RS-FISH: Precise, interactive and scalable smFISH spot detection using Radial Symmetry

Ella Bahry[1,*], Laura Breimann[1,2,*], Leo Epstein[3,*], Klim Kolyvanov[1], Kyle I. S. Harrington[3,&], Timothée Lionnet[4,&], Stephan Preibisch[1,5,&]

[1]Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany
[2]Department of Biology, Center for Genomics and Systems Biology, New York University, New York, NY, USA
[3]Helmholtz Imaging Platform, Max Delbrück Center for Molecular Medicine, Berlin, Germany
[4]Institute for Systems Genetics and Department of Cell Biology, NYU School of Medicine, New York, NY, USA
[5]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

* equal contribution
& corresponding authors

keywords: smFISH, image analysis, computer vision, transcription, microscopy

Abstract:
Studying transcription using single-molecule RNA-FISH (smFISH) is a powerful method to gain insights into gene regulation on a single cell basis, which relies on accurate identification of sub-resolution fluorescent spots in microscopy images. Here we present Radial Symmetry-FISH (RS-FISH), which can robustly and quickly detect even close smFISH spots in two and three dimensions with high precision, allows interactive parameter tuning, and can easily be applied to large sets of images.

Availability and implementation:
RS-FISH is an open-source implementation written in Java/ImgLib2 and provided as a macro-scriptable Fiji plugin. Source code, tutorial, documentation, and example images are available at: https://github.com/PreibischLab/RadialSymmetryLocalization

Contact:
kyle.harrington@mdc-berlin.de, Timothee.Lionnet@nyulangone.org, stephan.preibisch@mdc-berlin.de / preibischs@janelia.hhmi.org

-

Introduction

Sequencing and imaging methods have been developed to study transcription in single cells [1,2], which has been described as a complex, stochastic process with irregular bursts of transcription separated by long periods of inactivity [3–5]. Single-molecule fluorescence in situ hybridization (smFISH) that visualizes individual mRNA molecules in microscopy images gained much popularity and has been applied to a range of specimens [4,6]. Extracting information from smFISH experiments relies on single-molecule spot detection that is precise and easy-to-use for experimentalists. Since detection of more subtle transcriptional changes requires analysis of thousands of smFISH images [7–9] and imaging of entire specimens (**Fig. 1a, Supplementary Material**) opens new avenues for studying transcriptional regulation in the context of entire organisms, smFISH spot detection should also be fast and easy to parallelize. Several methods are available to the community, however, commonly used packages lack interactive data analysis functionality and show limited applicability to larger datasets due to slower processing times and challenging automation and distribution processes [6,10–12]. Here we present RS-FISH that extends Radial Symmetry [13] (RS) in order to identify smFISH spots, which we release as an open-source, scriptable Fiji plugin [14]. It robustly and quickly detects single-molecule spots in 3D with high precision even when they are in close proximity. Importantly, RS-FISH allows interactive parameter adjustment to efficiently identify correct parameters for any given type of smFISH image, which can subsequently be applied in a straight-forward way to large datasets consisting of thousands of images.

Method and Implementation

RS is an efficient non-iterative alternative to Gaussian fitting methods commonly used for point localization. RS was developed initially for 2D subpixel localization of circular objects by computing the intersection point of image gradients (**Fig. 1a**) [13]. Here, we extend RS to detect smFISH spots in anisotropic three-dimensional (3D) image datasets (**Fig. 1**). First, we derived a 3D version of RS that can be straightforwardly extended to higher dimensions that may have potential for spatiotemporal localization of blinking 3D spots (**Supplemental Material**). Second, we extended RS to detect axis-aligned, ellipsoid objects, enabling RS-FISH to account for typical anisotropy in 3D microscopy datasets that result from different pixel sizes and point spread functions (PSF) in lateral (x,y) compared to axial (z) dimensions (**Supplemental Material**). Third, the speed of the closed-form computation of RS allowed us to combine RS with robust outlier removal using Random Sample Consensus [15] (RS-RANSAC) in order to identify sets of image gradients that support the same ellipsoid object given a specific error for the gradient intersection point (**Supplemental Material**). This allows RS-FISH to: identify the set pixels that support a certain localization error for a specific smFISH spot (**Fig. 1a**), separate close detections (**Fig. 1b**), and ignore outlier pixels with high noise that would disturb the localization process (e.g. dead or hot camera pixels).

