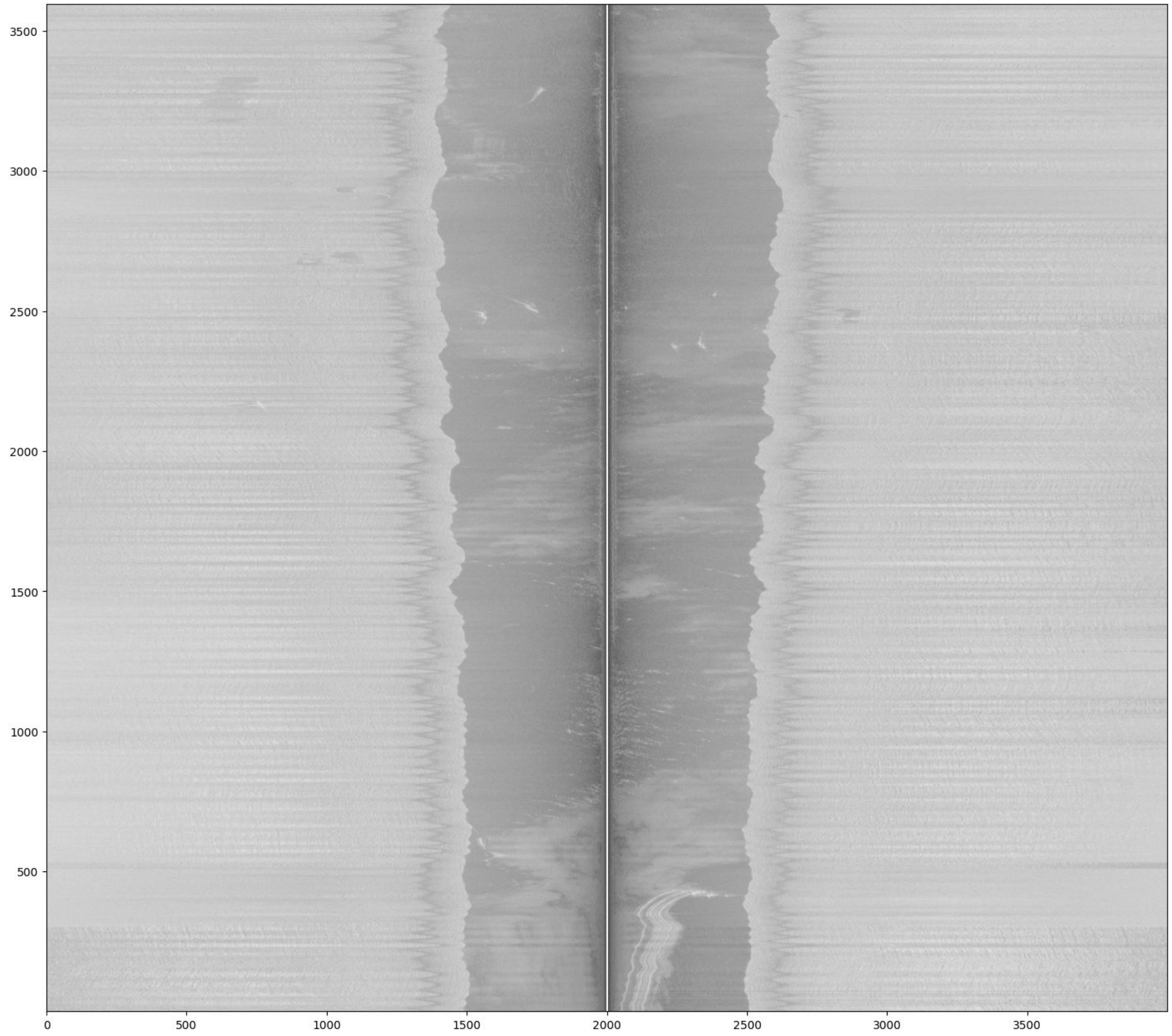
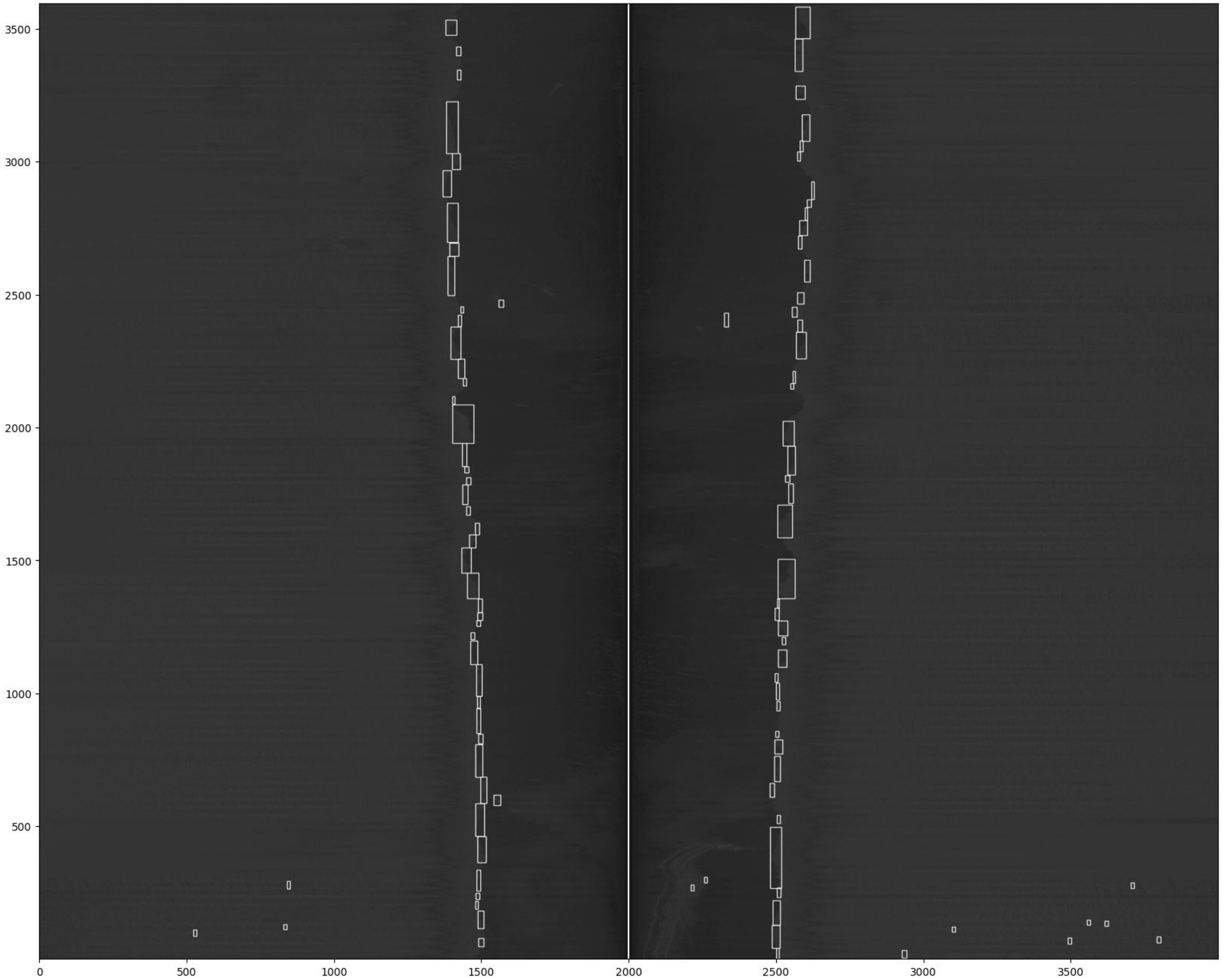
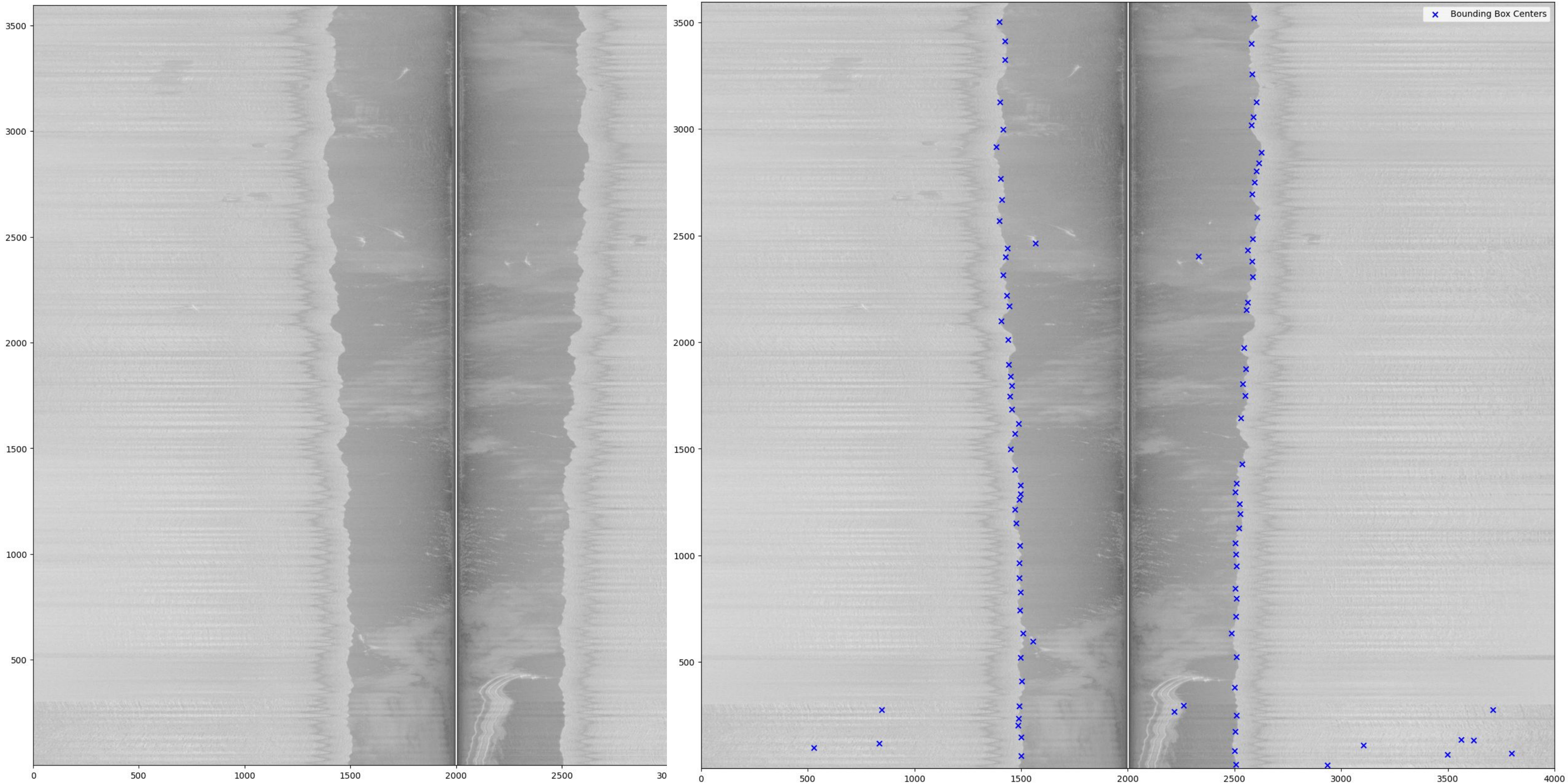
Scripps Summer Data

