Published in final edited form as:

Conf Rec Asilomar Conf Signals Syst Comput. 2018 October ; 2018: . doi:10.1109/ACSSC.2018.8645478.

Fusing multimodal microscopy data for improved cell boundary estimation and fluorophore localization of *Pseudomonas aeruginosa*

Scott Ward^{*}, Edward A.K. Cohen^{*}, Niall Adams^{*,†}

^{*}Department of Mathematics, Imperial College London, London, United Kingdom

Scott Ward: scott.ward12@imperial.ac.uk; Edward A.K. Cohen: e.cohen@imperial.ac.uk; Niall Adams: n.adams@imperial.ac.uk

[†]Data Science Institute, Imperial College London, London, United Kingdom

Abstract

With advances in experimental technologies, the use of biological imaging has grown rapidly and there is need for procedures to combine data arising from different modalities. We propose a procedure to combine yellow fluorescence protein excitation and differential interference contrast microscopy time lapse videos to better estimate the cellular boundary of *Pseudomonas aeruginosa* (*P. aeruginosa*) and localization of its type VI secretion system (T6SS). By approximating the shape by an ellipse, we construct a penalized objective function which accounts for both sources; the minimum of which provides an elliptical approximation to their cellular boundaries. Our approach suggests improved localization of the T6SS on the estimated cell boundary of *P. aeruginosa* constructed using both sources of data compared to using each in isolation.

I Introduction

Imaging data is abundant in biology and in order to gain a fuller understanding of microorganisms, multimodal microscopy is a popular and frequently implemented imaging technique. This involves imaging the same biological sample under different microscopes allowing for simultaneous visualisations of various cellular characteristics. The task is then to optimally combine these sources of data for superior and insightful quantitative analysis. Even with advances in biomedical imaging, procedures can still suffer from issues which severely reduce image quality, such as mechanical drift and photobleaching. Photobleaching is more precisely the phenomena of fluorophores moving to a permanently dark state [1]–[3]. Problems such as these adversely affect image quality which, in turn, can severely limit conclusions drawn from the data.

In this paper we propose an approach to extract the spatial pattern of the type VI secretion system (T6SS) on the cell boundary of *Pseudomonas aeruginosa* (*P. aeruginosa*). Our data comes from multimodal microscopy experiments with yellow fluorescence protein (YFP) excitation, highlighting the secretion system, and differential interference contrast (DIC) microscopy. Our YFP microscopy suffers from photobleaching and therefore fades with time, some *P. aeruginosa* even disappearing (see Figure 1). The *P. aeruginosa* are stationary and therefore any movement observed can be considered drift resulting from mechanical aberrations. Additionally over the timescale viewed (2-3 hours) the *P. aeruginosa* do not

reproduce and maintain a near constant size, displaying minimal signs of physical growth. In order to improve estimation of the cellular boundary and locate the secretion systems more accurately, we combine our sources. By constructing a penalized objective function we are able to balance the trade-off between both sources and thus enhance cell shape estimation.

Constructing a penalised objective function is a popular approach to image fusion. A general image fusion methodology is proposed by [4] where common structures between underlying models for the data are used to link and thus construct a penalized objective function. This methodology is adapted by [5] for applications in multifrequency electromagnetic scattering imaging data. Here they propose that the common structure between two images is the gradient of the refractive index of the scene under examination. The refractive index is inhomogeneous and a function of the location in the scene and each image of the same scene is taken using a different frequency of light. Thus the refractive indices in each image are different but the location of changes of the indices are the same, which is used as the *structure operator* [5]. Unlike [5], we suppose that the bacteria in the images are sufficiently well modeled by an ellipse and that differences between the sources arise due to the different techniques used to acquire the data.

Section II outlines the models for the shape of individual *P. aeruginosa* from each data source and how the characteristics of the bacteria, previously discussed, influence the modeling of their morphology, with an overview of our pipeline given in Figure 2. Section III demonstrates the results of our image fusion approach on our data set and makes qualitative comparisons between using each of the sources independently and cohesively.

ϕ and Φ denote the probability density and distribution function of the standard normal distribution respectively. $N_{(a,b)}(\mu, \sigma^2)$ is the truncated Gaussian distribution with mean μ and variance σ^2 restricted to the range (a, b) (similarly for closed and half open/closed intervals such as $[a, b]$ etc).

