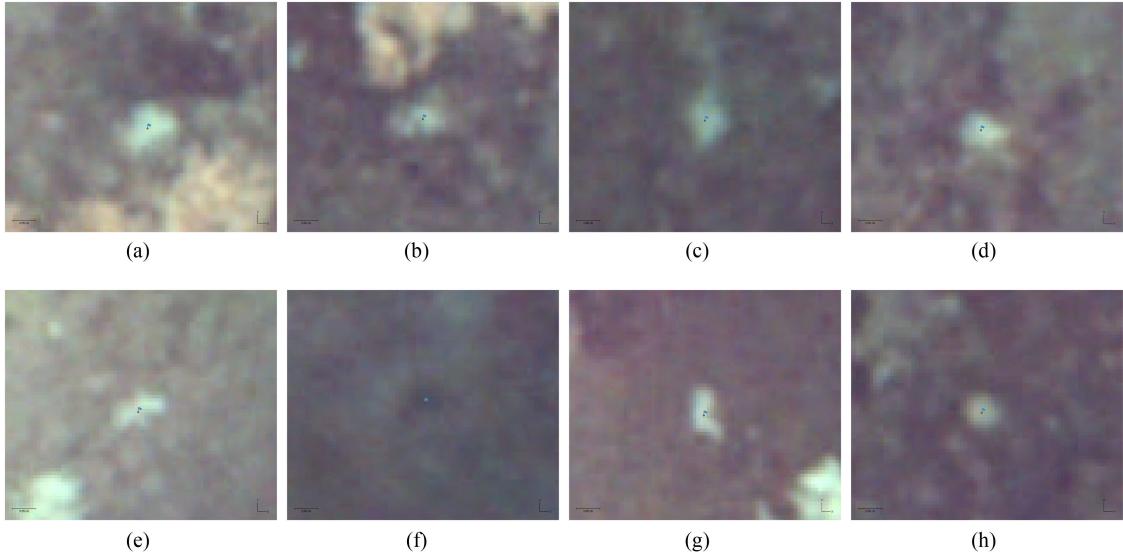


Fig. 15. Objects. The figures are cut-outs of the photomosaic. The control points are indicated in the center of each object. Each image has a width of 35 cm and a height of 30 cm. Object 6 was not used as shown in Table VII. (a) Object 1. (b) Object 2. (c) Object 3. (d) Object 4. (e) Object 5. (f) Object 6. (g) Object 7. (h) Object 8.

efficiencies [30] of the red (590-nm), green (530-nm), and blue (460-nm) channels. The pixel size of the hyperspectral mosaic was set to 1 cm. This was due to the along track vFoV of the HSI being $\approx 0.3^\circ$ (0.4° in air), which corresponded to ≈ 1 cm at 2-m range. The average across-track hFoV was $\approx 0.06^\circ$, or ≈ 2 mm at 2-m range.

In order to estimate the precision of the georegistration, we evaluated the georegistration of eight objects within the transects, shown in Fig. 15. The objects were bright objects of around 5-cm-diameter size, visible within two or three overlapping hyperspectral mosaics. Therefore, we had to use the mosaics prior to the blending step in Section II-G3. We chose small objects to minimize the error when manually putting a marker in the center of the objects. However, the objects had to be large enough that they were visible and identifiable in all overlapping transects. Moreover, adjacent transects going in opposite directions were used since errors in the geometric calibration will be most apparent for these cases. The georegistration precision, or error, for an object was computed as the Euclidean distance in the horizontal plane between the object markers in two overlapping hyperspectral mosaics and is shown in Table VI. Two of the objects (5 and 8) were visible in three overlapping transects, meaning that there were three errors to be computed. The average of the errors in Table VI was 0.88 cm. The same objects were marked in the photomosaic to evaluate the consistency between the photogrammetry model and the georegistration. The average distance, or the MAE, between the photomosaic marker and the marked points in the hyperspectral mosaics, was computed for each seabed object and is shown in Table VII. The average distance for all objects was 0.48 cm. This consistency metric measured the georegistration accuracy with respect to photogrammetry.

TABLE VI
ABSOLUTE ERROR BETWEEN OBJECT LOCATIONS IN OVERLAPPING PAIRS OF HYPERSPECTRAL MOSAICS.

