SeaGrassDetect

A Novel Method for the Detection of Seagrass from Unlabelled Underwater Videos

Sengupta, Sayantan; Ersbøll, Bjarne Kjær; Stockmarr, Anders

Published in:
Ecological Informatics

Link to article, DOI:
10.1016/j.ecoinf.2020.101083

Publication date:
2020

Document Version
Early version, also known as pre-print

SeaGrassDetect: A novel method for detection of seagrass from unlabelled under water videos

Sayantan Sengupta*, Bjarne Kjær Ersbøll, Anders Stockmarr

Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

Abstract

Benthic vegetation is arguably one of the most important indicators of the state of marine environment. Assessment of the status of eelgrass *(Zostera marina)* is commonly done using various remote sensing methods such as aerial photography or satellite images. These methods often fail to capture the true scenario beneath the surface of the water if the water is turbid or the satellite image is masked by cloud cover which makes it impossible to see beneath them. As a second line of defense, researchers have used under water videos (obtained either with scuba,snorkel or other visual observations) to assess the ground truth. Lots of man-hours are spent browsing through many hours of video data manually by an expert and assessing the status (presence/absence) which is a very common practice. Here we propose two methods for detection of eelgrass (presence/absence) from under water videos obtained from Roskilde Fjord in Denmark. We extend these methods to show that it can be used as a proxy to estimate coverage of eelgrass in a given area which match well with an expert's estimation. The benefit of using this method is that it is objective, less biased, cost efficient, robust to noisy environment, does not require pixel-level annotated ground truth images and can be used on existing video transects. This method can also detect rare errors from domain expert's visual estimation.

Keywords: ecological monitoring; coverage estimation; seagrass detection; ecological indicators; vegetation mapping; ground truth.

1. Introduction

Any study on assessment of marine ecosystems includes measuring the status of various ecological indicators which determines the state of the environment. The seagrass are important habitats in coastal ecosystems. It helps in carbon sequestration [1]. It also provides protection and nurseries for juvenile fishes and contributes to sediment stabilisation. Eelgrass(*Zostera Marina*), a kind of seagrass that grows in the sea bed has been considered one

*I am corresponding author
Email address: says@dtu.dk (Sayantan Sengupta)

of the most important ecological indicators [2]. The role of seagrass is often compared to that of a forest [3]. Seagrass among other benthic vegetation in the coastal marine ecosystem are under tremendous human pressure [4] and according to The International Union for Conservation of Nature (IUCN) Red list of Threatened species (www.iucnredlist.org), eelgrass is decreasing. If no precautions are taken, fish and sea products are estimated to reduce dramatically by the middle of 21st century [5] and all the world's oceans are said to be affected [6]. In northern temperate ecosystems, eelgrass is the dominant seagrass and in many areas constitutes the only rooted vegetation in soft sediments and constitute a key organism both in terms of ecosystem functioning and as indicator for environmental assessment and management of marine ecosystems. According to studies only 5-10% of the world's seafloor is mapped [7]. Due to it's vulnerability to anthropogeneic pressures, seagrass is used as an indicator of ecosystem health and state in many environmental regulations such as the European WFD [8]. There is a need to map the benthic vegetation in a easy and robust way to be able to assess the full impact of human exploitation of nature as well as continuous monitoring of the seabed to evaluate the positive steps taken to replenish the seagrass.

Monitoring of seagrass is inherently difficult due to large patchiness. Traditional methods are to assess the depth limit [9], biomass cover [10], etc. But these methods are time consuming, subjective and associated with large uncertainty due to diver-specific variations [11]. Other approaches include remote sensing methods like satellite images or aerial photography [12]. But remote sensing methods are also prone to errors when satellite images are masked by cloud cover and the turbidity of water prevent visibility of vegetation cover beneath the surface of the water. Due to this researchers have also traditionally relied on under water video transects and scuba divers as ground truth to assess the status of seagrass. Many studies such as [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], has used either videos, scuba divers or remote operating vehicles(ROV) to assess ground truth for different species of seagrass. [23] has reviewed 91 seagrass related studies which comprise 58 journal articles, 20 conference proceedings, 7 technical notes, 2 books, 3 Ph.D theses of which 52 studies have used scuba divers, video transects as method for ground truthing. Ground truthing is usually done by visual inspection of the videos by an expert, who keeps a track of the presence/absence or coverage of eelgrass along with the location(gps coordinates), time, height from the ground and other measured quantities, which is a very tedious task. Lot of person hours are spent on this exercise. It would save lot of time, if this whole process of detecting grass from videos be automatized. Also this process includes an element of expert judgement and can never be fully unbiased. Multiple experts may not agree on the same precision of estimated coverage from the video transects as no scientific standards are established. Therefore, our aim was to investigate if the assessment of traditional diver transects could be improved through automated image analysis, allowing larger areas to be covered without the time-consuming and error-prone post-analysis of video inspection by domain experts.

With the increasing popularity of machine learning in the past few years, this kind of problem, which falls in the domain of Computer Vision can now be efficiently tackled saving both time and commercial resources for the ecological researchers. Training a machine learning model requires labelled training data. [24] manually labelled a small

dataset(less than 200 images) to train a Support Vector Machine(SVM) and Artificial Neural Networks(ANN) to predict presence/absence of posidonia meadows from Autonomous underwater vehicles(AUV) and ROV. [25] has also used underwater vehicles to evaluate bottom coverage descriptors for assessing the good ecological status of seagrass meadows. However, to the author's knowledge, an autonomous detection and mapping of seagrass has not yet been achieved. [26] divide the images in small patches and classify each patch as seagrass or background. [27], [28] and [24] also follow a patch classification approach to segment images. [27] also use a pixel refinement method as post processing step to get better results than pure patch based classification. Due to the lack of availability of standard labelled seagrass images in public domain, there have been limited use of machine learning models on seagrass detection tasks. [29] in 2018 published the first publicly available pixel annotated seagrass dataset captured from an AUV at different depths and trained a Convolutional Neural Network(CNN) model to segment seagrass images. Using the dataset of [29], [30] trained different deep learning architectures, including the DeepLabv3Plus network and compared their results. In 2018 [31] came out with a method which use Visual Geometric Group network(VGGNet) as the encoder and a Fully Convolutional Network(FCN-8) as decoder plus skip connections to segment seagrass images collected from an Autonomous underwater vehicles(AUV). There exists concern regarding the annotation of ground-truth data of [29], which was done relatively roughly with certain inaccuracies using polygons [30]. In reality, the seagrass areas are defined by the fine leaves of the plants. This also raises doubts and uncertainity about the correctness of the labelling process as inaccurate labelling have significant negative effect on the metrics.

In this paper, we propose a novel method using an edge detection based approach as a feature extractor to discriminate between regions with eelgrass and no eelgrass without any need of labelled seagrass images. This will bypass the need of domain expert's requirement to estimate coverage or status of seagrass manually. Line detection is a classic technique that has been widely used in many areas like agriculture for task related to detection of crop rows [32], Optical Character Recognition(OCR) for detecting characters[33] and Biometrics[34]. We propose two methods, one which involves fusion of predictions obtained from two individual features to distinguish eelgrass frames from no eelgrass frames and another using a Gaussian mixture model to cluster video frames of eelgrass from no eelgrass. The methods also give a way to quantify the subjectivity involved in eelgrass estimation through the threshold. We extended this model further from presence/absence modelling to a proxy for coverage estimation in a given area and show that it follows the domain expert's estimation very well. It even detects and rectify rare human error from the domain expert, which makes it truly a robust tool for eelgrass detection. This method can also be extended to detect other species of seagrass or vegetation that has sharp edges (posidonia oceania meadows).

2. Materials and Methods

2.1. Field site

The data used for this research is obtained from Roskilde Fjord in Denmark. Roskilde Fjord is a 40-km long and narrow micro-tidal estuary forming four major basins. It is 10 km

Figure 1: The study site have been zoomed in to show the path traced by the diver of all the video transects. One of them have been further zoomed in.

wide and has a surface area of 122 km^2, a volume of ($360 \times 10^6 m^3$) and a mean depth of 3 m [35]. The water temperature vary from 0°C in the winter to 22°C in the summer. There is a strong hydraulic effect from the outer boundary where strong western wind events are able to create water levels of up to 170 cm above mean sea level, although the tidal range only is + 10 cm. There are several small freshwater inputs along the estuary. A large freshwater source in the inner part is absent which is unusual compared with many estuaries. Water residence is 90 days during maximum winter runoff [36]. Roskilde Fjord is connected to the estuary Isefjorden and through a narrow strait to the Kattegat forming the Baltic Sea/North Sea transition zone.

2.2. Data Description

We used the video transects that were recorded as part of the national monitoring program: NOVANA [37]. A scuba diver with a camera attached to his head was towed by a boat. The videos were recorded on different areas in the fjord (Figure 1), and other measurements such as co-ordinates, time, date, height from the bottom were measured and embedded in the video. The diver mostly maintained a height of 1 m from the seabed. A total of 14 video transects were recorded with an average length of 17 minutes each, specifications of which are given in Table 1.

