Pattern Recognition in Super Resolution Images

Cross, Daniel; Wand, Nathaniel O.; Styles, Iain B.; Thomas, Steven G.; Poulter, Natalie S.

DOI:
10.1364/BODA.2015.JT3A.7

Document Version
Peer reviewed version

Citation for published version (Harvard):
Cross, D, Wand, NO, Styles, IB, Thomas, SG & Poulter, NS 2015, Pattern Recognition in Super Resolution Images. in *Optics in the Life Sciences.* Optical Society of America, pp. JT3A.7, Bio-Optics: Design and Application 2015, Vancouver, Canada, 12/04/15. https://doi.org/10.1364/BODA.2015.JT3A.7

Pattern Recognition in Super Resolution Images

Daniel J. Cross[1], Nathaniel O. Wand[1], Iain B. Styles[1,2], Steven G. Thomas[3] and Natalie S. Poulter[3]
[1]PSIBS Doctoral Training Centre, University of Birmingham, Edgbaston, Birmingham, UK;
[2]School of Computer Science, University of Birmingham, Edgbaston, Birmingham, UK;
[3]Institute of Biomedical Research, University of Birmingham, Edgbaston, Birmingham, UK;
D.J.C.: E-mail: DJC919@bham.ac.uk,
N.O.W.: Email: NOW348@bham.ac.uk,
I.B.S.: E-mail: I.B.Styles@cs.bham.ac.uk,
S.G.T.: Email: S.Thomas@bham.ac.uk,
N.S.P.: Email: N.Poulter@bham.ac.uk

Abstract: Super resolution microscopy techniques allow for high precision localisation of individually labeled GPVI receptors. In this work we apply a quadtree-based pattern recognition scheme to identify patterns and structure from STORM images.
OCIS codes: (100.6640) Superresolution; (100.5010) Pattern Recognition;

1. Introduction

The GPVI collagen receptor has long been known to play a key role in the activation and aggregation of platelets following trauma. However, what is less well understood is whether the spatial localisation of the receptor is indicative of the underlying biochemistry of thrombus formation and evolution [1].

The activation of the GPVI receptor is typically the result of damage to a vessel wall, exposing the highly collagenous sub-endothelial matrix to the contents of the circulatory system. Initially, platelets are drawn out of circulation by interaction of the GPIb-IX-V receptor complex with von Willebrand Factor; following this, GPVI receptors present on the platelet surface bind to the exposed collagen matrix, resulting in spreading and adhesion of further platelets to form a thrombus.

Issues with elucidating the process of thrombus formation are mainly encountered when attempting to quantify the interactions of the multimeric complexes responsible, which often occur on spatial scales below that of the diffraction limit of light. With the development of super resolution imaging systems capable of resolving structure on the order of 20 nanometers, this limit has effectively been overcome, allowing for investigations into the spatial distributions of surface proteins and receptors to be undertaken, as can be seen in the STORM image and point localisations shown in figure 1 and 2.

Previous studies have studied the effectiveness of using a combination of Ripley's K-function and heat mapping to produce a density based threshold used to identify structure in super resolution images [1]. This has lead to investigations into the interactions between clusters of the Lat adaptor protein and T-cell antigen receptor (TCR), ultimately revealing that adaptor clusters assembled prior to signaling events do not undergo phosphorylation or transport to activation sites during TCR signaling [2].

In addition to identifying structures within images, clustering algorithms are capable of quantifying various different parameters associated with these structures. These include the spacing and density of points within clusters, in addition to their size and shape, allowing for automated extraction of parameters that would have previously required analysis by hand.

2. Background

Methods of clustering can be loosely grouped into two major families, those that rely on the distance metric between points to determine their cluster assignment, and those that utilise local point density.

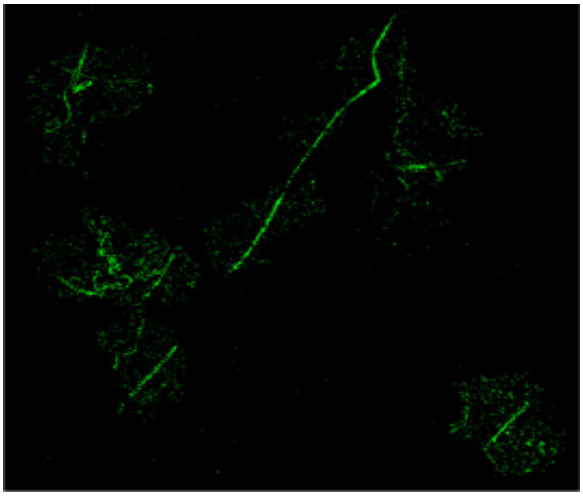

Fig 1: STORM image of fluorescently labelled GPVI molecules in platelets on a collagenous surface. Image courtesy of N.S. Poulter.

Fig 2: GPVI point localisations extracted from a STORM image of platelets adhered to a collagenous surface. Image courtesy of N .S. Poulter

Quadtree decomposition belongs to the family of clustering techniques which rely on evaluating the density of points to be clustered. This is particularly useful when dealing with long filamentary structures, such as the distribution of GPVI along collagenous fibres in adherent platelets, as there is no inherent bias towards forming specific cluster shapes.