Object	Error [cm]		
1	0.976406	-	-
2	0.865851	-	-
3	1.405102	-	-
4	0.882734	-	-
5	0.323098	1.236574	0.915757
6	1.727157	-	-
7	0.054562	-	-
8	0.587181	1.042662	0.590786

TABLE VII
MAE BETWEEN OBJECT LOCATIONS IN HYPERSPECTRAL MOSAICS AND OBJECT LOCATION IN PHOTOMOSAIC

Object	Error [cm]
1	0.55709724
2	0.40832466
3	0.59960039
4	0.63603072
5	0.32847161
6	-
7	0.45074828
8	0.40474771

IV. DISCUSSION

The quality of the RGB camera calibration is important both for the HSI calibrations and the quality of the photogrammetry model. The calibrations in the laboratory show that a calibrated RGB camera is a powerful tool for close-range pose estimation. For example, the estimation of altitudes during the tracks was highly consistent, both within each track and with respect to the manual measurements of altitude. The RMSE of 0.26 pix also indicates good overall performance. Regarding the photogrammetry model, the *in situ* adjustment of the RGB intrinsics in Table II shows an agreement between the photogrammetry model and laboratory calibrations.

A direct comparison of HSI calibration methods 1 and 2 in Table IV indicates a high agreement for most parameters, highlighting the potential for using method 2 for calibration. Table IV shows that both methods achieve low MAEs, with method 1 having the lowest error of 1.04 pix, with 0.39 and 0.88 pix in x^h and y^h , respectively. Comparing the calibrations from method 2 (especially the TRR algorithm) and method 1, there is a high degree of consistency for all six extrinsic parameters besides the roll angle $\delta\phi$. However, note that $\delta\phi$ is coupled with the principal point c_x^h . In reality, $\delta\phi$ and c_x^h in Table IV are highly consistent. This degree of consistency in parameter estimation suggests that method 2 is suitable for accurate calibration of all parameters in Table IV. However, the TRR algorithm did not converge well for the *in situ* HSI calibration, believed due to noisy gradients in the objective function. Therefore, the derivative-free Powell's algorithm was applied *in situ*. The algorithm was efficient at convergence in the laboratory, and the estimation of parameters was quite consistent. For the *in situ* calibration, Powell's algorithm mainly adjusted the relative orientation and increased the average luminance correlation significantly, as shown in Table V. As seen in the visualization of method 2 in Fig. 12(b), features in luminance coincide nicely after calibration, indicating the potential of using this method for *in situ* calibration. It also indicates that we can do georegistration in UHI that is highly consistent with the georegistration of RGB data in the photogrammetry model.

It is further argued that the underwater photogrammetry model described is accurate enough for the purposes of this article. First, because the ROV possesses a sophisticated sensor suite, believed to give highly reliable pose measurements. Second, the mapping has substantial overlap, yielding images of features from multiple transect lines. Third, the initial alignment of images was done using the RGB camera calibration from the laboratory. Combined with an RMSE of 0.44 pix for the tie points from $\approx 22\,000$ images, the model is assumed to be accurate enough. The DEM in Fig. 13 illustrates the 3-D model nicely. The DEM is made directly from the mesh used in the ray casting and appears smooth and seamless. Notably, we could have used the mesh with a higher resolution.

The square hyperspectral mosaic in Fig. 14 contains 12 mosaics that are not smoothed or averaged, but rather blended based on a shortest optical path criterion. We computed the georegistration precision to be 0.88 cm and, therefore, argue an absolute precision of the georegistration around 1 cm, similar to

the resolution of the hyperspectral mosaic. We also compared the hyperspectral mosaics with the photomosaic in Fig. 14, computing an average error, or accuracy of 0.48 cm in georegistration of eight objects. Therefore, we argue that our method enables georegistration that is highly consistent with photogrammetry. In total, the hyperspectral mosaic is the first spatially consistent mosaic from multiple transects in UHI history.

V. CONCLUSION

In this work, we have developed a methodology for precise and consistent georegistration in UHI using an RGB camera to build a photogrammetry model for estimating the pose and the 3-D model. To enable this, two methods for geometric calibration of the HSI with the RGB camera has been proposed. The calibration and georegistration have been demonstrated in a laboratory test tank and for *in situ* UHI of the Tautra cold-water coral reef in the Trondheim Fjord.