II Modelling *P. aeruginosa*

In this section we outline the model used to describe the shape of *P. aeruginosa* in the our microscopy time lapse videos. Appendix V-A and V-B outline the image preprocessing used to extract features from the DIC and YFP videos respectively.

A DIC Microscopy

Suppose that our DIC videos have $\tilde{n}$ frames and that we are able to isolate bacteria i in $n - \tilde{n}$ frames of the video (i.e. bacteria i is segmented successfully from neighboring bacteria in n frames). From each of the n frames where bacteria i is isolated we observe $a_{i,j}$ and $b_{i,j}$ for $j \in \{1, \dots, n\}$, the major and minor axis length of bacteria i respectively. We model bacteria i independent of all other bacteria $j \neq i$ and $a_{i,j}$ and $b_{i,j}$ as,

$$\begin{aligned}\frac{1}{a_{i,j}^2} &= \frac{1}{\tilde{a}_i^2} + \epsilon_{i,j,\tilde{a}_i} \\ \frac{1}{b_{i,j}^2} &= \frac{1}{\tilde{b}_i^2} + \epsilon_{i,j,\tilde{b}_i}\end{aligned}\quad (1)$$

where,

$$\begin{aligned}\epsilon_{i,j,\tilde{a}_i} \mid \epsilon_{i,j,\tilde{b}_i} &\sim N_{(-1/\tilde{a}_i^2, 1/\tilde{b}_i^2 - 1/\tilde{a}_i^2 + \epsilon]}^{(0, \sigma_{\tilde{a}_i}^2)} \\ \epsilon_{i,j,\tilde{b}_i} &\sim N_{(-1/\tilde{b}_i^2, \infty)}^{(0, \sigma_{\tilde{b}_i}^2)}\end{aligned}$$

and $\tilde{a}_i, \tilde{b}_i$ are parameters of the model representing the true major and minor axis lengths of bacteria i respectively. To simplify notation we remove the subscript i noting that we handle each bacteria independently. The use of truncated Gaussian random variables emulates the fact that the major axis of *P. aeruginosa* can never be smaller than its minor axis. It can be shown that the log likelihood under Equations 1 is,

$$\begin{aligned}\ell &= \sum_{i=1}^n \log \phi\left(\frac{y_i - \tilde{y}}{\sigma_{\tilde{y}}}\right) + \log \phi\left(\frac{z_i - \tilde{z}}{\sigma_{\tilde{z}}}\right) - \log \sigma_{\tilde{y}} \\ &\quad - \log \sigma_{\tilde{z}} - \log \left[\Phi\left(\frac{z_i - \tilde{y}}{\sigma_{\tilde{y}}}\right) + \Phi\left(\frac{\tilde{y}}{\sigma_{\tilde{y}}}\right) - 1 \right] \\ &\quad - \log \Phi\left(\frac{\tilde{z}}{\sigma_{\tilde{z}}}\right),\end{aligned}\quad (2)$$

where $y_i = 1/a_i^2$, $z_i = 1/b_i^2$, $\tilde{y} = 1/\tilde{a}^2$, $\tilde{z} = 1/\tilde{b}^2$, $\sigma_{\tilde{y}} = \sigma_{\tilde{a}}^2$, and $\sigma_{\tilde{z}} = \sigma_{\tilde{b}}^2$.

Over the course of the videos the bacteria do not exhibit significant changes in size. Additionally, the physical characteristics of *P. aeruginosa* means that the probability of a_i being *close* to b_i is negligible. This can be modeled adequately when $\sigma_{\tilde{a}}^2$ and $\sigma_{\tilde{b}}^2$ small. The following result provides analytic maximum likelihood estimates (MLEs) for $\tilde{a}$ and $\tilde{b}$ of Equations 1 in this scenario.

Proposition. Let $\mathbf{Z} = \{(a_1, b_1), \dots, (a_n, b_n)\}$ for $n \in \mathbb{N}$ be data following Equations 1. Then the log likelihood is given as Equation 2 and as $\sigma_{\tilde{a}}^2, \sigma_{\tilde{b}}^2 \rightarrow 0$, the likelihood is maximized at,

$$\hat{a}^2 = \frac{n}{\sum_{i=1}^n \frac{1}{a_i^2}}, \quad \hat{b}^2 = \frac{n}{\sum_{i=1}^n \frac{1}{b_i^2}}. \quad (3)$$

Proof. See Appendix V-C.

