Imperial College London
Department of Computing

Automatic Cell Tracking in Noisy Images for Microscopic Image Analysis

Pedro Damian Kostelec

September 2014

Supervised by Ben Glocker

Submitted in part fulfilment of the requirements for the degree of Master of Science in Computing (Artificial Intelligence) of Imperial College London

Abstract

Acknowledgement

I offer my sincerest gratitude to life, ...

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Contents

5 Data acquisition and annotation IN PROGRESS

This chapter describes the data that influenced the decisions of selecting the cell detection and tracking methods. In section 5.1 we briefly describe the imaging method used to acquire the image sequences and present some example datasets. Section 5.2 presents the data annotation tool that was developed to ease the data annotation process and in section 5.3 we discuss how the tools was used for annotating the datasets, and the difficulties that were encountered.

5.1 Data acquisition and example datasets IN PROGRESS

As discussed in the concluding section of chapter 2 the datasets heavily influence the choice of algorithms for cell detection and tracking. Many computer vision algorithms rely on heuristics to improve their accuracy. In cell detection methods, this is obvious from the fact that a method developed for a certain type of imaging method will likely perform poorly on an image sequence of different types of cells (e.g. different shape of cells). In cell tracking heuristics help adjust the algorithms to the specific behaviour of cells that are being analysed. For example, a different tracking method could be used for images with cells that move slowly (and there is a large overlap between cells in consecutive frames) than for cells that move quickly (and there is little overlap between cells in consecutive frames).

The data acquisition process is not part of this research. However, for the reasons stated above, it is important to understand how the images were obtained and know the characteristics of the datasets. Below, we outline the image acquisition process, and then present some of the datasets we wish to analyse.

The image sequences that inspired the development of the methods describe in this thesis were acquired *in vivo*. This means that the images are obtained on living rat specimen, in contrast to *in vitro* where cells are analysed on a tissue sample in a standard laboratory environment using petri dishes and other instruments. *In vivo* analysis is preferred over *in vitro* because it is better suited for observing the behaviour of cells in its natural environment.

Ask Leo: Info about the ventilator method. Is it two-photon mocroscopy? What camera was used to capture the images?

More recently, the introduction of microscopes allowering for thicher tissue penetraction and higher resolution (spinning-disc and two-photon microscopes), more complex tissuea dn organs, sucha as the skin, lives, brain and lung, can also be imaged. The observation of the lung was a challenge for a long time owing to motion artefacts.
The introbutio of fluorescence (confocal) microscopy in combination with spinning-disc and two-photon microscopes has allowed the use of fluorescent antibodies for labelling different cell populations on anatomical structures, as well as the use of transeenic mice with fluorescent leukocyte subsets.

All the data was provided by Dr. Leo Carlin from the Leukocyte Biology Section at the National Heart and Lung Institute (NHLI)[1].

5.1.1 Datasets NEW

From the datasets provided by Dr. Leo Carlin, five have been selected to use in the evaluation of this work because of their distinct characteristics.

For each dataset, describe what it is (cell, background), characteristics (density, length, quality, artefacts), how easy/hard would they be to track, dimensions

[1] http://www1.imperial.ac.uk/nhli/

Dataset A

This is series30green

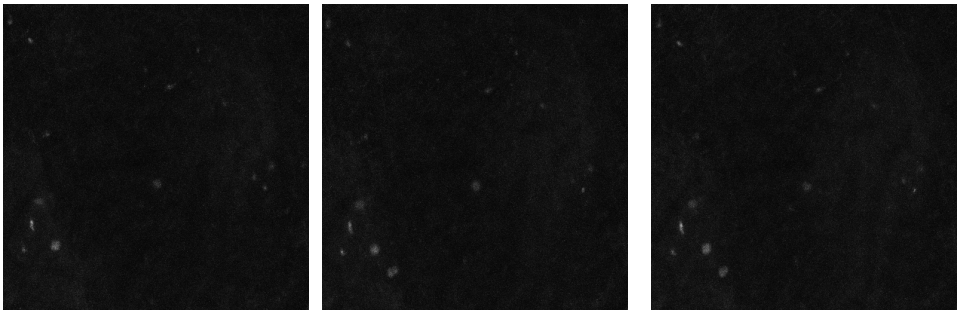

Figure 5.1: Three consecutive frames from dataset A.

This is a dataset obtained from the lung. This dataset contains a very low cell density (about 3 cells per frame). The image sequence contains 66 frames, all of which were annotated. The cells appear grey on a dark, relatively homogeneous, background. The cell boundaries smoothly blend into the background. The images are of constant quality, and there are few significant camera artefacts. The cells move slowly.

Lung for sure?

