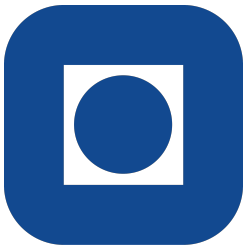

NTNU

TTK4255

Robotvision

Hyperspectral imaging

Mads Formo

April 2, 2020

Contents

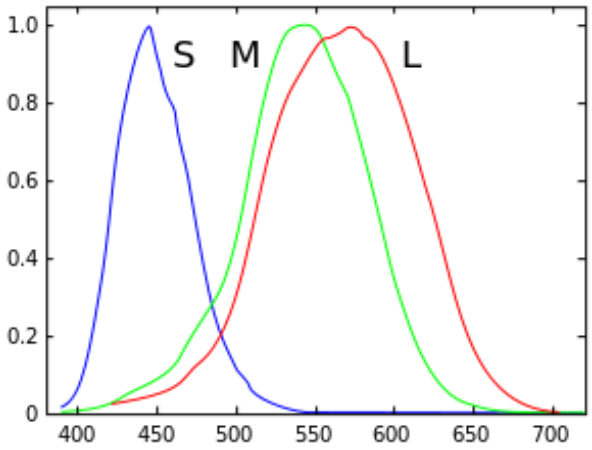

Figure 1: Graph for the human color sensitivity curves, according to Wikipedia [1]

1 Getting familiar with the data

1.1 Finding the spectral resolution

To find the spectral resoulution of the dataset, we load the *hico_wl* array, which contains the wavelength corresponding to band i. We loop through the array and compare each wavelength i with the the previous wavelenght *i-1* and we find that the average distance between the wavelengths is $5.728nm$, which seems to be constant between all wavelengths.

1.2 Relation to human color perception

The color sensitivity of the human eye is shown in fig. 1. As we can see, blue color has a peak around $450nm$ (S-curve), green peaks at $550nm$ (M-curve), and red at $600nm$ (L-curve).

1.3 Create a pseudo RGB image from the hyperspectral bands

From the *hico_wl* array, we find that Blue (450nm) is located at index $i = 8$, green (550nm) at $i = 25$, and finally red (600nm) at $i = 34$. We combine these indices from the HICO dataset and show it as an image to create a pseudo RGB image, shown in fig. 2.

1.4 Representative spectra for selected points

We want to look at the representative spectra of the points (20,20), (100,70) and (400,30), which is in deep water, shallow water and vegetation respectively. As we can see in fig. 3, we see that there is a clear difference in the spectra between water and vegetation. Both have amplitude peaks at the lower end of the spectra and then drop off in power as the wavelength increases. Vegetation however increases again in power at a wavelength of around 700nm, while the water is still decreasing. The findings here seem to agree to the findings of Lucke et al [2] as the general shape of the curves matches those of figure 12 in that report.

