Enhanced Detection and Quantification of DNA Hybridization in Fluorescence Images via Voronoi Diagram Clustering

Mohammad Asadi
Electrical Engineering
Stanford
Email: masadi@stanford.edu

***Abstract*—The precise quantification of DNA hybridization rates using fluorescence imaging is pivotal in molecular biology. This project introduces an innovative method combining Voronoi diagram clustering with convolution and interpolation for peak detection in fluorescence images. [1] Addressing challenges such as variable cluster formation, instrumental sensitivity, and noise, the methodology significantly improves the alignment, intensity extraction, and quantification of hybridization events. [2]**

I. Introduction

Fluorescence imaging is a pivotal technique in molecular biology, allowing for the visualization and quantification of biological phenomena such as DNA hybridization. [3] In this study, we focus on the analysis of fluorescence images obtained from a MiSeq imager system that captures the hybridization of target DNA strands to those attached to a flowcell surface [4]. A major challenge in this context is the accurate detection and quantification of fluorescence signal peaks amid varying experimental conditions and inherent noise. To address this, we integrate Voronoi diagram clustering with traditional methods like convolution and interpolation, aiming to enhance peak detection and signal analysis [1].

II. Image Alignment and Matching

The accurate alignment of fluorescence images is paramount to the success of hybridization rate quantification. This process involves matching the coordinates provided by the MiSeq imager with the actual locations of clusters on the flowcell. Given the high sensitivity of the imaging process, even minor misalignments can lead to significant errors in subsequent analyses. [3] The whole pipeline is shown in Figure 1.

A. Challenges

Several challenges complicate the alignment process:

- **Relative Cluster Positions:** The initial coordinates of DNA clusters are relative and can vary due to experimental conditions.
- **Additional Clusters:** Unanticipated clusters may form, necessitating their identification and incorporation into the alignment process.
- **Instrument Sensitivity:** Variations in temperature, distance, and experimental setup can alter the appearance of clusters, affecting their detectability and positional accuracy.

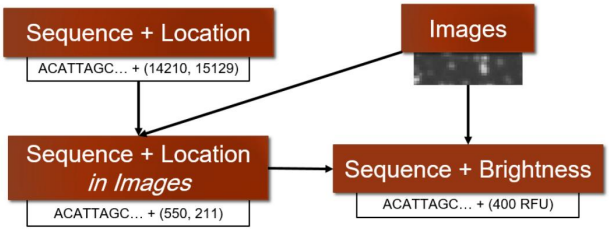

Fig. 1. A schematic of the pipeline from getting the flowcell along with the locations from Illumina MiSeq device to extraction of intensities and matching them to the corresponding DNA sequences.

B. Alignment Process

The alignment process is twofold, consisting of an initial approximation of cluster positions followed by fine-tuning to achieve precise alignment.

1) **Approximate Shift:** Using the coarse coordinates from the MiSeq imager, an initial approximation is made to align the images with the flowcell. This step involves correlating the provided coordinates with observed fluorescence signals, applying a global shift to minimize the overall displacement.
2) **Fine-Tuning Scaling and Rotation:** After the approximate shift, detailed adjustments are made for scaling and rotation. This phase relies on iterative optimization techniques, focusing on minimizing the discrepancy between the expected and observed positions of DNA clusters. Techniques such as least-squares optimization and machine learning models trained on known alignments are employed to refine the alignment parameters.

C. Outcome

The outcome of this meticulous alignment process is a series of fluorescence images accurately matched to their corresponding locations on the flowcell. This precision is crucial for the reliable extraction and quantification of fluorescence signals in later stages of the analysis.

III. Intensity Extraction Techniques

Following the alignment of fluorescence images, the next critical step is the extraction of intensity data from each DNA cluster [4]. A sample of the taken images from which the intensities are going to be extracted can be seen in Figure 2. This process is complicated by several factors, including the high density of clusters, their variable shapes, and the presence of noise.