The videos has been recorded at different times of the day over the period of 2 days. As can be seen from the Figure 2, the colour of the environment change very much based on the spatial location as well as the time of the day. Also can be seen in Figure 2, the

Compression	.mpg
Frame width	720
Frame height	576
Frame rate	25 fps
Data rate	10000 kbps
Total bit rate	10256 kbps

Table 1: Video specifications of the transects

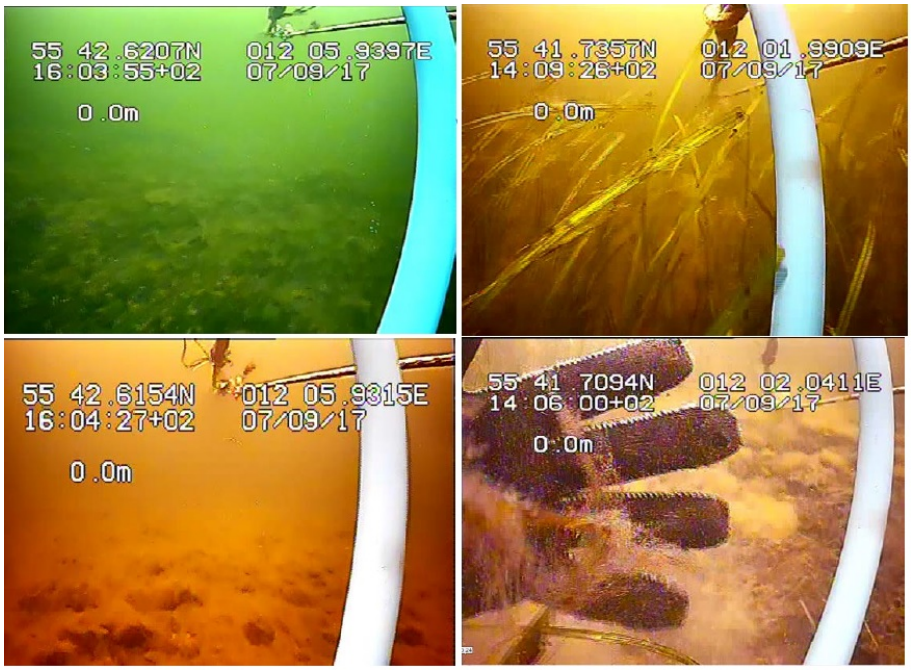

Figure 2: Snapshots of the environment at different times

presence of *Noisy objects* in the form of wires of the recording device, hand movement of the diver and also the embedded texts on the top half of the video. Two separate approaches, feature fusion and probabilistic clustering are used to detect eelgrass and will be discussed in details. To extract embedded information (location, timestamp, depth, date) from the video transects, a neural network has been used which gives a 98% accuracy on the text recognition and this will not be discussed in details.

2.3. Feature fusion approach for eelgrass detection

2.3.1. Core idea of edge detection

An important physical feature about eelgrass is sharp edges. Our algorithm is based on finding sharp edges or edge detection algorithms (as more commonly known in the domain of computer vision). An edge point is defined as the region in the image, where there is a significant gradient change(from white to black or vice versa). Then a contour of similar connected edge points with similar gradient change constitutes a line. The algorithm finds

all these connected regions as lines, which may vary in length. We have used a linear time Line Segment detector (LSD) giving sub-pixel accurate results [38]. This algorithm works on gray scale video frames by detecting number of edges of varying lengths in the video frames.

2.3.2. Why this works for eelgrass detection?

On a video frame with eelgrass, the algorithm should detect lot of edges and in return give more number of lines compared to a frame with no eelgrass, where we should ideally get very less number of lines (as the flat sea bed has no objects with sharp edges apart from eelgrass). Two quantities can be computed using this information:*Number of lines* and *Total length of lines* from each video frames. This also implies that the total length of the lines (cumulative sum of all the line length) in each frame with presence of eelgrass will be more than frames with no eelgrass.

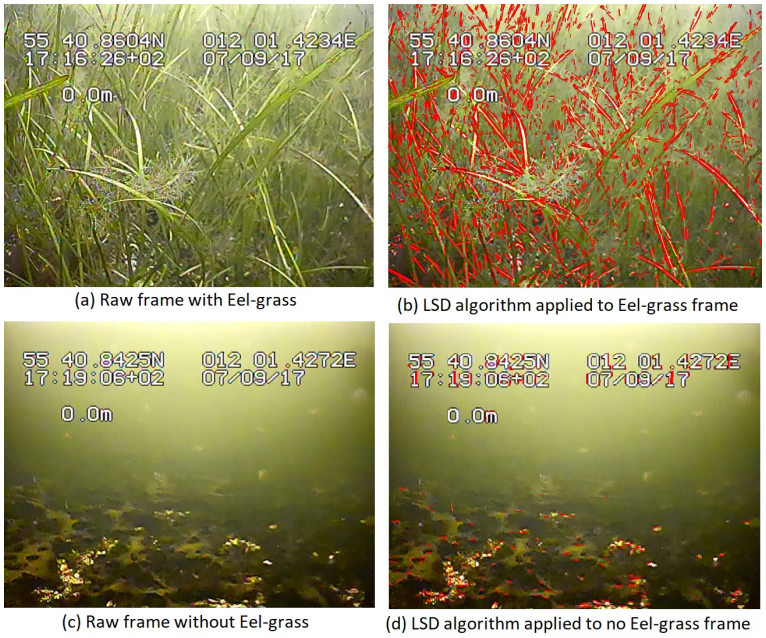

Figure 3: Effect of LSD algorithm on two different scenarios

We can choose to work with either of these features and define a suitable threshold. For a random video frame, if we consider *Total length of lines* as a feature of importance, then if the measured quantity is less than threshold, it is labelled as absence of eelgrass. It is labelled as presence of eelgrass if the quantity is more than the threshold. From figure 3 we can see that counting the *Number of lines* or *Total length of lines*, we can find a suitable decision boundary/threshold using visual inspection on frames with presence of eelgrass from absence of eelgrass.

2.3.3. Efficient filtering of the temporal data

In Fig 4, we have used one of the video transect, where we have shown the temporal data of *Number of lines* recorded as the diver navigates. We can see lot of fluctuations or spikes. These spikes are due to noisy objects, such as camera wires, hand movement of the diver in front of the camera, etc. Any filtering techniques require a choice of window size over which smoothing operation takes place. Instead of finding how many data points constitutes the window size, we found a way to map the window size to physical distance on the ground. To do this, we found the median distance between unique locations in each

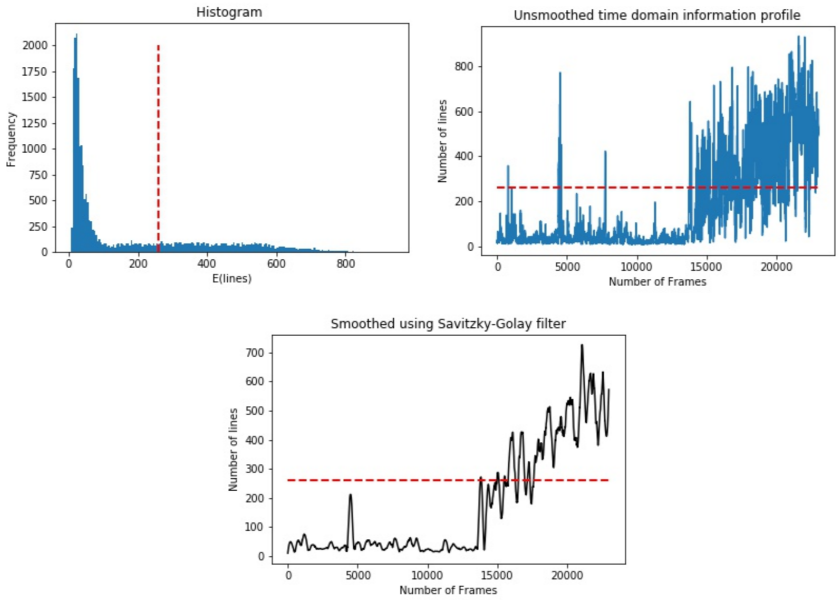

Figure 4: Summary of a video transect based on Number of lines(feature 1). Top left shows the histogram, Top right shows the raw time profile. Bottom shows the smoothed version of the time profile.

transects. For the 14 video transects, they have four unique median values as 0.2m (3 transects), 0.3m (7 transects), 0.4m (3 transects) and 0.435m (1 transect). We have used a physical distance of 7.5 meters on ground as the size of window, due to which the 17-35 data points are used in the filtering algorithm as window size. For eg. for a given transect if the median distance between successive unique location is 0.2m, then to cover a physical distance of 7.5m, window size is 37 samples. We have used Savitzky-Golay filter [39] with varying window size of 17-35 data points and polynomial order of 1, as a smoothing function to remove these noisy fluctuations. A threshold was chosen by visual inspection which best classified the eelgrass frames from non-eelgrass. This threshold is 260 lines for the video transect which is illustrated in Fig 4. As can be seen from the smoothed temporal profile, a small bump (which is actually noise due to the camera going above the surface of the water and captures non eelgrass objects) is greatly reduced from the unsmoothed temporal profile by the smoothing operation.

