RMIT
UNIVERSITY

Automatic Biological Object Segmentation and Tracking in Unconstrained Microscopic Video Conditions

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Xiaoying Wang

B. Eng - Jiangsu Normal University

M. Eng - China University of Mining and Technology

School of Engineering

College of Science Engineering and Health

RMIT University

November, 2018

Declaration of Authorship

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Signed: Xiaoying Wang

Date: 14/11/2018

Acknowledgements

I would like to acknowledge the enormous help and support I received during my four years of overseas study life, especially the thesis writing process. First of all, I wish to sincerely thank my supervisors Dr. Eva Cheng and Prof. Ian Burnett for their patience in supervising me, for providing a unique intellectual study environment by allowing me to attend many talks, seminars and academic conferences in the field related to my research topic, and for their kindness in treating me as family. I was guided to see the natural beauty of research and life through their immense knowledge. Their prompt and reliable feedback on anything asked is unequalled. As always, my appreciation to them, especially for accompanying and guiding me all the way through a difficult time during the last stage of PhD candidature. This thesis would not exist and my PhD study would not have been such a rewarding experience without their endless support and unshakeable confidence in me.

I extend my sincere gratitude to Dr. Richardt Wilkinson and Assoc. Prof. Margaret Lech for the privilege of learning from them, for their thoughtful consideration, and for their insightful questions and inspiring comments during my research. I am deeply thankful for their strong support and guidance, especially after my original supervisors left the university.

In addition, I have benefited tremendously from collaboration with Dr. Yushi Huang and Prof. Donald Wlodkowic, including our multidisciplinary discussions and the use of biological small organism video data collected and provided by them.

I would like to thank my colleagues who have encouraged, advised, and made this

PhD candidature experience so enjoyable over the last four years. I will never forget the valuable help that they provided through discussions and questions on my research.

I am very grateful to Dr. Eva Cheng, Dr. Katrina Neville, Dr. Glenn Matthews, and Dr. Samuel Ippolito for accepting me as a tutor for their courses, introducing me to an integral part of teaching practice, and developing my knowledge and skills about learning and teaching under their direction and supervision. I also thank the College Learning and Teaching and Academic Development Groups for providing expertise and services to support the development of my professional capabilities and teaching practices.

Huge thanks to the RMIT Foundation Studies staff for giving me the opportunity to work as a mentor for new students, providing flexibility to balance my research with my mentor duties, and providing an opportunity to explore cultural life by offering access to vibrant student communities and services and helping organise a number of social activities and events for new students.

I would like to show my appreciation to the RMIT Student Ambassadors Program for giving me the chance to be a student representative, which provided me a platform to challenge myself, develop new skills, and gain valuable experiences by building a greater connection to the RMIT community and provided training on communication, teamwork, presentations and event management skills.

Most of all, my thanks to my dear family and friends, who have always believed in me more than I do myself. They deeply understand me and know my happiness, my dreams, my pain, and my struggle. I thank them so much for their strong support and for always providing me a warm harbour to rest.

Abstract

Cell and small biological organism tracking research is of fundamental importance for the analysis of dynamic behaviour for assisting the development of many biomedical image related applications. With the rapid development of digitised imaging systems, the immense collections of experimental (microscopic) videos make it nearly impossible to manually analyse the obtained data. Therefore, recent research has drawn attention to building automatic tracking systems to track the movement of cells and small biological organism models using videos taken by microscopes.

Although general object tracking (such as traffic cars and pedestrians) has been studied for decades, existing general object tracking systems cannot directly be applied to cell and small biological organism tracking, due to the differences in the imaging devices and conditions of the targets. This research therefore investigates the novel application of computer vision techniques to reliably, accurately and effectively track the movement of cells and small biological organisms automatically.

Due to difficulties in generating video segmentation ground-truth, there is a general lack of segmentation datasets with annotated ground-truth (particularly for biomedical images). This work proposes an efficient and scalable crowdsourced approach to generate video segmentation ground-truth and develops a tracking ground-truth generation system. To illustrate the proposed approach, an annotated zebrafish larvae video segmentation dataset and three tracking datasets have been generated and made freely available online.

Automatic cell tracking techniques require accurate cell image segmentation; however, current general object segmentation techniques are susceptible to errors due to

the poor microscopic imaging conditions, which include low contrast typical of cell microscopic images. This work proposes a novel image pre-processing technique to enhance low greyscale image contrast for improved cell image segmentation accuracy. An adaptive, shifted bi-Gaussian mixture model is matched to the original cell image intensity histogram for greater differentiation between the cell foreground and image background, while maintaining the original intensity histogram shape.

Small biological organism videos taken by microscope imaging devices under realistic experimental conditions have more complex video backgrounds than cell videos. This work first investigates single zebrafish larvae tracking using dense SIFT flow and downsampling techniques. Many existing multiple small organism tracking systems require very strict video imaging conditions, which typically result in unreliable tracking results for realistic experimental conditions. Thus, this research further investigates the adaptation of advanced segmentation techniques to improve the performance of small organism segmentation under complex imaging conditions.

Finally, this work improves the multiple object association method based on the segmentation module for the proposed system, to address object misdetection and overlapping problems. This system is then evaluated on zebrafish videos, *Artemia franciscana* videos and *Daphnia magna* videos, under a wide variety of (complex) video conditions, including shadowing, labels, and background artefacts (such as water bubbles of different sizes). The tracking accuracy of the proposed system outperforms three existing tracking systems.

Thus, the work in this thesis has contributions in automatic cell and biological organism tracking, where the investigation studied the region-based segmentation dataset construction generalised for biological organisms, intensity contrast enhancement for micrographs, segmentation improvement by removing imaging constraints and the final tracking accuracy enhancement.

List of Publications

- X.Y. Wang, E. Cheng, I.S. Burnett, R. Wilkinson, and M. Lech. " Automatic tracking of multiple zebrafish larvae with resilience against segmentation errors," *to be presented at the Biomedical Imaging (ISBI 2018), IEEE International Symposium. April 4-7, 2018, Washington D.C., U.S.*

- X.Y. Wang, E. Cheng, I.S. Burnett, R. Wilkinson, and Y.S. Huang. "MSBOTS: a Multiple Small Biological Organism Tracking System robust against non-ideal detection and segmentation," *Submitted to Nature Protocols*, 2018.

- X.Y. Wang, E. Cheng, I.S. Burnett, Y.S. Huang and D. Wlodkowic. "Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions," *Scientific Reports (J)*, Vol. 7(1): 17596, pp. 1-8, 2017.

- X.Y. Wang, E. Cheng, I.S. Burnett. "Improved cell segmentation with adaptive bi-Gaussian mixture models for image contrast enhancement pre-processing." In *Proceedings of Life Sciences (IEEE LSC 2017), IEEE International Conference* on, pp. 274-277, Dec. 13-15, 2017, Sydney, Australia.

- X.Y. Wang, E. Cheng, I.S. Burnett, Y.S. Huang and D. Wlodkowic. "Crowdsourced generation of annotated video datasets: a zebrafish larvae dataset for video segmentation and tracking evaluation." In *Proceedings of Life Sciences (IEEE LSC 2017), IEEE International Conference* on, pp. 87-90, Dec. 13-15, 2017, Sydney, Australia.

- X.Y. Wang, E. Cheng, I.S. Burnett. "Improved (STEM) cell segmentation with histogram matching image contrast enhancement." In *Proceedings of Signal and Information (ChinaSIP 2015), IEEE China Summit and International Conference* on, pp. 816-820. 2015.

Contents

Declaration of Authorship **iii**

Acknowledgements **v**

Abstract **vii**

List of Publications **ix**

Contents **xi**

List of Figures **xvii**

List of Tables **xxi**

List of Abbreviations **xxiii**

List of Symbols **xxv**

1 Introduction **1**

1.1 Problem Statement 3

1.1.1 Lack of datasets annotated with ground-truth 3

1.1.2 Severe intensity contrast of cell microscopic images 4

1.1.3 Microscopic video imaging constraints 5

1.1.4 Individual identity ambiguity 6

1.2 Aim and Objectives 7

1.3 Contributions 8

1.4 Thesis Outline 10

2 Literature Review **15**

2.1 Background 16

2.1.1 Purpose and importance of cell segmentation and tracking . 17

2.1.2 Purpose and importance of zebrafish larvae tracking 19

2.1.3 Purpose and importance of other small organism tracking . . 21

2.2 Cell Tracking Methods 23

2.3 Zebrafish Larvae Tracking Methods 26

2.3.1 Adult zebrafish tracking vs. larval zebrafish tracking 28

2.3.2 Existing zebrafish larvae tracking systems 29

Single zebrafish larva tracking 29

Multiple zebrafish larvae tracking 32

2.4 Summary 35

3 Datasets **37**

3.1 Introduction 37

3.2 Existing Segmentation and Tracking Datasets 39

3.2.1 Cell datasets 40

3.2.2 Biological organism databases 41

3.3 Video Data Acquisition for Dataset Annotation 41

3.3.1 Zebrafish larvae video acquisition 42

3.3.2 *Artemia* and *Daphnia* video acquisition 44

3.4 Crowd-sourced Segmentation Ground-truth for Dataset Generation . 44

3.4.1 Interactive segmentation tool 46

3.4.2 Crowd-sourced worker training and selection 47

3.4.3 Accuracy verification of obtained segmentation ground-truth 48

First-pass: Visual assessment 49

Second-pass: Quantitative assessment 51

3.5 Segmentation Results and Discussion 52

3.5.1 The generated zebrafish larvae segmentation dataset 52

3.5.2 Segmentation accuracy evaluation metrics 54
3.5.3 Evaluation of the segmentation dataset 55
3.6 Tracking Ground-truth Generation 59
3.6.1 Tracking ground-truth generation software 59
3.6.2 Generated tracking datasets 61
Dataset 1: Zebrafish larvae tracking 62
Dataset 2: *Artemia* tracking 62
Dataset 3: *Daphnia* tracking 62
3.7 Summary 62

4 Improved Cell Segmentation **65**
4.1 Introduction 65
4.2 Proposed Cell Micrograph Pre-processing 68
4.2.1 First Gaussian component intensity histogram matching . . 69
4.2.2 Parameter estimation for the first Gaussian component . . . 72
4.2.3 Second Gaussian component intensity histogram matching . 74
4.2.4 Parameter estimation for the second Gaussian component . . 76
4.2.5 Relative weight ratio between the two Gaussian components 79
4.3 Results and Discussion 80
4.3.1 Cell database 80
4.3.2 Visual experimental results 81
Visual result example 1: 5^{th} frame of the first cell image sequence 81
Visual result example 2: 14^{th} frame of the first cell image sequence 83
Visual result example 3: 14^{th} frame of the second cell image sequence 84
Visual result example 4: 84^{th} frame of the second cell image sequence 86

4.3.3 Overall cell segmentation accuracy 88
4.3.4 Effect of relative weight on segmentation accuracy 89
4.4 Summary 90

5 Single and Multiple Zebrafish Larvae Tracking 93

5.1 Introduction 93
5.1.1 Single zebrafish larva tracking 95
5.1.2 Multiple zebrafish larvae tracking 95
5.2 Single Zebrafish Larva Tracking 96
5.2.1 SURF features 97
5.2.2 Dense SIFT flow 99
5.2.3 Image warping 101
5.2.4 Estimation of zebrafish object centroid 103
5.2.5 Improving the tracking accuracy 104
5.2.6 Results and discussion 105
Tracking accuracy evaluation 106
Effect of varying downsampling rate on tracking accuracy . 107
5.3 Multiple Zebrafish Larvae Tracking 108
5.3.1 Zebrafish segmentation in unconstrained microscopic video conditions 108
Background subtraction 110
Zebrafish larvae segmentation 113
5.3.2 Multiple zebrafish segmentation performance evaluation . . 114
Zebrafish larvae association between frames 115
Results and discussion 116
5.3.3 Improved zebrafish larvae association across video frames . 122
Representing detected zebrafish larvae 124
Cost matrix and initial assignment 125

Position estimation for misdetected and occluded zebrafish larvae . . . 127
Bridging trajectory gaps for misdetected and occluded zebrafish larvae . . . 128
5.3.4 Locomotion characteristic analysis . . . 129
5.3.5 Results and discussion . . . 131
Tracking accuracy evaluation . . . 131
Tracking trajectory example . . . 133
Analysis of zebrafish larvae movement characteristics . . . 135
5.4 Summary . . . 137

6 Generalised Multiple Small Biological Organism Tracking System (MSBOTS) 141
6.1 Introduction . . . 141
6.2 Applying MSBOTS to other small organisms . . . 142
6.2.1 Movement characteristics of tested organisms . . . 143
6.2.2 Evaluated microscopic videos . . . 144
6.3 Results and Discussion . . . 145
6.3.1 Overall tracking performance for *Artemia* videos . . . 147
6.3.2 Overall tracking performance for *Daphnia* videos . . . 147
6.4 Summary . . . 149