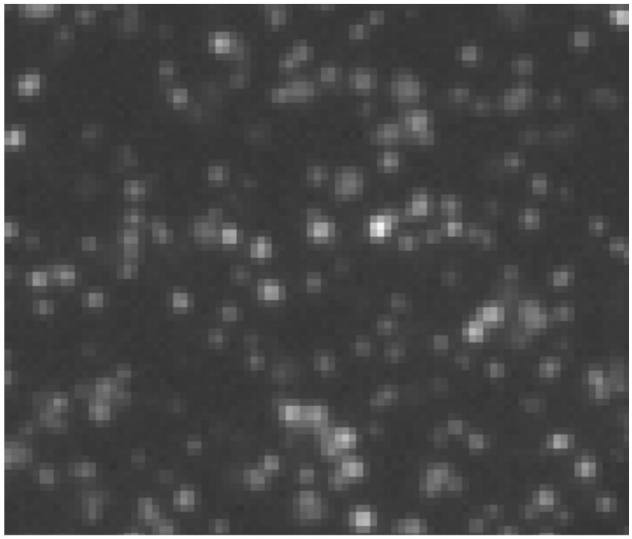

Fig. 2. A sample image taken from the flowcell during the hybridization experiment.

A. Extraction Challenges

The main challenges in intensity extraction include:

- **Off-Center Coordinates:** Approximate cluster coordinates may not align perfectly with the actual center of fluorescence signals.
- **Cluster Density:** High cluster densities make it difficult to distinguish between adjacent clusters, leading to potential signal overlap.
- **Variable Cluster Shapes:** The irregular shapes of clusters can complicate the extraction of intensity data, as assumptions about uniformity are often invalid.

B. Techniques Employed

To address these challenges, several extraction techniques are utilized:

1) **Foreground Pixel Summation:** This approach involves summing the intensity values of all pixels identified as part of a cluster's foreground. While robust against minor misalignments, it can be influenced by neighboring cluster signals.
2) **Deconvolution:** Applying deconvolution algorithms with an approximate blurring kernel helps to separate the signals of closely situated clusters. This technique enhances the signal-to-noise ratio but assumes uniform cluster shapes.
3) **Interpolation and Extraction:** Interpolation between pixels allows for the estimation of intensity values at precise locations, offering a balance between precision and susceptibility to neighboring cluster interference.

C. Voronoi-Based Extraction

Building upon these techniques, the utilization of Voronoi diagram clustering offers a novel method for intensity extraction. By creating polygons around each cluster, this method allows for the isolation of signals, significantly improving the accuracy of intensity measurements even in densely packed areas.

IV. Application of Voronoi Diagrams in Enhancing Fluorescence Image Analysis

Voronoi diagrams are a fundamental construct in computational geometry, offering a versatile method for partitioning a space into regions based on distances to a specified set of objects [6]. In the simplest form, given a set of points in a plane, the Voronoi diagram partitions the plane into regions where each region corresponds to the area closer to a particular point than to any other. These regions are referred to as Voronoi cells.

A. Introduction to Voronoi Diagrams

Voronoi diagrams have been employed in various fields for spatial analysis, partitioning spaces based on proximity to a set of points. Their application in image analysis, particularly for segmenting spatial data, offers a novel approach to delineating cluster boundaries in fluorescence images.

B. Significance in Image Analysis

In the context of image analysis, Voronoi diagrams are particularly useful for segmenting images based on spatial properties, distinguishing regions of interest, and facilitating the analysis of spatial distributions. Their ability to adaptively partition space based on the proximity to defined points makes them ideal for tasks where spatial organization and region assignment are crucial.

C. Application in Fluorescence Image Analysis for DNA Hybridization

The challenge in DNA hybridization analysis involves accurately identifying and quantifying signal from hybridized DNA clusters. Figure 3 shows a better view of the experimental setup resulting in the flowcells that are then used for taking the images. Traditional methods often struggle with these complexities, leading to inaccuracies in intensity extraction and quantification of the hybridization rate. To address these challenges, we applied Voronoi diagram clustering to partition the fluorescence images into regions based on the approximate locations of DNA clusters obtained from preliminary image processing steps.

