Technical Explanation of the Biovolume Calculation Algorithm

Kananat Siangsano

March 18, 2025

1 Overview

This document explains the implementation of the biovol_cal.py script to estimate biovolume from FlowCam images. The algorithm is an adaptation of biovolume calculation method for Imaging FlowCytobot (IFCB) detailed in [Moberg and Sosik, 2012].

2 Preprocessing Step

The primary purpose of the preprocessing step is to transform raw FlowCam images into binary image masks that accurately represent the shapes of organisms, which is the expected input of the Solid of Revolution (SOR) and Distance Map (Distmap) algorithm. These binary masks serve as the foundation for all subsequent biovolume calculations.

The preprocessing workflow consists of these key steps:

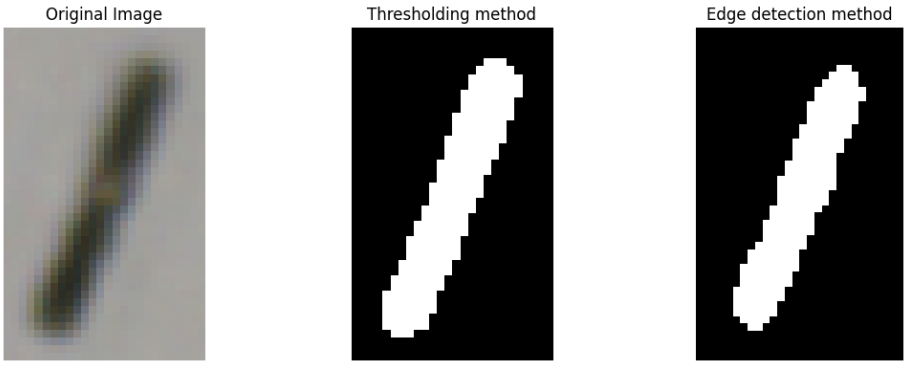
1. **Edge Detection**: The algorithm first applies Canny edge detection to identify organism boundaries
2. **Skeletonization**: Detected edges are thinned to single-pixel width, this step is crucial for defining the endpoints to be used in the next step
3. **Endpoint Connection**: Disconnected endpoints are joined with straight lines
4. **Flood Filling**: The closed contours are filled to create binary masks
5. **Noise Removal**: Small artifacts below a minimum size threshold are removed.

The algorithm includes fallback mechanisms: If Canny edge detection fails to detect sufficient object area (less than 20% of the image), it switches to Sobel-based edge detection.
[After some thinking, the Endpoint Connection step could probably be made more sophisticated by implementing the Hough transform [Ballard, 1981] instead of the current approach.]

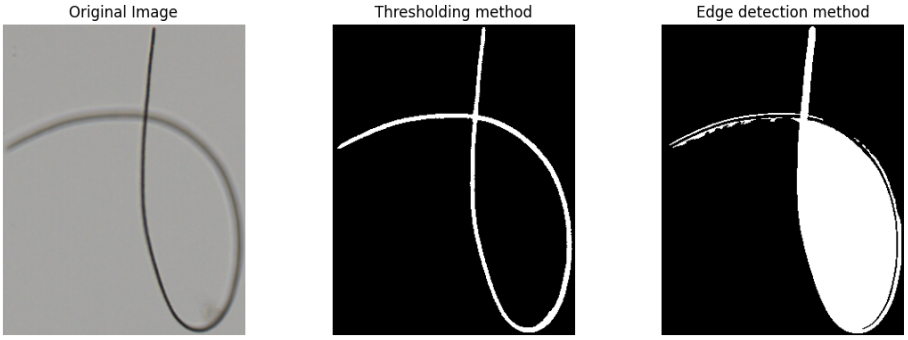
2.1 Comparison: Thresholding vs. Edge Detection

There are two main approaches for segmenting organisms from background: direct thresholding and edge detection. Each has strengths and weaknesses depending on image characteristics. The following section shows the comparison of the results from these approaches.