B YFP Microscopy

By using a temporal edge detection algorithm [7] (see Appendix V-B) we extract contours of constant pixel intensity surrounding individual bacteria. More precisely, for the i^{th} bacteria we observe $Z_i = \{(x_1, y_1)^T, \dots, (x_m, y_m)^T\}$, a set of m points lying on a contour surrounding a bacteria.

We then model these points as lying on an ellipse with added Gaussian noise,

$$\frac{((x_i - x_0)\cos(\theta) + (y_i - y_0)\sin(\theta))^2}{\tilde{a}^2} + \frac{((x_i - x_0)\sin(\theta) - (y_i - y_0)\cos(\theta))^2}{\tilde{b}^2} = 1 + \xi_i, \quad (4)$$

where $\xi_i \sim \mathcal{N}(0, \sigma^2)$, $i = 1, \dots, m$ are independent. The orientation of the ellipse, measured anti-clockwise from the horizontal axis, is given by θ whilst (x_0, y_0) represents the center of the ellipse. Defining the left hand side of Equation 4 as $g(x_0, y_0, \tilde{a}, \tilde{b}, \theta)$, this model can be fit by minimizing the sum of squares,

$$q(x_0, y_0, \tilde{a}, \tilde{b}, \theta) = \sum_{i=1}^m (g(x_0, y_0, \tilde{a}, \tilde{b}, \theta) - 1)^2, \quad (5)$$

which is also referred to as the algebraic distance in the conic fitting literature [6], [8].

C Combining the DIC and YFP Microscopy Data

Frequently, in conic fitting, a constraint is applied when optimising equations such as Equation 5, in order to remove trivial solutions and multiples of a solution [6]. By imposing a penalty which combines both our data sources we are able to estimate approximate ellipses and avoid trivial solutions simultaneously. In order to combine both sources of data we propose optimizing

$$\tilde{q}(x_0, y_0, \tilde{a}, \tilde{b}, \theta) = \sum_{i=1}^m (g(x_0, y_0, \tilde{a}, \tilde{b}, \theta) - 1)^2 + m\lambda(\tilde{a} - \hat{a})^2 + m\mu(\tilde{b} - \hat{b})^2, \quad (6)$$

where $\hat{a}$ and $\hat{b}$ are the estimates from the previous Proposition. The tuning parameters λ and μ are chosen using K -fold cross validation [9].

III Results

By combining both sources of data we can clearly see qualitative improvements of the elliptical fits by examining Figure 3. In particular, the contour points extracted from the pseudo-color images generated using our preprocessing steps (see Appendix V-B) do not accurately outline the bacteria, rather they suggest an elliptical approximation that is larger

than the bacteria size. By imposing constraints on the major and minor axis lengths through the MLEs of Proposition II-A in the penalized objective function 6, we avoid this over-sizing.

Conversely, we also see improvements when using both sources rather than the DIC microscopy data in isolation. Due to the experimental noise there is no guarantee that the DIC microscope will align exactly with the YFP microscope. This is apparent as the yellow ellipse does not overlay accurately with the underlying YFP pseudo-color image in Figure 3. By optimizing Equation 5 we allow for the center of the ellipse to adjust to the contours of the YFP microscopy data and position the estimate more appropriately, which in turn improves the localization of the T6SS with respect to the cell boundary.

IV Conclusions

In this work we have proposed an approach to combining multimodal microscopy data, in particular focusing on cellular videos of *P. aeruginosa* using DIC and YFP microscopy techniques. For these data, we observe superior elliptical estimation when combining both sources using our penalized objective function and have been able to alleviate issues which would otherwise arise when considering the data independently of each other.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We would like to thank Sara Planamente and Alain Filloux in the Department of Life Sciences, Imperial College London, and Paul Freemont in the Department of Medicine, Imperial College London for their ongoing discussions and providing the data used in this work.