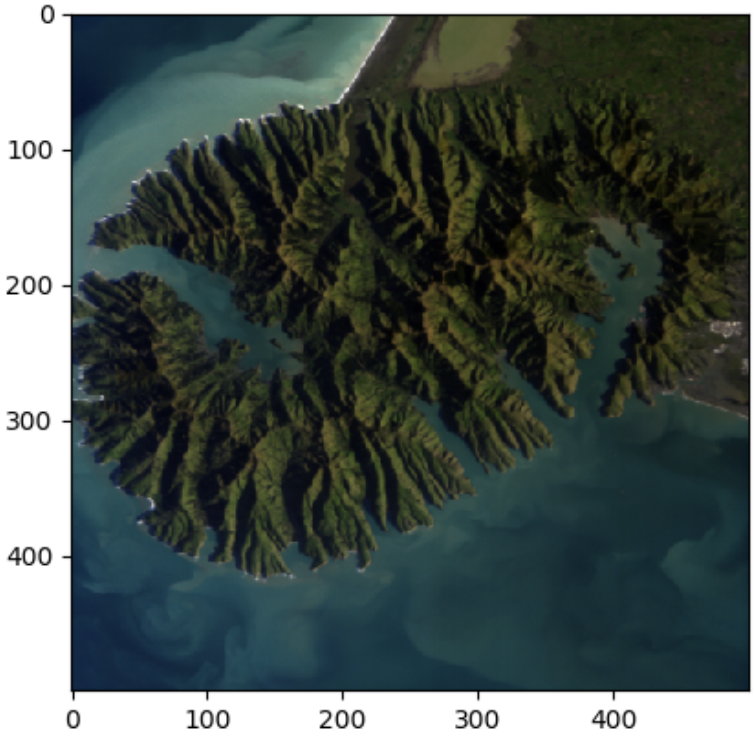

Figure 2: Pseudo RGB image, showing R (600nm), G (550nm), B (450nm)