The algorithm uses the following process to perform the decomposition:

- Divide the dataset into four subsets of equal dimensions.
- Count the population of each subset; if the count is greater than the minimum cluster size defined by the operator, label the subset for further division, else assign a label to each point in the subset corresponding to the division rank at which the subdivision check fails.
- Once all subsets of a given rank have been assessed, repeat step 2 until no further divisions can be made.

Once completely decomposed, elements of the dataset labelled similarly can be said to reside in areas of equal density (as shown in figure 3). However, in doing this all information regarding the horizontal relationships between data points is lost i.e. the membership of two separate points to the same density rank or even subset doesn't necessarily imply their membership of the same cluster.

In order to extract the relationships between elements, a percolation algorithm was applied to the decomposed dataset using the following method:

- Label all elements of a given density rank with identical labels, calculate euclidean distance between all points.
- Iterate through all points, determining the points that lie within a distance of half the associated subdivision. Check these points for an existing cluster label.
- If already clustered, apply the existing cluster label to the new point; if not, apply a unique label to both points. Repeat across all points.

3. Results

Figures 3 and 4 qualitatively demonstrate the ability of the quadtree decomposition algorithm to identify clusters in point localisation datasets. The filamentary structure visible in the highest density regions of figure 3, those coloured dark blue/purple, are characteristic of adherent platelets indicating this method's suitability for analysing thise particular type of data. Whilst figure 4 does not contain structures of filamentary nature, it does contain multiple small circular objects, which the algorithm readily identifies. This is important, as it shows that whilst the algorithm was selected for its non-bias towards circular structures, it is still capable of identifying them.

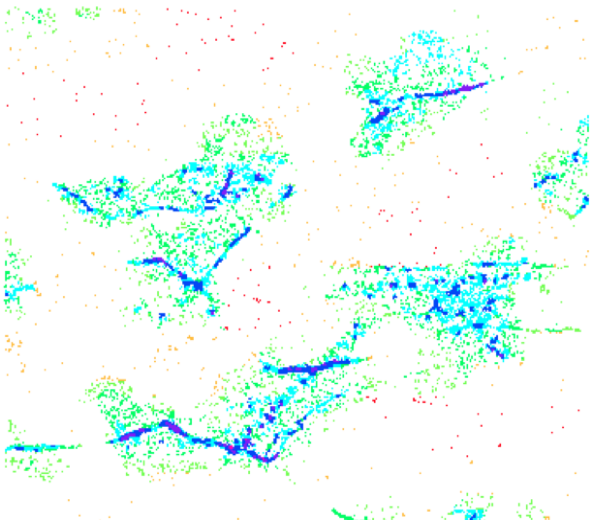

Fig 3: Results of the quadtree decomposition algorithm applied to the dataset shown in figure 2. Minimum cluster size is was defined as 100 points, and varying density ranks are denoted by a rainbow colour map, with red representing lowest density rank and purple representing highest.

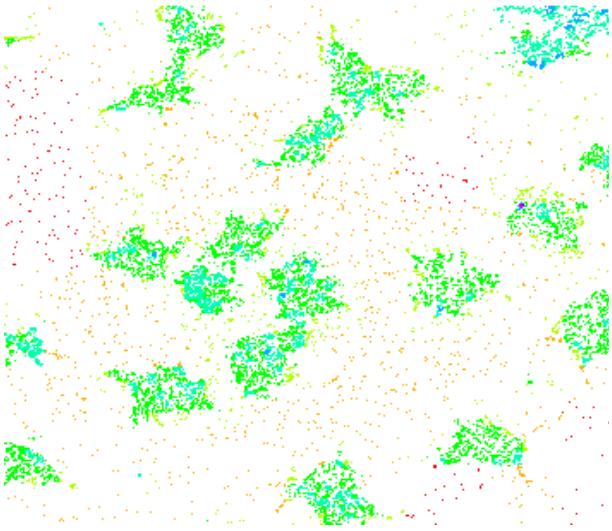

Fig 4: Results of quadtree decomposition on a wide field dataset, minimum cluster size defined as 100 points with a rainbow density colour map applied.

The promising nature of these observations indicates that the quadtree algorithm will likely be suitable for performing further pattern analysis on other structures. However, as mention in section 2, this will require the application of an algorithm capable of assembling clusters from the decomposed datasets. Future work on this technique will primarily focus on the development and application of such algorithms, as well as the fabrication of known structures in order to provide quantitative analysis of the technique.

References

[1] D.M. Owen, C. Rentero, J. Rossy, A. Magenau, D. Williamson, M. Rodriguez and K. Gaus, "PALM imaging and cluster analysis of protein heterogeneity at the cell surface." in *Journal of Biophotonics*, Vol. 3 (2010), pp. 446-454

[2] D.J. Williamson, D.M. Owen, J. Rossy, A. Magenau, M. Wehrmann, J.J. Gooding and K. Gaus, "Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events." in *Nature Immunology*, Vol. 12 (2011), pp. 655-662