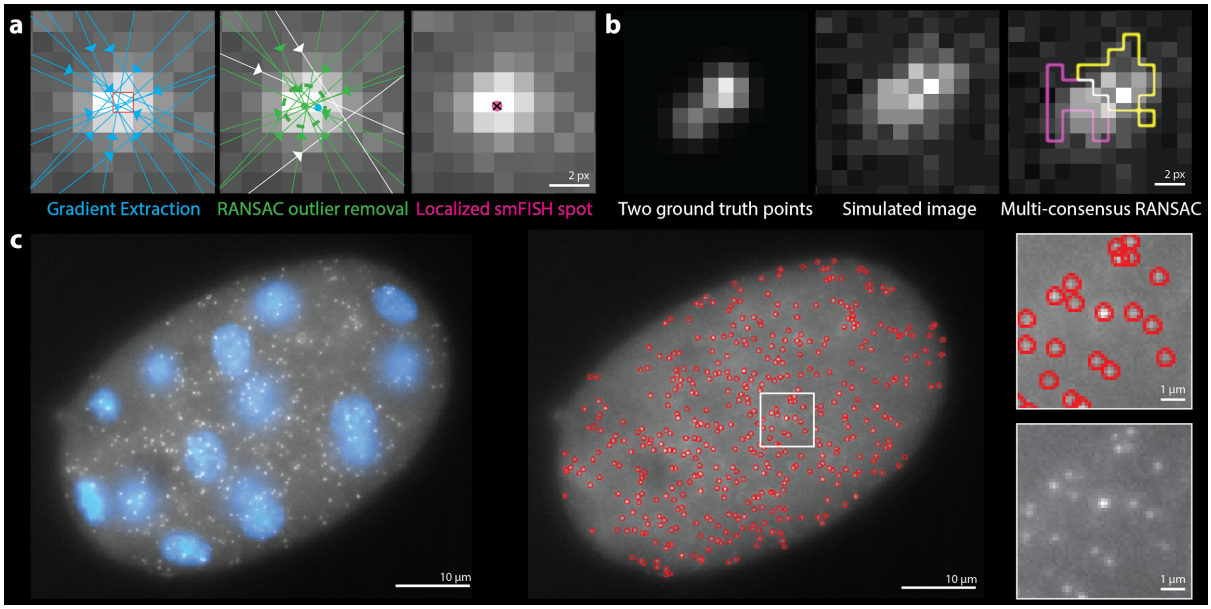

Figure 1: RS-FISH accurately detects smFISH spots. (**a**) Illustrates the concept of single smFISH spot detection using Radial Symmetry-based RANSAC. (left) Blue lines indicate the gradients that are calculated in a local patch around a DoG-detected location (red square) for radial symmetry fitting. The location for which example gradients are shown is indicated by arrowheads. (middle) Gradients that agree on a common center point (green gradients, green dot) given a defined error (green dotted line) are identified using RANSAC outlier removal. Using all gradients would lead to a different center point (blue). (right) RS-FISH detects the center of the smFISH spot by computing the intersection point of all inlier gradients (pink dot with the black cross). (**b**) Detection of two close points using multi-consensus RANSAC. (left) Two ground truth points in close proximity shown without noise. (middle) Both points are detected as a single spot during DoG due to simulated low SNR. (right) Using the multi-consensus RANSAC, the area was identified to consist of two independent detections. The two sets of gradients (i.e. pixels) used to detect each spot are highlighted in yellow and pink. (**c**) Max intensity projection of 5 central slices of an early *C. elegans* embryo with smFISH staining for dpy-23 mRNAs (white) and DAPI (blue) (Methods for smFISH staining in Supplemental Material). The image with red circles is a screenshot from the Fiji plugin and shows all smFISH spots detected by RS-FISH.