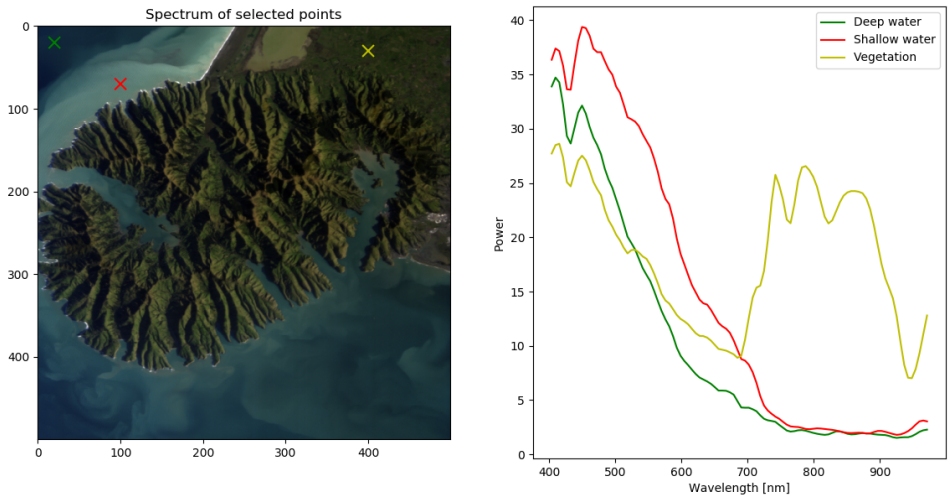

Figure 3: Representative spectra of specific points

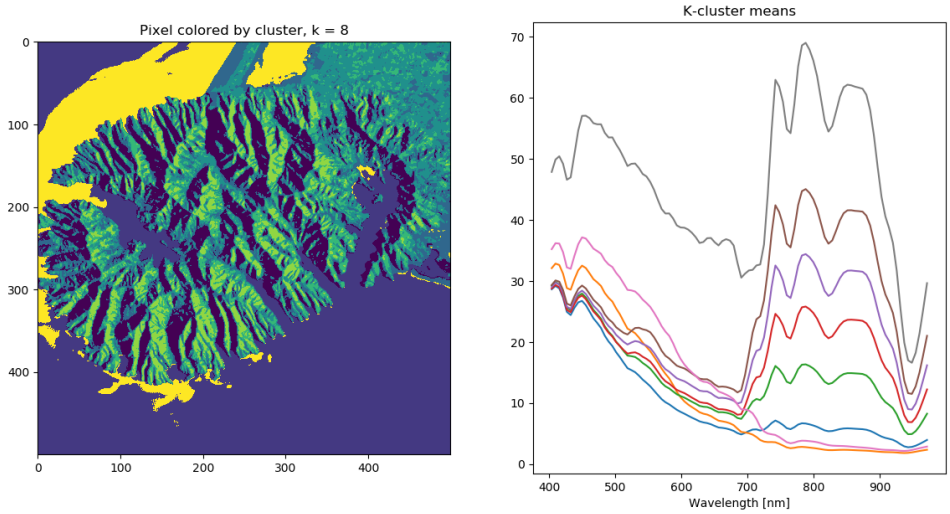

Figure 4: K-mean clusters of the image

2 Classification & Bio-geophysical Parameter Retrieval

2.1 Can we predict where there is chlorophyll through classification?

We will use *K-Means clustering* to classify the data. K-means clustering is a unsupervised learning algorithm that can be used to classify and cluster data into k different clusters. The data points are adjusted iteratively until all points are associated with the nearest cluster. We want to cluster each observation (pixels, with n spectral channels) into a specific cluster (environment class, ie. deep water, shallow water, vegetation).

As we can see from fig. 3, we know that those three different points have distinctly different spectra, thus it should be possible to classify them accordingly. The results of a K-mean clustering, run with Spectral Python's *kmeans* function [3], can be seen in fig. 4. We clearly see different classes for water, land, and vegetation, the latter containing lots of chlorophyll. We also see a very distinct class along the coast on the upper part of the image. This may very well be a collection of chlorophyll, but it might also just be shallow water, or more likely a combination of both.

2.2 How well can we directly estimate the chlorophyll content?

We use the NASA OBPG algorithm, defined in equation 4 in the assignment [4], as well as the parameters given there, to try to visualize the chlorophyll contents. Using the closest available wavelengths in the dataset, $\lambda_{green} = 553$ (i = 26) and $\lambda_{blue} = [444, 490, 507]$ (i = [7, 15, 18]). The results can be seen in fig. 5. We can clearly see high concentrations on the north west coast (assuming north is at the top of the image), same place as in fig. 4, but now we also see quite a bit on the southern coast as well. Thus it seems that this algorithm performs better than the k-means clustering.

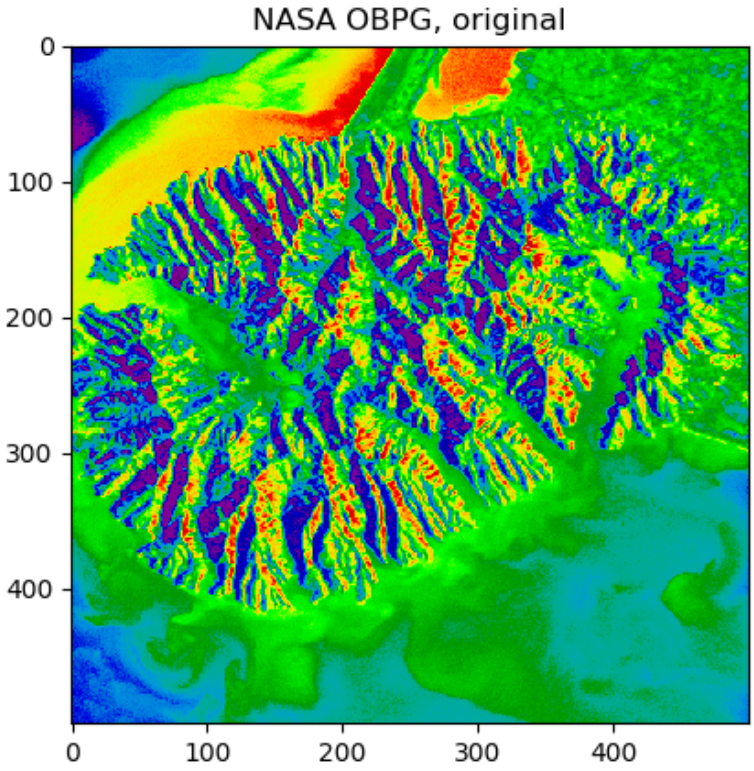

Figure 5: Results from the NASA OBPG algorithm, showing chlorophyll concentrations in the water