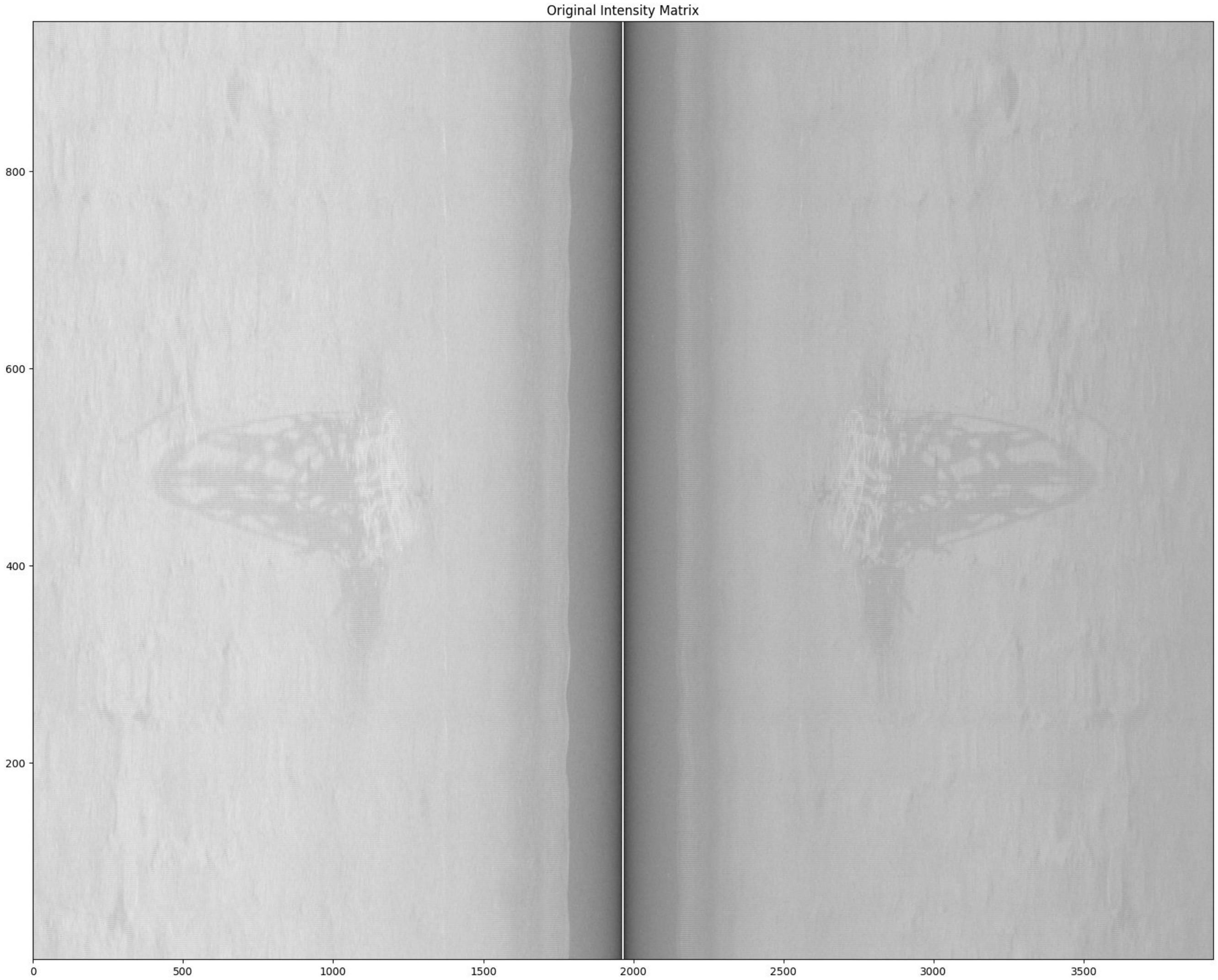


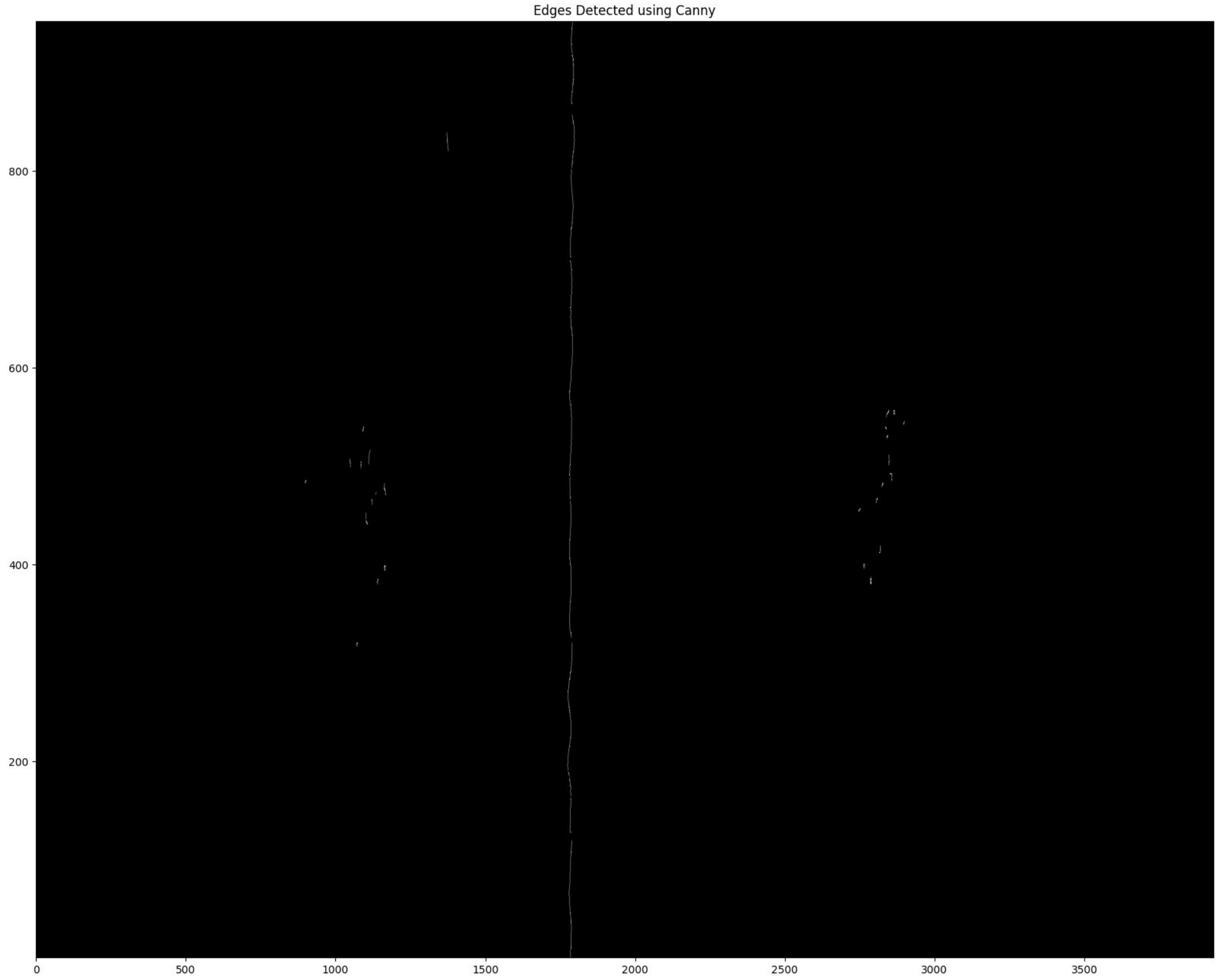

3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500

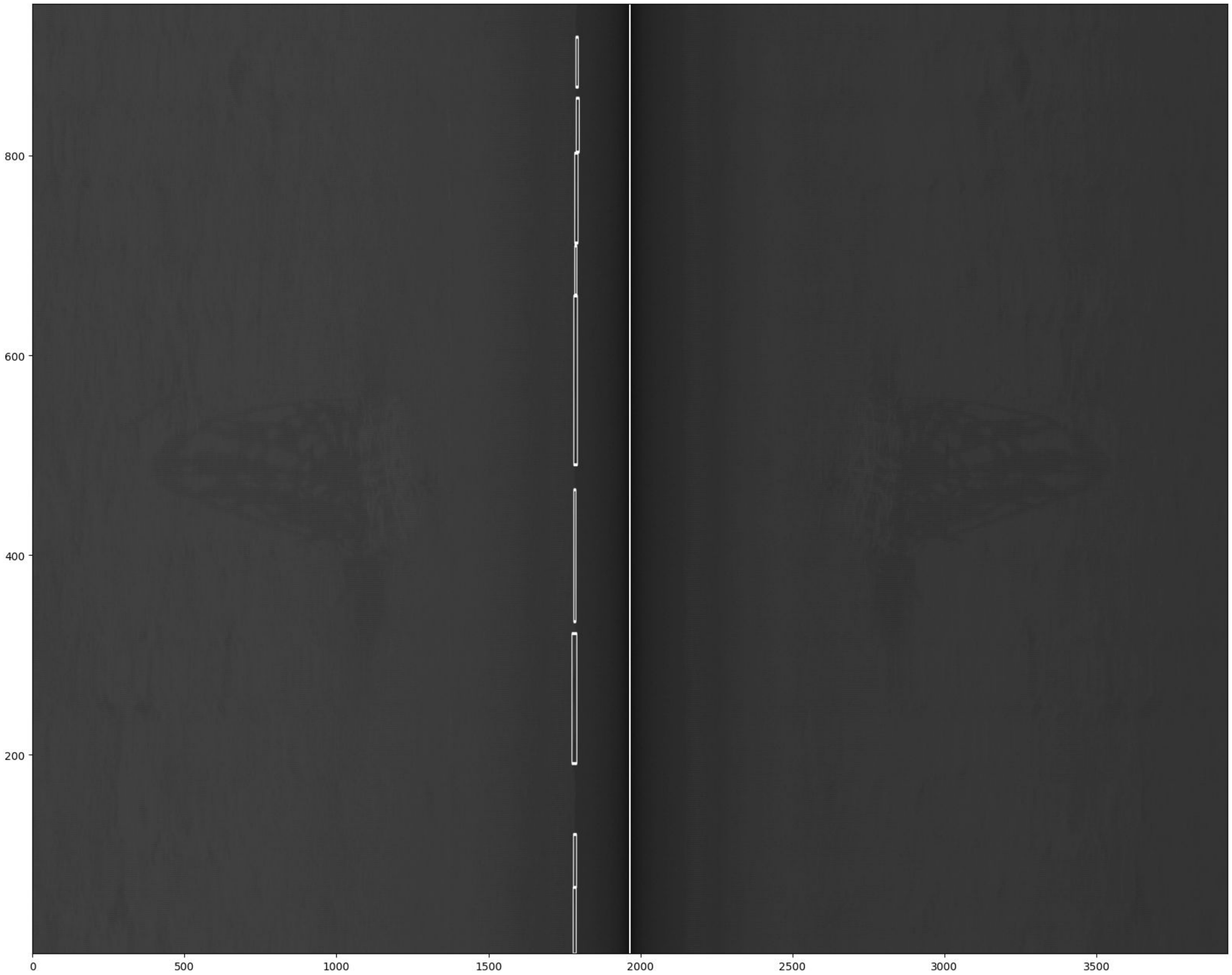


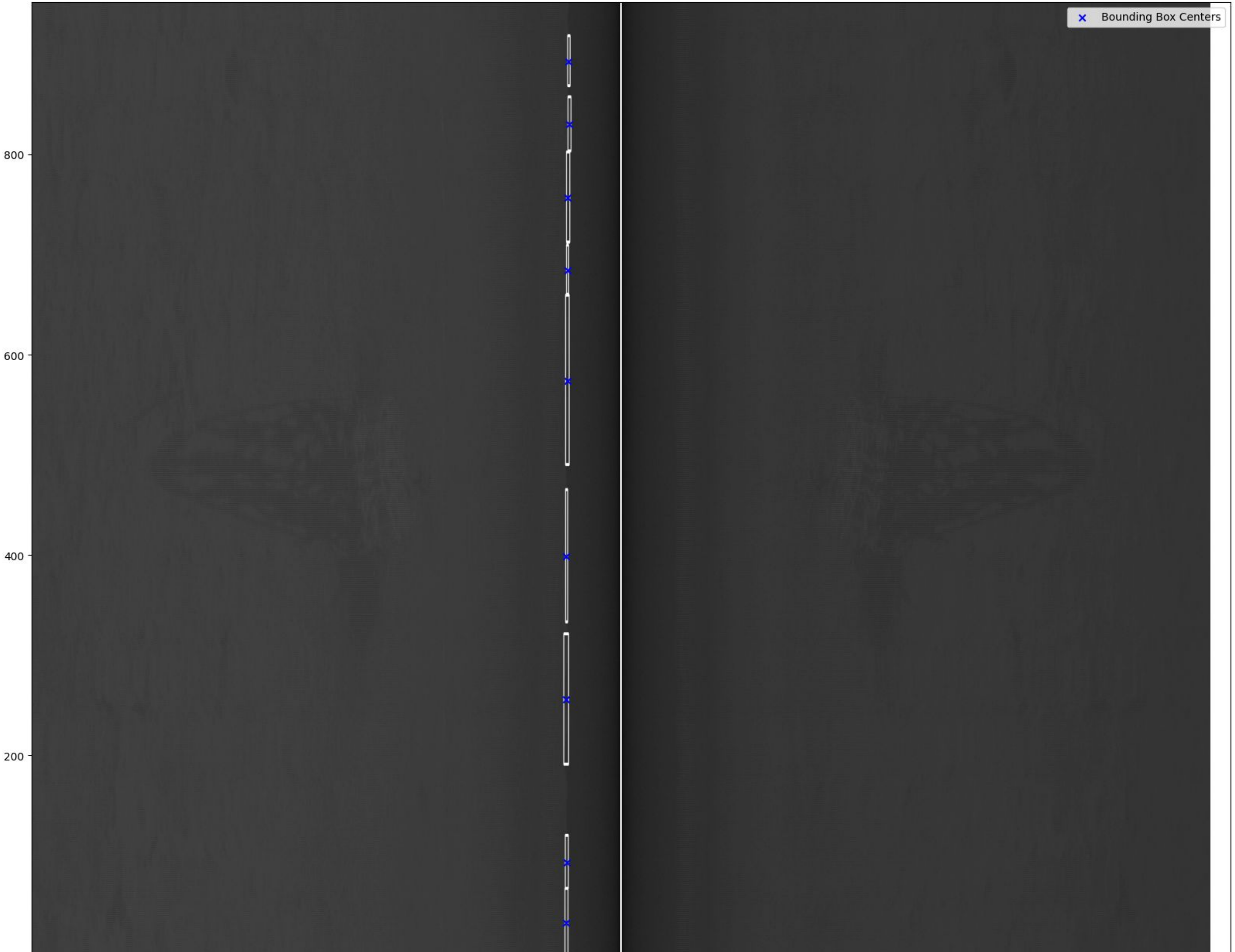
Bounding Box Centers
3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500
4000