2.3.4. Need for optimal combination of both features

Finding a suitable threshold independently for both the features (*Number of lines* and *Total length of lines*) leads to different outcomes. On closer inspection of those video frames, where the outcome from the independent features does not match, we found that there are no clear pattern. We manually checked the video transects for most of the locations of mismatched predicted frames (226 out 18727 unique location) and found out that number of lines (feature 1) is performing slightly better than total length of the lines (feature 2), although there was no clear indication of which feature works better. Each of them has it's own characteristics. From our analysis we found that Feature 2 (*Total length of lines*) is more sensitive to noise, but gives better prediction in eelgrass zones. Feature 1 (*Number of lines*) has a tendency to underestimate (as perceived from visual inspection), but more robust in noisy environment. Although these two features are correlated, as they are measured from the same quantity (detected lines), but they also carry different information in predicting presence/absence status of frames.

2.3.5. Optimal combination strategy

In our method, we will be using both these features(*Number of lines* and *Total length of lines*) and derive an optimal combination strategy to predict presence/absence status of each frames. To do this, first the optimal threshold is found for the feature *Number of lines* by visual inspection that gives the best prediction for all the video transects, which is 260. Since this falls in the purview of unsupervised machine learning problem (since the label: presence/absence of each frames are not known beforehand), the only evaluation of the results is by visual inspection. Visual inspection is specially important in the regions where there is a transition from present to absent or vice versa. Those regions were carefully analysed to fine tune the threshold. Then prediction using the feature *Total length of lines* is found for different threshold varying from 4000 to 7000. The optimal value of the threshold of *Total length of lines* is chosen as 6300, as it gives the maximum match with the threshold of *Number of lines* as 260. Using these two thresholds, we formulate a combination strategy of the form:

$$\text{Status} = \begin{cases} \text{Present} & \text{if} \lambda \cdot \frac{N.L}{260} + (1-\lambda) \cdot \frac{T.L.L}{6300} \geq 1 \\ & \text{where N.L : Number of Lines} \\ \text{Absent} & \text{if} \lambda \cdot \frac{N.L}{260} + (1-\lambda) \cdot \frac{T.L.L}{6300} < 1 \\ & \text{where T.L.L : Total length of Lines} \end{cases}$$

To find the optimal value of λ , we vary it from $\lambda = 0$(using just Total length of lines) to $\lambda = 1$(using just number of lines) at an interval of 0.02. This gives different weightage to the features and our job is to find out this optimal weightage in a heuristic way, which shall reflect the relative importance of each feature when making prediction, thus giving it more flexibility on noisy environments. A confusion matrix (a small subset shown in Table 2) is formed in which contains the number of predicted mismatched frames between different value of λ. The idea is to find a region of lowest value of mismatched predicted frames as well as low change of this value when λ is perturbed. As the golden standards of the

$\lambda=$	**.52**	**.54**	**.56**	**.58**	**.6**	**.62**	**.64**	**.66**	**.68**	**.7**
.52	0	2	5	12	20	26	28	35	40	47
.54		0	3	10	18	24	26	33	38	45
.56			0	7	15	21	23	30	35	42
.58				0	8	14	16	23	28	35
.6					0	6	8	15	20	27
.62						0	2	9	14	21
.64							0	7	12	19
.66								0	5	12
.68									0	7
.7										0

Table 2: Variation in number mismatched predicted frames with λ. These are the number of frames out of a total 18727 frames where the predictions mismatch. The diagonals are zero as comparing the same value of λ with each other will always give a perfect match.

predicted frames are not available, this is solely based on the fact that small changes will be the expected consequence of an optimal λ. This give a neighbourhood of optimal value in the function space as the gradient tends to be very small.

We found two such regions from Table 2 as $\lambda = 0.52$ and 0.62. On manual investigation of the disagreed prediction status among these two possible values, $\lambda = 0.62$ performed better than $\lambda = 0.52$ and thus chosen as the optimal value. $\lambda = 0.62$ also performed way better when compared to $\lambda = 0$ and with $\lambda = 1$, the performance was slightly better. This way of choosing the optimal λ is independent from domain expert's label information.

2.4. Probabilistic clustering approach for eelgrass detection

2.4.1. Gaussian Mixture Model

GMM is one of the most powerful and popular technique for unsupervised learning task. It is widely used in clustering of data [40] and has found use in many applications like Computer vision [41], signal processing [42], biometrics [43], natural language processing [44], bioinformatics [45] among many other field. The goal of the Gaussian mixture-model (GMM) is to derive a distribution for an M-dimensional vector $x \in R^M$ which we will write as *p(x)*. We wish this distribution to be potentially very flexible and a common strategy for obtaining this in a tractable manner is to make *p(x)* be a combination of more simpler and tractable elements. These tractable elements are multivariate normal densities. A normal distribution is defined as the density:

$$\mathcal{N}(x|\mu,\, \sigma^2) \;= \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

where $\mu \in R^M$ is the mean and σ is $M \times M$ covariance matrix. $M = 2$ in our case as we have two features. In GMM, we use normal distribution as a building block of more complicated

distribution:

$$p(x) = \sum_{z=1}^{Z} p(x, z) = \sum_{z=1}^{Z} p(x|z)p(z) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

where z is the number of cluster as defined by the user. The above equation shows that the density of the data can be expressed as weighted sum of many normal distributions. GMM works by maximising the log-likelihood of the data, given by:

$$\mathcal{L}(\pi, \mu, \Sigma) = \log p(X|\mu, \Sigma, \pi) = \sum_{i=1}^{N} \log p(x_i|\mu, \Sigma, \pi) = \sum_{i=1}^{N} \log[\sum_{k=1}^{K} \pi_k \mathcal{N}(x_i|\mu_k, \Sigma_k)]$$

In this case $Z = 2$ as we want to cluster the data in to 2 cluster, one where eelgrass is present and the other where eelgrass is absent. The goal is to solve for the above optimisation problem and estimate the three parameters namely μ (mean), Σ (covariance) and π_k (weight of each cluster).To solve the above optimisation of the log-likelihood, a popular method known as EM (Expectation maximisation) is used [46]. After the algorithm has converged successfully, we can substitute the three estimated parameters back in the previous equation to get the probability of each data point belonging to certain cluster. This kind of clustering is also known as soft clustering due to the probabilistic nature of assigning clusters. In comparison to hard clustering methods like K-means and hierarchical clustering algorithms where the assignment of a data point to a cluster is binary and not probabilistic, GMM provides more flexibility in the way data points are assigned to clusters.

2.4.2. Eelgrass detection using GMM

We have 18727 data points(unique locations), each measuring two features from the 14 video transects. First, the data is filtered suitably using the same filtering technique as discussed in section 2.3.3. We apply GMM to cluster the data in 2 clusters. The mean and the covariance is found and is shown in Figure 5. Each point in Figure 5 is a unique location from the video transects. It can be seen that the model has captured the elongated spread of the data. The two clusters are not well separated. Any data point in the scatter plot, whose probability of being in cluster 1 is greater than being in cluster 0, will be labelled as eelgrass present and probability of being in cluster 1 is less than being in cluster 0, will be labelled as eelgrass absent. This effectively corresponds to a decision boundary where the probability of eelgrass present is equal to probability of eelgrass absent which is equal to 50%.

3. Results

3.1. Comparison of feature fusion and probabilistic clustering approaches

Figure 6 shows the contours of different decision boundaries from the two different approaches on a scatter plot. Three linear decision boundaries (corresponding to the feature fusion approach) and one non-linear decision boundary (corresponding to GMM) can be seen. It is interesting to see that the decision boundaries from both these approaches are not so far away from each other.

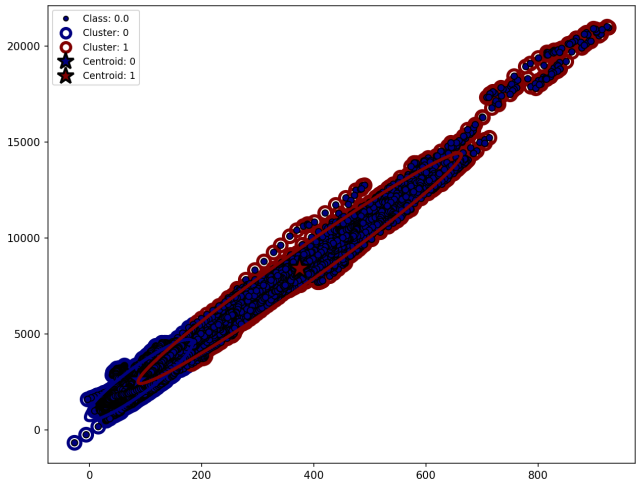

Figure 5: Two clusters are estimated using GMM in a scatter plot. The x-axis corresponds to *Number of lines* and y-axis corresponds to *Total length of lines*. The red cluster (cluster 1) corresponds to region of eelgrass and blue cluster (cluster 0) corresponds to region of no eelgrass. Corresponding mean and co-variances are also drawn for each clusters.

The GMM has slightly lower threshold (using the 50% cut point) than the combination strategy due to which it will tend to over-estimate some data points as eelgrass which the other method classifies as non-eelgrass. It can also be seen that comparing the threshold from individual features (*Number of lines* and *Total length of lines*) and the combination, angular trend of the scatter plot is followed by the combination decision boundary. It has already been discussed that $\lambda = 0.62$ also performed better when compared to $\lambda = 0$ and with $\lambda = 1$ by visually inspecting the video frames whose prediction does not match with each other.