This research is supported by Wellcome Trust under grant number 210298/Z/18/Z

References

- [1]. Bernas T, Rajwa BP, Asem EK, Robinson JP. Loss of image quality in photobleaching during microscopic imaging of fluorescent probes bound to chromatin. *Journal of Biomedical Optics*. 2005; 10(6)
- [2]. Peng X, Andegeko Y, Weisel LR, Lozovoy VV, Dantus M. Greater signal, increased depth, and less photobleaching in two-photon microscopy with 10 fs pulses. *Optics Communications*. 2008; 281(7):1841–1849.
- [3]. Hoebe RA, Van Oven CH, Gadella TWJ Jr, Dhonukshe PB, Van Noorden CJF, Manders EMM. Controlled light-exposure microscopy reduces photobleaching and phototoxicity in fluorescence live-cell imaging. *Nature Biotechnology*. 2007; 25(2):249–253.
- [4]. Haber E, Oldenburg D. Joint inversion: a structural approach. *Inverse Problems*. 1997; 13(1):63–77.
- [5]. Fatone L, Maponi P, Zirilli F. An image fusion approach to the numerical inversion of multifrequency electromagnetic scattering data. *Inverse Problems*. 2001; 17(6):1689–1701.
- [6]. Fitzgibbon AW, Fisher RB, Hill F, Eh E. Direct Least Squares Fitting of Ellipses. 1996:1–15.
- [7]. Mashanov GI, Molloy JE. Automatic detection of single fluorophores in live cells. *Biophysical Journal*. 2007; 92(6):2199–2211. [PubMed: 17208981]

- [8]. Gander W, Golub GH, Strebels R. Fitting of circles and ellipses-least squares solution. SVD and Signal Processing III. 1994; 34(4):558–578.
- [9]. Klipp, E, Herwig, R, Kowald, A, Wierling, C, Lehrach, H. Chapter 9 Systems biology in practice: concepts, implementation and application. John Wiley & Sons; 2008. 327
- [10]. Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979; 9(1):62–66.
- [11]. Chan TF, Vese LA. Active contours without edges. IEEE Transactions on Image Processing. 2001; 10(2):266–277. [PubMed: 18249617]
- [12]. Sohrab, HH. Chapter 3: Limits of functions Basic Real Analysis. 2nd ed. Birkhäuser; 2003. 109

Fig. 1.

Top row: (1) First frame in a YFP experiment of *P. aeruginosa* in which the intensity is high, (2) Last frame in a YFP experiment of *P. aeruginosa* in which the intensity is low, (3) First frame in a DIC experiment of *P. aeruginosa*, (4) Last frame in a DIC experiment of *P. aeruginosa* and shows that DIC microscopy does not suffer from similar issues. *Bottom row:* (1) Image of 2 adjacent *P. aeruginosa* in the first frame of a YFP experiment where the intensity is high, (2) Image of 2 adjacent *P. aeruginosa* in the last frame of a YFP experiment where the intensity is low, (3) Image of 2 adjacent *P. aeruginosa* in the first frame of a DIC experiment, (4) Image of 2 adjacent *P. aeruginosa* in the last frame of a DIC experiment. The bottom row highlights the effects of mechanical drift as the pair of bacteria have shifted marginally to the left.

Fig. 2.

Our image processing pipeline. Starting from the DIC microscopy data (1a), we binarise each frame and extract the major and minor axis lengths of each isolated bacteria (1b) using procedures outlined in Appendix V-A and then model this data using Equations 1. Starting from the YFP microscopy data (2a), we extract points along the contour as well as the fluorophore localisations (2b) using image processing procedures outlined in Appendix V-B and then model this data using the algebraic distance (2c) [6]. In step (3) we combine the

data sources using the estimates of (1c) to penalize the least squares objective function of step (2c).

Fig. 3.

Examples of elliptical fits to individual *P. aeruginosa*, red ellipses indicate fits using the YFP microscopy data only, yellow ellipses indicate fits using the DIC microscopy data only and purple give the fitted ellipses when combining both sources of data. The images on the left are the pseudo-colour images generated from the YFP microscopy videos using the temporal edge detection algorithm [7]. The white points are the contour points extracted from the pseudo-color images for each of the bacteria. The image on the right exemplifies the final result on an entire data stream, the ellipses represent the bacterial approximations using our penalized optimization approach whilst the asterisks indicate the position of the fluorophores.