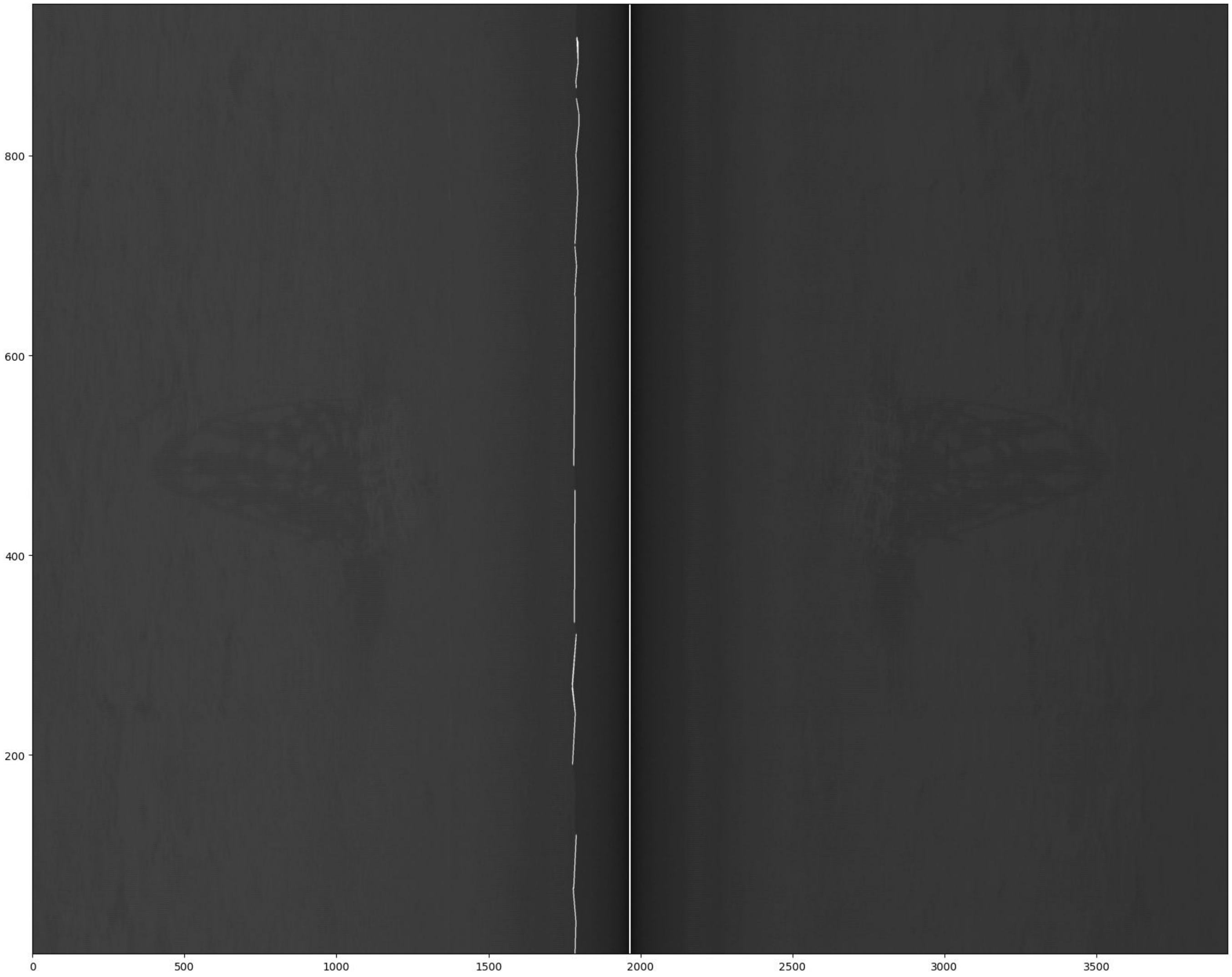
Marine Sonic Feature Data

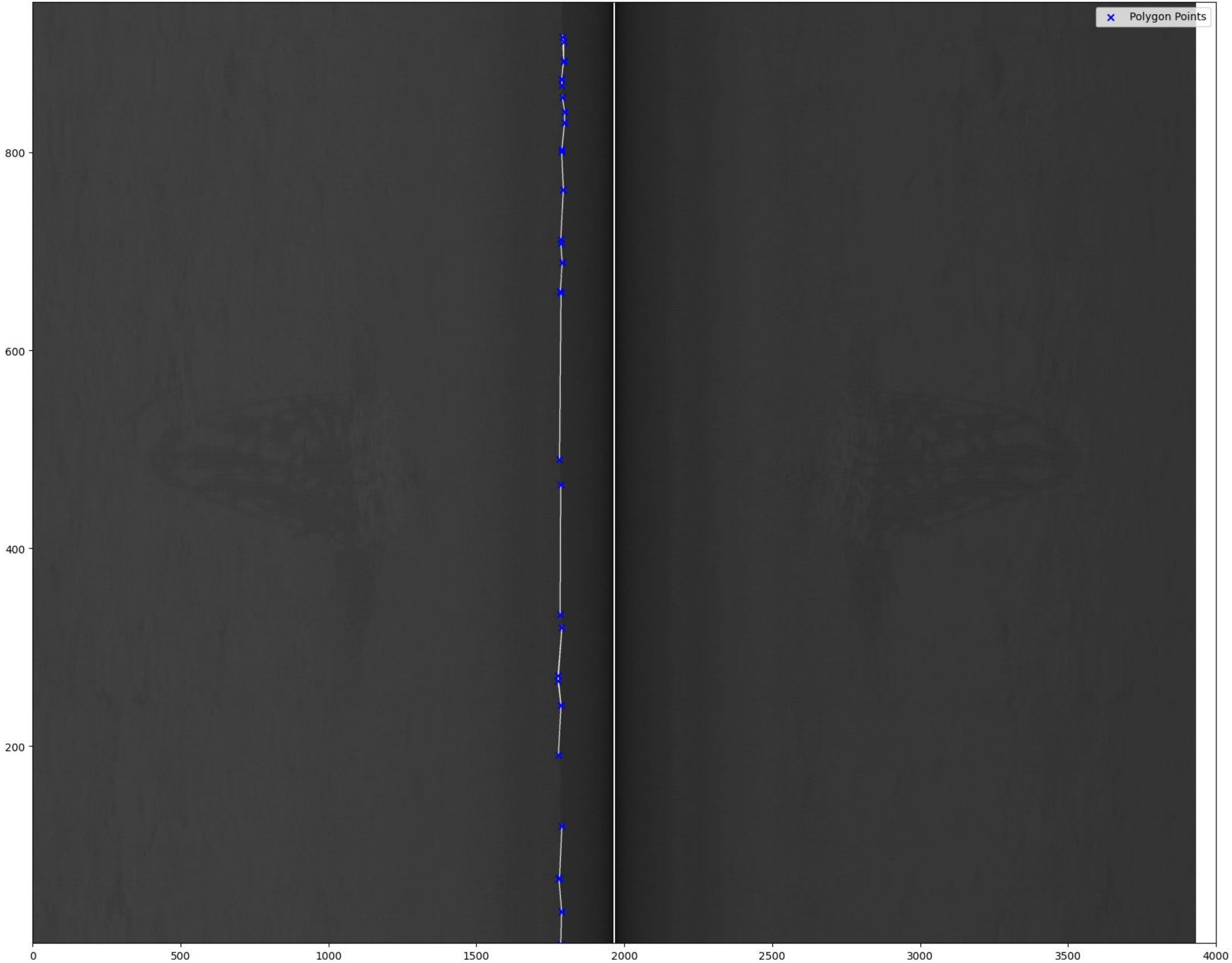



Edges Detected using Canny
800
600
400
200
0
500
1000
1500
2000
2500
3000
3500

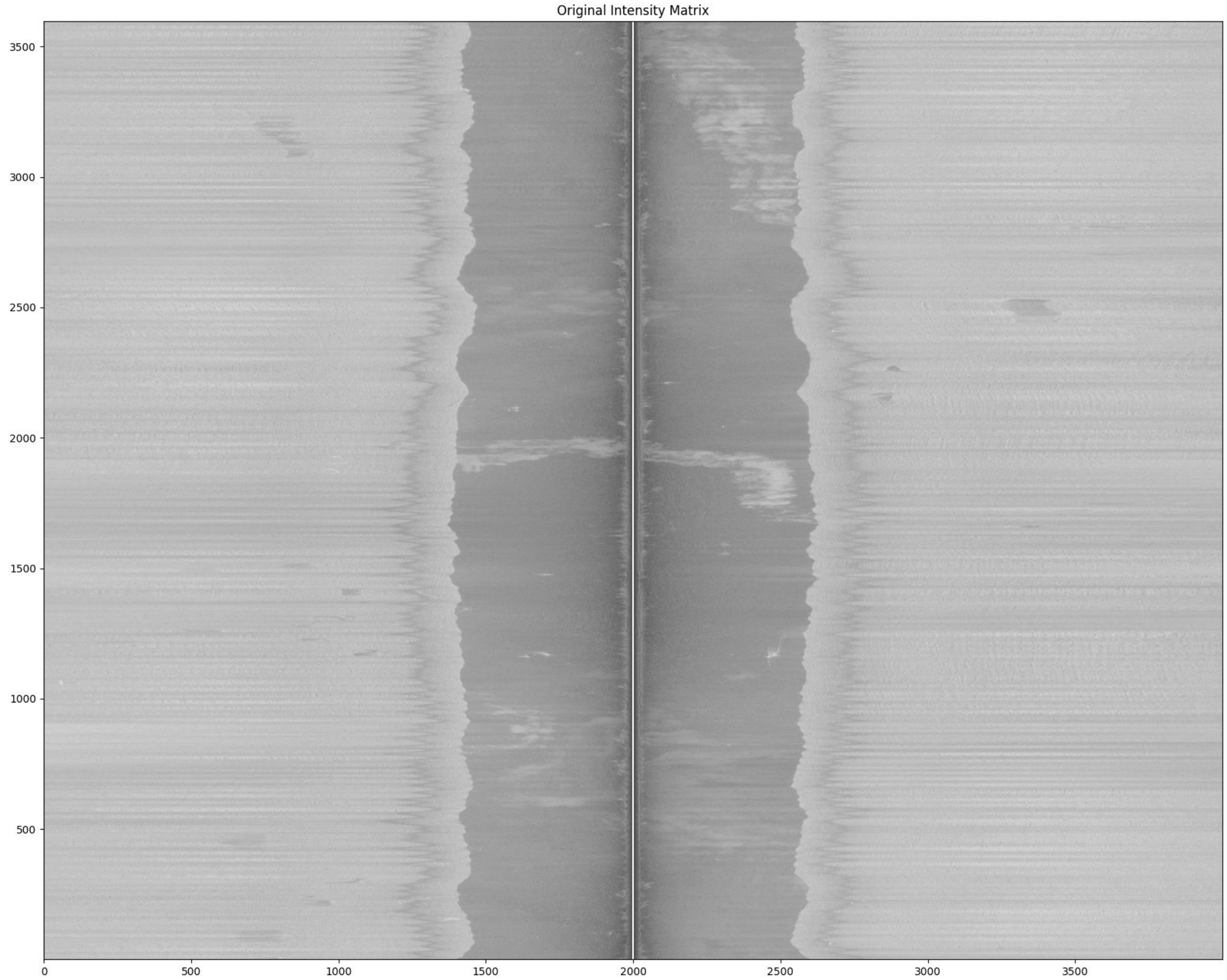

800
600
400
200
0
500
1000
1500
2000
2500
3000