7 Conclusions and Future Work 151
7.1 Conclusions . . . 151
7.2 Future Work . . . 156
7.2.1 Applying the proposed MSBOTS in behavioural ecotoxicology 156
7.2.2 Examination of detailed organism body movement . . . 157
7.2.3 Applying machine learning techniques to organism tracking 158

References 159

List of Figures

1.1 A typical cell micrograph example with low intensity contrast between the cells and the background. 5
1.2 Solutions for organism individual identity ambiguity. 6
1.3 Thesis outline 10

2.1 Micrograph examples of stem cell and zebrafish larvae. 17
2.2 A visual example of cell tracking between successive microscopic video frames. 18
2.3 Typical tracking errors by the LSRtrack system [5]. 31
2.4 Object detection based on pixel threshold using experimental data presented in Section 3.3. 32
2.5 Video frame examples for a variety of imaging conditions. 34

3.1 Overview of dataset management workflow. 45
3.2 Overview of workflow for workers to generate segmentation ground-truth. 45
3.3 Example of selecting target object against the background. 46
3.4 Adjustment example after object and background differentiation. . . 47
3.5 Visual examples of segmentation 49
3.6 Typical visual perceptible errors. 50
3.7 Visual example of the obtained segmentation ground-truth selected from the zebrafish segmentation dataset. 53
3.8 Average segmentation evaluation recall scores for each video sequence. 56

3.9 Average segmentation evaluation precision scores for each video sequence. 57
3.10 Average segmentation evaluation $F_{measure}$ scores for each video sequence. 58
3.11 Average segmentation evaluation SI scores for each video sequence. 58
3.12 Interface of the proposed tracking ground-truth generation software to select the centroid positions of target objects in microscopic videos 60
3.13 Flow chart of tracking ground-truth selection for whole video sequences. 61

4.1 Flow chart of the proposed pre-processing method. 66
4.2 Microscopic stem cell video frame example and its original intensity histogram 70
4.3 Bi-Gaussian mixture model example. 71
4.4 Shifted bi-Gaussian model. 73
4.5 Location of the (hidden) secondary Gaussian shape. 74
4.6 Close examination of the second Gaussian shape. 75
4.7 Fitted Gaussian model for the second Gaussian shape. 76
4.8 Resultant cell video frame example and its intensity histogram after processing by the proposed method. 78
4.9 Visual example 1 of a segmentation result. 81
4.10 Segmentation result of the 5^{th} frame from the first cell video sequence using the proposed approach. 82
4.11 Visual example 2 of a segmentation result. 83
4.12 Segmentation result examples of the 14^{th} frame from the first cell video sequence. 84
4.13 Visual example 3 of a segmentation result. 85
4.14 Segmentation result of the 14^{th} frame from the second cell video sequence using the proposed approach. 86

4.15 Visual example 4 of a segmentation result. 87
4.16 Segmentation result of the 84^{th} frame from the second cell video sequence using the proposed approach. 88
4.17 Relative modal amplitude effect on segmentation accuracy. 89

5.1 Flow chart of the proposed method. 97
5.2 Sparse feature points from various feature descriptors: 98
5.3 Visual demonstration of results of each step by the proposed single zebrafish larva tracking system. 100
5.4 Visual demonstration of results of each step by the proposed single zebrafish larva tracking system. 102
5.5 Zebrafish position correction when no movement is detected. 105
5.6 Overview of the proposed zebrafish larvae segmentation method with an existing tracking algorithm for further evaluation. 109
5.7 Building GMM training samples for a short video. 110
5.8 Segmentation accuracy over the 10 zebrafish video sequences. . . . 118
5.9 Tracking accuracy over the 10 zebrafish video sequences. 120
5.10 Overall multiple zebrafish larvae tracking system. 123
5.11 Storage structure of detected zebrafish larvae in a video sequence. . 124
5.12 Point calculation for occluded organisms 127
5.13 Tracking results evaluation among the methods compared using zebrafish larvae videos. 131
5.14 Visual example comparing tracking trajectories. 134
5.15 Zebrafish larvae movement acceleration. 135
5.16 Zebrafish larvae movement velocity. 136
5.17 Zebrafish larvae movement direction analysis. 137

6.1 Microscopic video frame examples of *Artemia franciscana*. 144
6.2 Microscopic video frame examples of *Daphnia magna*. 145

6.3 Tracking results evaluation among the tracking methods compared using *Artemia franciscana* microscopic videos. 146
6.4 Tracking results evaluation among the comparison methods using *Daphnia magna* microscopic videos. 148

List of Tables

2.1 Literature categorised according to cell tracking methods 24

3.1 Imaging conditions of the presented zebrafish segmentation dataset . 43

3.2 Verification of the generated ground-truth for the zebrafish segmentation dataset 52

4.1 Mean cell segmentation accuracy across the cell image database based on IRU 89

5.1 Tracking accuracy performance comparison. 106

5.2 Total number of swapped individual identities 121

List of Abbreviations

SNR	**S**ignal-to-**N**oise **R**atio
DIC	**D**ifferential **I**nterference **C**ontrast
VMR	**V**isual **M**otor **R**esponse
MM	**M**athematical **M**orphology
SE	**S**tructure **E**lement
SIFT	**S**cale **I**nvariant **F**eature **T**ransform
hpf	**h**ours **p**ost **f**ertilization
dpf	**d**ays **p**ost **f**ertilization
fps	**f**rames **p**er **s**econd
PMMA	**P**oly **M**ethyl **M**eth**A**crylate
IDS	**I**maging **D**evelopment **S**ystems
IST	**I**nteractive **S**egmentation **T**ool
VSB	**V**ideo **S**egmentation **B**enchmark
BMS	**B**erkeley **M**otion **S**egmentation
FBMS	**F**reiburg **B**erkeley **M**otion **S**egmentation
CI	**C**onfidence **I**nterval
SI	**S**imilarity **I**ndex
GMM	**G**aussian **M**ixture **M**odel
EM	**E**xpectation **M**aximisation
PDF	**P**robability **D**istribution **F**unction
HE	**H**istogram **Equalisation**
GFP	**G**reen **F**luorescent **P**rotein
OI	**O**ptical **I**maging
NIR	**N**ear **I**nfra**R**ed
DMSO	**D**i**M**ethyl **S**ulf**O**xide
MSBOTS	**M**ultiple **S**mall **B**iological**O**rganism **T**racking **S**ystem

List of Symbols

f	frame rate	fps
$^{\circ}$C	temperature	degree
$\times$	amplifier	times
r	radius	mm
pH	acidity scale	
p_{thr}	intensity threshold value	grey level

Chapter 1

Introduction

Cell and small biological organism tracking research is of significant importance in the analysis of dynamic cell and organism behaviour for assisting the development of cell studies, biomedical research and immune biology [1–4]. The automatic tracking of cells and small organisms, such as zebrafish larvae, *Artemia franciscana* and *Daphnia magna*, will provide efficient, reliable and well-validated approaches for quantitatively and statistically analysing their behavioural responses to different stimuli. However, with the rapid development of storage devices and imaging systems, the immense number of cell and small biological living organism microscopic time-lapse videos makes it nearly impossible to analyse these cell and organism video databases manually. Recent research has thus focused intensive attention on building effective and automatic cell and small biological organism tracking systems to segment and track cells and organisms in time-lapse microscopic video sequences.

There are difficulties specific to cell tracking, for example, mitosis and apoptotic events, especially when there are a large number of cells present in a culture. Thus, many challenges remain for accurate and automatic tracking of cells compared with general computer vision-based object tracking.

Zebrafish (*Danio rerio*) larvae are rapidly gaining popularity as a vertebrate and mammal model for many biomedical applications including screening for biochemical abnormalities [5] and behavioural science investigations [6–8]. Consequently,

tracking of these larvae, as example small biological organisms, has emerged as a challenge.

Manual quantification of cell and zebrafish larvae videos taken by microscope is labour intensive and frequently not feasible [9]. Recent research attention has therefore focused on the development of automatic multiple cell and zebrafish larvae tracking systems due to the increased availability of digital microscopy and video storage systems.

Computer-aided techniques are therefore necessary for cell and small biological organism tracking, and these techniques usually involve target detection and inter-frame association [9]. However, these automatic techniques face many challenges for both the biomedical and computer vision fields. First, the video frames taken by microscopes generally exhibit a poor Signal-to-Noise Ratio (SNR) due to the low intensity contrast and low imaging light intensities needed to limit fluorescent photobleaching (especially for videos taken by fluorescent microscopes) [9] or the high imaging light intensities required for optical imaging techniques [10]. Second, the video frame intensity profiles may vary over the course of imaging due to changes in biological structure and/or deviations in the intensities of the fluorophore labelling materials [9]. Third, biological organism models may have uneven movement [11], and their dynamic behaviour becomes unpredictable after biochemical testing chemicals or medicines are applied [12–14]. Furthermore, the well-known Brownian motion (the jittery, chaotic motion of microscopic particles suspended in water caused by the random collisions with water molecules) results in unpredictable moving trajectories due to the discrete nature of random collisions [15]. The Brownian motion affects the detection of microscopic objects in fluid media.

Considering these challenges in combination with the specific peculiarities of microscopic video and the dynamic movement characteristics of the tracked objects can therefore potentially provide significantly better tracking performance than general purpose tracking [9].

This work proposes techniques that build and extend upon existing cell tracking techniques and general computer vision-based object tracking algorithms applied to automatically track the movement of cells and small biological organisms such as zebrafish larvae (and potentially other organisms). The ultimate aim is to analyse experimental time-lapse microscopic video sequences (already and continuously obtained by applied science researchers) to investigate segmentation, individual trajectory tracking, and dynamic behavioural parameter estimation.

1.1 Problem Statement

Tracking techniques are vital for understanding the biology and ecology of moving organisms. Although tracking techniques have assisted the development of behaviour and interaction analyses of large organisms [16–18], such as mammals, birds and adult fish, the tracking of small organisms (at millimetre scale; most are considerably smaller than 1 mm [19]) is hampered by the constraints of existing tracking methods [20, 21] as further discussed below.

1.1.1 Lack of datasets annotated with ground-truth

Video segmentation research has emerged over the last decade for biomedical image and video processing, especially for biological organism tracking. However, due to the difficulties of generating the video segmentation ground-truth, there is a general lack of segmentation datasets with annotated ground-truth, which poses a challenge for improving biomedical image segmentation approaches and severely limits the evaluation of segmentation algorithms.

There are a number of standard benchmark datasets for still image segmentation evaluation in general computer vision, such as the Berkeley segmentation dataset (including BSDS300 and BSDS500) [22], which is widely used for used contour/boundary

evaluation. However, standard datasets for benchmarking moving objects in video sequences are still emerging in the field of computer vision [23–26].

Although there is a cell segmentation dataset annotated with ground-truth presented in Section 3.2, there is currently no video segmentation dataset publicly available for zebrafish larvae movement. The lack of such a dataset is a common bottleneck that constrains the development and evaluation of biomedical image research.

1.1.2 Severe intensity contrast of cell microscopic images

The tracking of moving biological cells in time-lapse video sequences is fundamental to further understand biological processes. Automatic cell tracking techniques require accurate cell image segmentation; however, current segmentation techniques are susceptible to errors due to non-ideal but realistic cell image conditions, including the low contrast typical of cell microscopic images. The whole-cell micrograph shown in Figure 1.1 is very dark, which hinders the separation of the cell boundaries from the image background, and the intensity values of cells vary (some cell regions are hardly observable, whereas other cell regions are clear).

Effective and accurate cell segmentation poses many unique challenges to general object segmentation techniques in computer vision [27] (such as the segmentation of pedestrians or traffic from natural scenes), due to the low intensity or greyscale contrast, illumination changes, cell cluster events (ambiguous segmentation boundaries), severe image noise and cell shape diversity in typical time-lapse cell videos. The direct application or adaptation of general object segmentation methods [28–30] to cell segmentation generally does not take into account the characteristics of time-lapse videos taken by microscopic imaging systems under laboratory experimental conditions.