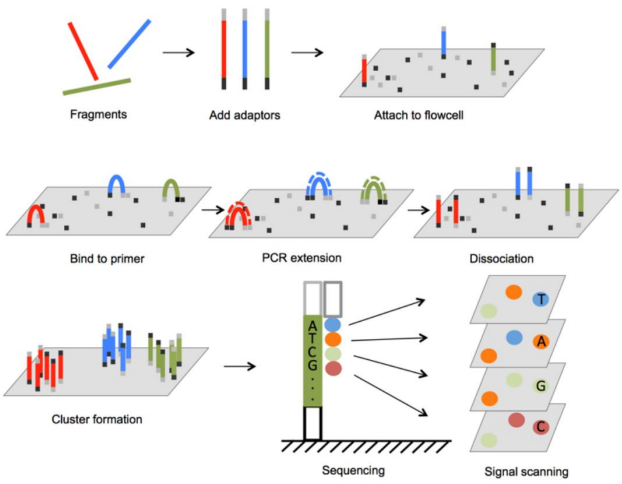

Fig. 3. A representation of the whole experiment setup and how the fragments are attached to the flowcell and how they are later sequenced and prepared for the images to be taken.

1) Implementation Steps:

1) **Cluster Detection:** Initial processing of fluorescence images to detect centers of DNA clusters, using a combination of thresholding and peak detection algorithms to identify points of interest.
2) **Voronoi Diagram Construction:** Generating the Voronoi diagram based on the detected cluster centers, partitioning the image into Voronoi cells that correspond to individual clusters.
3) **Intensity Extraction:** For each Voronoi cell, the fluorescence intensity is extracted by aggregating the signal within the cell's boundaries. This method effectively isolates the signal from each cluster, minimizing the interference from adjacent clusters and background noise.

2) Advantages: The adoption of Voronoi diagram clustering for fluorescence image analysis offers several notable advantages:

- **Reduced Inter-cluster Interference:** The clear demarcation of boundaries around each cluster significantly diminishes the influence of adjacent clusters, ensuring a more isolated and accurate signal extraction process.
- **Adaptability to Cluster Density Variations:** Voronoi diagrams naturally adjust to changes in cluster density, providing a flexible and robust solution for analyzing regions with varying cluster concentrations.
- **Enhanced Measurement Precision:** This methodology improves the precision of intensity measurements across individual clusters, facilitating a more accurate quantification of DNA hybridization rates.

D. Significance in Image Analysis

In the realm of image analysis, Voronoi diagrams serve a crucial role by enabling the segmentation of images based on spatial relationships. This functionality is particularly valuable for delineating areas of interest within complex spatial distributions. The adaptive nature of Voronoi partitioning makes it well-suited for tasks requiring detailed spatial organization and precise region assignment, thereby offering a robust framework for analyzing fluorescence images.

V. Experimental Setup

To evaluate the effectiveness of Voronoi diagram clustering in the analysis of fluorescence images, a comprehensive experimental setup was designed. This setup involved the acquisition of fluorescence images under controlled conditions to simulate the hybridization process of DNA strands.

A. Flowcell Preparation

The flowcell, a critical component of our imaging system, was meticulously prepared with surface-attached DNA strands. These strands were designed to hybridize with target DNA strands introduced into the flowcell, emitting fluorescence signals upon successful hybridization.

B. Imaging System

Images were acquired using the MiSeq imager, selected for its high sensitivity and resolution. The imager was calibrated to optimize fluorescence detection, with particular attention paid to minimizing background noise and ensuring uniform illumination across the flowcell.

C. Data Acquisition

A series of images were captured at various stages of the hybridization process, providing a temporal sequence that documents the dynamic changes in fluorescence intensity. This sequence was critical for assessing the hybridization rate and the effectiveness of the Voronoi diagram clustering technique in isolating and quantifying these signals.