We have suggested a geometric sensor model for the HSI and demonstrated the performance of two HSI calibration methods in the laboratory. One of these methods was also designed for *in situ* use. The laboratory trials showed that both calibration methods yielded high accuracy in georegistration using a calibration pattern and a calibrated RGB camera for pose estimation. The georegistration model was further developed for *in situ* use, where a photogrammetry model yielded the RGB camera pose and the 3-D seabed model. Using the proposed ray casting procedure for intersecting the HSI rays with the 3-D model, we were able to do *in situ* georegistration and calibration of the HSI. After the calibration, all the hyperspectral imagery was georegistered, creating a point cloud of measurements from each transect. Finally, the hyperspectral measurements were IOP/AOP corrected to spectral reflectance using the optical paths of the HSI rays, before gridding the point clouds into a hyperspectral reflectance mosaic of the seabed.

The georegistration method, supported by the calibrations, allowed for achieving unprecedented precision in georegistration of UHI, demonstrated by an average error in georegistration of 0.88 cm across hyperspectral transect mosaics. Therefore, the precision of georegistration is argued to be around the spatial resolution of the mapping, which is 1 cm. The georegistration method facilitated UHI that was consistent with photogrammetry, with an average georegistration error of 0.48 cm between the photomosaic and the hyperspectral transect mosaics. This resulted in a hyperspectral mosaic that was highly consistent with the RGB photomosaic spatially. The hyperspectral mosaic was also the first-ever properly georegistered UHI mosaic to contain multiple transects, enabling us to build a spatially consistent hyperspectral map of the seabed.

We consider the next step in research to be using the georegistration method for developing improved IOP/AOP correction methods for UHI. Potential improvements in IOP/AOP correction include accurate geometric modeling of the light field from the light sources, using a beam attenuation meter for attenuation measurements, estimation and correction for backscatter from the water column, and correction for the orientation of surface normals on the seabed. For this research, it is suggested to use

a higher resolution mesh model or potentially combining UHI with a stereo camera for improved depth estimation.

APPENDIX A IOP/AOP ESTIMATION

Assume that constant $C(\lambda)$ is an AOP inversely proportional to the light source spectrum, while $K(\lambda)$ comprises the IOP called beam attenuation and AOPs related to the geometric loss from the light sources. The method for obtaining $K(\lambda)$ is inspired by Lyzenga [31], who uses airborne multispectral measurements of the same bottom type to compute the relationship between attenuation coefficients. Essentially, the full procedure estimates reflectance using *in situ* reflectance targets, with correction for attenuation. Assume the following.

- 1) We select a set of radiance spectra from the mosaic cells, $L_{n,m}^{0,i}(\lambda)$, of an OOI with a known reflectance spectrum, $R_c(\lambda)$. Superscript i denotes sample number. These mosaics are computed by simply performing the gridding step in Section II-G2 on the raw radiance data cube. The pixels should have varied corresponding $d_{n,m}^i$. In our samples, $d_{n,m}^i$ ranges from 1.25 to 3.25 m.
- 2) We select a reference wavelength λ_{ref} with high SNR to define the relative attenuation $\hat{K}(\lambda) = K(\lambda) - K(\lambda_{\text{ref}})$.
- 3) We assume $\hat{K}(\lambda)$ and $C(\lambda)$ to be the only wavelength-dependent factors.

With the assumptions made, we rewrite (11) to

$$\ln(\hat{L}_{n,m}^{0,i}(\lambda)) = \hat{K}(\lambda) \cdot (2d_{n,m}^i) - \ln(\hat{L}_{n,m}(\lambda)) \quad (13)$$

where $\hat{L}_{n,m}^{0,i}(\lambda) = L_{n,m}^{0,i}(\lambda)/L_{n,m}^{0,i}(\lambda_{\text{ref}})$ is the measured spectral radiance normalized by the radiance at λ_{ref} . In the same way, $\hat{L}_{n,m}(\lambda) = L_{n,m}(\lambda)/L_{n,m}(\lambda_{\text{ref}})$ is the normalized attenuation-corrected radiance. Using the normalized spectra removes the effects of wavelength independent factors, meaning that we only consider the change in spectral shape. Using regular least squares (LS) regression, (13) is solved for the slope $\hat{K}(\lambda)$ and intercept $-\ln(\hat{L}_{n,m}(\lambda))$ for all $\lambda \neq \lambda_{\text{ref}}$. Finally, $K(\lambda_{\text{ref}})$ and $L_{n,m}(\lambda_{\text{ref}})$ is computed using measured radiance $L_{n,m}^0(\lambda_{\text{ref}})$:

$$\ln(L_{n,m}^{0,i}(\lambda_{\text{ref}})) = K(\lambda_{\text{ref}}) \cdot (2d_{n,m}^i) - \ln(L_{n,m}(\lambda_{\text{ref}})) \quad (14)$$

where (14) is solved like (13) with LS regression. Then, we compute the diffuse attenuation coefficient $K(\lambda)$ and light source constant $C(\lambda)$ using the following expressions:

$$K(\lambda) = \hat{K}(\lambda) + K(\lambda_{\text{ref}}) \quad (15a)$$

$$C(\lambda) = \frac{R_c(\lambda)}{\hat{L}_{n,m}(\lambda) \cdot L_{n,m}(\lambda_{\text{ref}})}. \quad (15b)$$

REFERENCES

- [1] I. Dumke *et al.*, "First hyperspectral imaging survey of the deep seafloor: High-resolution mapping of manganese nodules," *Remote Sens. Environ.*, vol. 209, pp. 19–30, 2018.
- [2] 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 Sens.*, vol. 11, no. 6, 2019, Art. no. 685.
- [3] Ø. Ødegård, A. A. Mogstad, G. Johnsen, A. J. Sørensen, and M. Ludvigsen, "Underwater hyperspectral imaging: A new tool for marine archaeology," *Appl. Opt.*, vol. 57, no. 12, pp. 3214–3223, 2018.

- [4] B. Liu *et al.*, "Underwater hyperspectral imaging technology and its applications for detecting and mapping the seafloor: A review," *Sensors*, vol. 20, no. 17, 2020, Art. no. 4962.
- [5] P. Doucette, J. Antonisse, A. Braun, M. Lenihan, and M. Brennan, "Image georegistration methods: A framework for application guidelines," in *Proc. IEEE Appl. Imag. Pattern Recognit. Workshop*, 2013, pp. 1–14.
- [6] R. Szeliski, *Computer Vision: Algorithms and Applications*. New York, NY, USA: Springer, 2010.
- [7] G. Johnsen, M. Ludvigsen, A. Sørensen, and L. M. S. Aas, "The use of underwater hyperspectral imaging deployed on remotely operated vehicles—Methods and applications," *IFAC-PapersOnLine*, vol. 49, no. 23, pp. 476–481, 2016.
- [8] F. Foglini *et al.*, "Application of hyperspectral imaging to underwater habitat mapping, southern adriatic sea," *Sensors*, vol. 19, no. 10, 2019, Art. no. 2261.
- [9] A. A. Mogstad *et al.*, "Mapping the historical shipwreck figaro in the high arctic using underwater sensor-carrying robots," *Remote Sens.*, vol. 12, no. 6, 2020, Art. no. 997.
- [10] 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," *J. Field Robot.*, vol. 35, no. 3, pp. 312–329, 2018.
- [11] S. B. Williams *et al.*, "Monitoring of benthic reference sites: Using an autonomous underwater vehicle," *IEEE Robot. Autom. Mag.*, vol. 19, no. 1, pp. 73–84, Mar. 2012.
- [12] D. L. Bongiorno, A. J. Fairley, M. Bryson, and S. B. Williams, "Automatic spectrometer/RGB camera spatial calibration," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2013, pp. 4431–4434.
- [13] J. M. Roberts, A. Wheeler, A. Freiwald, and S. Cairns, *Cold-Water Corals: The Biology and Geology of Deep-Sea Coral Habitats*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [14] J.-Y. Bouguet, "Camera calibration toolbox for Matlab," 2004. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html
- [15] B. Sun, J. Zhu, L. Yang, S. Yang, and Z. Niu, "Calibration of line-scan cameras for precision measurement," *Appl. Opt.*, vol. 55, no. 25, pp. 6836–6843, 2016.
- [16] C. D. Mobley, *Light and Water: Radiative Transfer in Natural Waters*. New York, NY, USA: Academic, 1994.
- [17] F. Dukan, "ROV motion control systems," Ph.D. dissertation, Dept. Marine Technol., Norwegian Univ. Sci. Technol., Trondheim, Norway, 2014.
- [18] S. M. Nornes, "Guidance and control of marine robotics for ocean mapping and monitoring," Ph.D. dissertation, Dept. Marine Technol., Norwegian Univ. Sci. Technol., Trondheim, Norway, 2018.
- [19] F. Dukan and A. J. Sørensen, "Integration filter for APS, DVL, IMU and pressure gauge for underwater vehicles," *IFAC Proc. Vol.*, vol. 46, no. 33, pp. 280–285, 2013.
- [20] M. Bryson, M. Johnson-Roberson, O. Pizarro, and S. B. Williams, "True color correction of autonomous underwater vehicle imagery," *J. Field Robot.*, vol. 33, no. 6, pp. 853–874, 2016.
- [21] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.*, vol. 310, no. 1, pp. 1–26, 1980.
- [22] T. Möller and B. Trumbore, "Fast, minimum storage ray-triangle intersection," *J. Graph. Tools*, vol. 2, no. 1, pp. 21–28, 1997.
- [23] S. Gottschalk, M. C. Lin, and D. Manocha, "OBSTree: A hierarchical structure for rapid interference detection," in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn.*, 1996, pp. 171–180.
- [24] *Ray Casting With Python and VTK: Intersecting Lines/Rays With Surface Meshes*, 2014. [Online]. Available: <https://pyscience.wordpress.com/2014/09/21/>
- [25] A. A. Mogstad and G. Johnsen, "Spectral characteristics of coralline algae: A multi-instrumental approach, with emphasis on underwater hyperspectral imaging," *Appl. Opt.*, vol. 56, no. 36, pp. 9957–9975, 2017.
- [26] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, vol. 6, no. 2, pp. 418–445, 1996.
- [27] T. F. Coleman and Y. Li, "On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds," *Math. Program.*, vol. 67, no. 1, pp. 189–224, 1994.
- [28] M. A. Branch, T. F. Coleman, and Y. Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 1–23, 1999.
- [29] M. J. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 2, pp. 155–162, 1964.