3500


Bounding Box Centers
800
600
400
200

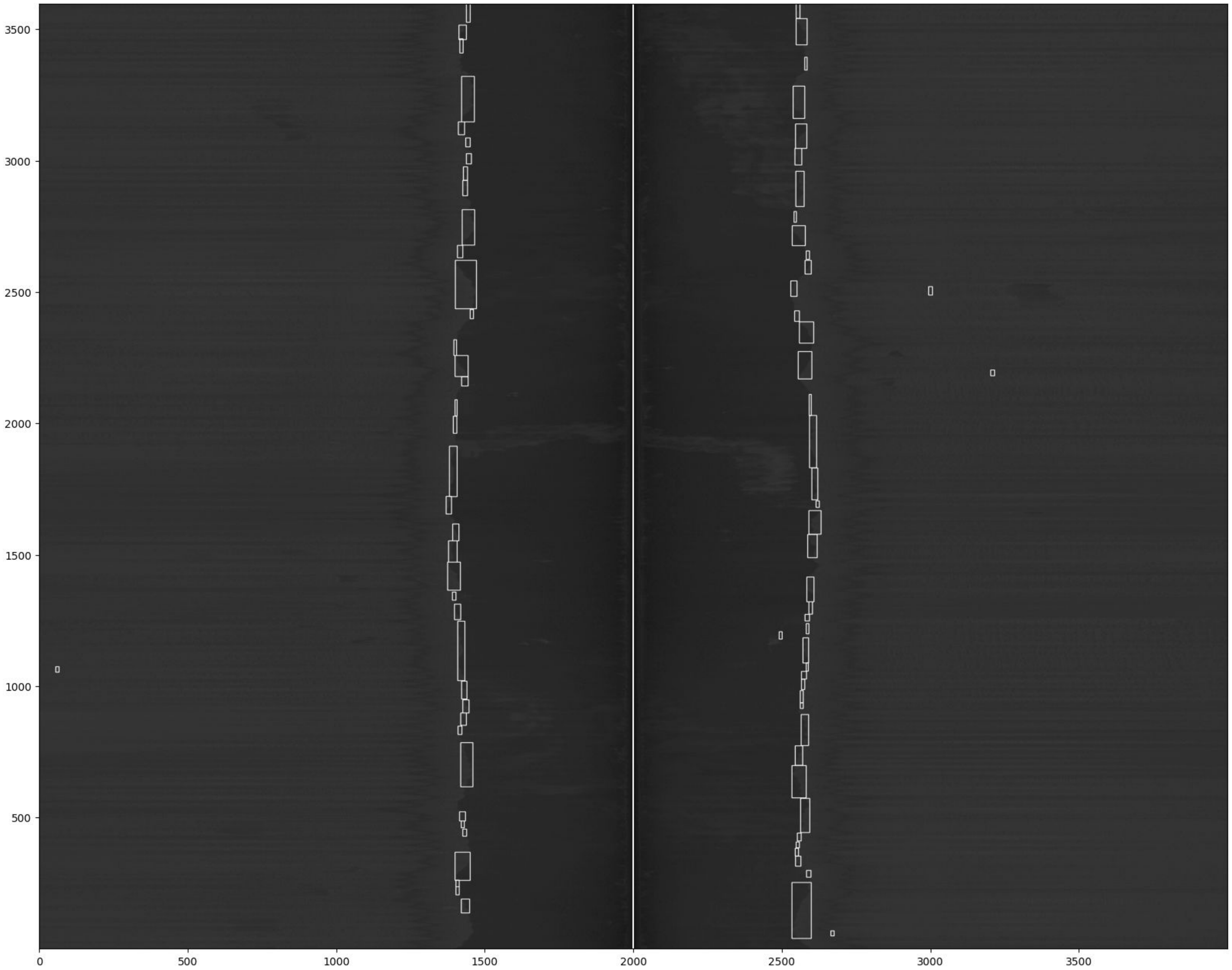

800
600
400
200
0
500
1000
1500
2000
2500
3000
3500

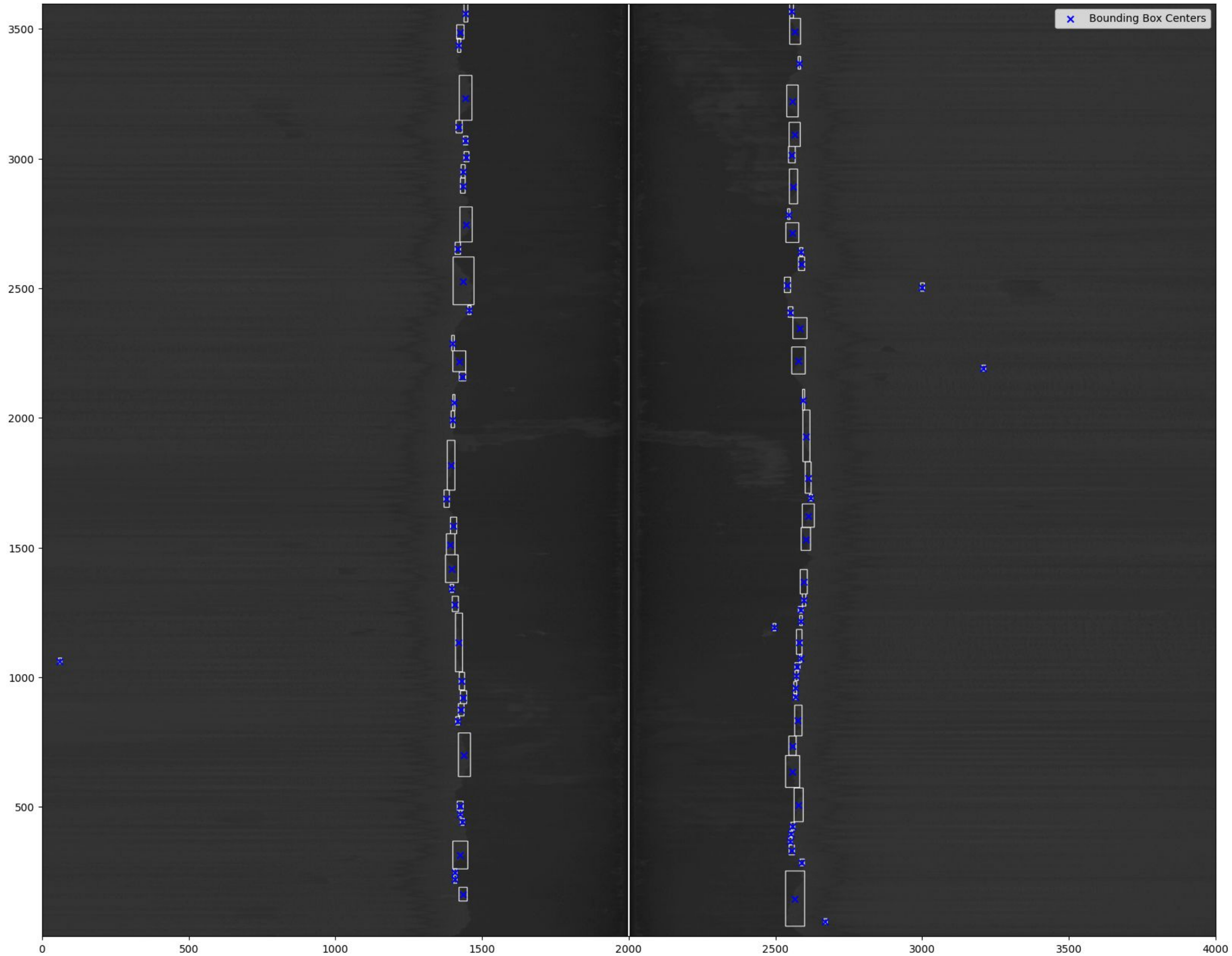

Polygon Points
800
600
400
200
0
500
1000
1500
2000
2500
3000
3500
4000

Scripps Summer Data (Noisy)