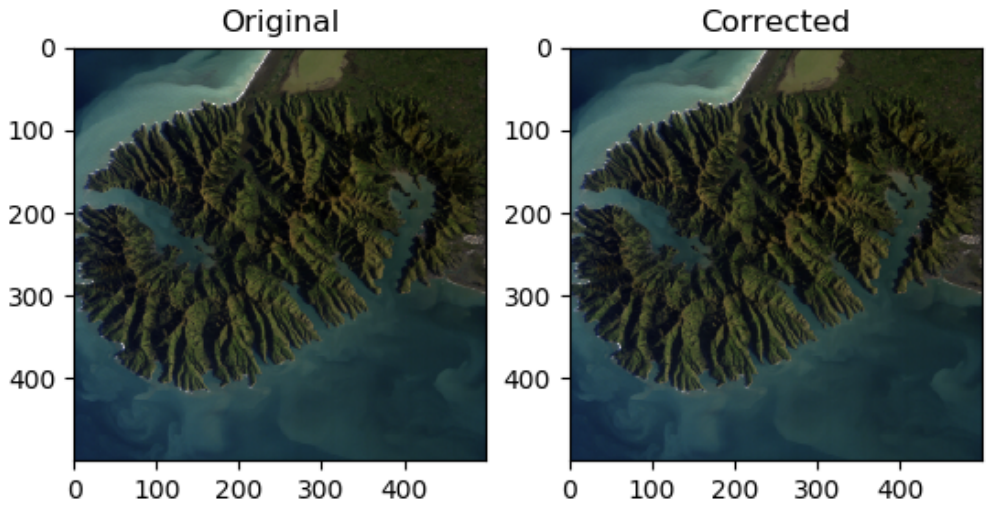

Figure 6: Pseudo RGB image of the atmosphere corrected image

2.3 How can we estimate the reflectance from the surface of the ocean?

The data in the HICO dataset actually contains the measurement of radiance exiting the top of the atmosphere, and not the radiance of the water directly. Therefore we must recover the radiance R_{rs} of the water from the top of atmosphere (TOA) measurements. This is done with the empirical line (ELM) method, as described in [4], on the form eq. (1). Where a and b are terms that model the absorption of light in the atmosphere, which is to be estimated, and L is the measured value in the HICO dataset.

$$R_{rs}(\lambda) = \frac{L(\lambda) - b(\lambda)}{a(\lambda)} \tag{1}$$

After performing the atmospheric correction, the resulting image is shown in fig. 6. It is very difficult to see a clear difference between the two by only looking at the image, but if we look at the pixel values for the red, green and blue channels separately, we see that the blue color channel is only about 2% of the original, uncorrected pixel value, while the red and blue channels are about 9-10% of their original values. Thus it would seem that the atmospheric correction removes some of the blue color of the image.