- [30] F. Sigernes *et al.*, "The absolute sensitivity of digital colour cameras," *Opt. Exp.*, vol. 17, no. 22, pp. 20 211–20220, 2009.
- [31] D. R. Lyzenga, "Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and landsat data," *Int. J. Remote Sens.*, vol. 2, no. 1, pp. 71–82, 1981.



Håvard Snefjellå Løvås received the M.Sc. degree in marine technology from the Norwegian University of Science and Technology, Trondheim, Norway, in 2019, where he is currently working toward the Ph.D. degree in underwater hyperspectral imaging techniques with the Centre for Autonomous Marine Operations and Systems.



Aksel Alstad Mogstad received the M.Sc. degree in marine biology from the Norwegian University of Science and Technology, Trondheim, Norway, in 2017, where he is currently working toward the Ph.D. degree in marine biological applications of underwater hyperspectral imaging with the Centre for Autonomous Marine Operations and Systems.



Asgeir J. Sørensen (Senior Member, IEEE) received the M.Sc. degree in marine technology and the Ph.D. degree in engineering cybernetics from Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1988 and 1993, respectively.

He has previously been employed at MARINTEK, ABB, and Marine Cybernetics. Since 1999, he has been a Professor of Marine Control Systems with the Department of Marine Technology, NTNU, where he is the Director and one of the key scientists with the Centre for Autonomous Marine Operations and Systems. He has authored more than 250 scientific articles and book chapters on marine control systems.



Geir Johnsen received the M.Sc. degree in marine biology and the Ph.D. degree from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1989 and 1994, respectively.

He is currently a Professor of Marine Biology with the Centre for Autonomous Marine Operations and Systems, Department of Biology, NTNU. He is also an adjunct Professor with the University Centre in Svalbard, Longyearbyen, Norway. His research interests include marine ecology and biodiversity, bio-optics, photosynthesis, pigment chemotaxonomy, underwater robotics, and sensor development for *in situ* identification, mapping, and monitoring of biogeochemical objects of interest in the marine environment.