Cell detection and segmentation are conducted on each frame independently without taking into account any high-level information, such as the underlying biological

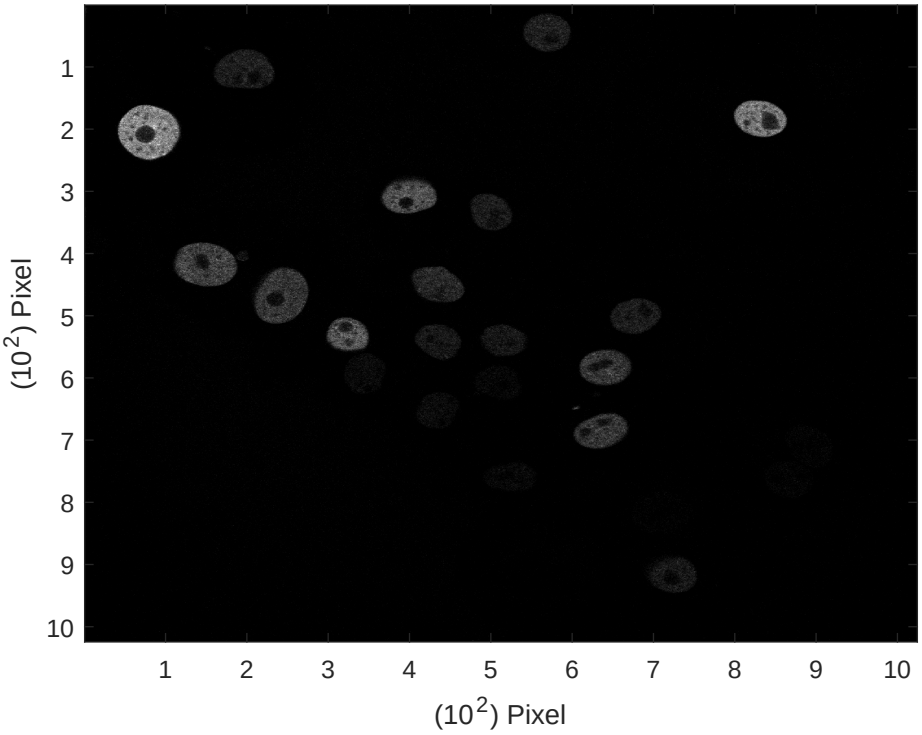


Figure 1.1: A typical cell micrograph example with low intensity contrast between the cells and the background.

structures. The segmentation result for one frame may be near perfect (when all cell regions in a micrograph are clear and have clear boundary outlines against the image background), but segmentation errors over the span of multiple video frames would result in severe tracking errors due to loss of the tracking trajectories of the studied targets [9].

1.1.3 Microscopic video imaging constraints

Many available zebrafish tracking systems are based on the assumption of high segmentation accuracy [31, 32] or require strict video imaging conditions for a clear video input background [5, 31, 33, 34].

The modifications of experimental conditions required to obtain the necessary imaging conditions in existing tracking systems can affect the reliability of the experimental results. For example, the observed swimming behaviour and kinematics may

be affected by observational bias potentially introduced by modifying environmental conditions, such as the high light intensities that are usually required for optical imaging techniques [10]. The influence of high-intensity light on light-sensitive organisms (e.g., phototaxis) may not be negligible [10]. However, an adequate light source is a central requirement for obtaining sufficient contrast between objects and the background [10].

More importantly, zebrafish larvae will disappear from the tracking system [5, 35] if the larvae stop moving. Detection loss of zebrafish larvae is not uncommon for the pixel displacement thresholding method because of the 'bursty' movement characteristics of larval zebrafish, as explained in Section 2.3.1.

1.1.4 Individual identity ambiguity

Individual identity ambiguity following crossing, cluster and overlapping events between connected microscopic video frames is a challenge for multiple biological object tracking.

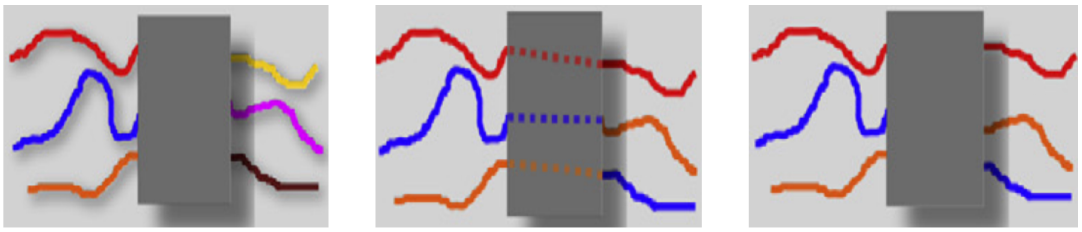


Figure 1.2: Solutions for organism individual identity ambiguity.

Multiple biological organism tracking systems [36–41] cannot maintain the identities of individuals following crossing, cluster and overlapping events. New tracking trajectory fragments are created after organisms cross or overlap with each other, as shown in Figure 1.2a.

Although some systems [42–48] attempt to maintain individual identities using probabilistic models to predict the position of individuals in the following frame after

crossing, cluster and overlapping, these techniques are not reliable for biological organism tracking (such as zebrafish larvae) because of their complex, unpredictable movement patterns under exposure to different chemicals, new medicine samples and toxicants. Identity swapping might occur (as shown in Figure 1.2b) when the moving trajectories of multiple organisms overlap.

The recently developed idTracker [31] is a multiple biological organism tracking system that is reported to maintain individual identities (as shown in Figure 1.2c) over a long period of time without identity error propagation by extracting "fingerprint" information for each individual. However, the core function is to extract intensity and contrast maps based on pixel intensity values, which are not robust to illumination changes through frames and are limited by the number of targets that can be detected at one time and the specific ratio between the organism size and the container size [6, 49].

1.2 Aim and Objectives

The aim of this thesis is applying advanced computer vision techniques to improve the accuracy, reliability and efficiency of automatic multiple biological object tracking by addressing the challenges discussed in Section 1.1.

- This work will study a segmentation and tracking dataset construction approach for efficiently and economically generating the segmentation and tracking ground-truth for general small biological organisms, to facilitate the numerical evaluation of segmentation and tracking algorithms and to address the challenge presented in Section 1.1.1.
- A zebrafish larvae segmentation and tracking dataset annotated with ground-truth is built using this approach.

- Since segmentation-based automatic cell tracking methods [28] require cell segmentation results for each frame, improving cell segmentation accuracy can enhance the overall performance of a tracking system. Thus, this work proposes to develop methods to increase the contrast differentiation between the target cells and the microscopic video background/noise to, in turn, achieve improved segmentation accuracy and address the challenge discussed in Section 1.1.2.

- Finally, this work aims to develop an automatic, accurate, and efficient zebrafish larvae tracking system under unconstrained microscopic video imaging conditions to systematically assess the behavioural parameters of small living aquatic organisms (according to the challenges in Section 1.1). The system can be used to test the behavioural responses of zebrafish larvae under laboratory conditions and different chemical exposures and allows experimental replication for statistical analysis, and address the challenges discussed in Section 1.1.3 and Section 1.1.4, e.g., input video constraints and multiple object tracking trajectories, respectively.

1.3 Contributions

The main technical contributions of this work and the associated publications are listed below:

- Proposing an efficient and scalable crowdsourced approach using interactive segmentation methods to generate video segmentation datasets annotated with ground-truth to facilitate database generation and supervised numerical evaluation of segmentation algorithms, particularly for general biological organisms. In addition, this work presents an annotated zebrafish larvae video segmentation dataset generated by the proposed approach and three tracking datasets annotated with centroid positions of different small organisms. The datasets

and ground-truth generation, verification, and numerical measurement software have been made freely available online. Furthermore, this work proposes a two-pass verification process to verify the segmentation results manually generated by the crowdsourced workers, and a set of metrics to enable the numerical evaluation of segmentation algorithms using such annotated datasets (Chapter 3, [50]).

- Presenting a novel image pre-processing technique to enhance the typical low greyscale microscopic image contrast for improved cell image segmentation accuracy. An adaptive, shifted bi-Gaussian mixture model is proposed to match the original cell image intensity histogram for greater differentiation between the cell foreground and microscopic image background, while maintaining the original intensity histogram shape. Rather than using a model with fixed parameters across an entire video sequence as in existing approaches, this work proposes the derivation of parameters of the mixture model from the characteristics of the microscopic images to match the intensity histogram for each video frame to adaptively address changes in the video background (Chapter 4, [51, 52]).

- Developing a segmentation method combining background subtraction, a Gaussian mixture model with adaptation of the number of Gaussian components and component parameters to changes in video background, and morphological operation to address the unresolved zebrafish larvae detection and segmentation problem and remove the strict constraints on microscopic video imaging conditions of existing small organism tracking systems. (Section 5.3.1 in Chapter 5, [53].)

- Developing an accurate, reliable and efficient automatic single and multiple zebrafish larvae tracking system based on the developed segmentation method. This system can deal with non-ideal detection and segmentation situations and can be applied to other small organisms, such as *Artemia franciscana*, and

Daphnia magna. (Section 5.3.3 in Chapter 5, [54, 55].)

1.4 Thesis Outline

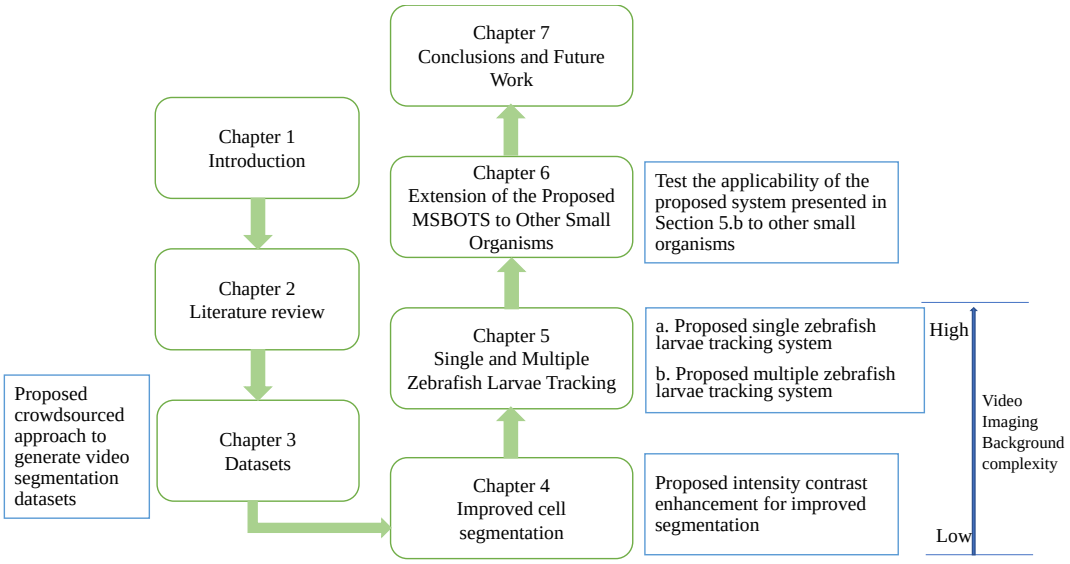


Figure 1.3: Thesis outline

To document the details of this research, the outline of the thesis (as shown in Figure 1.3) is organised as follows:

Chapter 1 introduces the background of this work including cell tracking and small biological organism tracking (i.e., zebrafish larvae, *Artemia franciscana*, and *Daphnia magna* tracking). It also presents the problems and challenges facing existing cell and small organism tracking (especially zebrafish larvae tracking due to the difficulties introduced by their 'bursty' movement characteristics), followed by the introduction of the aim and objectives of this work. Finally, Chapter 1 summarises the main technical contributions of this work associated with the related publication outcomes.

Chapter 2 reviews the literature on cell tracking, zebrafish larvae tracking, and selected, advanced computer vision techniques that can be applied to cell and small organism tracking research.

Chapter 3 presents the investigation of segmentation and tracking dataset generation. Due to the general lack of datasets for small organism segmentation and tracking in general image and video segmentation and tracking research, especially in biomedical image processing research, Chapter 3 proposes an efficient and scalable crowdsourced approach to generate video segmentation ground-truth to facilitate database generation for general biological organism segmentation and tracking algorithm evaluation. In addition, Chapter 3 presents a tracking ground-truth generation software with a user-friendly interface for selecting the centroid position of each organism in video frames. To illustrate the proposed crowdsourced approach, an annotated zebrafish larvae video segmentation dataset is generated. In addition, Chapter 3 provides a set of segmentation evaluation metrics to enable the evaluation of segmentation algorithms against the manually generated ground-truth. Finally, Chapter 3 uses metrics to evaluate the segmentation performance of five leading segmentation algorithms on the generated zebrafish video segmentation dataset. To further facilitate the quantitative and objective measurement of the overall tracking performance of the single and multiple organism tracking systems, another three datasets annotated with tracking ground-truth based on organism centroid positions are built and presented in this Chapter. The evaluation and comparison of the proposed MSBOTS (presented in Chapter 5) with existing automatic organism tracking systems are based on these datasets. The datasets and related software are publicly accessible.