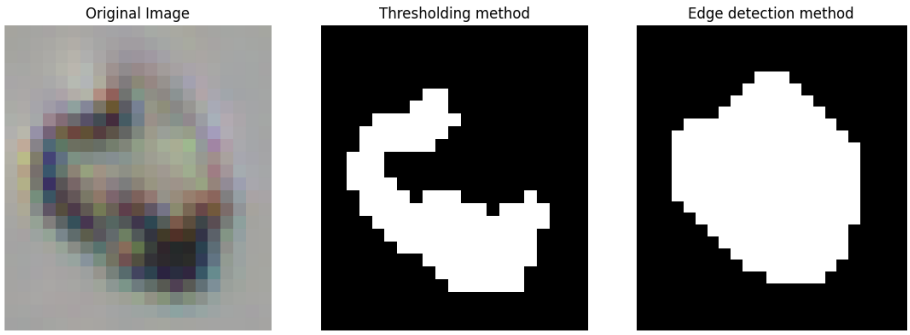
2.1.1 Both Succeed

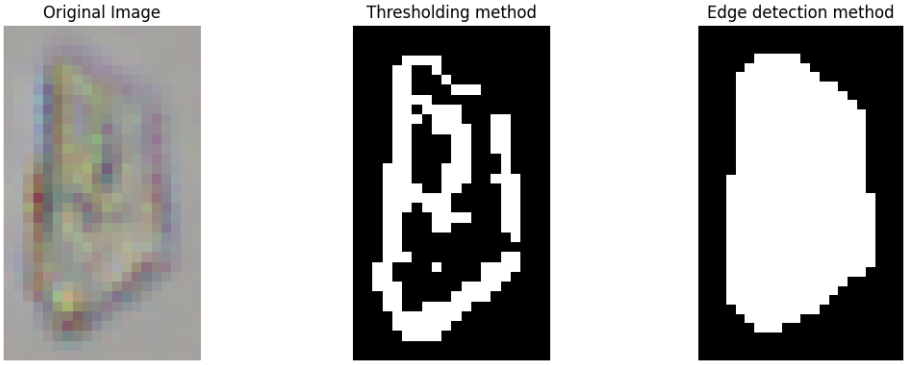


2.1.2 Thresholding Better

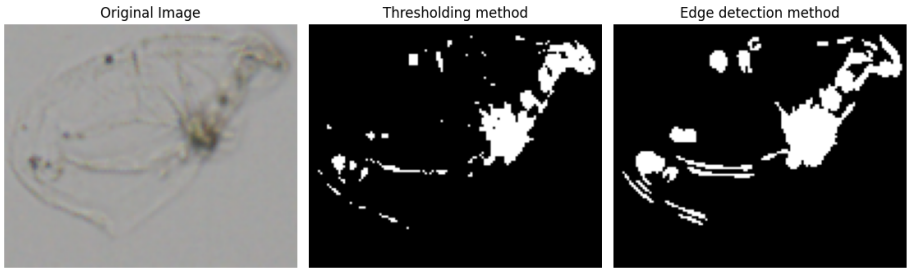


2.1.3 Edge Detection Better





2.1.4 Both Fail



Edge detection has potential to perform better because:

- It is less sensitive to uneven illumination across the field of view
- It can detect organisms with internal structures that might confuse thresholding

3 Morphological Analysis

3.1 Shape Parameters

For each detected organism, several morphological parameters are calculated to be used as method selection criteria.

1. **Convex Hull**: The smallest convex polygon containing all points of the organism
2. **Area Ratio**: The ratio between convex hull area and actual area
3. **Eccentricity**: Calculated from eigenvalues of the covariance matrix of coordinates
4. **Equivalent Diameter**: The diameter of a circle with the same area as the organism
5. **Major Axis Length**: The length of the primary axis from eigenvalue decomposition

3.2 Method Selection Criteria

The algorithm chooses between SOR and Distmap methods based on the following criteria:

$$\text{Use SOR if: } \begin{cases} \text{Area Ratio} < 1.2 \\ \text{OR} \\ (\text{Eccentricity} < 0.8 \text{ AND } p > 0.8) \end{cases} \tag{1}$$

where p is the ratio of equivalent diameter to major axis length. These criteria identify roughly circular or moderately elongated organisms that are suitable for the SOR approach.