To compute RS-FISH, we first generate a set of seed points by thresholding the Difference-of-Gaussian (DoG) filtered image that identifies potential locations of smFISH spots. Image gradients are then extracted from local patches around each of these spots (**Fig. 1a**), which are optionally corrected for non-uniform fluorescence backgrounds. Before RS localization, gradients are transformed to match the anisotropy of the dataset using an anisotropy coefficient that depends on pixel spacing, resolution, and PSF. The anisotropy coefficient does not change as long as these parameters are held constant. RS-FISH provides a tool to compute the anisotropy coefficient from any smFISH image directly using RS-RANSAC or Gaussian fits (**Supplemental Material**). Optionally, RS-RANSAC can be run in multi-consensus/multi-model mode, which allows detecting individual smFISH spots that were too close for the DoG detector to separate them during seed point generation (**Fig. 1b**). This is achieved by consecutively running RANSAC on all gradients of a local patch until all sets of pixels that support a certain localization error are identified (**Supplemental Material**). Finally, all detections are filtered by their distance to each other to avoid potential double detections, where detections with a distance below 0.5 pixels are disregarded. Each spot's associated intensity

value is computed using linear interpolation or by fitting a Gaussian only to the pixels that support each spot as identified by RS-RANSAC.

RS-FISH pixel operations are implemented in ImgLib2 [16], RS fitting and RANSAC are built on top of the image transformation framework *mpicbg* [17]. Importantly, RS-FISH's main parameters can be interactively tuned within the Fiji plugin to identify the best set of parameters for each type of smFISH experiment (**Supplemental Material**). Once the right set of parameters is identified, RS-FISH can be run in a scriptable advanced mode that also allows straight-forward parallel execution on compute clusters, for which we provide example scripts (**Supplemental Material**).

Results and Discussion

To validate and benchmark RS-FISH results, we performed comparisons against the established FISH-Quant software [10], and its new Python version BigFish where necessary. We use (a) simulated smFISH images with varying noise levels to assess detection performance and (b) real datasets of entire *C. elegans* embryos we acquired for runtime measurements. For (a), we perform parameter grid-searches that maximize the number of correctly identified spots (**Table 1, Supplemental Material**), which show that RS-FISH and FISH-Quant/BigFish show similar detection and localization performance. For (b), we run both RS-FISH and FISH-Quant with reasonable settings to identify smFISH spots, where we show that RS-FISH processing speed is roughly 10-fold improved (**Table 1, Supplemental Material**).

	RS-FISH	**FISH-Quant / BigFish**
smFISH spots missed (avg/median/stdev)	1.10% / 0.00% / 1.92%	2.99% / 0.00% / 7.75%
smFISH spot localization error (avg/median/stdev)	0.40px / 0.21px / 0.44px	0.34px / 0.27px / 0.27px
smFISH processing speed increase (relative speedup factor)	13.7x	1x

Table 1: Localization and Speed comparison. The first row compares the percentage of smFISH spots that were missed in simulated datasets, and the second row compares the localization error for the spots that were identified correctly. The last row compares the computation speed for real 3D volumes of smFISH acquisitions of entire *C. elegans* embryos. See supplement for details and all raw numbers underlying these statistics.

Importantly, we believe that RS-FISH has four major strengths compared to existing solutions. First, it allows interactive adjustment of all key parameters, thus making the task of correctly detecting most smFISH spots in any microscopy image as accessible as possible to experimentalists. Second, processing speed is significantly improved while achieving similar localization performance. Third, RS-FISH brings smFISH detection to the Fiji ecosystem, allowing simple installation and running while providing macro-recording functionality to easily automate smFISH detection on many images using local computers or the cluster/cloud. Finally, we developed a generic derivation of 3D Radial Symmetry for anisotropic objects that

can scale to the n-dimensional case and is efficiently implemented for a powerful fitting framework supporting robust outlier removal. This enables using Radial Symmetry for identifying arbitrary round objects such as nuclei, thus expanding the scope of Radial Symmetry beyond single-molecule detection.

Acknowledgments:

We would like to thank Stephan Saalfeld for providing the *mpicbg* framwork, the System Biology imaging platform and Andrew Woehler for technical support, Dhana Friedrich for discussions about the *C. elegans* smFISH imaging protocol, and Florian Mueller and Arthur Imbert for their advice on running BigFish. E.B. was supported by HFSP grant RGP0021/2018-102 and MDC-Berlin, L.B. was supported by MDC-Berlin and the Joachim Herz Foundation (#850022), L.E. and K.H were supported by MDC-Berlin and the Helmholtz Imaging Platform, K.K. was supported by MDC-Berlin, T.L. was supported by NIH grant R01 GM127538 and Melanoma Research Foundation Team Science award # 687306, and S.P. was funded by MDC-Berlin and HHMI Janelia.