Chapter 4 presents the pre-processing investigation of stem cell tracking using microscopic time-lapse videos. The accurate detection and segmentation of cells from time-lapse microscopic video sequences provide a critical foundation for understanding dynamic cell behaviours and cell characteristics. However, general object segmentation methods in computer vision are susceptible to errors due to the severe microscopic imaging conditions in time-lapse microscopic cell videos. To address the low image intensity contrast typical in microscopic cell images, Chapter 4 of this work investigates the novel use of an adaptive, shifted bi-Gaussian mixture model

to enhance the intensity contrast prior to cell segmentation. Rather than using a model with fixed model parameters across an entire video sequence as in existing approaches, it proposes the adaptive derivation of the mixture model parameters to match the intensity histogram for each video frame to address changes in the video background. To maintain the dominant characteristics of the original intensity histogram of microscopic cell video frame, the mixture model parameters are calculated from the two dominant features in each intensity histogram. The experimental results for the cell database show improved segmentation accuracy compared with existing image contrast enhancement methods. The pre-processed cell image exhibits greater differentiation between the cell foreground and video background while maintaining the original intensity of the histogram features.

Chapter 5 presents the development of two zebrafish larvae tracking systems: a single zebrafish larva tracking system under a clear video background and the more advanced multiple zebrafish larvae tracking system for challenging video imaging conditions and complex organism movement characteristics. The microscopic video conditions of zebrafish larvae are more complex than those of cell microscopic videos. Thus, Chapter 5 first presents the investigation of the dense SIFT flow technique for single zebrafish larva tracking. Single zebrafish larva tracking is simpler than multiple zebrafish tracking because there is no individual identity ambiguity problem for single zebrafish tracking. Chapter 5 subsequently develops an accurate, efficient and reliable multiple zebrafish larvae tracking system by exploring the adaptation of background subtraction and object segmentation techniques and multiple object tracking algorithms. Section 5.3.1 presents zebrafish larvae detection/segmentation from the microscopic video background without constraints on the microscopic video imaging background. Furthermore, Section 5.3.2 presents the association process in the detected zebrafish larvae of the multiple zebrafish larvae tracking system based on the improved segmentation result.

Chapter 6 presents the extension of the application of the proposed multiple zebrafish

larvae tracking system to *Artemia franciscana* and *Daphnia magna* using microscopic videos. The testing results for the 10 videos under typical experiment imaging conditions confirm the generalisable automatic tracking ability of the proposed system (presented in Section 5.3), and thus the system is denoted the Multiple Small Biological Organism Tracking System (MSBOTS).

Chapter 7 finally concludes the thesis with suggested future work to further the impact of the research conducted in this thesis.

Chapter 2

Literature Review

Typical manual tracking approaches are tedious and commonly require significant periods of manual observation and labelling of the image features to represent the activity for a single experimental task [56, 57]. Furthermore, as a subjective manual task, the results are difficult to reliably repeat and reproduce.

However, recent advances in (microscopic) medical imaging and the ubiquitous availability of multimedia storage capacity have led to increasingly large collections of time-lapse cell video sequences, and thus the traditional approach of manual cell analysis by visual inspection is time-consuming, tedious and subjective (often unrepeatable), requiring skilled analysis and inspection of cell images.

Effective and accurate cell analysis poses challenges for existing techniques in cell tracking, and in some cases, it is even impossible to manually track cell image sequences to obtain robust results, especially when there are a large number of cells in long image sequences. Manual cell tracking is commonly employed and is a tedious and unrepeatable task that often requires more than 24 hours of manual analysis due to the skilled inspection required to obtain dynamic and structural cell information. The shape deformability of cells, image contrast changes, cell clutter and severe image noise further add to the difficulty of manually tracking cells [39].

Recent research attention has therefore focused on the development of automatic biological object detection and tracking systems due to the increased availability of

digital microscopy and video storage systems.

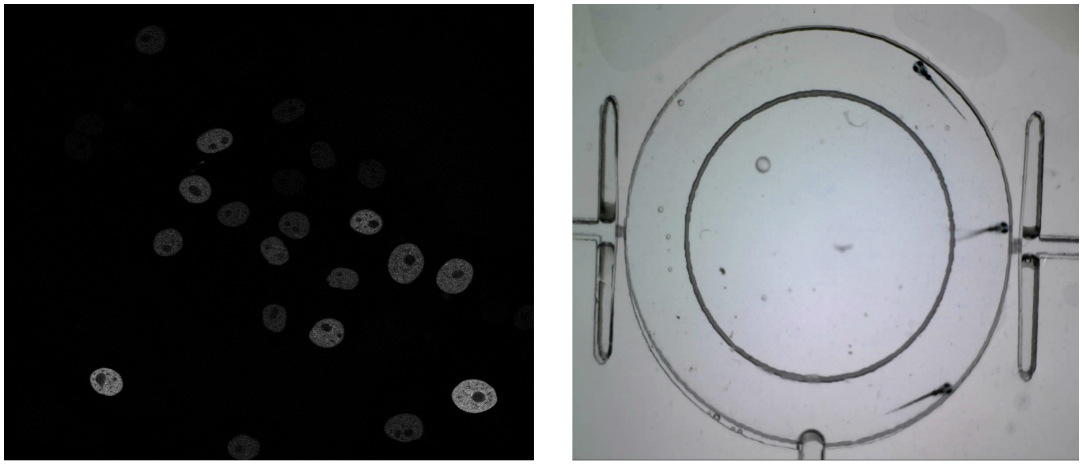
Specifically, three main challenges differentiate optical microscopy tracking from computer vision object tracking. First, it is not uncommon for a cell culture to contain thousands of cells. It is difficult to accurately track a certain target in a large population to obtain the movement features (such as the movement velocity and direction), dynamic behaviour and structural characteristics; such movement features are analysed and explored in Sections 5.3.4, 5.3.5 and 6.2.1. Furthermore, the edges of closely connected targets will inevitably overlap with each other, resulting in the formation of clusters in the culture, especially when a large number of targets exist under experimental conditions. It is a challenge for computer vision algorithms to automatically segment such ambiguous target object boundaries instead of a large target region cloud. This is a key research challenge for cell and organism tracking algorithms, addressed in Section 5.3.3. Finally, cell tracking algorithms must be robust against cell mitosis, in which one mother cell splits into two daughter cells with the same number of chromosomes of stem cells. To solve this problem, a detection-based cell tracking approach is investigated in this work and is presented in Chapter 4.

This Chapter reviews the literature on cell tracking (with an emphasis on the low intensity contrast problem of segmentation-based cell tracking methods), zebrafish larvae and other small organism tracking and the related computer vision techniques that can be applied to this research.

2.1 Background

Digital images or photographs taken by microscope imaging devices are called micrographs or photomicrographs. In contrast to macrographs of items that are visible to the naked human eye taken by normal cameras, micrographs are used to magnify an item to display its microstructure in extensive detail [58–60]. Figure 2.1 shows

two micrographs of stem cells (Figure 2.1a) and zebrafish larvae (Figure 2.1b) taken by optical microscopes. A time-lapse microscopic video is a series of such micrographs taken at a fixed frame rate.



(a) (b)

Figure 2.1: Micrograph examples of stem cell and zebrafish larvae.
(a) Scale 63× [61]; (b) scale information was not provided in the video, which was provided by a science application lab.

To illustrate the process of segmentation-based cell (or other small biological organism) tracking, Figure 2.2 shows an example of the tracking of detected/segmented cells among three successive microscopic video frames. $D_{i,j}$ indicates the label of the detected cell j in the frame at time i ($i = 1, 2, 3$). The aim of automatic cell tracking is to accurately link the corresponding cell $D_{i+1,j}$ for every cell $D_{i,j}$, as visually illustrated by the solid yellow and dashed red lines linking, for example, cells 1 and 3, respectively. And thus, the tracking/linking step is based on the object segmentation (detecting/segmenting objects from original video frames as example shown in Figure 2.1).

2.1.1 Purpose and importance of cell segmentation and tracking

Cell tracking can provide critical information on dynamic cell behaviours and average cell characteristics for many biomedical research and commercial applications

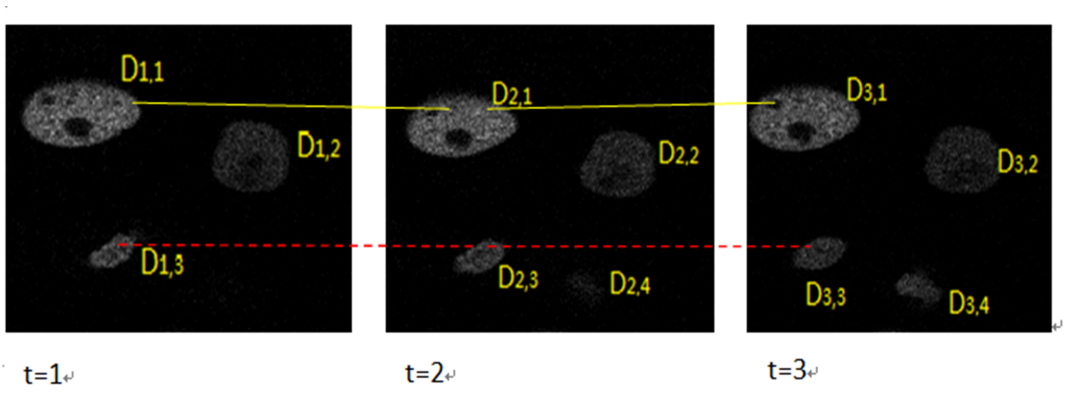


Figure 2.2: A visual example of cell tracking between successive microscopic video frames.

[1]. Automatic tracking of biological cells has attracted significant research attention in recent years.

In the last decade, the use of optical micrographs in cell biology has received dramatic attention in biomedicine, drug-cell interactions and drug discovery[58, 59, 62] because cells are an important object for toxicology studies, studies of patient-drug interactions and new drug development in pharmaceutical research [59]. Current cell tracking techniques are applied to a wide range of cells, including not only embryonic stem cells and tumour cells but also epithelial and endothelial cells [29]. Furthermore, strategies for microscopic analysis have been expanded by the automation of sample preparation, with the ability to constantly provide a large collection of digitised microscopic images [59, 63].

The study of cell tracking provides information on dynamic cell behaviours and average cell characteristics [1]. Almost all cells in the human body undergo significant motion, and the characteristics of the cell motion typically convey critical pathological information [2]. For example, the migration mechanisms of tumours and how they kill the host organism remain to be elucidated [64]. However, the migration mechanisms of tumours within tissues are similar; thus, such investigations of cellular and molecular migration can further current knowledge on the spreading and invasion of tumour cells, with a view towards developing new treatment strategies

[65]. The importance of studying cell tracking is also illustrated by studies of the clinical application of immune cell migration tracking, which aids the identification of infection points to facilitate the removal of the invading pathogen.

2.1.2 Purpose and importance of zebrafish larvae tracking

Zebrafish are an ideal vertebrate model system to study the genetics and biological mechanism of many human diseases [3, 66]. In recent years, zebrafish have become one of the most prominent vertebrate model organisms [67, 68] successfully used to study developmental genetics [69–72], neural systems [7, 73], and the genetics of behaviour and brain function [3, 74–76]. Such wide use of zebrafish is due to the optical transparency of the zebrafish body (which enables detailed anatomical characterisation, direct investigation of tissue movement, cell migration and interaction during neurogenesis [3, 73]), ease of genetic manipulations, physiological similarity to mammals, robust behaviour, low cost to incubate and raise, and potential for high-throughput screening [7].

Zebrafish are an especially suitable model for testing drug addiction, e.g., alcoholism. The effect of alcohol on behaviour and brain function can be studied by mixing alcohol with the water in the fish tank because of the simple alcohol delivery mechanism in zebrafish [3, 74, 77]. The alcohol in the water is absorbed through the gills and skin of the fish, and thus the alcohol level in zebrafish vessels quickly reaches equilibrium with the external alcohol concentration [3]. The results obtained in [3] imply that zebrafish are a potentially successful model system for studying the effects of drug addiction on behavioural characteristics. Thus, behavioural tests with zebrafish offer a useful tool for facilitating the identification of genes involved in addiction mechanisms.

The amenability of zebrafish to genetic change by gene knock-down or overexpression has resulted in the isolation of a variety of mutants for studying developmental

mechanisms [66, 73, 78]. Zebrafish genes and mutations have syntenic correspondence relationships with the human genome [66, 78]. Since the introduction of zebrafish in biological and medical laboratories, many milestones have been achieved, positioning zebrafish as a prominent genetic model organism [79].

Accurate movement characteristic analysis from the automatic tracking of zebrafish larvae is crucial for many biomedical image-related research and applications. Locomotive behaviour analysis of zebrafish larvae can be divided into three main categories based on their basic applications as vertebrate models.

First, the movement pattern of zebrafish larvae can be determined by simple pixel quantification based on the number of pixels whose values change beyond a threshold value within consecutive video frames [80]. By classifying the activity pattern as burst (very large movement) and freeze (no movement) classes, the sleep and awake states of zebrafish larvae can be analysed and correlated with inactive states and detectable movement states. This sleep/awake pattern research based on classification of zebrafish larvae movement has been widely used as a model for the study of the genetic mechanisms that regulate human sleep [80, 81]. For example, more than 10% of the human population suffer from chronic sleeping disturbances, and the identification of defective hypocretin/orexin (Hcrt) signalling as a cause of mammalian sleep disorders highlighted the potential of genetic approaches to sleep research [80]. One hypocretin/orexin gene has been identified in zebrafish [80, 82], and its gene structure and expression pattern and the axonal projections of larval zebrafish Hcrt neurons are strikingly similar to those in mammals [80, 81]. Thus, zebrafish larvae have been introduced as a model system to study the sleep/awake regulator Hcrt and to determine the effect of long-term expression of this signal [80, 81].