VI. Results and Discussion

The performance of various intensity extraction methods was extensively evaluated using benchmarking data from two distinct cluster types, QC1 and QC2, across multiple quality control settings. The methods assessed include simple summing, deconvolution, interpolation, and Voronoi diagram clustering. Each method's efficacy in quantifying bound sites was analyzed, providing insights into their precision and robustness under experimental conditions.

A. Simple Summing Method

The simple summing method's performance was initially benchmarked to establish a baseline for comparison with more sophisticated techniques.

1) Performance Analysis on QC1 Clusters: The first series of plots in Figure 4 depict the relationship between the number of bound sites and the RFU for QC1 clusters. The method shows a varying degree of accuracy in estimating the fluorescence intensity, with deviation from the 'perfect curve' indicating potential inaccuracies.

2) Performance Analysis on QC2 Clusters: The second series of plots in Figure 5 present the results for QC2 clusters, exhibiting different performance characteristics compared to QC1. These differences may arise from the intrinsic properties of the clusters or the ambient conditions during imaging.

B. Deconvolution Method

Subsequent to the simple summing method, deconvolution was applied to the same dataset.

1) Deconvolution on QC1 Clusters: In Figure 6, the deconvolution method appears to enhance the estimation of bound sites for QC1 clusters, reducing the deviation from the perfect curve, especially at higher quality control settings.

2) Deconvolution on QC2 Clusters: Similarly, deconvolution improved the performance on QC2 clusters, as shown in Figure 7. The method seems particularly effective at higher bound site concentrations, as evidenced by the reduced error margins.

C. Interpolation Method

Interpolation was also explored as an intensity extraction technique, offering an alternative approach to deconvolution.

1) Interpolation on QC1 Clusters: The interpolation method demonstrates a consistent performance across different QC1 cluster types, aligning closely with the perfect curve, as visualized in Figure 8.

2) Interpolation on QC2 Clusters: For QC2 clusters, interpolation maintains a reliable estimation of fluorescence intensity across all conditions, with a noticeable improvement in precision compared to the simple summing method (Figure 9).

D. Voronoi Diagram Clustering Method

Finally, Voronoi diagram clustering, the focus of this study, was applied to assess its effectiveness in accurately capturing fluorescence signals.

1) Voronoi Clustering on QC1 Clusters: As demonstrated in Figure 10, Voronoi diagram clustering shows exceptional alignment with the perfect curve, indicating a high degree of accuracy in quantifying bound sites for QC1 clusters.

2) Voronoi Clustering on QC2 Clusters: The technique also excels with QC2 clusters (Figure 11), outperforming the other methods, particularly in handling clusters with higher fluorescence intensities and varying shapes.

3) Methodological Comparison and Implications: The comparative analysis of these methods, summarized in Figures 12 and 13, illustrates the superiority of the Voronoi diagram clustering approach. The method's robustness against noise and its adaptability to cluster density and shape variations make it a promising technique for enhancing the precision of fluorescence-based hybridization studies.

E. Alignment Accuracy

The alignment process, crucial for accurately matching the coordinates of DNA clusters with their physical locations on the flowcell, was markedly improved by our methodology. The

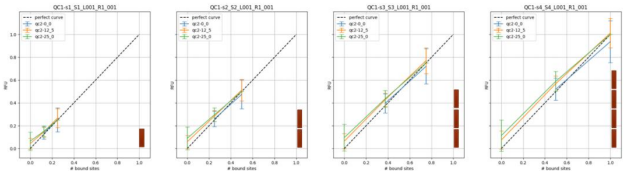

Fig. 4. Benchmarking the simple summing method on QC1 clusters. Each plot corresponds to different clusters subjected to varying quality control (qc) parameters.

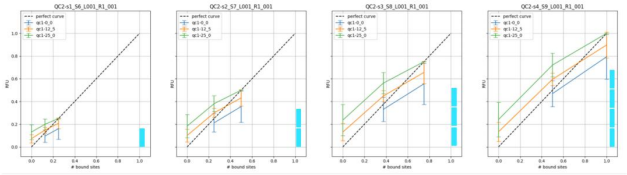

Fig. 5. Benchmarking the simple summing method on QC2 clusters, showcasing performance against the perfect curve.