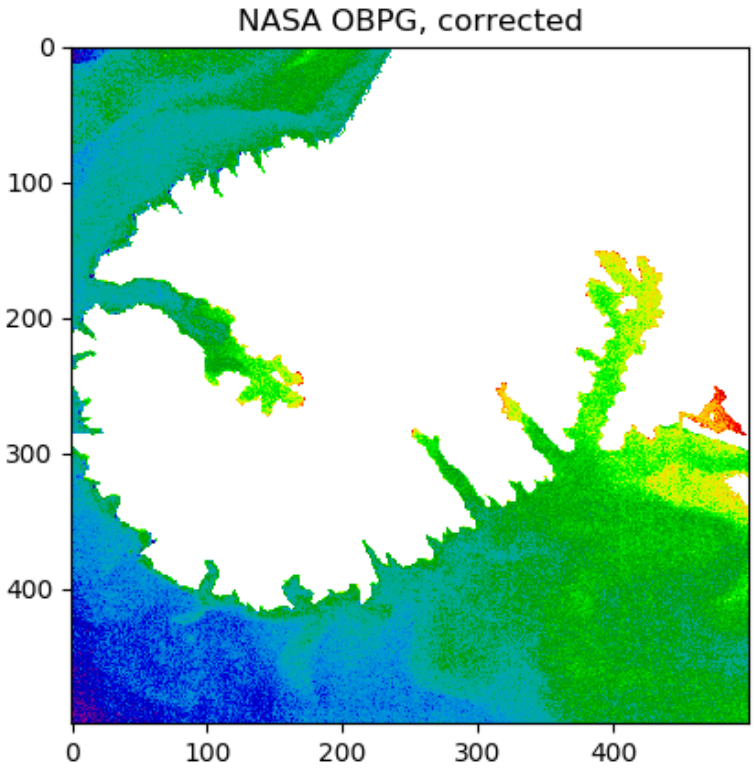

Figure 7: Results from the NASA OBPG algorithm using the atmospheric correction described in section 2.3, showing chlorophyll concentrations in the water

2.4 Compute chlorophyll concentration using atmosphere-corrected data

After performing the atmospheric correction, as described in section 2.3, we perform the NASA OBPG algorithm again on the corrected image cube. As we can see from the results in fig. 7, it is quite different from the original results from fig. 5. The strong concentration we found at the north west coast of the original (fig. 5), is no longer to be seen in the corrected image. Instead, we see a high concentration of chlorophyll on the south eastern coast, as well as a small patch in the bay on the western part of the landmass.

Plot figures before and after side-by-side?

2.5 Classify land versus water

We want to classify only the chlorophyll content of the water, we're not interested in looking at the land. Therefore we want to mask out the landmass and show only the water. We again use K-means clustering to classify the data into different classes. By performing it multiple times for varying numbers of classes and comparing the result from fig. 4, fig. 8, as well as all other iterations from 1-10 classes on both corrected and original data, we find that performing k-means clustering with 10 clusters on the atmospherically corrected image cube gives the best performance regarding classifying

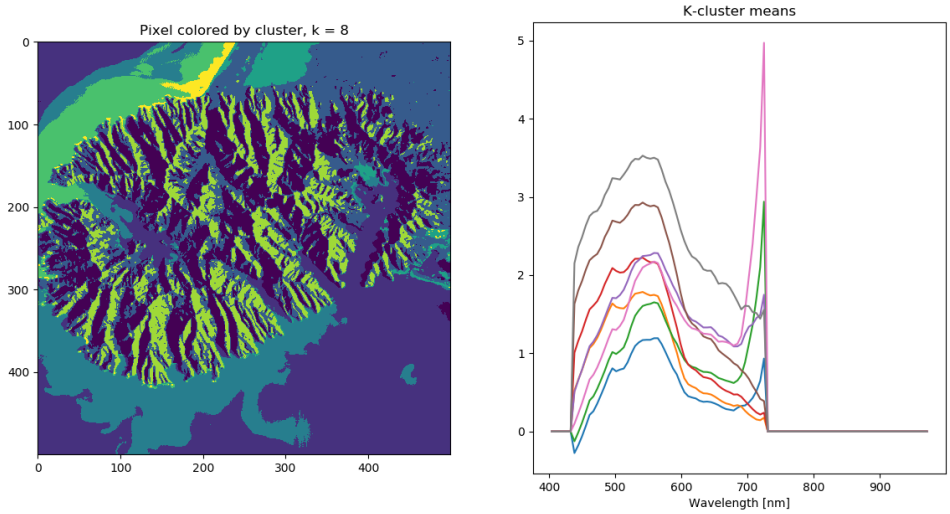

Figure 8: K-mean clustering performed on the atmospheric corrected image cube

water vs land. Setting the max number of iterations to 100, the algorithm converges after around 80 iterations, giving a good result.

After the k-means clustering is performed, we look at the spectral plot of the different classes. We know from section 1.4 roughly how the spectra of water looks compared to land/vegetation. We the pick the spectral lines from the k-means result that most resemble the spectral lines of water, and we remove any lines that resemble those of land/vegetation. We then end up with a handful of different classes that represents water, and the rest is land. We set the pixel values of all points inside the water classes to 1, and all other pixels to 0 to create the mask. The mask itself is shown in fig. 9. When put over the image, we get fig. 10.