Second, the zebrafish tracking algorithm detects the coordinate information of the zebrafish larvae and then calculates the kinetic parameters of the subject, such as positional preference, velocity and displacement, based on changes in the centroid coordinates of the zebrafish larvae in successive video frames [83]. There is little

quantitative information on how spontaneous propulsive movement develops, and little is known on how genetic and environmental factors affect the environmental behaviour of zebrafish larvae. This information is essential for understanding the source of variability of movement behaviours of zebrafish and investigating novel treatments of neurological diseases associated with movement disorders of zebrafish larvae as a model of human systems.

The final category of movement analysis of zebrafish larvae is kinematic measurement, which identifies anatomic landmarks of zebrafish and detects how their spatial relationships change during consecutive frames. This can be used to determine the detailed movements of individual zebrafish larvae, such as trunk curvature and tail beating frequency. This information is also used for differentiating the turning and spontaneous propulsive movement of zebrafish. The neural basis of behavioural choice for vertebrates is largely unknown, and zebrafish are an attractive model for vertebrates. The kinematics of motor events are studied for the measurement of tail beating frequency for different manoeuvres; this information is further used to analyse the response of zebrafish larvae to different stimuli [84].

2.1.3 Purpose and importance of other small organism tracking

Small (often microscopic) organisms such as *Daphnia magna* and *Artemia franciscana* are commonly used for ecotoxicity studies, group behaviour studies, and ecosystem research [4, 85, 86].

Groups of animals exhibiting collective behaviour have been studied intensively, particularly their self-organisation and leadership capabilities [86]. Analysing and modelling this group behaviour constantly inspires the development of engineering systems, such as the design of distributed control systems [86, 87]. Leadership and decision-making processes are crucial in animal groups because there is limited

knowledge for processing information such as choosing the directions of food or migration paths, and these processes allow a small proportion of individuals to lead the group [86]. Thus, the dynamic behaviour characteristics of *Artemia* are studied and analysed as a model for developing external leadership by robot agents to interact with or induce a collective response of the group [86].

Observations of the activity and viability of *Artemia* in water samples by microscopes are used to obtain information on state and environmental conditions under aquatic conditions [4]. In addition, the presence and variability of *Artemia* as biological indicators provide information on the ecosystem effects of pollutants for water analysis [4]. Further, *Artemia franciscana* can provide orders of magnitude higher (10 - 1000 times higher [88]) sensitivity than conventional lethal biomonitor organisms (which are predominantly applied based on the mortality of these species) to evaluate the toxic effects of environmental contaminants and stressors [57, 88, 89]. Thus, *Artemia franciscana* have emerged as a sensitive biomonitor organism model to rapidly detect adverse ecosystem changes such as toxins [57]. For example, changes in the swimming patterns (such as average velocity) of *Artemia* are widely used as behavioural toxicity end-points to monitor aquatic pollution and statistically assess swimming activity [57].

Daphnia, freshwater relatives of *Artemia franciscana* [90], are also frequently used as model organisms to assess the hazards and risks of chemicals released in the aquatic environment (e.g., freshwater rivers) due to their high sensitivity to many chemicals and their representation of freshwater organisms [91]. Behavioural tests can provide earlier responses than traditional standard acute and chronic tests because toxic stress will induce rapid movement pattern changes among *Daphnia* under toxic levels lower than acute toxic levels [85, 91].

2.2 Cell Tracking Methods

The traditional technique (still used currently in some laboratories) for observing biomolecules or live cells or small organisms is tagging the target by fluorescent labelling or injection [73, 92] or so-called quantum dots [93], which produce dynamic and structural information on the targets via fluorophore tracking [73, 92].

Various fluorescent and bioluminescent contrast agents have been used for cell labeling and tracking (such as green fluorescent protein (GFP) [94–96], red, orange and yellow FP [97, 98], and bioluminescent proteins [99–101]), in combination with optical imaging (OI) techniques. OI employs a light source (often a monochromatic laser) that excites a fluorophore to emit in the near infrared (NIR) spectrum, and a detector (often a highly sensitive CCD device) captures the fluorophore's emission [100, 102]. Fluorescent imaging techniques are widely used for biological cells and tissues due to their various advantages, including high detection sensitivity and high light penetration depth in biological tissues [102].

However, these methods are not reliable for tracking living cells over a long period of time due to the severe photobleaching of fluorophore materials and their broad overlapping spectra [103]. In addition, different fluorescence labelling techniques may lead to diverse appearances, even for different samples of the same biological subject [9]. In addition, fluorescence microscopy is subject to practical physical limitations caused by the low image contrast and light diffraction within the optical system, resulting in the fundamental problem of image resolution [104]. Furthermore, fluorescence microscopes are much more expensive than light microscopes. The entire data collection process is also more complex for fluorescence microscopy than for optical microscopy [73].

Thus, the automatic tracking of biological cell movement provides an efficient approach for investigating their dynamic behaviour characteristics in biomedical research and applications [1, 105] compared with traditional manual visual inspection

approaches.

Existing work [11, 28, 39] has explored the application of general object tracking to cell tracking, but cell tracking differs from general object tracking in computer vision (e.g., of people, cars, faces etc.). General object tracking does not need to consider issues such as the large number of cells in cell image frames, cluster events among connected cells, severe noise in cell image sequences, shape diversity between different types of cells, inevitable mitosis events with stem cells, and contrast change in cell image sequences [61, 106].

Table 2.1: Literature categorised according to cell tracking methods

Model-based evolution methods		**Detection-based methods**	
Model	**Update**	**Segmentation**	**Association**
Level set [29, 107]	**Kalman filter [108]**	**Intensity thresholding [109, 110]**	**Viterbi-based search**
Active contour [111, 112]	**Particle filter [44]**	**Gradient detection [113, 114]**	**Feature matching [115]**
Mean-shift [11, 29]	**Bayesian method [116]**	**Morphological operations [117, 118]**	**Nuclei merging [119]**
Watershed [120, 121]		**Chan-Vese segmentation [30]**	**Nearest-neighbor linking [122]**

Table 2.1 briefly summarises the literature on existing cell tracking methods. As shown, there have been two main categories of cell tracking methods in the last decade: model-based evolution methods and detection-based methods [11, 28]. The basic approach of model-based algorithms is the creation of a model for each cell in the first frame of a video sequence. The information for these models is updated in the following two or more consecutive frames over time [11]. Representative models include the level-set [29, 107], active contour [111, 112] and mean-shift models [11, 29]. Chan-Vese segmentation [30] further combines the level sets and Gaussian Mixture Model (GMM) by iterative Expectation Maximisation (EM) estimation, attempting to implicitly represent cell boundaries. The model-based approach is robust against topological changes during cell mitosis and can easily be extended to higher dimensional processing for multiple cell tracking problems [28, 29]. However, the disadvantage of these approaches is the inability to deal with cells entering\leaving the observation area and cells without regions of overlap in two successive frames.

The reinitialisation required in these cases significantly adds to the computational complexity.

A common strategy for detection-based cell tracking algorithms is the segmentation of all cells in each frame. The detected cells in consecutive frames are then associated based on certain criteria over time [123]. The advantage of this method is the computational efficiency at the segmentation stage. However, it also places more tasks in the next association stage because no information from previous frames is used in the segmentation step. The other shortcoming of this method is the uncertainty in establishing a one-to-many or many-to-one correspondence due to high cell intensity, cell mitosis events and cells entering/leaving the observation area. To solve this problem, a large number of features are trained to obtain the criteria to determine the splitting and merging of some trajectory fragments in the association process [29].

Many advanced object segmentation techniques in computer vision have been applied to enhance cell segmentation accuracy. Mathematical morphological operations [117, 118], gradient (edge) detection [113, 114], intensity thresholding [109, 110] and watershed algorithms [120, 121] have also been employed as additional strategies to help improve cell segmentation accuracy. However, the efficiency and applicability of these techniques are limited by the image artefacts generated and over- or under-segmentation, especially for cells whose boundaries have intensity values similar to the background.

Histogram Equalisation (HE) [124] is one of the most commonly used basic contrast enhancement techniques due to its simple implementation [125, 126]. However, HE inherently flattens the intensity Probability Distribution Function (PDF), which can lead to over- or under-enhancement of image regions [124, 127, 128]. Washed-out images, patchiness and visual artefacts are drawbacks that can affect traditional HE methods and their variants [126]. Salihah [129] reviews another three contrast enhancement methods: partial contrast [130], bright stretching and dark stretching

[131]. Partial contrast [130] uses a linear mapping function to stretch narrow intensity levels to a wider range such that the whole intensity range [0, 255] is occupied. Thus, the contrast and brightness of the original image will be enhanced. Bright stretching and dark stretching [131] are all based on autoscaling, while bright stretching is to increase the brightness level of images and dark stretching is the reverse procedure of bright stretching. Intensity transformation, such as logarithmic transformation [132], power-law transformation and bit-plane slicing [133], and spatial filtering including smoothing spatial filters [134] and sharpening spatial filters [132, 135] have also been applied in contrast enhancement for cell microscopic images. Intensity transformation and spatial filtering directly manipulate pixels on cell image plane. More detailed technical discussion on cell segmentation and contrast enhancement will be presented in Chapter 4.

2.3 Zebrafish Larvae Tracking Methods

As zebrafish (*Danio rerio*) larvae have increasingly been used as a vertebrate and mammal model for many biomedical applications including screening for biochemical abnormalities [5] and behavioural science investigations [3, 6], larvae tracking has emerged as a challenge [136]. Recent research attention has therefore focused on the development of automatic multiple zebrafish larvae tracking systems.

Similar research challenges to cell tracking are also found in small organism tracking, such as the time-consuming, unrepeatable results of manual tracking and the increasing amounts of acquired data requiring analysis. Recent research has found that the nervous system of zebrafish provides an opportunity to study the fundamental basis of neural pathways in mammals; in particular, zebrafish larvae are attractive vertebrate experimental models because of the limited complexity of their neural circuits, which are similar in structural function to those of mammals [5, 80]. In addition,

zebrafish larvae have gained popularity in recent years as a vertebrate model to understand the development of human tissue and organs [137] and as a model of human disease in new medicine discovery and molecular pathology [80, 83]. Conventional movement analysis of zebrafish requires labour-intensive and time-consuming manual analysis; however, few techniques are available that enable automatic immobilisation and analysis of movement behaviour of a large number of zebrafish larvae in real time [138]. Automatic and reliable methods for analysing the movement of zebrafish larvae are therefore fundamental for facilitating biomedical applications in the fields of neural research, genetics and drug research [5].

Automatic zebrafish larvae tracking is usually conducted by recording the zebrafish larvae using microscope imaging devices for a period of time. Computers are then used to automatically detect and quantify the movement of the objects in the resultant time-lapse videos. The movement of zebrafish larvae can be analysed individually or in a group by observing a single zebrafish or multiple zebrafish swimming in the wells (called a petri dish in applied science) used to house the zebrafish larvae subjects.

Compared with cell tracking and general object tracking, the detailed difficulties of zebrafish tracking can generally be classified as segmentation and tracking problems. With respect to detection difficulties, the similar size and non-rigid shape of zebrafish are difficult to represent by one or a set of topological templates [6, 33, 49, 139]; texture information for zebrafish larvae is not sufficient for detecting fish location because of the transparent zebrafish body [140–142]. The tracking difficulties can be divided into three classes. The movement pattern complexity (especially upon exposure to test chemicals or medicines) [33, 75, 143, 144] of zebrafish makes it difficult for existing probabilistic-based prediction models [42, 43, 145, 146] to fully simulate all moving activities. In addition, similarities in size and shape complicate the differentiation of individuals in a long video sequence. Finally, occlusion and crossing events trigger detection errors, leading to fragments in tracking trajectory

and swapping of individual identities, which can also be propagated to the following sequences [6, 36, 49].

2.3.1 Adult zebrafish tracking vs. larval zebrafish tracking

Many automatic single and multiple tracking systems have been recently developed for adult zebrafish [5, 31–34, 147, 148], such as state-of-the-art methods based on deep learning [32], particle filtering [148], and the well-known idTracker [31], and outstanding tracking performance has been reported for adult zebrafish.

