Automatic Identification of Cell Motion, Splitting, and Death in a 4D Dataset

Gareth Cairns
School of Computer Science, University of St. Andrews

Objectives

The objectives for this project were as follows:

- To create a literature review on the current state-of-the-art in the field of cell tracking and cell biology to aid in furthering the field.
- To create an end-to-end project to perform automated tracking of cells over time in a 3D space. This must be able to detect cells in the 3D space of a Z-Track of images, and track these cells over time.
- To capture critical interesting moments within the cell data such as cell death and mitosis.

Abstract

The aim of this project is to provide a program to be used by the Biology Department of the University of St. Andrews. The program will be used to monitor cell movement, cell splitting and cell death as they develop in the embryo; this enables researchers to understand the processes as cells start to form in organs. Biologists interested in tracking cell motion, splitting, and death, currently use a time consuming manual method of tracking the cells using bespoke software, pinpointing all cells location in 3D space at a given time, the time this takes to conduct limits the number of experiments they can conduct. This project seeks to automate this cell tracking process using computational methods, therefore enabling researchers to perform more detailed experiments and free up researchers time.

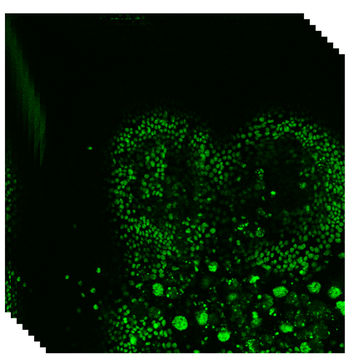

Figure 1: An Example of a Z-Stack of Cell Microscopy Images

Design

To detect cells within the 3D dataset, I created a 3D matrix of images from the Z-Stack and processed them using various image transformation tools to prepare them for watershed segmentation to detect the cells in the images (Figure 2).

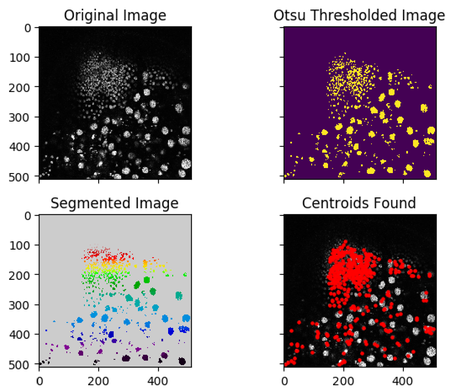

Figure 2: The Process of Finding Centroids of Detected Cells within our system.

Tracking cells involved matching detected cells between time steps using a nearest-neighbour matching algorithm creates a trail in space which can be seen in Figure 3, this location history is saved for every cell, matching it to the closest cell in the next time interval.

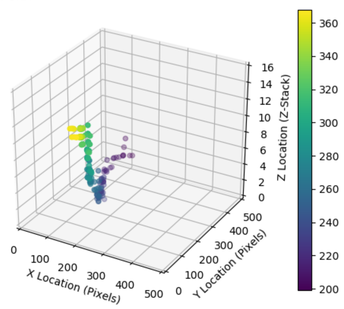

Figure 3: A Cell Tracked over Time in 3D Space by the system. Where time is represented by the colour of the cell.

Results

The outputs of this system show that the cell detection pipeline employed by this project can achieve high accuracy and is capable of detecting hundreds of closely spaces cells in 3D space (Figure 4).

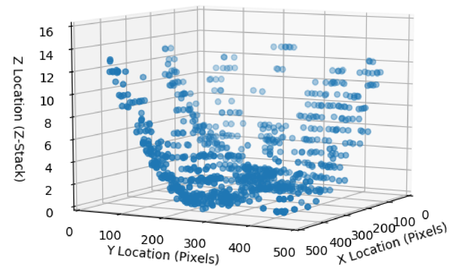

Figure 4: View of cells detected in the 3D Z-Stack.

Using the nearest-neighbour algorithm to track cells over time leads to coherent trails as seen in Figure 3, however, cell mismatching can lead to the trails diverging from the correct tracking data given from manually tracking the cells. (Figure 5)

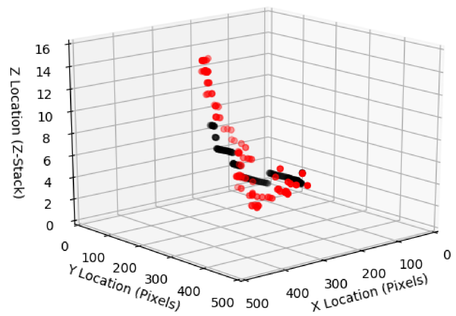

Figure 5: Example of a Cell which diverges from the ground-truth manually tracked cell data. (Manual Data: Black, Automated: Red)

Further work will need to be done to give a perfect match on the tracking data, making use of different matching algorithms for cells over time.

Important Result

In this project, I discovered that the methods of 3D Watershed Segmentation along with cluster detection is a very useful tool for locating cells within 3D space. Using nearest neighbour matching, however, did not give us a level of accuracy to be a useful tool for researchers to track cells over time in a 3D space.

Technologies Used

The technologies employed by this project to provide the features are listed below.

- Python - Used to construct Program in its entirety.
- Sci-Kit Learn and Sci-Kit Image - Used to perform image transformations and Watershed transformation for cell detection.
- OpenCV2 - Used to load images from the Z-Stack to be processed.

```
0 0 0 0 genName
13 0 -1 -1 genName2
13 0 -1 -1 0 -1 9
58
14 47 188 9 -1 -1 -1
17 41 192 10 -1 -1 -1
18 62 196 9 -1 -1 -1
19 55 189 8 -1 -1 -1
20 47 184 7 -1 -1 -1
22 48 191 7 -1 -1 -1
23 44 185 7 -1 -1 -1
```

Figure 6: Sample of the System's Output SMD File, generated through the tracking progress, showing that the structure is in place for the file to be interpreted by the SIMI Biocell software.

Conclusion

This system is successfully able to detect cells in all frames of the 4D data set from within 3D space. These cells can then be tracked over time, attempting to match the manually tracked data as closely as is possible. It has been shown that the tracking system in place could be improved, however, this project has provided a solid ground for further research to take place on.

Overall, I assess this project to have been completed to a satisfactory degree, providing useful functionality to the users of the system, while providing a dramatic increase in data retrieval speed and quantity through automatic methods.