What is the difference between these cells and the ones in dataset B? They are taken simultaneously... on is series30red the other series30green

Dataset B

This is series30red

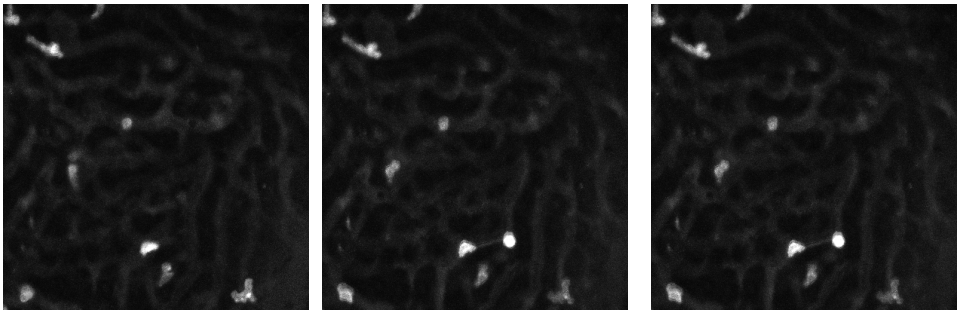

Figure 5.2: Three consecutive frames from dataset B.

This dataset is also obtained from the lung. In fact, it was obtained

Lung for sure?

simultaneously with dataset A, but represents a different types of cells .The cells appear brighter than in dataset A, but their shapes vary. Some are round and others elongated because of the tight blood vessels in which they move. In the background we can clearly discern the blood vessels in a darker grey colour. Cells in this dataset are more active in movement. The images are of constant quality, and there are few significant camera artefacts.

Ask Leo to help me specify

Are thes blood vessels?

Dataset C

This is series13green

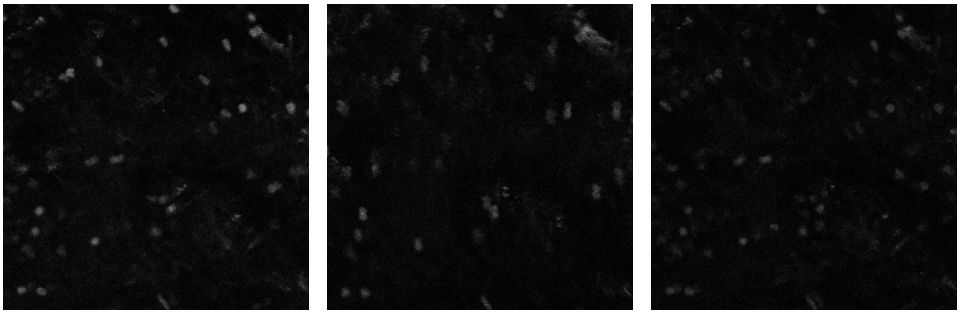

Figure 5.3: Three consecutive frames from dataset C.

Dataset D

This is series14croppedclean

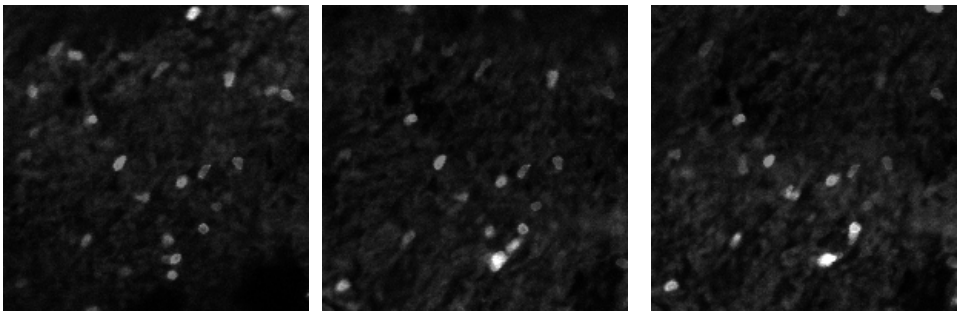

Figure 5.4: Three consecutive frames from dataset D.

Dataset E

This is seriesm170_13cropped

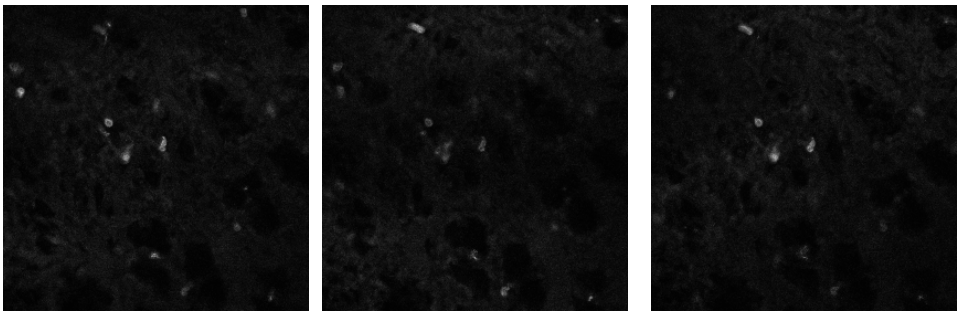

Figure 5.5: Three consecutive frames from dataset E.

rewrite about he type of cells i am tracking briefly, and focus a lot on the imaging technique. Provide examples of different images from different datasets, illustrate the problem of out of focus, the out-of-sync shutter, etc.