However, the locomotive characteristics of zebrafish larvae are dramatically different from those of adult zebrafish. Adult fish swim continuously, whereas zebrafish larvae may display little or no movement over time [149], and thus their dynamic responses can be imbalanced. Zebrafish larvae exhibit a mean proportion of activity of less than 0.075 over time, according to statistics reported in [149]. This inactivity is the primary cause of tracking failure in these systems and traditional statistical tests based on movement features for tracking and analysing larval behaviour. In addition, the intensity contrast between adult fish and the water background is also greater than that for zebrafish larvae due to the transparent nature of the larval peripheral body. Moreover, the adult zebrafish tracking systems reported in [31, 32] are both based on the assumption of high-intensity contrast, a commonly required imaging condition constraint for existing zebrafish tracking systems.

The factor of small size poses many challenges, particularly for aquatic organisms [57]. Radiofrequency identification chips, also called u-chips or transponders, are widely used in individual organism identification application, but the smallest available chip size is currently approximately 0.4 mm [150]. Devices of this size will dramatically affect the natural dynamic behaviour of organisms at the millimetre scale [20, 93]. In addition, general object tracking is already a complicated problem due to object occlusion, overlapping, non-rigid object structures (object rotation and

changes in scale), or changes in motion patterns, among other factors [57, 151], and the difficulty level increases when the tracking targets are small in size because small organisms provide less information on the imaging noise.

3D systems with multiple cameras [20] or super-resolution images built from multiple low-resolution images [20] have been proposed to obtain more information for accurate tracking of small organisms. These systems increase computational complexity, change the detection and association tracking system structure and require further object location registration and association among cameras or images.

2.3.2 Existing zebrafish larvae tracking systems

Single zebrafish larva tracking

Single zebrafish larva tracking is comparatively simpler than multiple zebrafish tracking due to the absence of identity ambiguity or scenarios of occlusion of multiple objects. However, single zebrafish larva tracking is still very important in many behavioural studies and neurological research [35, 152]. Automatic tracking of single zebrafish larva has been used as a model system in many behavioural studies [35, 152]. For example, studies of the Visual-Motor Response (VMR) use single zebrafish larva placed in each well to determine their locomotor response to light increments and decrements [153] and study eye relaxation in vertebrates. Single zebrafish larva tracking has also been applied to discover and characterise the complex behaviour of psychoactive drugs [35]. Another application of single zebrafish larva is the identification of the genetic basis of seizure resistance [152].

To enable high-throughput studies using single zebrafish larva, petri dish plates with multiple wells containing a single larval zebrafish in each well are widely used in single larva tracking systems to record and compare a group of tests simultaneously [5, 33, 147]. Permitting only a single zebrafish larva in each petri dish also avoids the overlapping and swapping of trajectories that can result from housing multiple

zebrafish larvae in one container. However, experiments with one zebrafish per dish strictly constrain research applications in studies of interaction and grouping behaviour.

The View Point tracking system has been used extensively for single zebrafish larva tracking [35]. To achieve accurate zebrafish larvae detection performance by the View Point video-tracking system, care must be taken during recording to avoid glare from light illumination on the water surface, which may interfere with the imaging devices and manual control of the system, while aligning the software grid with the zebrafish larva container (misaligning the video-tracking system grid could lead to loss of detection of fish movement) [153]. The pixel displacement threshold between successive frames for distinguishing the zebrafish larvae from the background is empirically determined in the setup based on the specific imaging cameras and light setup [153].

The LSRtrack system [5] is another popular technique for single zebrafish larva tracking and has gained wide usage for single zebrafish tracking research. However, the authors reported tracking failure as a result of unreliable selection of the tracking contrast threshold (as illustrated in Figure 2.3) [81]. LSRtrack uses a similar object detection mechanism as the View Point tracking system [35]. The pixel displacement is compared with a pre-selected threshold value within consecutive video frames. However, image pixel values are unreliable as a determining factor due to changes in lighting, perspective and noise [154].

Filter banks, mutual information, gradient and phase features are therefore applied to the detection of zebrafish larvae to improve reliability over simple pixel values. However, these methods fail if significant changes occur (a common event in zebrafish movement) [154]. Reference [140] proposed the application of affine correlation for serial zebrafish registration and texture tracking, emphasising image texture content and generating stable and smooth motion estimation. However, this strategy is only suitable for specific textural imagery, e.g., in differential interference contrast (DIC)

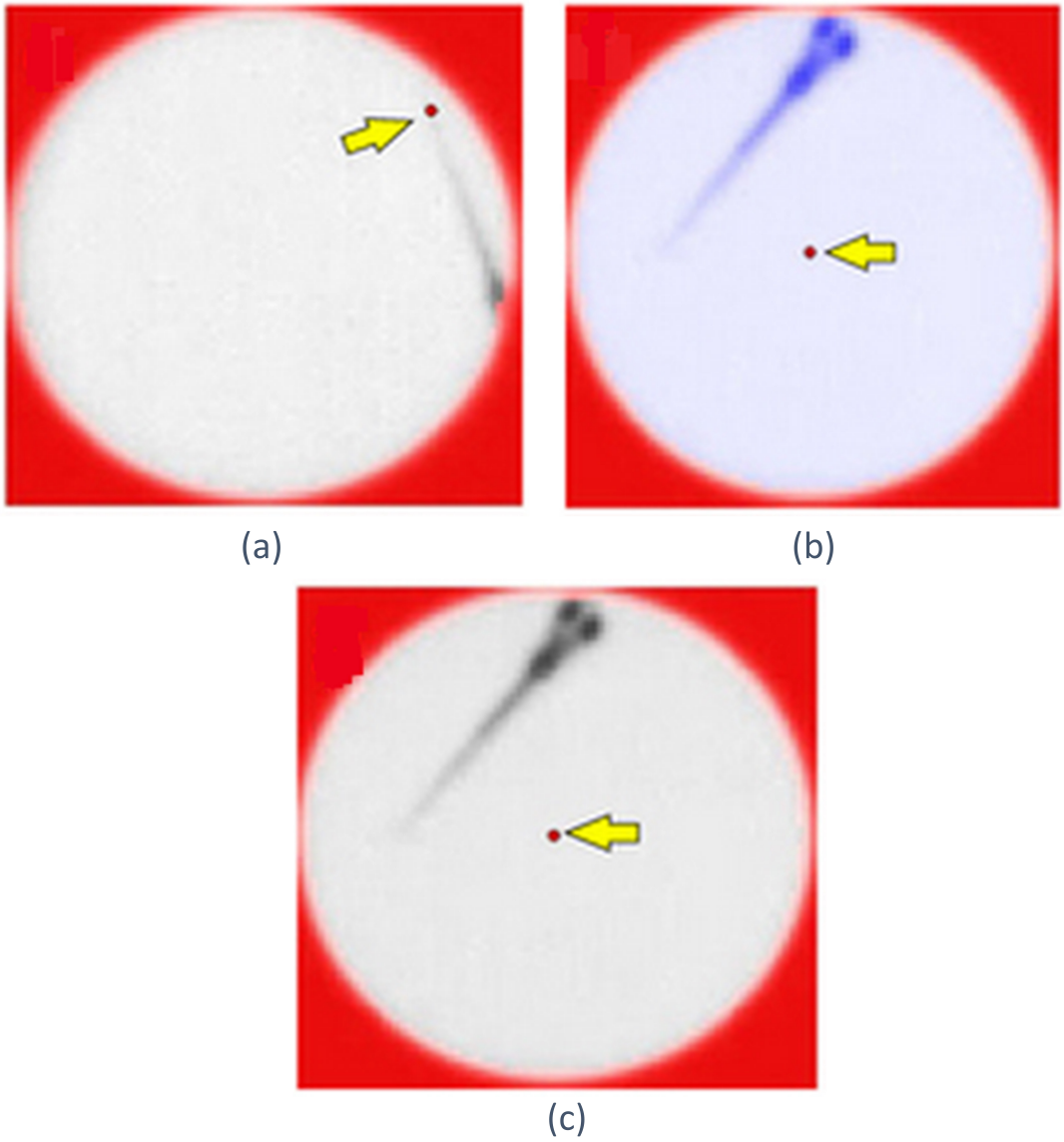


Figure 2.3: Typical tracking errors by the LSRtrack system [5].

(a) The zebrafish is hidden by the Petri dish wall; (b) mistracking caused by a high-intensity threshold value; (c) typical error caused by a low-intensity threshold value.

microscopy images. Furthermore, the measurements may also be biased if no initial deformation exists [140].

Multiple zebrafish larvae tracking

Recent research attention has focused on the development of automatic multiple zebrafish larvae tracking systems, which is much more complex than single zebrafish larva tracking research due to the existence of identity ambiguity and scenarios of multiple zebrafish occlusion [6].

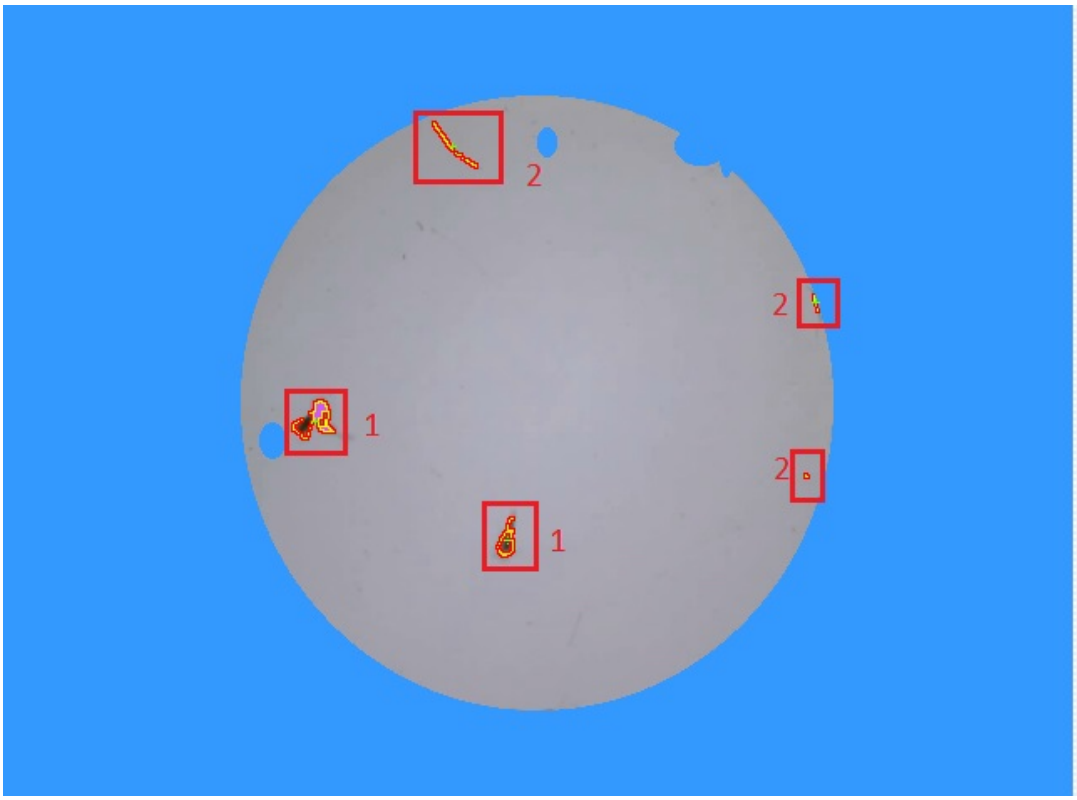


Figure 2.4: Object detection based on pixel threshold using experimental data presented in Section 3.3.

The LoliTrack system [34] is a widely used commercial system for multiple small organism tracking, including zebrafish larvae tracking [57, 155]. Figure 2.4 shows the object selection results of LoliTrack based on the manual adjustment of a threshold pixel intensity value. Regions labelled with a number '1' in the red box are the object target. Due to their similar intensity values, water particles or impurities (as

highlighted by the red boxes labelled by the number '2') are also selected by LoliTrack as target objects. Further manual adjustments are therefore needed to remove these typical water impurities.

Although many systems have been developed to explore multiple zebrafish larvae tracking [20, 33, 148], the video input must be obtained under strict imaging condition constraints. As reported in [5] and [32], even small impurities in the water (as shown in Figure 2.5a) and lighting reflections (as shown in Figure 2.5b) will affect the tracking result. In addition, the small size difference between the zebrafish larvae and the petri dish (as shown in Figure 2.5c) and between the adult zebrafish and the fish tank (as shown in Figure 2.5d) can result in the detection of water impurities such as water bubbles (as shown by the red circles in Figure 2.5e), excretions, and small particles (as shown by the red triangles in Figure 2.5f) that are not usually detectable in adult fish experiments but inevitably affect the detection of zebrafish larvae.