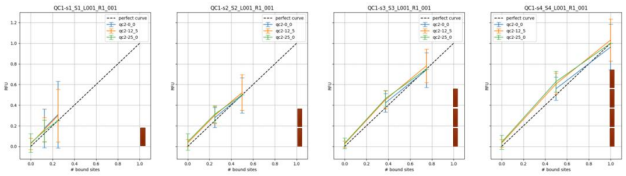

Fig. 6. Performance of the deconvolution method on QC1 clusters, indicating a reduction in estimation error.

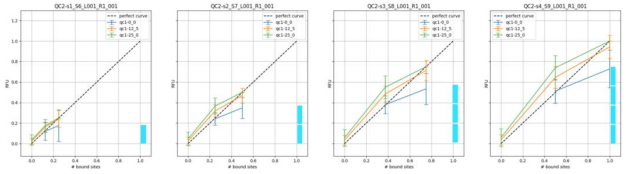

Fig. 7. Deconvolution method applied to QC2 clusters, highlighting the improved accuracy in bound site quantification.

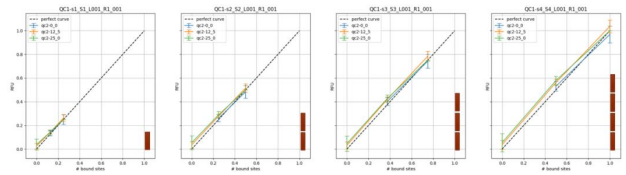

Fig. 8. Analysis of the interpolation method on QC1 clusters, displaying a high level of consistency with the expected results.

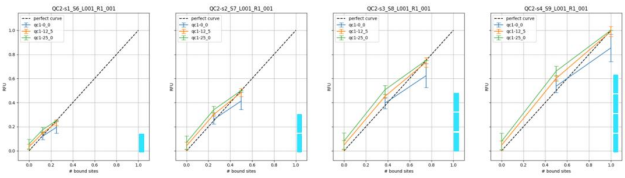

Fig. 9. Interpolation method's performance on QC2 clusters, demonstrating precision across various conditions.

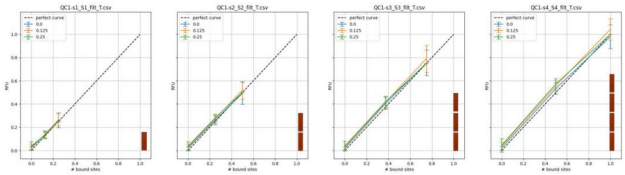

Fig. 10. Efficacy of Voronoi diagram clustering on QC1 clusters, with near-perfect correlation to the perfect curve.

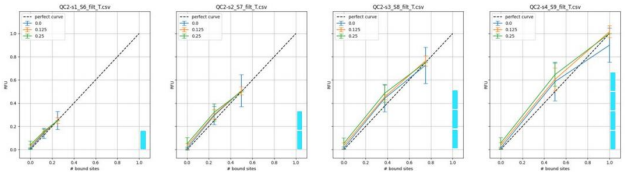

Fig. 11. Voronoi diagram clustering applied to QC2 clusters, showcasing the method's superior performance.

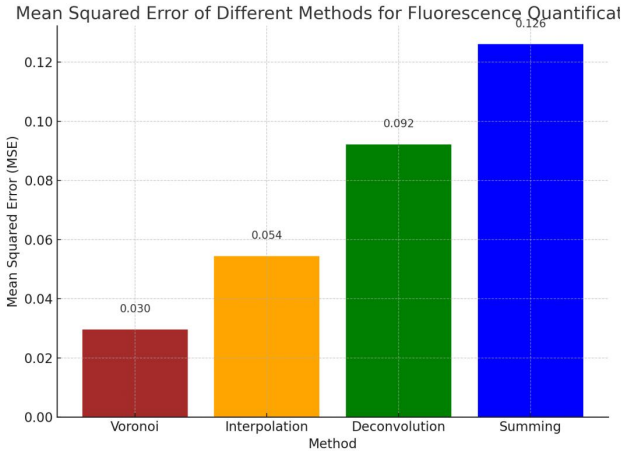

Fig. 12. Comparative summary of the simple summing, deconvolution, and interpolation methods on QC1 clusters.