Original images are 512-by-512.

TODO: I need more data on the different labeled cells (red, green)
TODO: I need more data on the exact technique and aparatus used to take the images (camera, etc)

Describe the future possible improvements of the imaging technique

Write what parameters were used for the tracker

5.2 The annotation tool NEW

some notes on the importance of accurate annotation,s advantages and disadvantages of dot annotations.

Describe the requirements of an efficient cell annotation tool, such as multipreview, linking, zooming, correct interpretation and saving of the dataformat

An overview of the annotation GUI

the multiple displays filter tools for additing/deleting dots and links simultaneous display of detections and links

5.3 Annotating cell images NEW

What follows is a description of the process of image annotation for use in the machine learning algorithm to detect cells. The image annotation, as required by Arteta's [1] algorithm, are dots on each cell of the image. The algorithm uses these dots as positive examples, and all the remaining pixels as negative examples of a cell.

We have annotated a subset of frames on the Lung dataset provided by Dr. Leo Carlin. The entire dataset is composed of 150 frames, and is divided into two channels, one for each type of cell (). We have marked 10 cropped frames on each channel (frame 1, 14, 25, 46, 81, 115, 131, 143 and 150) of dimensions about 128×118 pixels.

The annotation was performed using a Fiji [24] tool called PointPicker [25] which is accessible from *Analyse/Tools/PointPicker*. The annotation is done by manually clicking on each identified cell.The tool outputs a *txt* file containing x and y coordinates of each annotation, the image number, as well as some other metadata that is not important for us.

It must be noted that the images are very noisy, and it is often hard to distinguish cells from non cells. Figure 5.6 displays an example of an image that was annotated. It is therefore questionable how accurately the learning method will be able to learn the idea of a cell, given that the annotations are far from perfect. It would have been much easier to perform the learning using a synthetic dataset.

The data is then loaded into MATLAB and converted into the format required by the algorithm. The data tidying is performed by the script *prepareTrainData.m*.

The feature of the tool

A user guide is provided in the appending

Briefly explain who made the annotation

Explain who reviewed them

Explain how good the annoations are

Explain if more detailed annotation would result in better results

Rewrite: This describes annotation using Fiji, which was only used at the beginnig

I need to learn more about the types of cells I am tracking

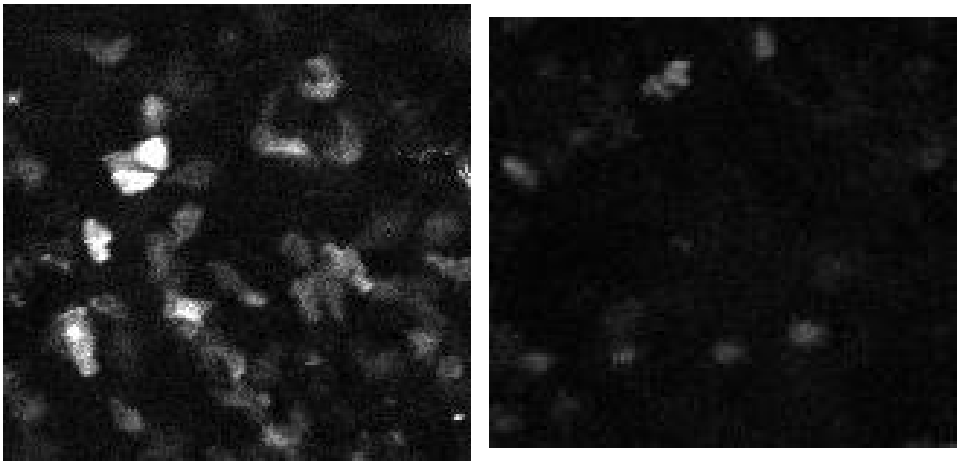

Figure 5.6: Examples of a cropped frame that was annotated for the cell detection machine learning algorithm. Each frame belongs to a different channel of the dataset.