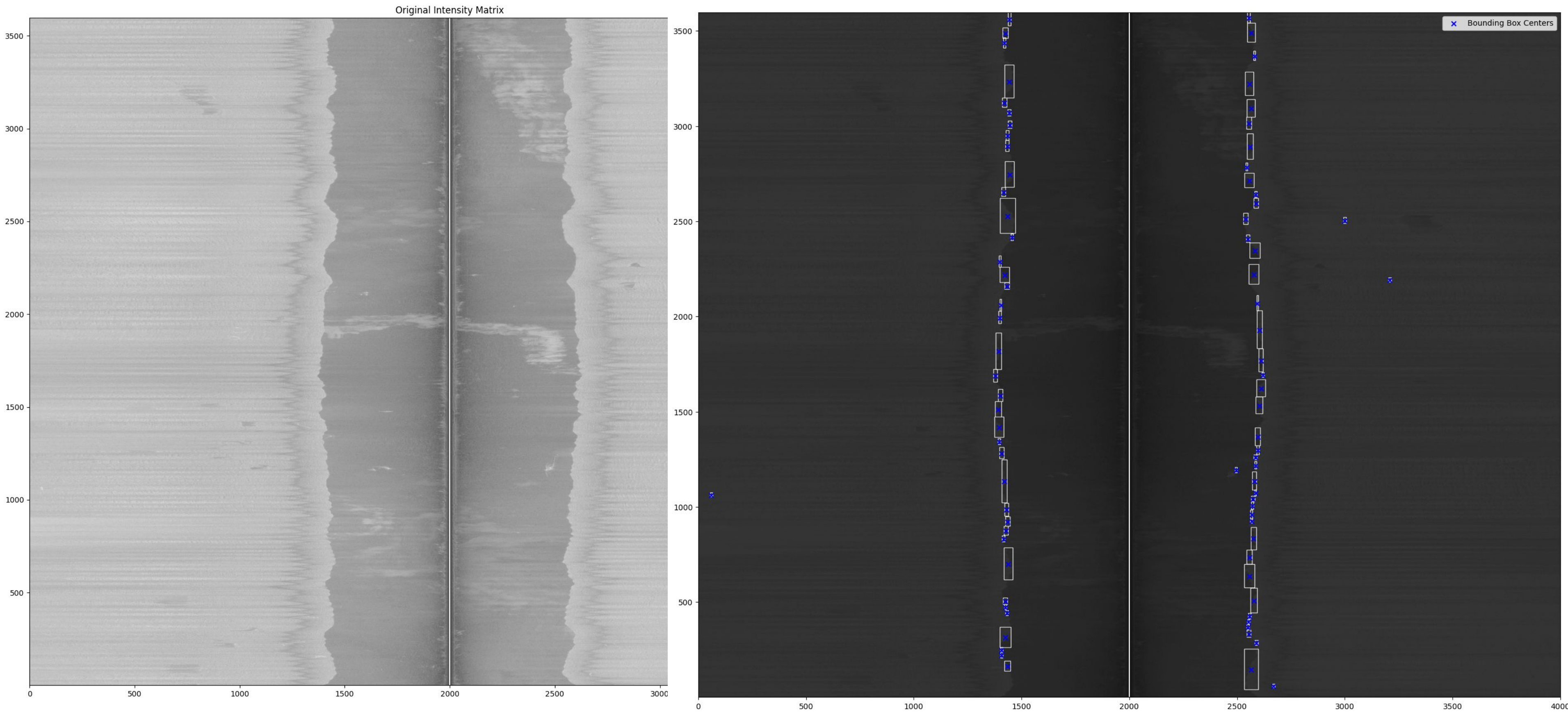


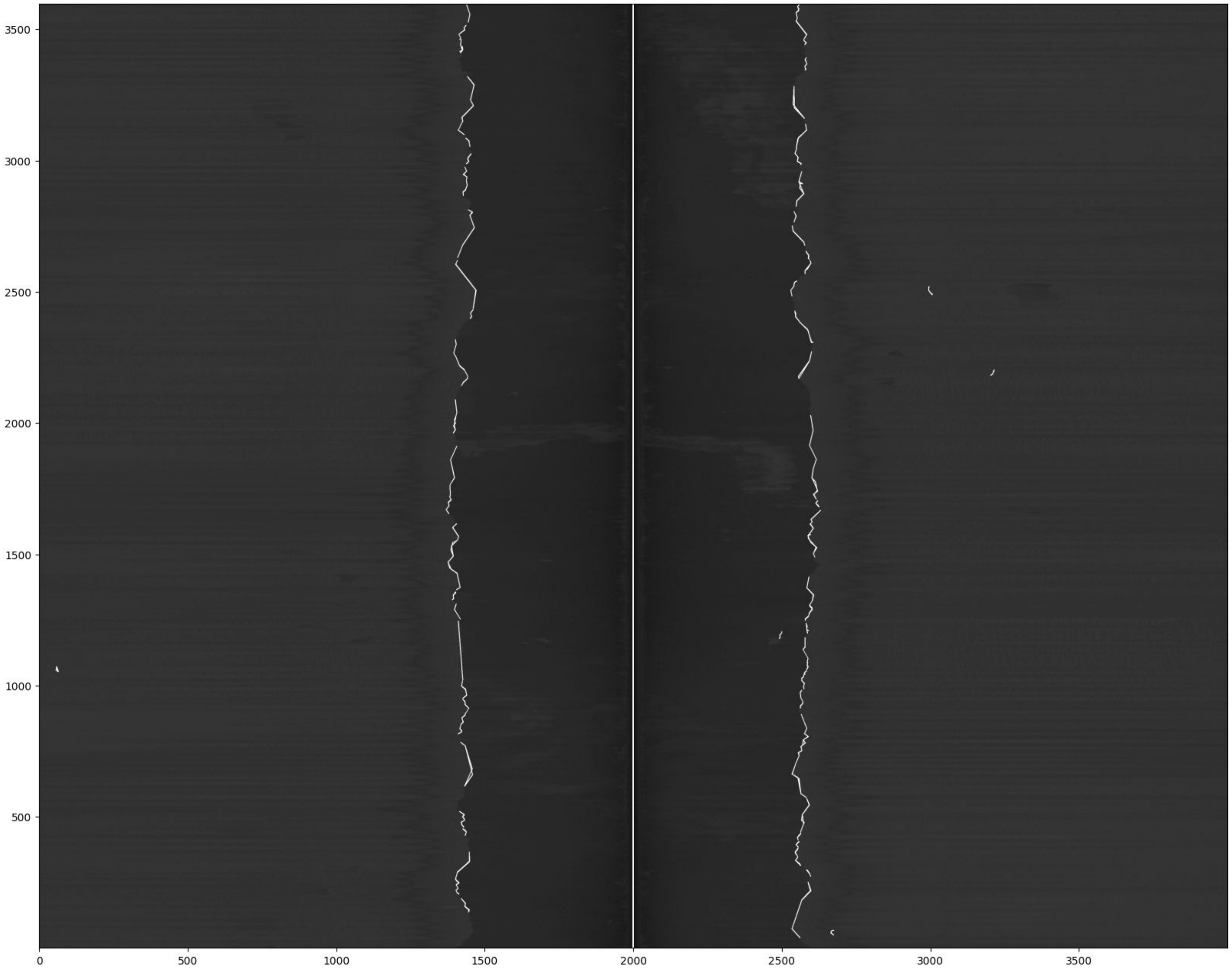
Edges Detected using Canny

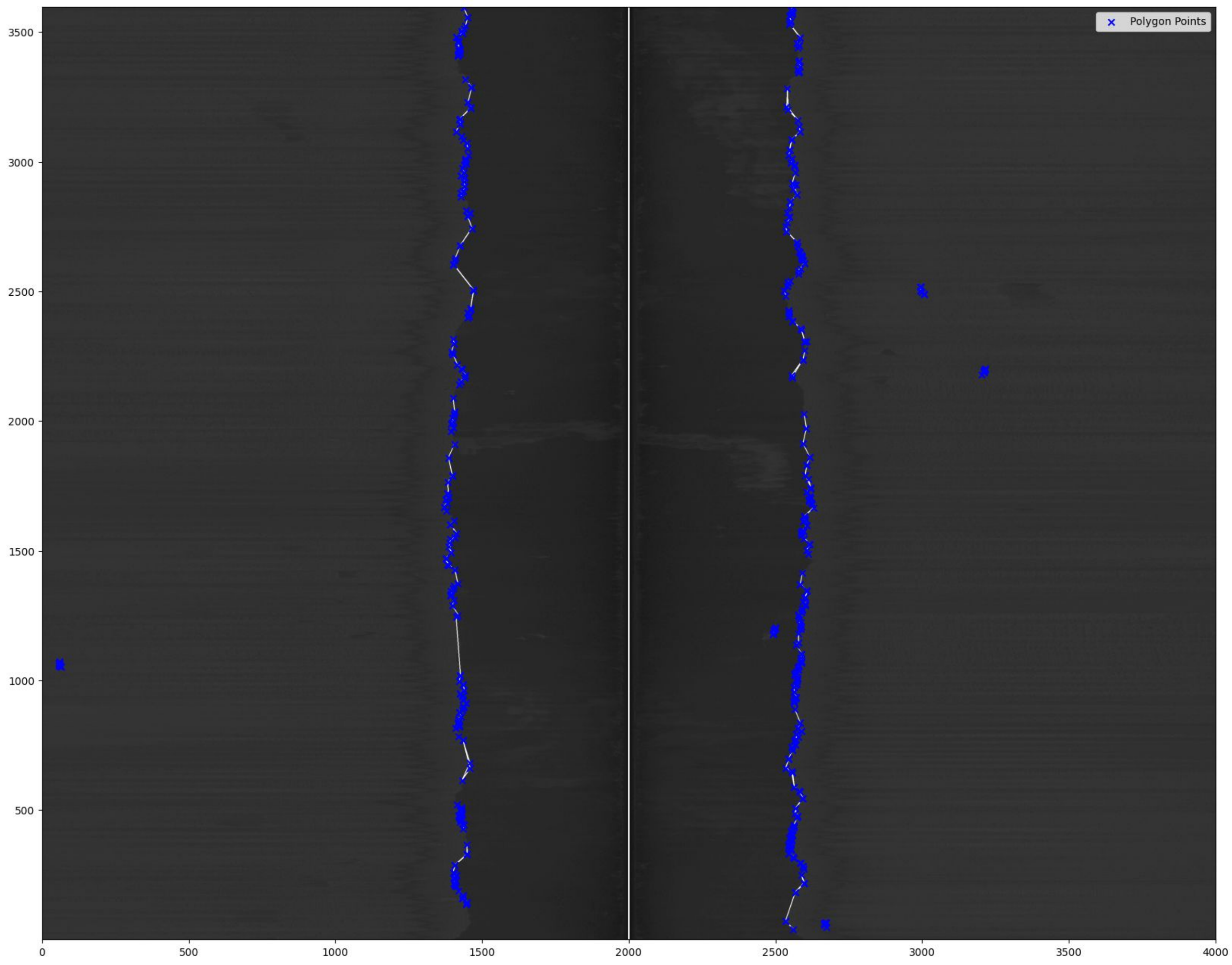

3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500



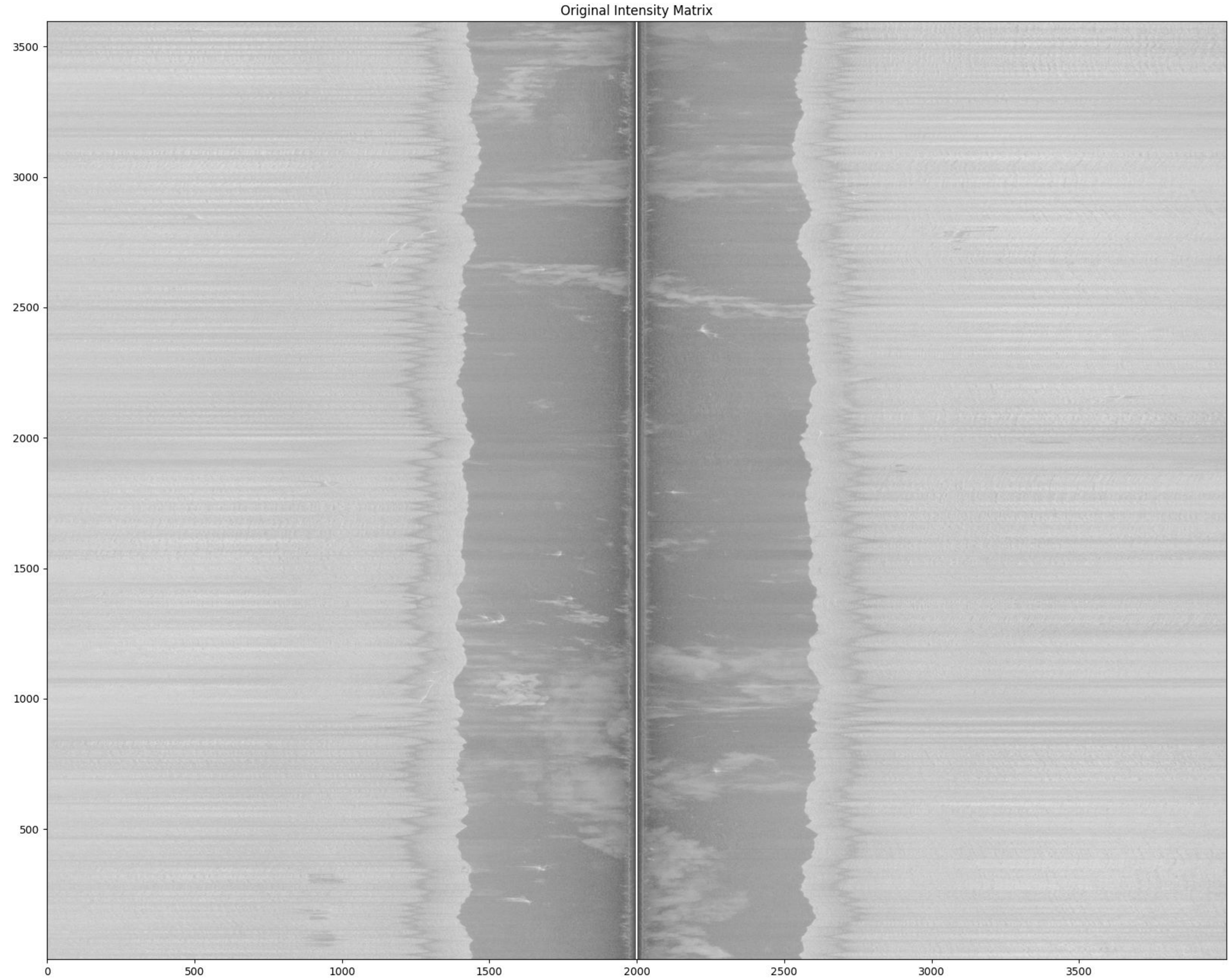
Bounding Box Centers
3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500