3.2. Comparison with Expert's estimation

We have the diver's evaluation from the above area of study from the SeaStatus Project as ground truth to compare the performance of our method. The diver's estimation is given in the form of percentage coverage of eelgrass from a start point to an end point. The presence/absence status of all unique location falling inside the range of a given area (from diver's estimation) were found and the number of *Present* and *Absent* frames were found. We give the algorithm's output as the ratio of *Present* frames to the total frames detected inside a given start point and end point as given by the diver. It can also be thought of as a proxy to percentage coverage. In Table 3, we have given a comparison of the algorithm's output for three different λ. $\lambda = 0$ corresponds to prediction using just *Total length of Lines* and $\lambda = 1$ corresponds to prediction using just *Number of Lines*.

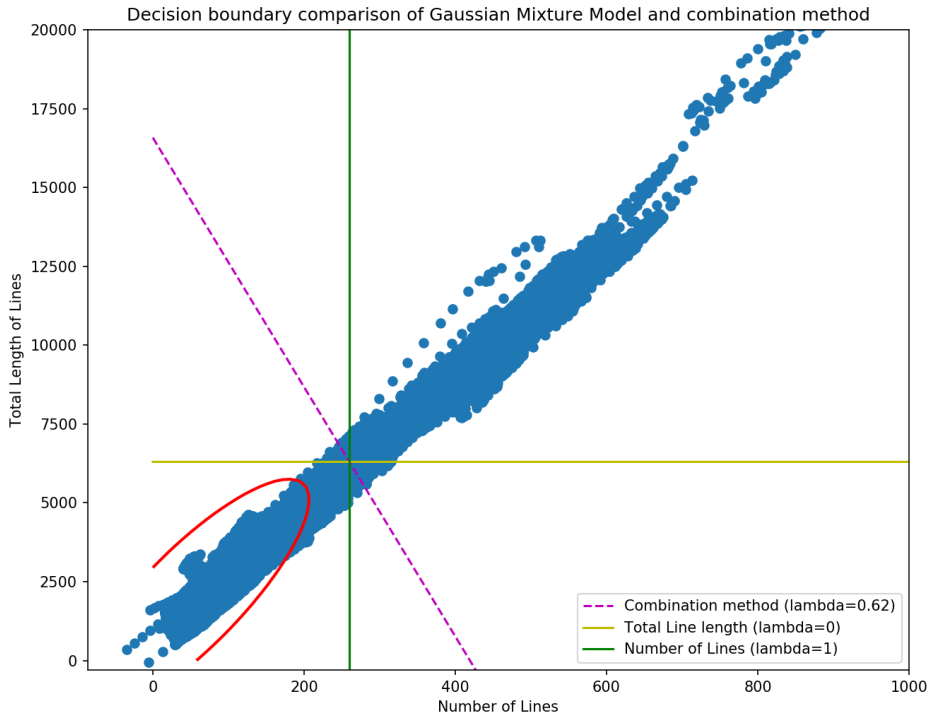

Figure 6: Comparison of decision boundaries between GMM and feature fusion method. The red curve corresponds to GMM decision boundary. Data point located on the left of the boundary is classified as eelgrass absent.

It is divided in to two sections to divide between observations which match and does not match with the expert's estimation. It can be seen that in the top part of the table, the algorithm's prediction in terms of coverage matches decently well with the expert's estimation(which is given as a percentage coverage). It is interesting to observe that the ratio of *Present* frames to total frames in a particular region can also be used as a proxy for the coverage. All the area where the expert have estimated as either very less $(0 - 20\%)$ or very high $(80\% - 100\%)$, both the approaches, feature fusion and GMM predicts similar estimates which shows that both the methods perform very well in dense and sparse region of eelgrass. In less dense regions, the performance varies. The second row of the table has the expert's estimate as 50% coverage, which the feature fusion also validates as 12 *Present* frames out of total 21 frames $(12/21 = 57.1\%)$ whereas the GMM estimates it to be 85%. Similar close comparison can also be seen from other rows as well in the top part of the table, except for the last observation in which the feature fusion is under-estimating the coverage compared to GMM. λ equal to 0.62 has a slight better performance than using just either one of the features ($\lambda = 1$ and $\lambda = 0$). This is also intuitive from Figure 6, the GMM over-estimate than the feature fusion due to lower threshold.

North(start)	East(start)	North(end)	East(end)	Expert	$\lambda = 0$	$\lambda = 0.62$	$\lambda = 1$	GMM
5542.649	1206.088	5542.656	1206.060	0%	0%	0%	0%	0%
5542.654	1206.006	5542.650	1205.985	50%	57.1%	57.1%	52.38%	85.7%
5542.649	1205.979	5542.645	1205.968	90%	100%	91.6%	75%	100%
5542.615	1205.932	5542.599	1205.916	0%	0%	0%	0%	0%
5541.713	1202.014	5541.729	1201.991	0%	0%	0%	0%	0%
5542.453	1200.592	5542.458	1200.598	95%	100%	100%	97.82%	100%
5542.440	1200.575	5542.445	1200.581	80%	100%	100%	100%	100%
5542.445	1200.581	5542.453	1200.592	85%	100%	100%	93.93%	100%
5542.42	1200.563	5542.428	1200.571	90%	59.45%	48%	55.4%	77%
5542.640	1205.959	5542.635	1205.954	0%	100%	100%	98.14%	100%
5542.393	1200.522	5542.398	1200.528	50%	0%	0%	0%	0%
5542.398	1200.528	5542.404	1200.537	70%	0%	0%	0%	0%

Table 3: Comparison with expert's estimation. First two columns are the coordinates of starting point, next two columns are the coordinates of the ending point. Next is the expert's prediction of coverage in percentage. The last four columns represent the ratio of number of *PRESENT* frames to Total frames in that region for different values of λ and GMM.

3.3. Region where algorithm's prediction and expert's estimation does not match

The bottom part of the table seem to have a complete mismatch of the prediction from the expert's estimation of the coverage of eelgrass. For further investigation, the video transects were inspected manually in the location/region of mismatch and some snapshots are shown in Figure 7. Figure 7(a) and (b) are two selected frames from the region, where the expert has estimated it as 0% coverage and our algorithm found it as a region of high percentage of *present* frames of eelgrass. In this case, both of the algorithms are correct. Figure 7(c) and (d) corresponds to the region where the expert estimated it as 50% coverage of eelgrass, whereas the algorithm found it to be devoid of any eelgrass (0%), which is also quite evident looking at Figure 7. Consulting with the domain expert, it was mutually concluded that the first case(Figure 7(a) and (b)) is not the correct estimation from the expert's end. It is possible for an expert to make few rare mistakes while labelling hundreds of hours of such videos. For the second case (Figure 7(c) and (d)), it is possible that there are some patches of eelgrass on the sides not captured on the video which is visible to the diver. The diver has more information about the environment, which is one of the reason why some prediction from the algorithm did not match the expert labels. It is possible, looking at Figure 7(d) that an eelgrass patch is covered with a layer of organic matter due to which the sharp edges of eelgrass is not detected through the camera. This is the only limitation as this method predicts the status based on the sharp edges it detect.

4. Discussion

4.1. Previous studies and necessity of automation

There exist many methods of collecting ground truth of eelgrass coverage. These include but not limited to videos, scuba diver, grab, aerial image, buoy and vegetation sampler.

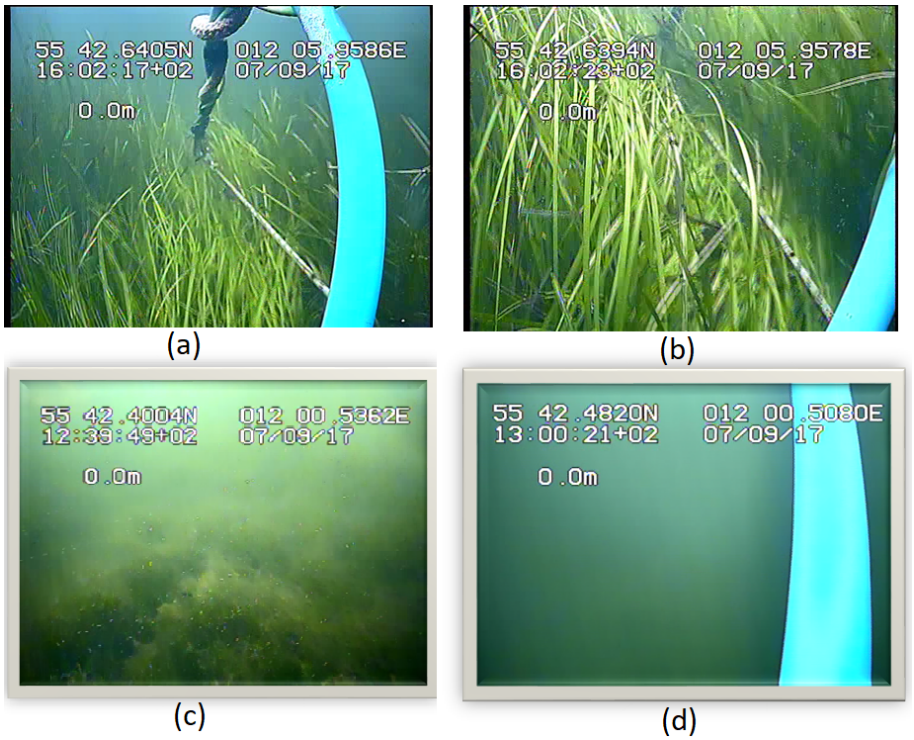

Figure 7: Snapshots of frames where expert's estimation of coverage mismatch with algorithm prediction.(a) and (b) are the snapshot of the regions from the first row of the bottom part of Table 3. (c) and (d) are the snapshot of regions from the second row of the bottom part of Table 3.