References

1. Zenklusen, D.; Larson, D.R.; Singer, R.H. Single-RNA Counting Reveals Alternative Modes of Gene Expression in Yeast. *Nature structural & molecular biology* **2008**, *15*, 1263–1271, doi:10.1038/nsmb.1514.
2. Itzkovitz, S.; Oudenaarden, A. van Validating Transcripts with Probes and Imaging Technology. *Nature Methods* **2011**, *8*, S12–S19, doi:10.1038/nmeth.1573.
3. Li, G.-W.; Xie, X.S. Central Dogma at Single-Molecule Level in Living Cells. Nature **2011**, *475*, 308–15.
4. Raj, A.; Peskin, C.S.; Tranchina, D.; Vargas, D.Y.; Tyagi, S. Stochastic MRNA Synthesis in Mammalian Cells. *PLoS Biology* **2006**, *4*, e309, doi:10.1371/journal.pbio.0040309.
5. Tunnacliffe, E.; Chubb, J.R. What Is a Transcriptional Burst? *Trends in Genetics* **2020**, *36*, 288–297.
6. Raj, A.; Bogaard, P. van den; Rifkin, S.A.; Oudenaarden, A. van; Tyagi, S. Imaging Individual MRNA Molecules Using Multiple Singly Labeled Probes. *Nature Publishing Group* **2008**, *5*, 877–879.
7. Urban, E.A.; Johnston, R.J. Buffering and Amplifying Transcriptional Noise During Cell Fate Specification. *Frontiers Genetics* **2018**, *9*, 591, doi:10.3389/fgene.2018.00591.
8. Shah, S.; Lubeck, E.; Schwarzkopf, M.; He, T.-F.; Greenbaum, A.; Sohn, C.H.; Lignell, A.; Choi, H.M.T.; Gradinaru, V.; Pierce, N.A.; et al. Single-Molecule RNA Detection at Depth by Hybridization Chain Reaction and Tissue Hydrogel Embedding and Clearing. *Development* **2016**, *143*, 2862–2867, doi:10.1242/dev.138560.
9. Sylwestrak, E.L.; Rajasethupathy, P.; Wright, M.A.; Jaffe, A.; Deisseroth, K. Multiplexed Intact-Tissue Transcriptional Analysis at Cellular Resolution. Cell 2016, *164*, 792–804, doi:10.1016/j.cell.2016.01.038.
10. Mueller, F.; Senecal, A.; Tantale, K.; Marie-Nelly, H.; Ly, N.; Collin, O.; Basyuk, E.; Bertrand, E.; et al. FISH-Quant: Automatic Counting of Transcripts in 3D FISH Images. Nature Methods **2013**, *10*, 277–278.
11. Tsanov, N.; Samacoits, A.; Chouaib, R.; Traboulsi, A.-M.; Gostan, T.; Weber, C.; Zimmer, C.; Zibara, K.; Walter, T.; Peter, M.; et al. SmiFISH and FISH-Quant – a Flexible Single RNA Detection Approach with Super-Resolution Capability. *Nucleic Acids Research* **2016**, *44*, e165–e165, doi:10.1093/nar/gkw784.
12. Wu, A.C.-Y.; Rifkin, S.A. Aro: A Machine Learning Approach to Identifying Single Molecules and Estimating Classification Error in Fluorescence Microscopy Images. *Bmc Bioinformatics* **2015**, *16*, 102.
13. Parthasarathy, R. Rapid, Accurate Particle Tracking by Calculation of Radial Symmetry Centers. Nature Methods **2012**, *9*, 724–726, doi:10.1038/nmeth.2071.
14. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. Nature Methods **2012**, 9, 676–682.
15. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. In; Morgan Kaufmann: San Francisco (CA), **1987**; 726–740.
16. Pietzsch, T.; Preibisch, S.; Tomancak, P.; Saalfeld, S. ImgLib2--Generic Image Processing in Java. Bioinformatics **2012**, *28*, 3009–3011, doi:10.1093/bioinformatics/bts543.
17. Saalfeld, S.; Fetter, R.; Cardona, A.; Tomancak, P. Elastic Volume Reconstruction from Series of Ultra-Thin Microscopy Sections. *Nat Methods* **2012**, *9*, 717–720, doi:10.1038/nmeth.2072.