4000


Original Intensity Matrix
Bounding Box Centers

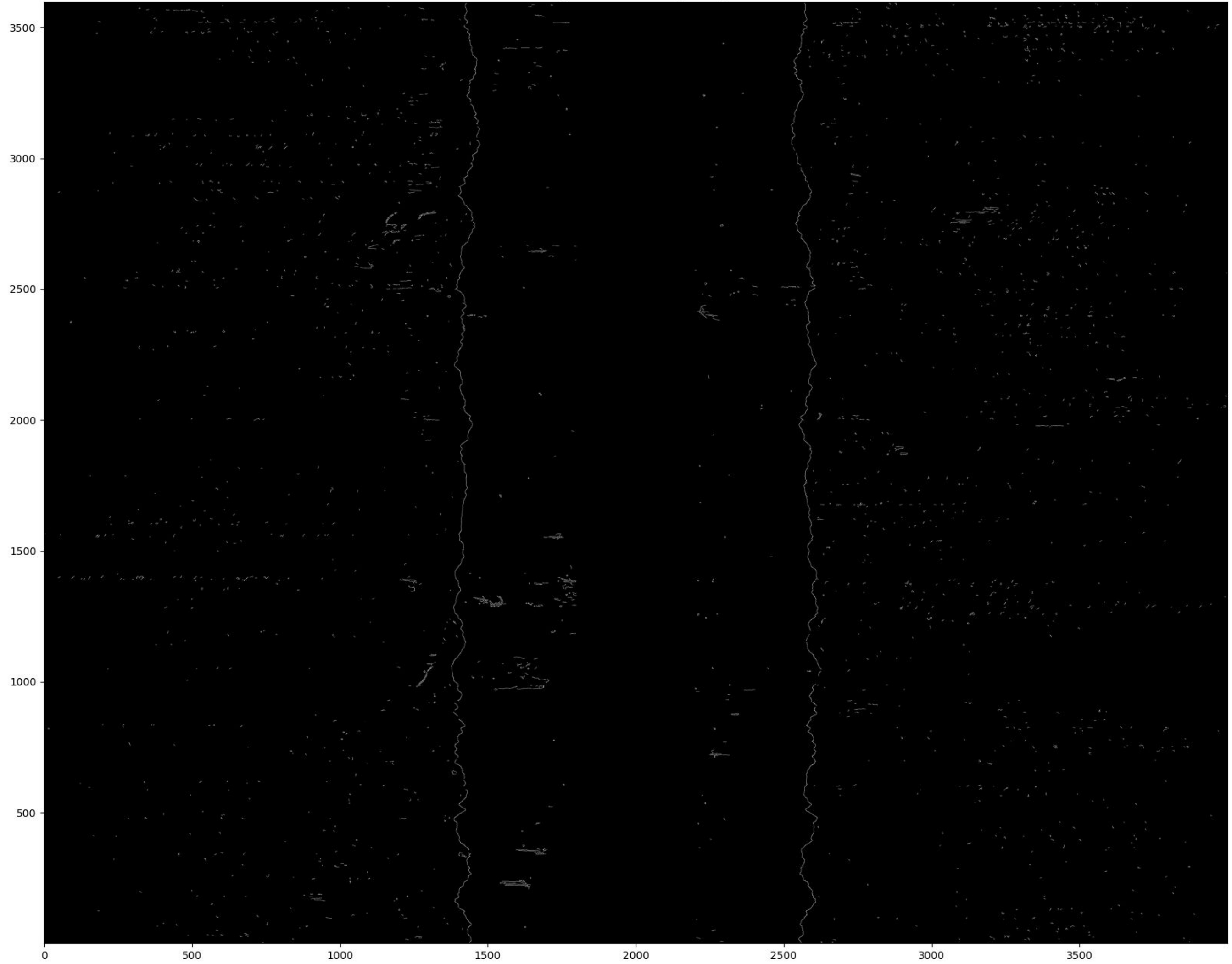

3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500

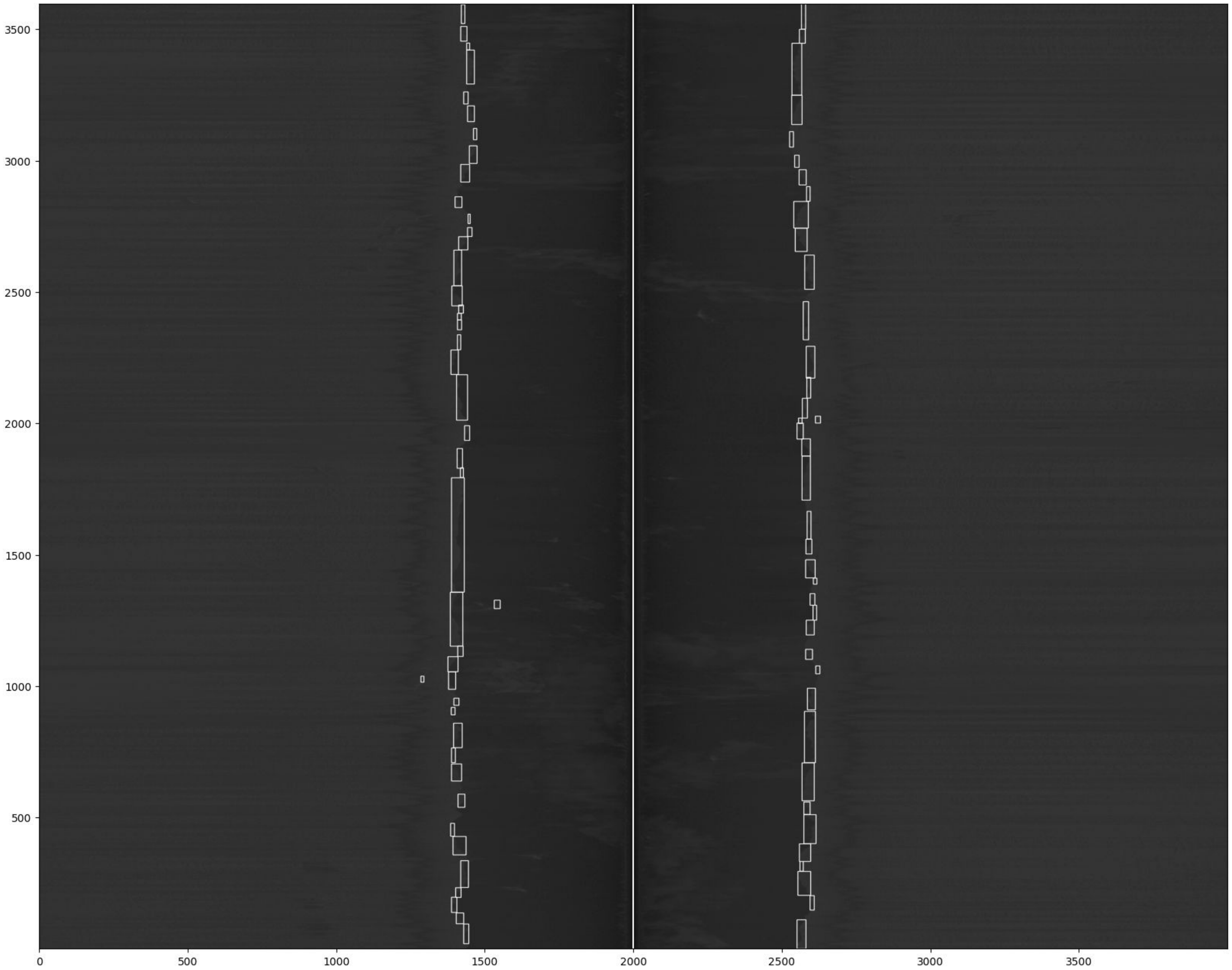

Polygon Points
0
500
1000
1500
2000
2500
3000
3500
4000
500
1000
1500
2000
2500
3000
3500