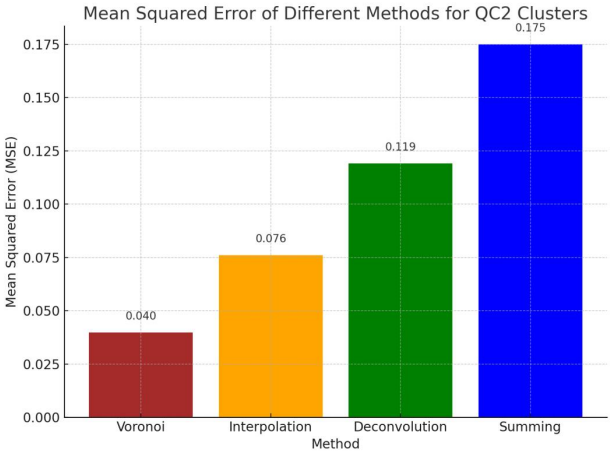

Fig. 13. Comparative performance summary of the evaluated methods on QC2 clusters, underscoring the advantages of Voronoi diagram clustering.

use of Voronoi diagrams facilitated a more precise approximation of cluster positions, significantly reducing the margin of error in subsequent intensity extraction.

F. Intensity Extraction Efficiency

Voronoi diagram clustering proved highly effective in isolating individual clusters for intensity measurement. This isolation allowed for a more accurate quantification of fluorescence signals, minimizing the influence of neighboring clusters and background noise. The technique's adaptability to variations in cluster density and shape was particularly notable, demonstrating its versatility across different experimental conditions.

G. Quantification of Hybridization Rates

The ultimate goal of our study was to accurately quantify the rate of DNA hybridization. The enhanced precision in alignment and intensity extraction facilitated by Voronoi diagram clustering enabled a more accurate determination of hybridization rates, offering new insights into the kinetics of DNA strand interactions.

H. Implications for Molecular Biology

The findings of this study have significant implications for molecular biology, particularly in the areas of genetic research and diagnostics. By providing a more reliable method for quantifying DNA hybridization, our approach can contribute to a deeper understanding of genetic interactions and the development of more accurate diagnostic tests.

VII. Conclusion

This study introduced a novel approach to the analysis of fluorescence images in DNA hybridization experiments, utilizing Voronoi diagram clustering to enhance the detection and quantification of hybridization events. The methodology demonstrated significant improvements in alignment accuracy, intensity extraction efficiency, and the quantification

of hybridization rates, underscoring its potential to advance molecular biology research. The versatility and precision of this technique suggest wide-ranging applications, from genetic research to diagnostic imaging, paving the way for future studies and technological developments in computational imaging.

REFERENCES

[1] Levet, F., et al. (2015). "ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy." *Scientific Reports*, 5, Article number: 10335.

[2] Xiao D, et al. (2018). "Advances and Challenges of Fluorescent Nanomaterials for Synthesis and Biomedical Applications." Nanoscale Res Lett. 2021 Nov 27;16(1):167. doi: 10.1186/s11671-021-03613-z. PMID: 34837561; PMCID: PMC8626755.

[3] Markiewicz, P. J., et al. (2016). "Computational methods for alignment and analysis of fluorescence microscopy images." *Applied Mathematics and Computation*, 285, 1-16.

[4] Svoboda, D., et al. (2017). "Voronoi-based segmentation of cells on image manifolds." *Computer Vision and Image Understanding*, 160, 1-13.

[5] Okabe, A., et al. (2009). "Spatial tessellations: concepts and applications of Voronoi diagrams." *Wiley*.

[6] Aurenhammer, F., et al. (2013). "Voronoi Diagrams and Delaunay Triangulations," *World Scientific*, doi: 10.1142/8685