2.6 Other bio-geophysical parameters

In the available spectral measurements between 400-1000nm we may find several other kinds of bio-geophysical parameters. In general, bio-geophysical parameters include biological (plant species, interactions in the ecology, biotic productivity), geological (soil types, erosion), and physical (light, heat) [5]. Examples include, phytoplankton pigments and gelbstoff [6], or land degradation processes in semi-arid geographical areas [7].

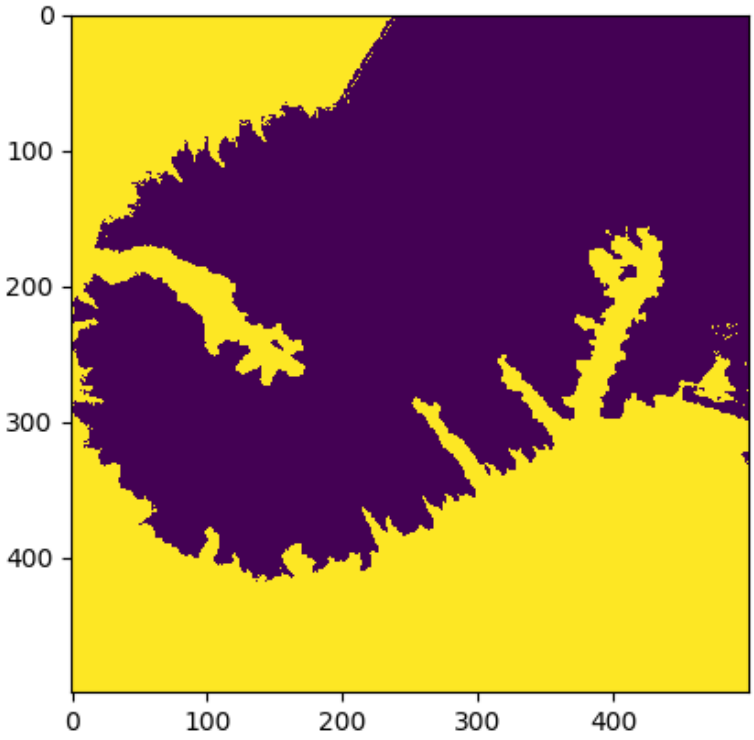

Figure 9: The mask found in section 2.5

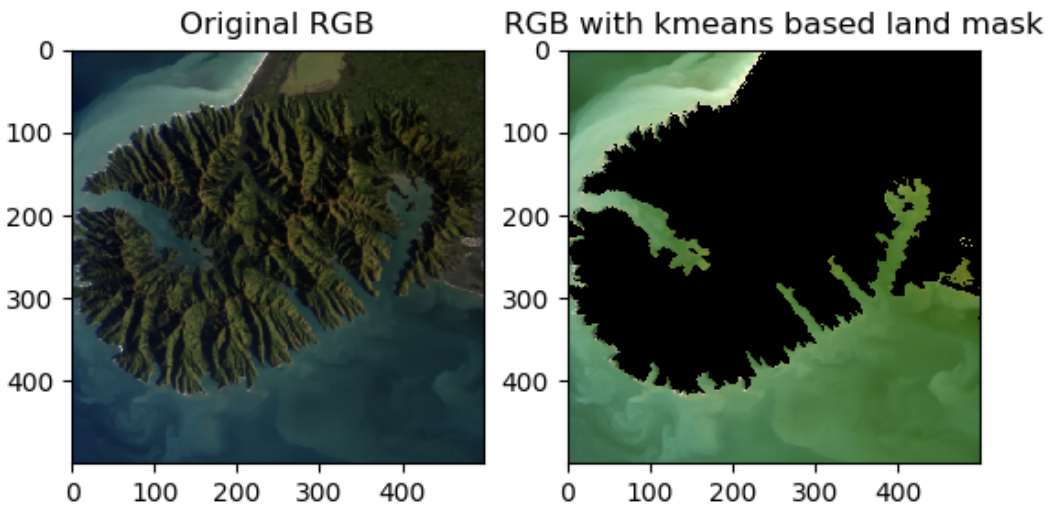

Figure 10: Masking away the land area

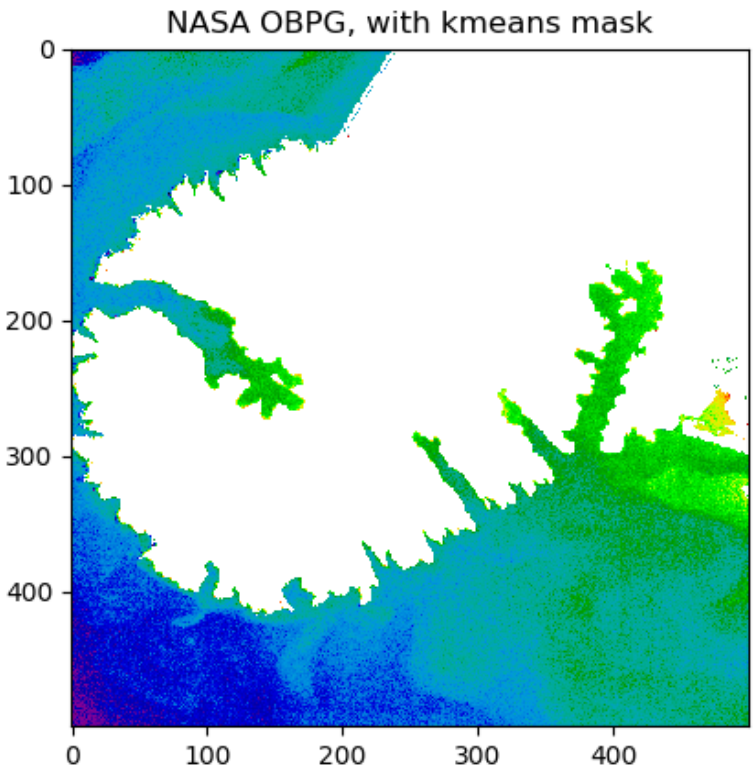

Figure 11: Running the OBPG algorithm on the masked image

References

[1] Wikipedia. *Spectral sensitivity*. Jan. 2020. URL: https://en.wikipedia.org/wiki/Spectral_sensitivity.

[2] Robert L. Lucke et al. "Hyperspectral Imager for the Coastal Ocean: instrument description and first images". In: *Appl. Opt.* 50.11 (Apr. 2011), pp. 1501–1516. DOI: 10.1364/AO.50.001501. URL: http://ao.osa.org/abstract.cfm?URI=ao-50-11-1501.

[3] Thomas Boggs. *SpectralPython*. 2014. URL: https://www.spectralpython.net.

[4] Sivert Bakken, Joe Garret, and Simen Haugo. "Hyper Spectral Imaging Project". In: *TTK4255 Robotic Vision, NTNU* (Feb. 2020). URL: https://ntnu.blackboard.com/bbcswebdav/pid-881058-dt-content-rid-25343529_1/xid-25343529_1.

[5] ESA Sentinel Online. *Bio-geophysical Variable Mapping*. URL: https://sentinel.esa.int/web/sentinel/thematic-areas/land-monitoring/bio-geophysical-variable-mapping.

[6] ZhongPing Lee, Kendall L. Carder, and Robert A. Arnone. "Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters". In: *Appl. Opt.* 41.27 (Sept. 2002), pp. 5755–5772. DOI: 10.1364/AO.41.005755. URL: http://ao.osa.org/abstract.cfm?URI=ao-41-27-5755.

[7] H. Kaufmann et al. "SAND - a hyperspectral sensor for the analysis of dryland degradation". In: *IEEE International Geoscience and Remote Sensing Symposium*. Vol. 2. 2002, 986–988 vol.2.