Strict input imaging conditions can be impossible to maintain in practice. For example, even if a clean environment is originally used to house the organism, excretions produced by the organisms during the experiment can render it impossible to maintain a completely clean and transparent container background during long-term organism observation. IdTracker [31] even explicitly defines the smallest acceptable size ratio between the zebrafish and the tank for creating the clear background environment required for video data. In addition, these larvae tracking systems [5, 33, 147] use a petri dish plate to separate individual zebrafish larvae. Only one zebrafish larva is allowed in each petri dish to avoid overlapping and swapping of trajectories due to housing of multiple zebrafish larvae in one container (as shown in Figure 2.5c). However, limiting experiments to one zebrafish per dish strictly constrains the research application because interaction and grouping behaviour cannot be studied.

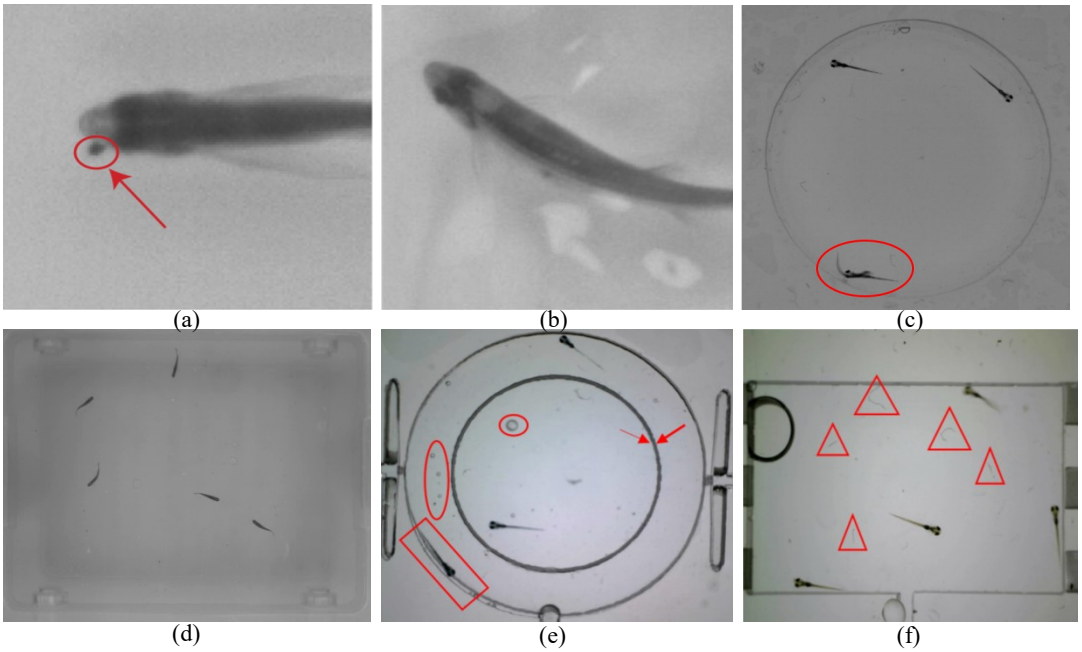


Figure 2.5: Video frame examples for a variety of imaging conditions.

(a, b) Small water impurities as indicated by the red circle in (a) and water reflections or ripples in (b) can affect the head detection. (c) Frame example with larvae occlusion, which is not observed when the larvae are separated in petri dish plates. (d) frame input with clear tank edges and a large size ratio between the adult fish and the container. (e) Frame example with labelling as indicated by the red arrows, water bubbles as highlighted by the red circles, and larvae with low intensity contrast between the well edge shadow as shown by the red rectangle. (f) Frame example with small water particles as shown by the red triangles.

2.4 Summary

Cell and small organism tracking is different from general object tracking in computer vision due to their specific challenges. This Chapter reviewed the literature on cell and small organism tracking. Section 2.2 discussed the background of cell tracking, especially the low intensity contrast problem hampering the enhancement of cell segmentation accuracy. Section 2.3 then presented the state-of-the-art in single and multiple zebrafish larvae tracking. The microscopic video conditions of zebrafish larvae are more complex than those of cell microscopic videos due to 1) the agitation of water impurities and water bubbles due to beating of the larval tail, 2) the ambiguous region between the zebrafish larvae and the video background due to the transparency of the zebrafish larvae peripheral body, and 3) the water surface shadow and ripples projected by illumination on the petri dish. This Section then reviewed the existing automatic single and multiple zebrafish larvae tracking systems.

The segmentation and tracking datasets will be discussed in Chapter 3, followed by the presentation, application and evaluation of the proposed methods to address the problems mentioned in Section 2.2 and Section 2.3, respectively in Chapter 4 and Chapter 5. Therein, the proposed multiple zebrafish larvae tracking system is developed and evaluated in Chapter 5, while the extended application, dealing with small biological organisms in general, is presented in Chapter 6.

Chapter 3

Datasets

3.1 Introduction

The Chapter 3 introduces the datasets upon which this work presents, applies and evaluates the proposed crowd-sourced segmentation ground-truth generation approach for general small biological organism segmentation evaluation.

Accurate target object segmentation from the time-lapse microscopic videos is of fundamental importance to the overall performance of whole tracking systems. However, whilst image segmentation has attracted intensive research attention, the evaluation of video segmentation algorithms aimed specifically at biomedical image processing is not as mature, and it remains difficult to compare and benchmark the different segmentation approaches [26, 156]. Reference [26] summarises the main difficulties for segmentation evaluation, including limited datasets for comparison (especially in biomedical image processing), lack of evaluation metrics or performance statistics, and difficulties in establishing manual ground-truth segmentation. The latter is due to poor reproducibility and the time, resources and labour required to train 'experts' and for them to generate the ground-truth segmentation [157].

Due to these challenging difficulties, subjective evaluation remains a common segmentation evaluation method in biomedical image segmentation. However, human

visual comparison of segmentation results is susceptible to both human error and subjective opinion variations. A common and less labour-intensive approach (for users) is supervised evaluation compared with unsupervised approaches [156], where the segmentation results are compared against an annotated ground-truth. References [158] and [159] proposed supervised evaluation approaches, whilst an interactive segmentation tool has been proposed in [160] and included in other imaging software (such as ImageJ or Fiji) to generate segmentation ground-truth. However, building a dataset with annotated ground-truth using these approaches is very time and resource consuming, where after the 'ground-truth' generation, manual visual assessment is typically conducted to validate the results [156] and the quantitative reliability of the manually generated 'ground-truth' is under-explored.

Section 3.4 proposes a general crowdsourced approach to generate segmentation ground-truth for time-lapsed videos for any biological organism. The proposed approach uses an online crowdsourcing platform, where workers are screened and trained to generate the ground-truth using the Interactive Segmentation Tool (IST) [160]. The reliability and consistency of the 'ground-truth' generated by workers are examined in a two-pass verification process proposed and applied in this Chapter: following a fast initial manual visual assessment, comparing the segmentation result to the original frame, a joint mask is automatically generated to objectively quantify under or over-segmentation errors. The proposed approach is applied to generate a dataset annotated with segmentation ground-truth using a series of zebrafish larvae videos.

To facilitate the evaluation of segmentation algorithms on the segmentation ground truth annotated database, Section 3.5.2 presents a novel metric, Similarity Index (SI), based on three standard metrics (Recall, Precision and $F_{measure}$) to quantify the segmentation accuracy. The zebrafish dataset generated in Section 3.4 is evaluated using these metrics, with the segmentation accuracy of five leading segmentation techniques compared. This dataset and code for validation and calculating evaluation

metrics are freely available online[1] for repeatable research usage.

To extend upon this segmentation dataset and evaluation, Section 3.6.2 presents the tracking datasets annotated with centroid ground-truth generated by the proposed centroid position selection method for single and multiple objects. The tracking accuracy evaluation in Chapters 5 and 6 is based on these tracking datasets.

3.2 Existing Segmentation and Tracking Datasets

General computer vision techniques have catalyzed the development of biomedical image processing research [157], and general segmentation datasets may be used to provide the indicative performance of biomedical image segmentation approaches. Whilst there are a few standard benchmark datasets for evaluating single still-image segmentation in computer vision (e.g., Berkeley segmentation dataset [22]), standard datasets for benchmarking moving objects for video segmentation are still emerging. Video segmentation datasets in general computer vision have only been presented in recent years, such as the Video Segmentation Benchmark (VSB100) [23], and Berkeley Motion Segmentation Dataset (BMS-26), which was extended into the Freiburg-Berkeley Motion Segmentation Dataset (FBMS-59) [24, 161].

Currently available datasets for region segmentation evaluation are limited. Reference [25] presents an influential dataset for occlusion boundary detection using natural scenes, but only one frame annotated with boundary ground-truth provided per video. However, the evaluation based on one frame per video may not provide reliable segmentation evaluation results. Further, the boundary segmentation result cannot support object temporal information, such as the object centroid location, as the segmented boundary may not guarantee the successful generation of a closed object region.

[1] https://github.com/Xiao-ying/-moving-zebrafish-larvae-segmentation-dataset-/tree/master/Data

In contrast, the most recent video segmentation dataset [23] divided frames into different regions with numbered annotation for each region in the ground-truth images; however, no discrimination between the object and background was specified, with ground-truth only generated approximately once per 20 frames. Thus, the segmentation performance evaluations based on this dataset required manual intervention for specifying the detection target and the background.

3.2.1 Cell datasets

Due to the lack of publicly available datasets, a small number of private segmentation datasets of real cells have been generated to validate biomedical image processing approaches [61]: Feng et al. [157] proposed image repositories for cell segmentation evaluation with three groups of cell squences with 60, 30, and 10 images, respectively. Yeast Protein Localization databases (YPL.db) were presented in [162, 163] to help determine location of certain protein in yeast cells, and the Dynamic Proteomics database has provided the levels and location information of proteins from living human lung cells for anti-cancer drug studies [164, 165]. However, all of these databases used fluorescent materials to tag or label cells for clearer observation [157, 162, 164]. Two reference time-lapse video datasets without fluorescent labelling were used in [166]. The raw videos are provided for public usage, whilst the manually generated binary object region masks are not included in the repository.

The N2DH-GOWT1 group cell sequences in the cell segmentation dataset [106] were not labeled by fluorescent materials presented in [61]. The group consists of two cell video sequences of 182 images, which are 2D images of GOWT1 mouse embryonic stem cells, taken by a Leica TCS SP5 laser scanning confocal microscope. The proposed cell image contrast enhancement approach presented in Chapter 4 is applied to this group N2DH-GOWT1 stem cell videos (of the training dataset, which has segmentation ground-truth annotated) for evaluation.

3.2.2 Biological organism databases

Due to the difficulties in generating the biological organism video segmentation ground-truth, the general lack of segmentation datasets with annotated ground-truth severely limits the evaluation of segmentation algorithms for biological and biomedical image processing. Standard datasets for benchmarking moving objects in video sequences are still emerging. Furthermore, the evaluation of segmentation approaches for a given biological organism is difficult if there is no suitable dataset for the specific organism under study.

Zebrafish larvae have emerged as a common vertebrate model for biomedical research; however, no publicly available zebrafish larvae video segmentation dataset has yet been discovered, largely due to the time and manual labour required to generate ground-truth segmentation and tracking. Thus, in this work zebrafish larvae videos are used as an example to illustrate the proposed crowdsourced segmentation dataset generation approach introduced in Sections 3.3 and 3.4.

3.3 Video Data Acquisition for Dataset Annotation

This Section proposes an efficient and scalable crowdsourced approach to generate video segmentation ground-truth to facilitate database generation for general biological organism segmentation evaluation. To illustrate the proposed approach, an annotated zebrafish larvae video segmentation dataset has been generated.

To further facilitate the evaluation of the overall tracking performance of tracking systems and to investigate the effect of enhanced segmentation accuracy on the following tracking performance, this work also explores the object location selection (indicated by object coordinate centroid position as tracking ground-truth). Since the tracking ground-truth generation is less time-consuming than that of segmentation ground-truth, except for videos of zebrafish larvae, an extra 10 videos of another

two types of small biological organisms: marine *Artemia franciscana* and freshwater *Daphnia magna*, are also applied to generate the tracking ground-truth. All the segmentation and tracking datasets annotated with ground-truth are made freely available online [1].

3.3.1 Zebrafish larvae video acquisition

Wild zebrafish embryos (*Danio rerio*) were incubated at 28°C in a Petri dish filled with an E3 medium. Any debris and unfertilized embryos were manually removed three hours post-fertilization (hpf). Five days post-fertilization, the larvae were obtained from hatched zebrafish embryos, which normally hatch around 48 hpf, and larvae were not fed before the video data acquisition. For data acquisition, zebrafish larvae were transferred to poly (methyl methacrylate) (PMMA) housing wells.

Low frame rate videos were recorded with a Dino-Lite AD7013MT microscope at frame rates of 14 or 15 fps. An 8-LED light provided cold white illumination to capture videos at a resolution of up to 2592×1944 pixels and magnification power up to $240\times$. The microscope is equipped with a CMOS sensor for imaging in full detail according to the DirectShow imaging standards [167]. The microscope is connected via USB to a PC running Windows 7 to acquire videos using the DionCapture 2.0 imaging software working together with the other microscope's package (such as Carry Pouch, Standard Calibration Target, and Open Cap etc.) and providing a user-friendly interface to directly manipulate time-lapsed video capture.