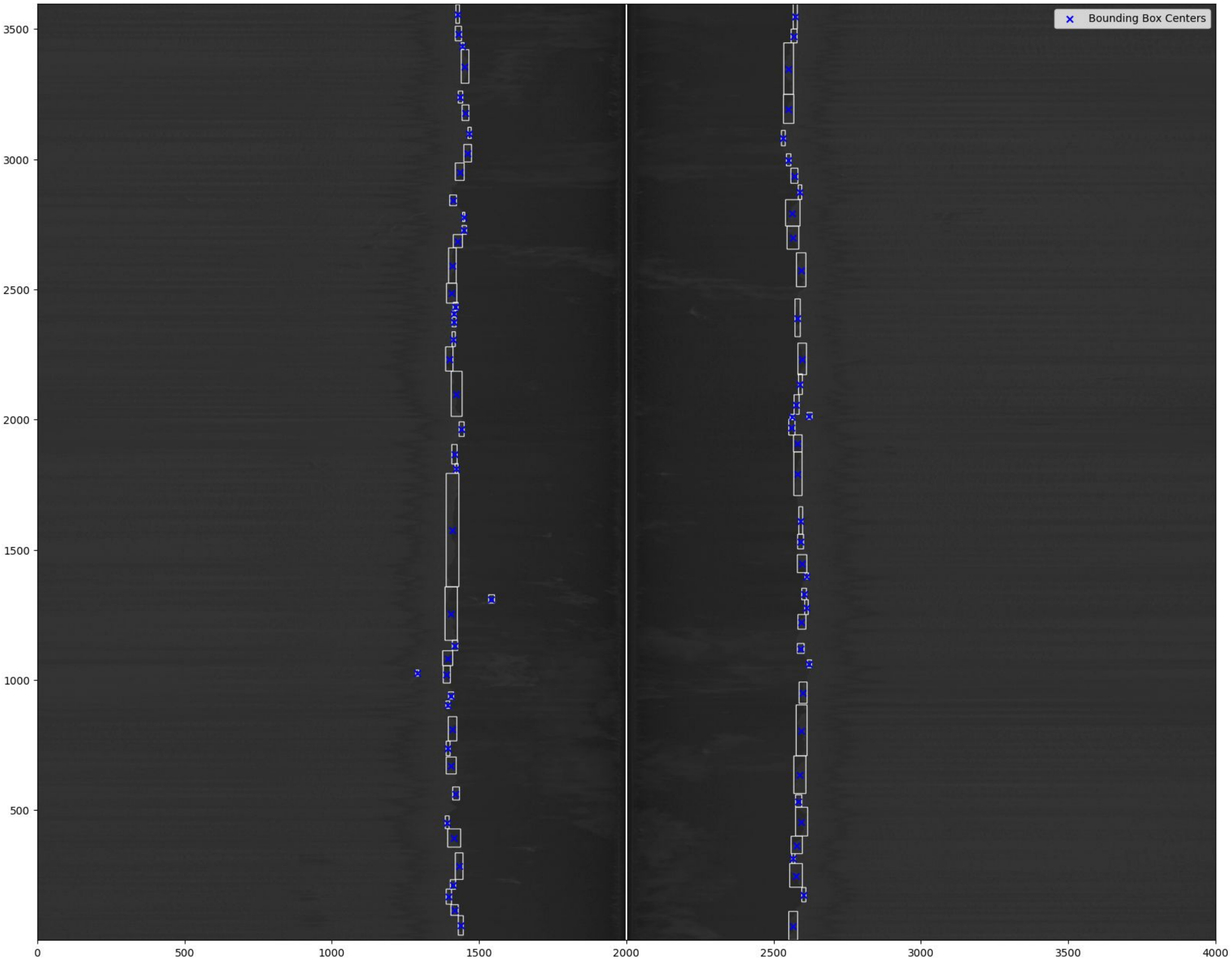
Scripps Summer Data (Noisy)

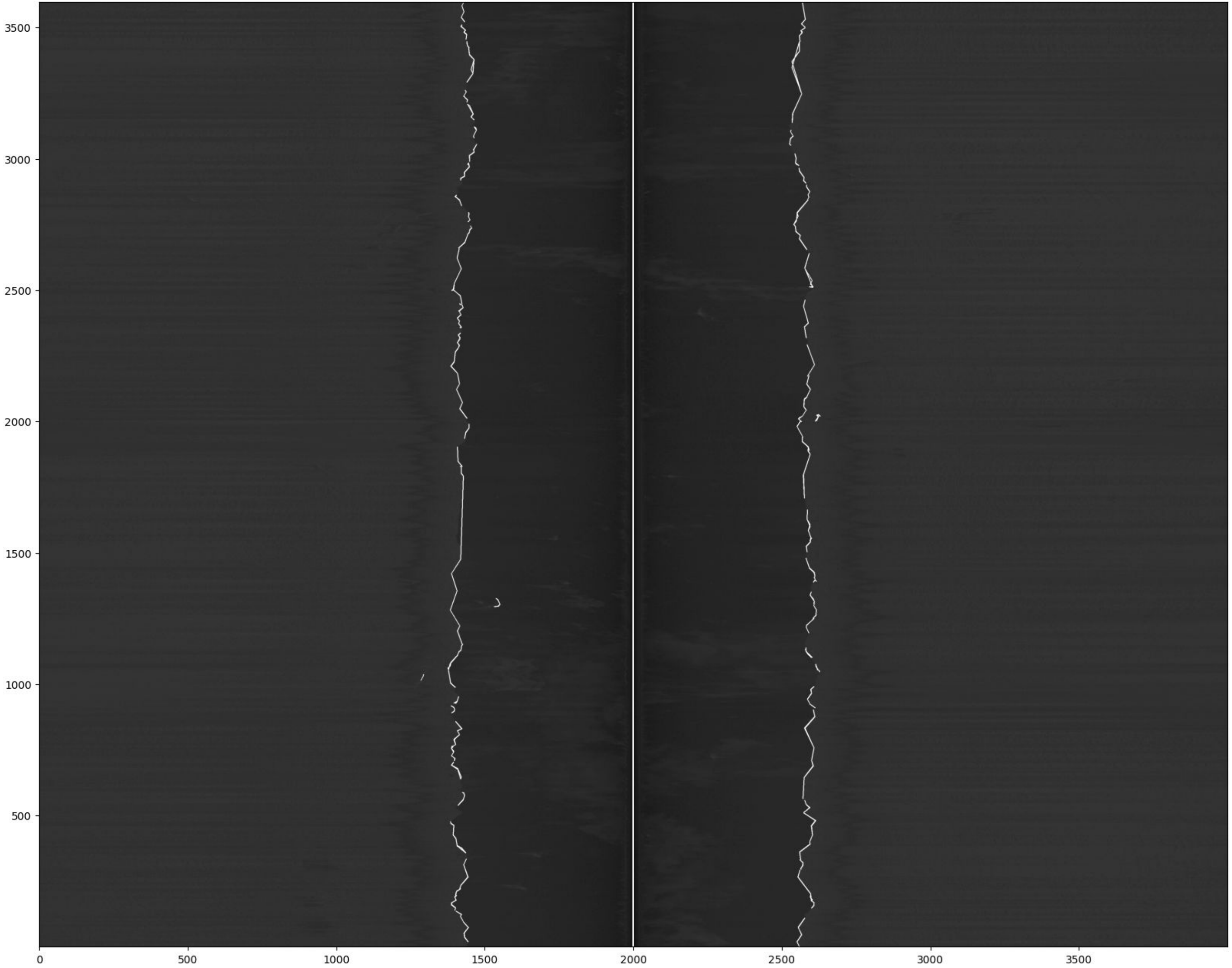


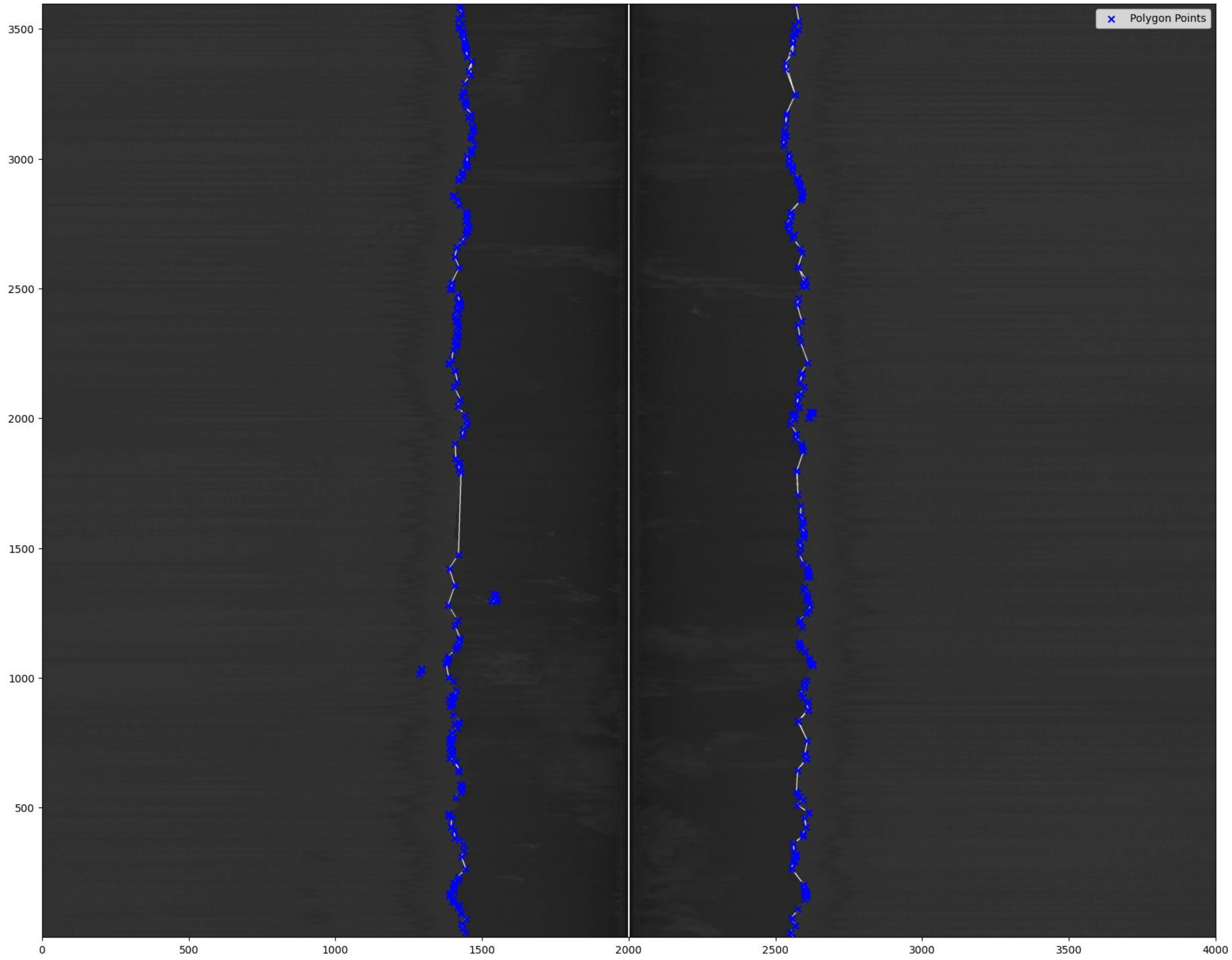
Edges Detected using Canny




3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500


Bounding Box Centers
0
500
1000
1500
2000
2500
3000
3500
4000
500
1000
1500
2000
2500
3000
3500