4 Volume Calculation Methods

4.1 SOR Method (Sosik and Kilfoyle Algorithm)

The SOR method [Sosik et al., 2003] models the organism as a solid of revolution around its major axis. The implementation:

1. **Rotates the binary mask** to align the major axis with the vertical axis
2. **Computes radius of each slice** by counting pixels and dividing by 2
3. **Represents the shape in 3D** by computing coordinates on 721 angles (selectable number of angles) between 0-180°

4. **Calculates surface areas** of the top and bottom quadrilaterals
5. **Computes volume** using frustum cone approximation

The SOR approach works well for organisms with rotational symmetry around their major axis.

4.2 Distmap Method (Moberg & Sosik Algorithm)

For more irregular shapes, we use the Distmap method [Moberg and Sosik, 2012]:

1. Computes the Euclidean distance transform from the perimeter
2. Each interior pixel is assigned a value representing distance to the nearest boundary
3. Applies correction factors to transform distance map to height map
4. Summation of the corrected height map for volume

The one shot formula for volume is:

$$V = c_1 \cdot \pi \cdot \sum D \tag{2}$$

where c_1 is a diamond correction factor defined as:

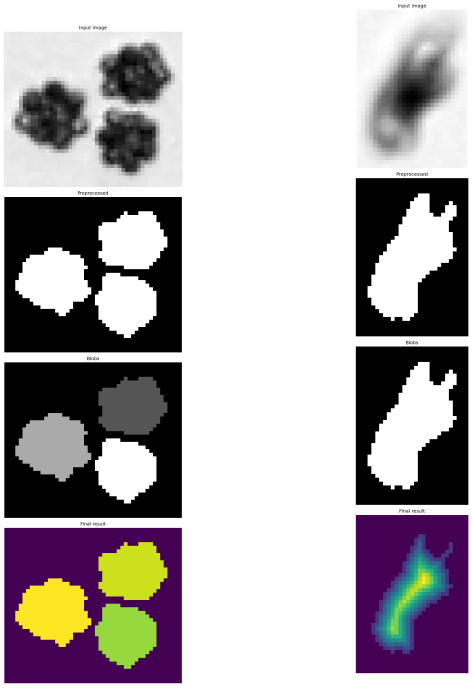
$$c_1 = \frac{x^2}{x^2 + 2x + 0.5} \tag{3}$$

and x is the representative transect, defined as $4 \cdot \text{mean}(D) - 2$.

5 Debug mode

The code includes debugging functionality that, when enabled, produces visualizations of:

- The grayscale input image
- The binary mask after preprocessing
- Labeled individual organisms
- The final result with intensity proportional to biovolume (uniform color if used SOR method)



SOR

Distance map

Figure 1: Examples of the debug mode output

6 Calibration and Unit Conversion

6.1 Pixel to Micrometer Conversion

The algorithm converts from pixel units to real-world measurements using calibration beads:

$$\text{length\_per\_pixel\_ratio} = \frac{\text{bead\_size}}{\text{bead\_equiv\_diameter}} \tag{4}$$

$$\text{pixel\_cube\_unit} = \text{length\_per\_pixel\_ratio}^3 \tag{5}$$

$$\text{real\_volume} = \text{pixel\_volume} \times \text{pixel\_cube\_unit} \tag{6}$$

Default values in the code are:

- `bead_size` = 50 μm (real bead diameter)
- `bead_equiv_diameter` = 30.25 pixels (measured diameter in image)

References

Dana H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981. doi: 10.1016/0031-3203(81)90009-1.

Emily A. Moberg and Heidi M. Sosik. Distance maps to estimate cell volume from two-dimensional plankton images. *Limnology and Oceanography: Methods*, 10:278–288, 2012. doi: 10.4319/lom.2012.10.278.

Heidi M. Sosik, Robert J. Olson, Michael G. Neubert, Alexi Shalapyonok, and Andrew R. Solow. Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. *Limnology and Oceanography*, 48(5):1756–1765, 2003. doi: 10.4319/lo.2003.48.5.1756.