[23] showed that underwater video and scuba diver are the most popular choice of collecting ground truth. Many previous studies that has used underwater videos and scuba diver for ground truth estimation of seagrass has been done manually by a domain expert. We did not find many papers where the ground truth estimation process has been automated except [47] and [29]. Seagrass can be broadly classified into 5 families with each one of them with their own species. Some of the popular species of seagrass that has been studied extensively are *P.oceania*, and *Z.marina*(eelgrass). [48] has used remotely operated vehicle(ROV) as ground truth to map *Z.marina*. [49],[50] has used scuba as ground truth for *Z.marina*. [51], [52], [53], [54], [55], [56], [57], [58], has used video transects as ground truth to map *Z.marina*. [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], has used either scuba or videos as ground truth to map *P.oceania*. [71], has used scuba as ground truth to map *H.verticillata*. [72], [73], [74], has used videos as ground truth to map *P.sinuosa* and *P.australis*. [75] has used video and ROV as ground truth to map *P.oceania*. [76], [77], has used scuba as ground truth to map *Z.caulescence*. [78] have used AUV videos as ground truth to map *A.antartica*. In all these different approaches of collecting and assessing ground truth, it still remain a manual task requiring a domain expert. Moreover, these studies have been carried out at different parts of the world with different domain experts and there are no information available about the comparison and evaluation between different domain expert's estimation. Thus there exists a need to efficiently carry out the ground truthing

process not only with respect to time, but also to provide a common framework which quantifies the idea of *presence* of eelgrass. This paper attempts to address both the above issues.

4.2. Prior assumptions

These methods have few assumptions. One of them is that there are no other kind of benthic vegetation in this area of study (Roskilde fjord, Denmark) apart from eelgrass, after confirming with the marine biologists. So any lines detected from the algorithm is assumed to be a proxy for eelgrass disregarding the artefacts. The path taken by the diver to record the transects will not affect the precision of the algorithm's output. In this case, the diver started off from the shore in a straight path until he encountered a dense patch of eelgrass. Then he moved along the boundary of the eelgrass patch due to which we encountered many video frames where there is a dense region of eelgrass covering half of a given video frame and no eelgrass in the other half of the same video frame. The threshold of the algorithm was tuned to label these regions as eelgrass present. It is important for the recording device to be maintained at a distance of 1 to 2 meter above the seabed ideally for the current setting of parameters of the model to work in other regions. For distance more that 2 meters from the ground, the parameter estimates of the model needs to be recomputed along with a new decision boundary. If the recording device is held quite high from the seabed, then eelgrass will appear smaller and consequently less number of lines will be detected, which will affect the choice of the threshold. This methodology can be extended to detect any other vegetation (both marine and land) whose physical characteristics include sharp edges.

4.3. Interpretation of GMM decision boundary

We presented two methods to detect the status (*Present*/*Absent*) of eelgrass from under water video transects and extended this further to estimate the percentage coverage in a given area. The feature fusion strategy requires the threshold of *number of lines* be chosen by visual inspection. The probabilistic clustering approach use GMM to cluster the data into two regions, thereby bypassing any human intervention during the process. The threshold of the GMM is defined as the contour where the probability of a data point belonging to either clusters have a specific value which could be 50% each. This value can be changed to favour a certain cluster. For e.g, in a study where it is more important to assess dense eelgrass regions in a region of dominant empty patches, the threshold could be set by contour where probability of a data point belonging to *Present* cluster is 40% and *Absent* cluster is 60%. This gives more weight to cluster of *Present* by including more data points in this cluster due to lower threshold of the probability. This situation corresponds to the case where finding any speck/patch of eelgrass is absolutely necessary, even at the cost of labelling some non-eelgrass region as *Present*.

Setting this threshold for GMM is a way to add bias in the model as we prefer one cluster over another. While the GMM method is tractable in the sense that it offers a framework where a single value (the probability used to separate the clusters) need to be set, it is also a method where the validity of this probability is harder to justify. When setting a *Present*/*Absent* threshold, one may relate to the data and the likelihood that a specific

threshold results in a sensible definition of *Present*, through information (visual and other) from the data. The GMM model gives a different opportunity to verify the validity of the parameter setting by comparing with the threshold obtained from the other method. This can give insights on the meaning of *Present* (found from visual inspection) to a probabilistic notion. It also give a choice to the researchers according to the varied scope of their study. It is interesting to see from Figure 6, the threshold chosen by visual inspection for the feature (*number of lines*=260) is slightly higher than GMM decision boundary (probability of a data point belonging to either clusters are equal i.e. 50% each). From our experiment, we found that in order to match the coverage estimation with feature fusion method, the GMM decision boundary should be set as 95% probability of each frame to be in *Present* cluster and 5% of being in *Absent* cluster. This also illustrates our own inherent bias towards *Absent* cluster. While choosing the threshold of *number of lines*, a data point (video frame) will be labelled as *Present* only when it is quite certain, even at the cost of some doubtful *Present* frames being put into *Absent* cluster.

4.4. Application of Line segment detection in other domains

We used a Line detection algorithm to extract relevant features. Line based features are useful in many applications. There are different ways of detecting line features, the most popular being using the Hough transform. They have been particularly useful in precision agriculture [79], [80], [81],[82] in detecting crop lines and weed edges using the Hough Transform and its variations such as the Randomized Hough Transform. The Hough Transform is computationally heavy and slow and has seldom been used in real-time system [32]. The line segment detector used in this paper is relatively fast and can be used on a real time system. Our method of eelgrass detection will also work with forward facing cameras mounted on a UAV and ROV, if maintained at a decent height from the sea bottom.

4.5. Effect of Noisy artefacts

The performance of the algorithm is slightly dependent of the noise introduced during the recording (hand movements, wires and cables from the camera, e.t.c), which to some extent affected the precision of the method due to falsely detected lines from non-eelgrass objects. Although filtering eliminates noisy fluctuations to a great extent, still noisy objects will force the threshold of the features (*number of lines* and *total line length*) to be set slightly higher because more lines detected from artefacts on a *Absent* region will push the boundary so that algorithm can classify it as *Absent* region. We increased the threshold of the feature 1 (number of lines) from 240 initially (which would have been the optimal) to 260 to account for artefacts in some of the frames. It may explain estimation in one region where the feature fusion method has lower value of the percentage coverage. This problem can be solved if a video camera is towed/dragged by a boat without the need for a diver, which shall also be a less expensive alternative for data collection. The video transects used for this method were recorded on afternoon with decent lighting conditions. It is therefore hard to predict, how efficient this method might be if tested on videos recorded under artificial lights in a night time. Although we believe that if a domain expert can detect presence/absence from such

videos by distinguishing sharp edges of eelgrass from the background, then the algorithm would also perform well.

4.6. Limitations

This method will not be useful to distinguish between different vegetation covers, both having sharp edges. That would require tweaking the line segment detection algorithm to match the expected length of a single blade of grass and the expected orientation to distinguish between different sharp edged vegetation. A Gabor filter will be an ideal choice for feature extraction instead of line segment detection in this case as it captures different orientation much better. Although our methods are designed to predict *present/absent* status looking at each video frames, it can very well be extended to be used as a proxy for finding eelgrass coverage by computing the ratio of number of *present* frames to total number of frames in a given area shown in Table 3 which match well with the domain expert's estimation. This has a certain limitation in terms of determining the true density of the eelgrass patches. For example, consider two region in the estuary. The first region has medium dense eelgrass uniformly spread through the area with some patch of low dense region as well. The second region has extremely high density in one half and completely no eelgrass in the other half. Computing the ratio of number of *present* frames to total number of frames will lead to higher estimation of eelgrass in region 1 compared to region 2, as region 1 has more number of frames with *present* status even though the length of the eelgrass and the density of eelgrass patch is lower in region 1 than in region 2. The choice of threshold will determine how dense the patch of eelgrass should be to be labelled *present*.

Eelgrass cover estimation is a task that carries a certain subjectivity with it. Multiple experts may have different precise estimation due to different perception and bias. Dense regions and empty regions are easy to identify (by the expert as well as by the algorithm). The regions with less dense and thin patches of coverage are the hardest challenge. The feature fusion method not only gives flexibility to the user/scientist to adjust the threshold based on the circumstances of the experiment (the way transects are recorded), but also a common framework to quantify the bias through the value of threshold of the features combined with visual inspection.

Another limitation of this algorithm is that the whole prediction of eelgrass is based on what actually is visible in the video. It is like a small window looking in one direction in the vast 3-dimensional space. The scene not captured by the video camera is not taken into account while predicting the status of eelgrass presence. In that sense, a diver who actually recorded the video will see the surrounding environment in more details and has more information. This problem was encountered in our analysis where we found in some of the regions, our algorithm is predicted lower values of the coverage compared with the expert's estimates. The possible explanation for this could be that the diver see eelgrass in the sides which is not captured in the videos. One way to get around this problem is to deploy a set of 3-4 cameras in different directions (forward, left, right and down) and do this analysis using information obtained from all the videos.