3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500



Polygon Points
3500
3000
2500
2000
1500
1000
500
0
500
1000
1500
2000
2500
3000
3500
4000

A. *Image Pre-processing*

Because the matrix I is a record of the reflected acoustic wave strength, it is not a standard image matrix. Pre-processing is necessary to enhance the measurements for further processing. Also, since the original sonar data are over-sampled, the first step in pre-processing is to down-sample sonar matrix data by a user-defined factor, d. This reduces the computational complexity of the problem. Then, the down-sampled image I_d becomes a matrix of size of $\lfloor \frac{n_s}{d} \rfloor \times n_t$, where $\lfloor \cdot \rfloor$ is the floor operator. Next, the down-sampled image is normalized linearly to obtain the grayscale image matrix,

$$I_g(i,j) = \frac{I_d(i,j)}{\max_{i,j}\left[I_d(i,j)\right]} \tag{4}$$

Then $I_g(i,j) \in [0,1],\ \forall i,j$. Finally, in order to adjust image intensities to enhance contrast, the histogram equalization technique [13] is applied to the gray image, and a transformed image, I_h, is obtained. Examples of the grayscale image I_g and the histogram equalized image I_h are shown in Fig. 3. It can be seen that the objects in I_h are better observed I_h than in I_g due to enhanced contrast.

B. Segmentation based on Matched Filters

In sonar images used in this study, all objects of interest have a similar structure comprised of a highlight area followed by a shadow. This is because the object reflects the sonar waves causing a the sonar to pick up a strong signal for that location, while location behind the object is blocked, results in a weak signal registration. The direction of the shadow area is always in line with the sonar scan direction, facing away from the sonar. To recognize and segment seafloor objects in the sonar images, a matched filter is designed, as shown in Fig. 4. W_m is the width of the matched filter and L_h and L_b are the lengths of the head and body parts, respectively. The matched filter is mathematically described by introducing a $(W_m \times (L_h + L_b))$ match filter matrix,

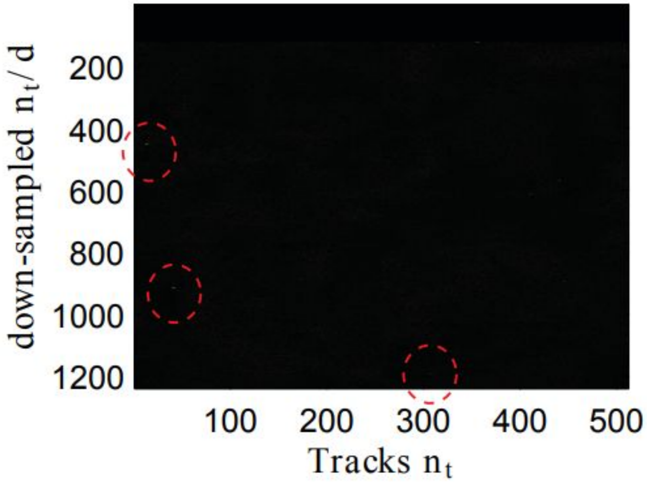
$$I_m(i,j) = \begin{cases} 1, \text{for } i \in [1, W_m] \text{ and } j \in [1, L_h] \\ -1, \text{for } i \in [1, W_m] \text{ and } j \in [L_h + 1, L_h + L_b] \end{cases} \tag{5}$$

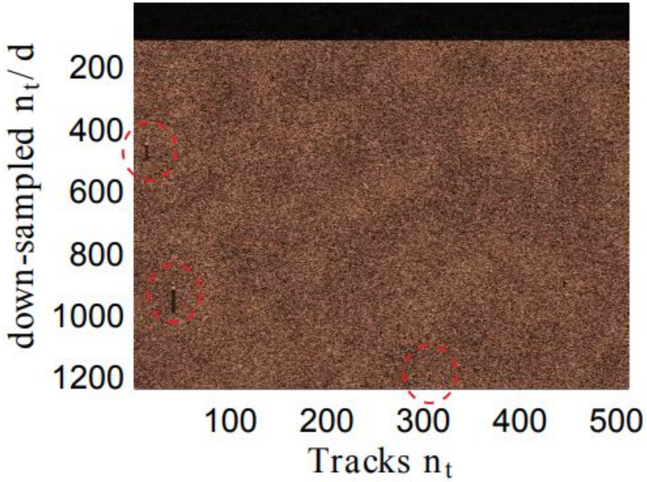
The grayscale image I_g is then converted to the binary image matrix

$$I_b(i,j) = \begin{cases} 1, & \text{if } I_g(i,j) \geq \theta_b \\ -1, & \text{if } I_g(i,j) < \theta_b \end{cases} \tag{6}$$

where θ_b is the binary threshold for the pixels. The normalized output of the matched filter is expressed as

$$I_n(i,j) = \frac{\sum_{\iota=1}^{W_m} \sum_{\zeta=1}^{L_h+L_b} I_b(i+\iota, j+\zeta) I_m(\iota, \zeta)}{\sum_{\iota=1}^{W_m} \sum_{\zeta=1}^{L_h+L_b} I_m^2(\iota, \zeta)}, \tag{7}$$


down-sampled n_t/d
200
400
600
800
1000
1200
100
200
300
400
500
Tracks n_t


down-sampled n_t/d
200
400
600
800
1000
1200
100
200
300
400
500
Tracks n_t

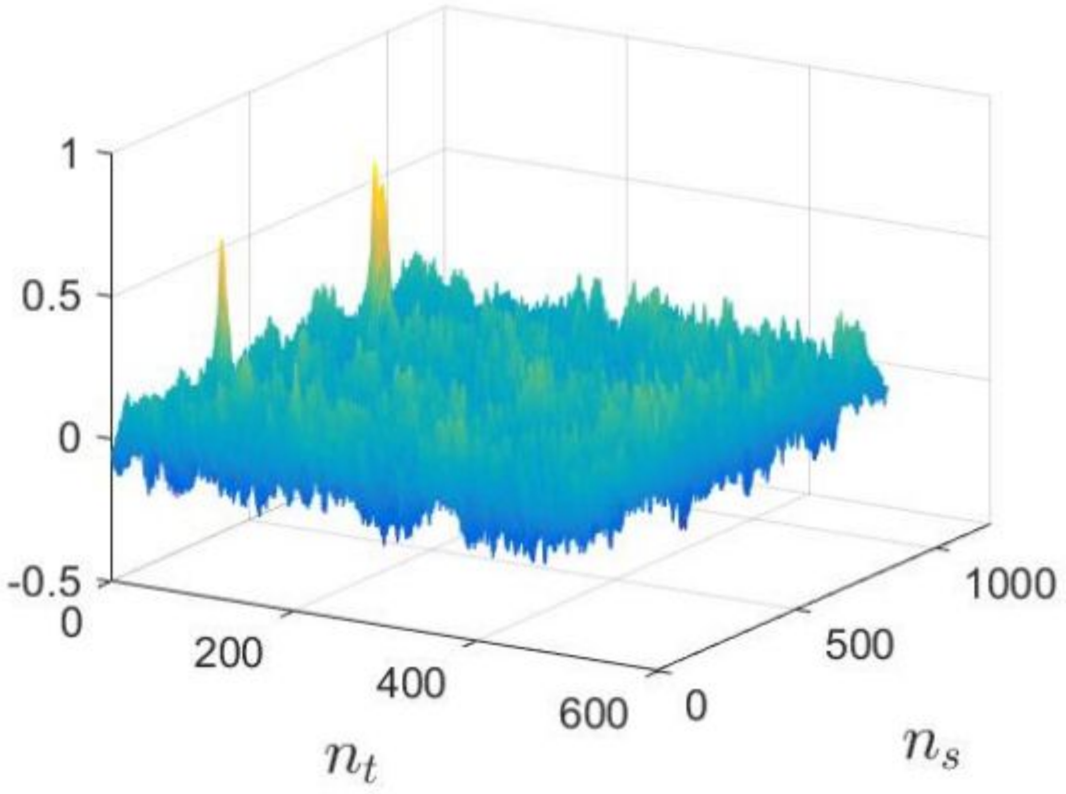


Fig. 5. The normalized output of the matched filter.

Labeling from Paper

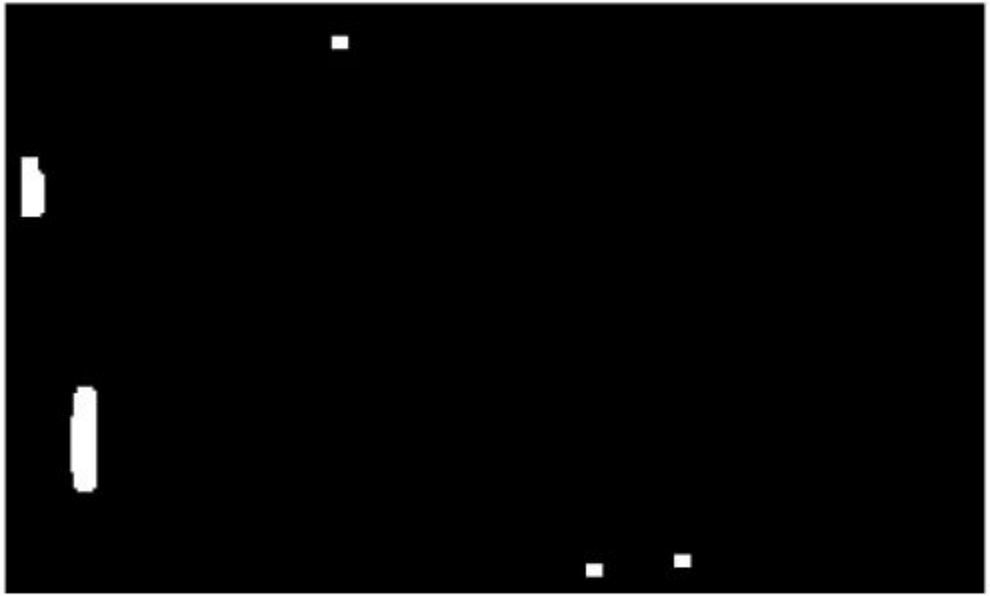


Fig. 6. Example σ set for a sonar image represented by white segments.