Add appending:
User guide for the annotation tool.
User guide for the detector/tracker

Bibliography

[1] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman, "Learning to detect cells using non-extremal regions," in *Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention - Volume Part I*, MICCAI'12, (Berlin, Heidelberg), pp. 348–356, Springer-Verlag, 2012. 9, 13, 22, 24, 26, 27, 28, 51

[2] Y. Chen, K. Biddell, A. Sun, P. Relue, and J. Johnson, "An automatic cell counting method for optical images," in *[Engineering in Medicine and Biology, 1999. 21st Annual Conference and the 1999 Annual Fall Meetring of the Biomedical Engineering Society] BMES/EMBS Conference, 1999. Proceedings of the First Joint*, vol. 2, pp. 819 vol.2–, Oct 1999. 11

[3] X. Chen, X. Zhou, and S.-C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *Biomedical Engineering, IEEE Transactions on*, vol. 53, pp. 762–766, April 2006. 12, 16

[4] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *Image Processing, IEEE Transactions on*, vol. 2, pp. 176–201, Apr 1993. 12

[5] J. Serra, *Image Analysis and Mathematical Morphology*. Orlando, FL, USA: Academic Press, Inc., 1983. 12

[6] D. Mukherjee, N. Ray, and S. Acton, "Level set analysis for leukocyte detection and tracking," *Image Processing, IEEE Transactions on*, vol. 13, pp. 562–572, April 2004. 12, 15

[7] C. Tang, Y. Wang, and Y. Cui, "Tracking of active cells based on kalman filter in time lapse of image sequences of neuron stem cells." 13, 17

[8] D. Xu and L. Ma., "Segmentation of image sequences of neuron stem cells based on level-set algorithm combined with local gray threshold.," Master's thesis, Harbin Engineering University, 2010. 13

[9] C. Arteta, V. S. Lempitsky, J. A. Noble, and A. Zisserman, "Learning to detect partially overlapping instances.," in *CVPR*, pp. 3230–3237, IEEE, 2013. 13, 14, 25

[10] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *Proceedings of the British Machine Vision Conference*, pp. 36.1–36.10, BMVA Press, 2002. doi:10.5244/C.16.36. 13

[11] T. Joachims, T. Finley, and C.-N. J. Yu, "Cutting-plane training of structural svms," *Mach. Learn.*, vol. 77, pp. 27–59, Oct. 2009. 13

[12] R. Bise, T. Kanade, Z. Yin, and S. il Huh, "Automatic cell tracking applied to analysis of cell migration in wound healing assay," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pp. 6174–6179, Aug 2011. 14, 31

[13] S. Huh, *Toward an Automated System for the Analysis of Cell Behavior: Cellular Event Detection and Cell Tracking in Time-lapse Live Cell Microscopy*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, March 2013. 14, 16

[14] D. House, M. Walker, Z. Wu, J. Wong, and M. Betke, "Tracking of cell populations to understand their spatio-temporal behavior

in response to physical stimuli," in *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*, pp. 186–193, June 2009. 16

[15] B. Xu, M. Lu, P. Zhu, Q. Chen, and X. Wang, "Multiple cell tracking using ant estimator," in *Control, Automation and Information Sciences (ICCAIS), 2012 International Conference on*, pp. 13–17, Nov 2012. 17

[16] K. Li and T. Kanade, "Cell population tracking and lineage construction using multiple-model dynamics filters and spatiotemporal optimization," in *Proceedings of the 2nd International Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB)*, September 2007. 18

[17] A. Massoudi, D. Semenovich, and A. Sowmya, "Cell tracking and mitosis detection using splitting flow networks in phase-contrast imaging," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 5310–5313, Aug 2012. 18

[18] L. Zhang, Y. Li, and R. Nevatia, "Global data association for multi-object tracking using network flows," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, June 2008. 19, 34

[19] C. Huang, B. Wu, and R. Nevatia, "Robust object tracking by hierarchical association of detection responses," in *Computer Vision - ECCV 2008* (D. Forsyth, P. Torr, and A. Zisserman, eds.), vol. 5303 of *Lecture Notes in Computer Science*, pp. 788–801, Springer Berlin Heidelberg, 2008. 19, 34

[20] R. Bise, Z. Yin, and T. Kanade, "Reliable cell tracking by global data association.," in *ISBI*, pp. 1004–1010, IEEE, 2011. 19, 20, 34, 38

[21] H. Kuhn, "The hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, pp. 83–97, 1955. 19

[22] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761 – 767, 2004. British Machine Vision Computing 2002. 22, 23

[23] I. Tsochantaridis, T. Hofmann, T. Joachims, and Y. Altun, "Support vector machine learning for interdependent and structured output spaces," in *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, (New York, NY, USA), pp. 104–, ACM, 2004. 25

[24] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9(7), pp. 676–682, 2012. 51

[25] S. F. I. o. T. L. Philippe Thévenaz, Biomedical Imaging Group, "Point picker: An interactive imagej plugin that allows storage and retrieval of a collection of landmarks," May 2014. 51