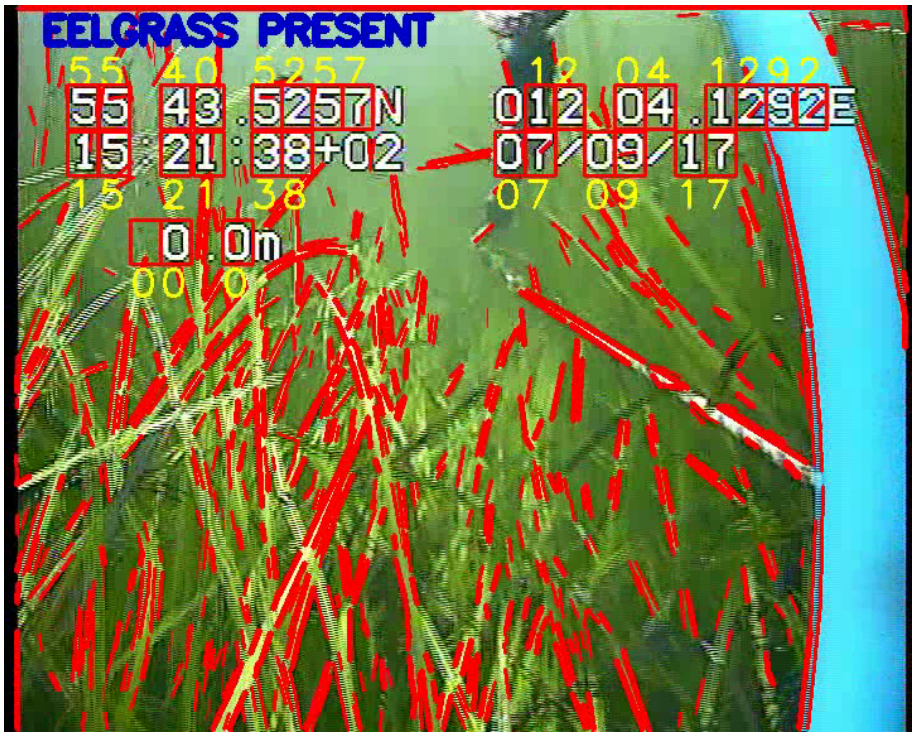

Figure 8: Short video showing the method detecting eelgrass in real time. Lot of lines can be seen to be detected due to the high density of eelgrass in the region. Embedded information are also extracted using a neural network.

5. Future Scope

The *present/absent* status obtained with this method can be used as a ground truth and extrapolated on unknown regions of satellite images where no video transects are available. On a satellite image, the transect location would appear as the few labelled pixels which would provide information to extrapolate or propagate the label information to other areas. Semi-supervised learning methods are useful tools to solve this problem and will be the main part of our future work.

Another interesting problem arises about pseudo labelling of video frames. There are concerns regarding the accuracy of labelled dataset from [29], which use coarse nature of polygons to label seagrass regions in a video frame. Ideally the labelled seagrass areas in a video frame should be represented with thin lines, which resembles the physical structure of the seagrass more. Therefore, it would be interesting to see if the output of line segment detection algorithm on raw video frames (Figure 3) could be used as a proxy labels to train classifiers to detect and estimate coverage of seagrass without the requirement of manual labelling of video frames. This also opens up many possibilities of using Transfer Learning methods to use the prior knowledge of the trained model from dataset of [29], retrain on our dataset and analyse if the prediction get any better than training a model from the scratch [83],[84]. This methods would allow more data-sets of seagrass to be used to fine tune models for better predictions in different scenarios. Ensemble methods [85] is another

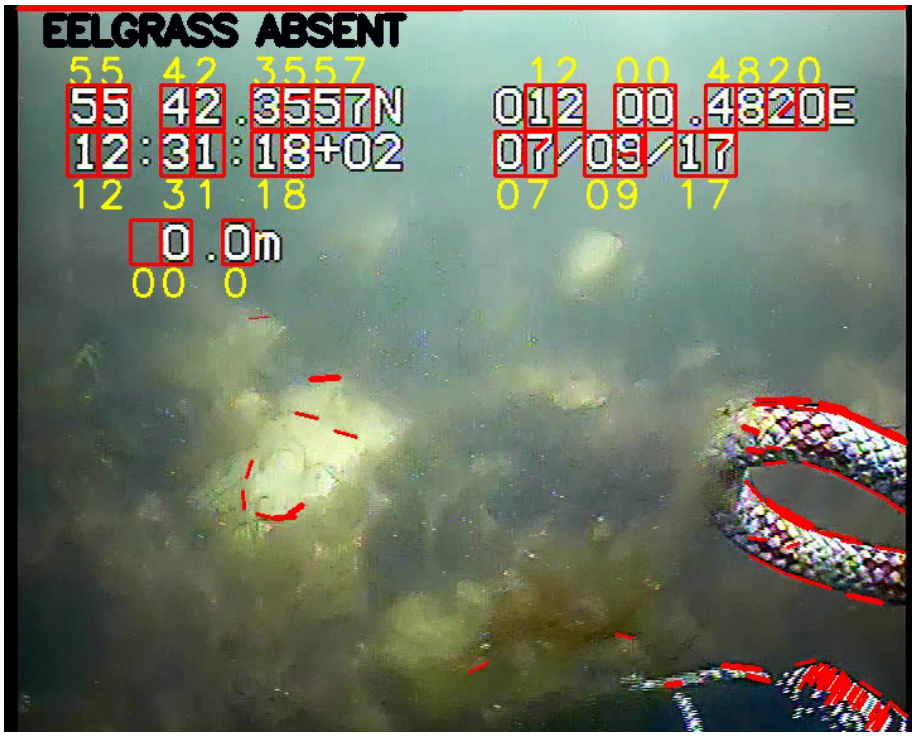

Figure 9: Short video showing the method working in real time in a no-eelgrass region. No significant lines are detected due to which the frames are classified as no-eelgrass locations. Embedded information are also extracted using a neural network.

interesting approach in which, instead of using the best classifier among many, it would be more robust to combine all the models in a intelligent way such that it performs better than the prediction of the best individual model.

Another major part of our future ongoing work is moving from presence/absence modelling towards robust quantification of seagrass. Instead of just using the frequency of *present* frames in a given area as a proxy for coverage estimation, we are looking into different probabilistic methods to get a better estimate of coverage which can overcome the limitation discussed before.

6. Conclusion

Here we presented two methods which detect eelgrass from underwater video transects. We showed how our methods are robust to noise introduced during the recording of the transects and how these methods can be extended nicely to be used as a proxy to find the percentage coverage in a given area which matches very well with a domain expert's estimation of coverage. A big advantage of our methods is that, it reduces the cost and complexity of conducting this kind of study anywhere as it does not require labelled video frames to predict eelgrass regions. It also saves lot of person-hours for an expert to visually inspect and assess many hours of videos to estimate the status of *presence/absence*. Eelgrass

detection is a subjective task that requires an expert judgement and the proposed methods reduces that by objective quantification of eelgrass presence through a threshold. Experts also make rare mistakes with this kind tedious task, which we also encountered during our analysis and we showed how our methods detected and rectified the errors from domain expert, which make these methods a robust tool free from human errors. Two demo videos showing the algorithm in action are available and accompanies the electronic version of this manuscript. Simply click on the 8 for a case where eelgrass is present. 9 corresponds to the case with no eelgrass.

7. Acknowledgement

We are thankful to Jacob Carstensen (Professor, Department of Bioscience, Aarhus University) and Karen Timmermann (Senior scientist, Department of Bioscience, Aarhus University) for their valuable input and ideas which shaped this paper into its current form. We are also thankful to Lars Boye Hansen (Head of Projects, DHI GRAS) and Mikkel Lydholm Rasmussen (Remote sensing specialist, DHI GRAS) for the insightful discussions about the data and Emil Guddal Larsen for telling us about the way diver records the status of the environment and the path taken in the transects.

8. Funding

This research is part of the SeaStatus Project (seastatus.dhigroup.com) which is funded by the Innovation Fund Denmark (IFD). The data was provided by DHI (Danish Hydraulics Institute) and the research was carried out at the Department of Applied Mathematics and Computer Science, Technical University of Denmark.

References

[1] J. W. Fourqurean, C. M. Duarte, H. Kennedy, N. Marbà, M. Holmer, M. A. Mateo, E. T. Apostolaki, G. A. Kendrick, D. Krause-Jensen, K. J. McGlathery, et al., Seagrass ecosystems as a globally significant carbon stock, Nature geoscience 5 (7) (2012) 505.

[2] J. Barrell, Does eelgrass (zostera marina) meet the criteria as an ecologically significant species?

[3] C.-F. Boudouresque, G. Bernard, P. Bonhomme, E. Charbonnel, G. Diviacco, A. Meinesz, G. Pergent, C. Pergent-Martini, S. Ruitton, L. Tunesi, Protection and conservation of posidonia oceanica meadows. (2012).

[4] C. J. Brown, S. J. Smith, P. Lawton, J. T. Anderson, Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques, Estuarine, Coastal and Shelf Science 92 (3) (2011) 502–520.

[5] B. Worm, E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, et al., Impacts of biodiversity loss on ocean ecosystem services, science 314 (5800) (2006) 787–790.

[6] B. S. Halpern, S. Walbridge, K. A. Selkoe, C. V. Kappel, F. Micheli, C. D'agrosa, J. F. Bruno, K. S. Casey, C. Ebert, H. E. Fox, et al., A global map of human impact on marine ecosystems, Science 319 (5865) (2008) 948–952.

[7] D. J. Wright, W. D. Heyman, Introduction to the special issue: marine and coastal gis for geomorphology, habitat mapping, and marine reserves, Marine Geodesy 31 (4) (2008) 223–230.

[8] J. Borum, C. M. Duarte, T. M. Greve, D. Krause-Jensen, European seagrasses: an introduction to monitoring and management, M & MS project, 2004.
[9] D. Krause-Jensen, T. M. Greve, K. Nielsen, Eelgrass as a bioindicator under the european water framework directive, Water Resources Management 19 (1) (2005) 63–75.
[10] J. Carstensen, D. Krause-Jensen, T. J. Balsby, Biomass-cover relationship for eelgrass meadows, Estuaries and coasts 39 (2) (2016) 440–450.
[11] T. J. Balsby, J. Carstensen, D. Krause-Jensen, Sources of uncertainty in estimation of eelgrass depth limits, Hydrobiologia 704 (1) (2013) 311–323.
[12] M. Frederiksen, D. Krause-Jensen, M. Holmer, J. S. Laursen, Spatial and temporal variation in eelgrass (zostera marina) landscapes: influence of physical setting, Aquatic Botany 78 (2) (2004) 147–165.
[13] B. M. Riegl, S. J. Purkis, Detection of shallow subtidal corals from ikonos satellite and qtc view (50, 200 khz) single-beam sonar data (arabian gulf; dubai, uae), Remote Sensing of Environment 95 (1) (2005) 96–114.
[14] D. Ierodiaconou, L. Laurenson, S. Burq, M. Reston, Marine benthic habitat mapping using multibeam data, georeferencedvideo and image classification techniques in victoria, australia, Journal of Spatial Science 52 (1) (2007) 93–104.
[15] D. A. Ryan, B. P. Brooke, L. B. Collins, G. A. Kendrick, K. J. Baxter, A. N. Bickers, P. J. Siwabessy, C. B. Pattiaratchi, The influence of geomorphology and sedimentary processes on shallow-water benthic habitat distribution: Esperance bay, western australia, Estuarine, Coastal and Shelf Science 72 (1-2) (2007) 379–386.
[16] J. Collier, S. Humber, Time-lapse side-scan sonar imaging of bleached coral reefs: A case study from the seychelles, Remote Sensing of Environment 108 (4) (2007) 339–356.
[17] A. W. Stevens, J. R. Lacy, D. P. Finlayson, G. Gelfenbaum, Evaluation of a single-beam sonar system to map seagrass at two sites in northern puget sound, washington, Tech. rep., Geological Survey (US) (2008).
[18] T. J. Malthus, E. Karpouzli, On the benefits of using both dual frequency side scan sonar and optical signatures for the discrimination of coral reef benthic communities, Advances in sonar technology (2009) 165–186.
[19] R. H. Chamberlain, P. H. Doering, B. Orlando, B. M. Sabol, Comparison of manual and hydroacoustic measurement of seagrass distribution in the caloosahatchee estuary, florida, Florida Scientist (2009) 386–405.
[20] H. Legrand, P. Lenfant, I. Sotheran, R. Foster-Smith, R. Galzin, J. Maréchal, Mapping marine benthic habitats in martinique (french west indies), Caribbean journal of science 46 (2–3) (2010) 267–283.
[21] A. Micallef, T. P. Le Bas, V. A. Huvenne, P. Blondel, V. Hühnerbach, A. Deidun, A multi-method approach for benthic habitat mapping of shallow coastal areas with high-resolution multibeam data, Continental Shelf Research 39 (2012) 14–26.
[22] E. Munday, B. Moore, J. Burczynski, Hydroacoustic mapping system for quantitative identification of aquatic macrophytes, substrate composition, and shallow water bathymetric surveying, in: 2013 OCEANS-San Diego, IEEE, 2013, pp. 1–3.
[23] M. U. Gumusay, T. Bakirman, I. Tuney Kizilkaya, N. O. Aykut, A review of seagrass detection, mapping and monitoring applications using acoustic systems, European Journal of Remote Sensing 52 (1) (2019) 1–29.
[24] Y. Gonzalez-Cid, A. Burguera, F. Bonin-Font, A. Matamoros, Machine learning and deep learning strategies to identify posidonia meadows in underwater images, in: OCEANS 2017-Aberdeen, IEEE, 2017, pp. 1–5.
[25] F. S. Rende, A. D. Irving, A. Lagudi, F. Bruno, S. Scalise, P. Cappa, M. Montefalcone, T. Bacci, M. Penna, B. Trabucco, et al., Pilot application of 3d underwater imaging techniques for mapping posidonia oceanica (l.) delile meadows, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 40 (5) (2015) 177.
[26] M. Massot-Campos, G. Oliver-Codina, L. Ruano-Amengual, M. Miró-Juliá, Texture analysis of seabed images: Quantifying the presence of posidonia oceanica at palma bay, in: 2013 MTS/IEEE OCEANS-

Bergen, IEEE, 2013, pp. 1–6.
[27] F. Bonin-Font, A. Burguera, J.-L. Lisani, Visual discrimination and large area mapping of posidonia oceanica using a lightweight auv, Ieee Access 5 (2017) 24479–24494.
[28] A. Burguera, F. Bonin-Font, J. L. Lisani, A. B. Petro, G. Oliver, Towards automatic visual sea grass detection in underwater areas of ecological interest, in: 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, 2016, pp. 1–4.
[29] G. Reus, T. Möller, J. Jäger, S. T. Schultz, C. Kruschel, J. Hasenauer, V. Wolff, K. Fricke-Neuderth, Looking for seagrass: Deep learning for visual coverage estimation, in: 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), IEEE, 2018, pp. 1–6.
[30] F. Weidmann, J. Jäger, G. Reus, S. T. Schultz, C. Kruschel, V. Wolff, K. Fricke-Neuderth, A closer look at seagrass meadows: Semantic segmentation for visual coverage estimation, in: OCEANS 2019-Marseille, IEEE, 2019, pp. 1–6.
[31] M. Martin-Abadal, E. Guerrero-Font, F. Bonin-Font, Y. Gonzalez-Cid, Deep semantic segmentation in an auv for online posidonia oceanica meadows identification, IEEE Access 6 (2018) 60956–60967.
[32] R. Ji, L. Qi, Crop-row detection algorithm based on random hough transformation, Mathematical and Computer Modelling 54 (3-4) (2011) 1016–1020.
[33] B. Chaudhuri, U. Pal, A complete printed bangla ocr system, Pattern recognition 31 (5) (1998) 531–549.
[34] N. Miura, A. Nagasaka, T. Miyatake, Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification, Machine vision and applications 15 (4) (2004) 194–203.
[35] M. R. Flindt, L. Kamp-Nielsen, J. C. Marques, M. A. Pardal, M. Bocci, G. Bendoricchio, J. Salomonsen, S. N. Nielsen, S. E. Jørgensen, Description of the three shallow estuaries: Mondego river (portugal), roskilde fjord (denmark) and the lagoon of venice (italy), Ecological Modelling 102 (1) (1997) 17–31.
[36] B. Rasmussen, A. Josefson, Consistent estimates for the residence time of micro-tidal estuaries, Estuarine, Coastal and Shelf Science 54 (1) (2002) 65–73.
[37] NERI, National Monitoring and Assesment Programme for the Aquatic and Terrestrial Environment nov, `https://www2.dmu.dk/1_viden/2_Publikationer/3_fagrapporter/rapporter/FR537.PDF`, accessed: 2019-10-10.
[38] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, G. Randall, Lsd: a line segment detector, Image Processing On Line 2 (2012) 35–55.
[39] A. Savitzky, M. J. Golay, Smoothing and differentiation of data by simplified least squares procedures., Analytical chemistry 36 (8) (1964) 1627–1639.
[40] G. J. McLachlan, K. E. Basford, Mixture models: Inference and applications to clustering, Vol. 84, M. Dekker New York, 1988.
[41] Z. Zivkovic, et al., Improved adaptive gaussian mixture model for background subtraction., in: ICPR (2), Citeseer, 2004, pp. 28–31.
[42] D. A. Reynolds, T. F. Quatieri, R. B. Dunn, Speaker verification using adapted gaussian mixture models, Digital signal processing 10 (1-3) (2000) 19–41.
[43] D. Reynolds, Gaussian mixture models, Encyclopedia of biometrics (2015) 827–832.
[44] P. A. Torres-Carrasquillo, D. A. Reynolds, J. R. Deller, Language identification using gaussian mixture model tokenization, in: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, IEEE, 2002, pp. I–757.
[45] T. L. Bailey, C. Elkan, et al., Fitting a mixture model by expectation maximization to discover motifs in bipolymers.
[46] A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the em algorithm, Journal of the Royal Statistical Society: Series B (Methodological) 39 (1) (1977) 1–22.
[47] M. Moniruzzaman, S. M. S. Islam, P. Lavery, M. Bennamoun, Faster r-cnn based deep learning for seagrass detection from underwater digital images, in: 2019 Digital Image Computing: Techniques and Applications (DICTA), IEEE, 2019, pp. 1–7.
[48] T. Domico, P. S. BioSurvey, Assessment of digital sonar technology to map eelgrass (zostera marina) in the san juan islands, Puget Sound BioSurvey (2001) 1–10.

[49] J. D. Warren, B. J. Peterson, Use of a 600-khz acoustic doppler current profiler to measure estuarine bottom type, relative abundance of submerged aquatic vegetation, and eelgrass canopy height, Estuarine, Coastal and Shelf Science 72 (1-2) (2007) 53–62.
[50] E. M. McCarthy, B. Sabol, Acoustic characterization of submerged aquatic vegetation: military and environmental monitoring applications, in: OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158), Vol. 3, IEEE, 2000, pp. 1957–1961.
[51] L. Reshitnyk, M. Costa, C. Robinson, P. Dearden, Evaluation of worldview-2 and acoustic remote sensing for mapping benthic habitats in temperate coastal pacific waters, Remote Sensing of Environment 153 (2014) 7–23.
[52] H. Vandermeulen, Bay-scale assessment of eelgrass beds using sidescan and video, Helgoland marine research 68 (4) (2014) 559.
[53] J. Barrell, J. Grant, Detecting hot and cold spots in a seagrass landscape using local indicators of spatial association, Landscape ecology 28 (10) (2013) 2005–2018.
[54] J. P. Barrell, Quantification and spatial analysis of seagrass landscape structure through the application of aerial and acoustic remote sensing, Ph.D. thesis (2016).
[55] M. Paul, A. Lefebvre, E. Manca, C. L. Amos, An acoustic method for the remote measurement of seagrass metrics, Estuarine, Coastal and Shelf Science 93 (1) (2011) 68–79.
[56] H. Van Rein, C. Brown, R. Quinn, J. Breen, D. Schoeman, An evaluation of acoustic seabed classification techniques for marine biotope monitoring over broad-scales (¿ 1 km2) and meso-scales (10 m2–1 km2), Estuarine, Coastal and Shelf Science 93 (4) (2011) 336–349.
[57] K. Collins, A. Suonpää, J. Mallinson, The impacts of anchoring and mooring in seagrass, studland bay, dorset, uk, Underwater Technology 29 (3) (2010) 117–123.
[58] A. Lefebvre, C. E. Thompson, K. Collins, C. L. Amos, Use of a high-resolution profiling sonar and a towed video camera to map a zostera marina bed, solent, uk, Estuarine, Coastal and Shelf Science 82 (2) (2009) 323–334.
[59] V. Pasqualini, P. Clabaut, G. Pergent, L. Benyoussef, C. Pergent-Martini, Contribution of side scan sonar to the management of mediterranean littoral ecosystems, International Journal of Remote Sensing 21 (2) (2000) 367–378.
[60] L. PIAZZI, S. ACUNTO, F. Cinelli, Mapping of posidonia oceanica beds around elba island (western mediterranean) with integration of direct and indirect methods, Oceanologica acta 23 (3) (2000) 339–346.
[61] W. J. L. Long, W. J. L. Long, Preliminary evaluation of an acoustic technique for mapping tropical seagrass habitats, no. 52, Great Barrier Reef Marine Park Authority, 1998.
[62] V. Pasqualini, C. Pergent-Martini, P. Clabaut, G. Pergent, Mapping ofposidonia oceanicausing aerial photographs and side scan sonar: Application off the island of corsica (france), Estuarine, Coastal and Shelf Science 47 (3) (1998) 359–367.
[63] G. Ardizzone, A. Belluscio, L. Maiorano, Long-term change in the structure of a posidonia oceanica landscape and its reference for a monitoring plan, Marine Ecology 27 (4) (2006) 299–309.
[64] A. Leriche, V. Pasqualini, C.-F. Boudouresque, G. Bernard, P. Bonhomme, P. Clabaut, J. Denis, Spatial, temporal and structural variations of a posidonia oceanica seagrass meadow facing human activities, Aquatic Botany 84 (4) (2006) 287–293.
[65] M. Montefalcone, G. Albertelli, C. Nike Bianchi, M. Mariani, C. Morri, A new synthetic index and a protocol for monitoring the status of posidonia oceanica meadows: a case study at sanremo (ligurian sea, nw mediterranean), Aquatic Conservation: Marine and Freshwater Ecosystems 16 (1) (2006) 29–42.
[66] N. Sánchez-Carnero, D. Rodríguez-Pérez, E. Couñago, S. Aceña, J. Freire, Using vertical sidescan sonar as a tool for seagrass cartography, Estuarine, Coastal and Shelf Science 115 (2012) 334–344.
[67] G. Di Maida, A. Tomasello, F. Luzzu, A. Scannavino, M. Pirrotta, C. Orestano, S. Calvo, Discriminating between posidonia oceanica meadows and sand substratum using multibeam sonar, ICES Journal of Marine Science 68 (1) (2010) 12–19.
[68] P. Descamp, F. Holon, L. Ballesta, A. Guilbert, M. Guillot, P. Boissery, V. Raimondino, J. Deter, Fast and easy method for seagrass monitoring: Application of acoustic telemetry to precision mapping of

posidonia oceanica beds, Marine pollution bulletin 62 (2) (2011) 284–292.
[69] G. De Falco, R. Tonielli, G. Di Martino, S. Innangi, S. Simeone, I. M. Parnum, Relationships between multibeam backscatter, sediment grain size and posidonia oceanica seagrass distribution, Continental Shelf Research 30 (18) (2010) 1941–1950.
[70] R. Ferretti, M. Bibuli, M. Caccia, D. Chiarella, A. Odetti, A. Ranieri, E. Zereik, G. Bruzzone, Machine learning methods for acoustic-based automatic posidonia meadows detection by means of unmanned marine vehicles, in: OCEANS 2017-Aberdeen, IEEE, 2017, pp. 1–6.
[71] M. J. Maceina, J. V. Shireman, K. A. Langeland, D. Canfield Jr, Prediction of submersed plant biomass by use of a recording fathometer., Journal of Aquatic Plant Management 22 (1984) 35–38.
[72] G. Kendrick, E. Harvey, J. McDonald, C. Pattiaratchi, M. Cappo, J. Fromont, M. Shortis, S. Grove, A. Bickers, K. Baxter, et al., Characterising the fish habitats of the recherche archipelago, Fisheries Research and Development Corporation report project (2001/060).
[73] Y.-T. Tseng, Recognition and assessment of seafloor vegetation using a single beam echosounder, Ph.D. thesis, Curtin University (2009).
[74] A. Jordan, M. Lawler, V. Halley, N. Barrett, Seabed habitat mapping in the kent group of islands and its role in marine protected area planning, Aquatic Conservation: Marine and Freshwater Ecosystems 15 (1) (2005) 51–70.
[75] P. Colantoni, P. Gallignani, E. Fresi, F. Cinelli, Patterns of posidonia oceanica (l.) delile beds around the island of ischia (gulf of naples) and in adjacent waters, Marine Ecology 3 (1) (1982) 53–74.
[76] T. Komatsu, C. Igarashi, K. Tatsukawa, S. Sultana, Y. Matsuoka, S. Harada, Use of multi-beam sonar to map seagrass beds in otsuchi bay on the sanriku coast of japan, Aquatic Living Resources 16 (3) (2003) 223–230.
[77] T. Sagawa, A. Mikami, T. Komatsu, N. Kosaka, A. Kosako, S. Miyazaki, M. Takahashi, Mapping seagrass beds using ikonos satellite image and side scan sonar measurements: a japanese case study, International Journal of Remote Sensing 29 (1) (2008) 281–291.
[78] D. Ierodiaconou, A. C. Schimel, D. Kennedy, J. Monk, G. Gaylard, M. Young, M. Diesing, A. Rattray, Combining pixel and object based image analysis of ultra-high resolution multibeam bathymetry and backscatter for habitat mapping in shallow marine waters, Marine Geophysical Research 39 (1-2) (2018) 271–288.
[79] K. Ramesh, N. Chandrika, S. Omkar, M. Meenavathi, V. Rekha, Detection of rows in agricultural crop images acquired by remote sensing from a uav, International Journal of Image, Graphics and Signal Processing 8 (11) (2016) 25.
[80] M. Basso, E. P. de Freitas, A uav guidance system using crop row detection and line follower algorithms, Journal of Intelligent & Robotic Systems (2019) 1–17.
[81] W. Winterhalter, F. V. Fleckenstein, C. Dornhege, W. Burgard, Crop row detection on tiny plants with the pattern hough transform, IEEE Robotics and Automation Letters 3 (4) (2018) 3394–3401.
[82] M. Li, M. Zhang, H. Huan, G. Liu, A new navigation line extraction method for agriculture implements guidance system, in: International Conference on Computer and Computing Technologies in Agriculture, Springer, 2012, pp. 299–308.
[83] J. Wohlert, A. M. Munk, S. Sengupta, F. Laumann, Bayesian transfer learning for deep networks, 2018.
[84] S. J. Pan, Q. Yang, A survey on transfer learning, IEEE Transactions on knowledge and data engineering 22 (10) (2009) 1345–1359.
[85] S. Sengupta, S. Sanyal, Multi-hypothesis classifier (2019). arXiv:1908.07857.