High frame-rate videos were captured by an Imaging Development Systems (IDS, GmbH, Germany) UI-3360CP-C-HQ microscope camera, with a high-resolution 12.5mm focal lens GMHR412514MCN (F1.4 – 16 iris range, Goyo optical Inc., Japan). Videos can be recorded at up to 152 fps at a full resolution of 2048×1088 pixels, using an illumination sensitive CMV2000 CMOS sensor. High frame-rate videos were

[1] https://github.com/Xiao-ying/

recorded at 117 fps due to the experimental requirements of the research collaborators bioMEMS Research Group [1], who collected and provided all of the zebrafish, *Artemia franciscana* and *Daphnia magna* microscopic videos. The microscope is connected via USB interface to the PC to acquire videos using the IDS Software Suite, enabling 420 MByte/s high data rate responding to the storage of the acquired high frame rate videos.

Table 3.1: Imaging conditions of the presented zebrafish segmentation dataset

Seq. No.	Frame rate/ fps	Imaging condition
1	15	Clear square well, adult seq.
2	14	Clear round well, well edge shadowing
3	15	Round clear well, labels, water particles
4	15	Round clear well, water particle
5	15	Round well, labels
6	117	Round clear well, well edge shadowing
7	15	Round well, labels, water bubbles and particles
8	15	Round well, labels, water particles
9	15	Square well, labels, water particles
10	117	Round well, water particles, edge shadowing

The dataset contains 3056 video frames across a total of 10 videos, with various frame rates and imaging conditions as listed in Table 3.1, presented in the order that the first sequence has the clearest background and the 10th (last) sequence has the most complex background.

Sequences no. 6 and 10 are high frame-rate microscopic videos, with the rest of the video sequences in the dataset at low frame rates; thus, the frame rates of sequences no. 6 and 10 are higher than the other videos. High frame-rate videos provide clear object posture information, such as the tail beating shape of the zebrafish, which make the future analysis of movement kinematics possible. However, high frame-rate videos have many more frames with little movement difference between successive frames. Thus, the large amount of frames can also dramatically increase the computational complexity.

[1] https://www.rmit.edu.au/news/all-news/2016/may/labonachip-technologies-pioneered-at-rmit

3.3.2 *Artemia* and *Daphnia* video acquisition

Cysts of the marine crustacean *Artemia franciscana* and freshwater *Daphnia magna* were hatched and cultured according to the Artoxkit-M [168] and Daphtoxkit-F [169](MicroBioTests Inc., Belgium) standard operating protocols. *Artemia franciscana* were hatched in a petri dish filled with sea water (pH 8.0 ± 0.5) at $24 \pm 0.5^{\circ}$C under exposure to 3000-4000 lux light source for 30 hours. *Artemia* were put in a group of 10 in a miniaturised Lab-on-a-Chip (LOC) chamber [170] when shooting videos with microfluidic infused at a flow rate of 5.25 mL/h. Five *Daphnia magna* neonates were randomly selected and transferred into a petri dish monitored temperature at $20.0 \pm 0.5^{\circ}$C. Both the microscopic time-lapse videos of marine *Artemia franciscana* and freshwater *Daphnia magna* were captured at low frame-rate of 15 fps with the same illumination conditions and imaging software as that of the low-frame rate zebrafish larvae videos.

3.4 Crowd-sourced Segmentation Ground-truth for Dataset Generation

Crowdsourced generation as a general approach to dataset annotations has been previously investigated, but was mainly used for simple annotation such as text classification or point location [171, 172], rather than applied to region based image segmentation, which is usually performed by trained experts. This work proposes to combine the interactive segmentation tool IST [160] with the crowdsourced platform Mechanical Turk[1] to generate segmentation ground-truth for videos of biological organisms. To illustrate the proposed crowdsourced segmentation dataset generation approach, segmentation ground-truth of 10 zebrafish larvae videos were annotated.

[1] https://www.mturk.com/mturk

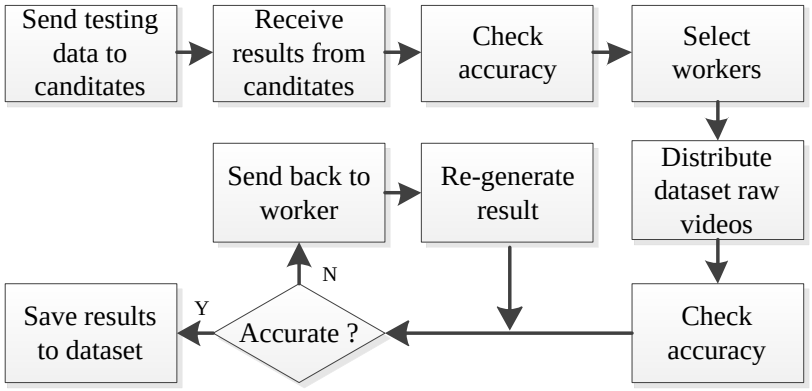


Figure 3.1: Overview of dataset management workflow.

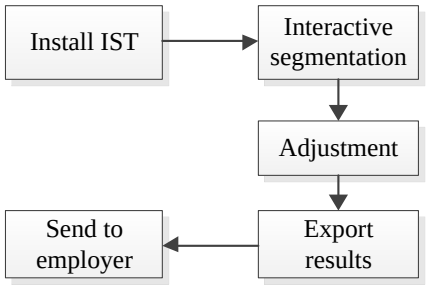


Figure 3.2: Overview of workflow for workers to generate segmentation ground-truth.

Figure 3.1 and Figure 3.2 show the management workflow and the mechanism used to select crowdsourced workers to generate segmentation ground-truth for a dataset, respectively. To ensure that an appropriate crowdsourced workforce is employed, the proposed approach prefers workers who have a similar experience or are in a similar research area as identified on the crowdsourcing platform. Workers with demonstrated ability are then assigned to manually generate the ground-truth segmentation for video sequences, which are then annotated for the segmentation dataset.

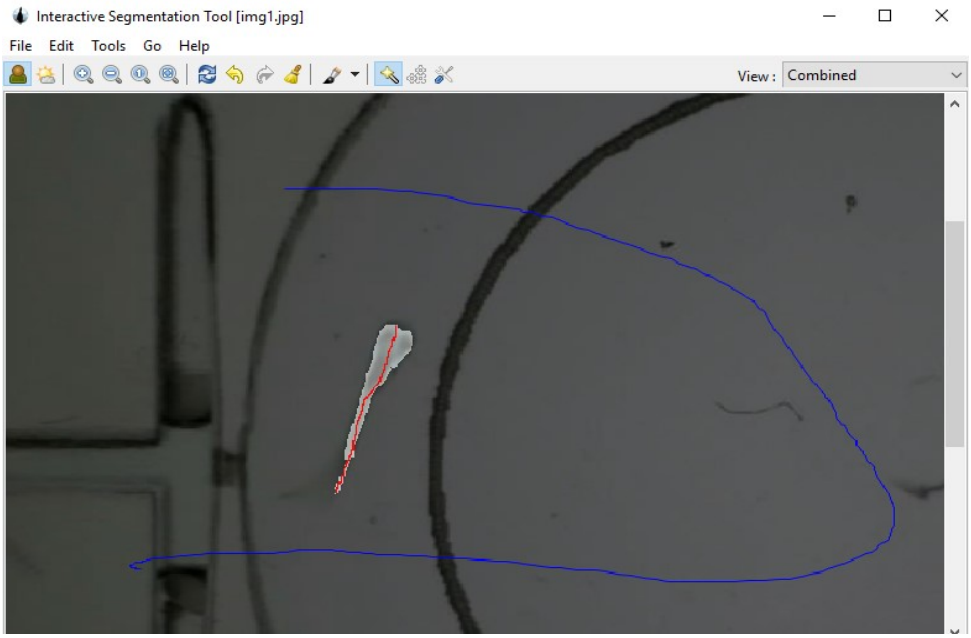


Figure 3.3: Example of selecting target object against the background.

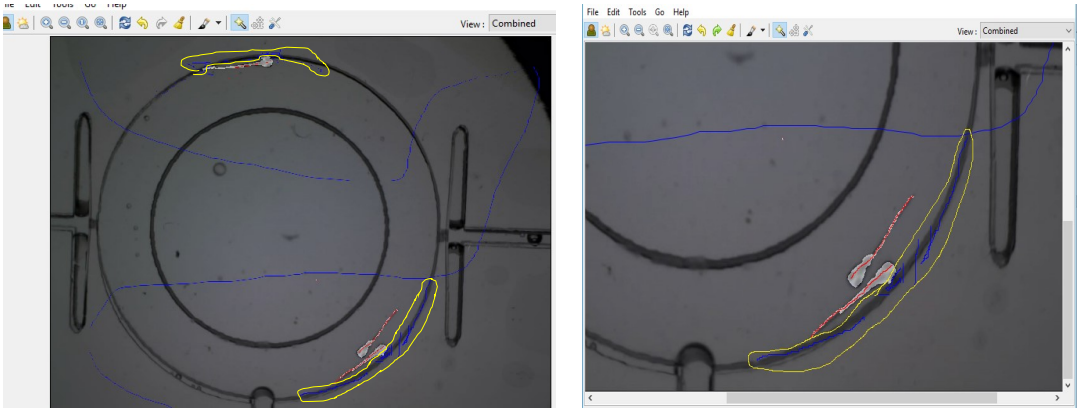
In the following discussion, the crowdsourcing platform utilised is Mechanical Turk; however, the proposed methodology has been designed to be generalisable across various crowdsourcing platforms and image/video content e.g., any object for computer vision applications or any biological organisms.

3.4.1 Interactive segmentation tool

To perform manual ground-truth segmentation, the IST [160] is used. The tool supports four segmentation algorithms: seeded region growing [173, 174], graph cut [175, 176], object extraction [177], and binary partition tree [178].

For the zebrafish larvae segmentation dataset generated in this Chapter, the seeded region growing segmentation algorithm [173, 174] was selected, due to the smooth outline generated by the algorithm for the small zebrafish larvae body and low intensity contrast between the zebrafish body periphery and the microscopic video background. To perform the segmentation, the worker drags a frame into the IST, and the foreground mask is automatically computed after the worker manually indicates on the image the background region/s (indicated by the blue line drawn by clicking and

holding the right key of mouse as shown by Figure 3.3) and foreground object/s of interest (shown by the red line on the fish drawn by the left mouse key as per Figure 3.3).



(a) (b)

Figure 3.4: Adjustment example after object and background differentiation.

(a) Direct adjustment based on the resultant segmentation; (b) Further detailed adjustment by zooming into the frame.

To improve the segmentation accuracy, the generated foreground mask usually requires further (iterative) manual adjustment by the worker. Especially when the object is near the edge of a well, more differentiation lines need to be carefully selected, as shown in Figure 3.4 (highlighted by the yellow lines). Figures 3.4a and 3.4b display examples of further manual segmentation adjustment based on the obtained foreground mask with and without zooming in the video frame.

3.4.2 Crowd-sourced worker training and selection

To generate the zebrafish larvae dataset, an instructional document[1] and samples[1] of segmentation results from a short zebrafish larvae video sequence segmented by the authors using the IST were supplied to candidate workers for training, and a demonstration video provided by [160] was also included to illustrate how the IST is used with a few image examples.

[1] https://github.com/Xiao-ying/-moving-zebrafish-larvae-segmentation-dataset-/tree/master/Data

For the candidate screening test, three zebrafish larvae test images (extracted from a zebrafish video sequence) were sent to the candidate workers. The segmentation results were manually visually inspected and then quantified using the proposed two-pass verification process described in Section 3.4.3 for segmentation accuracy, which determines whether the candidate workers were required to re-generate the segmentation or not further selected to work on the dataset. The final dataset only consists of segmentation results from workers whose segmentation accuracy exceeded 95%.

Workers were selected using the criteria:

1. Having similar image segmentation experience or can complete the training activity according to the provided instructions;
2. Binary segmentation results with the same image size as the original frame;
3. Each 'ground-truth' image having less than 5% error as determined by the proposed two-pass validation process;
4. High consistency with generated results, determined by the 95% confidence interval of the calculated error scores as described in Section 3.4.3.

   Criteria 2, 3, 4 were also used to verify the workers' segmentation results prior to inclusion in the final dataset build.

3.4.3 Accuracy verification of obtained segmentation ground-truth

Workers are assigned one video sequence at a time to segment. When more than 95% of the frames in a video sequence were considered to be accurately segmented, another unique video sequence was sent to the worker for segmentation. To verify the accuracy of workers' segmentation results, a two-pass verification process is proposed: an initial quick manual visual assessment, followed by an objective quantitative assessment.

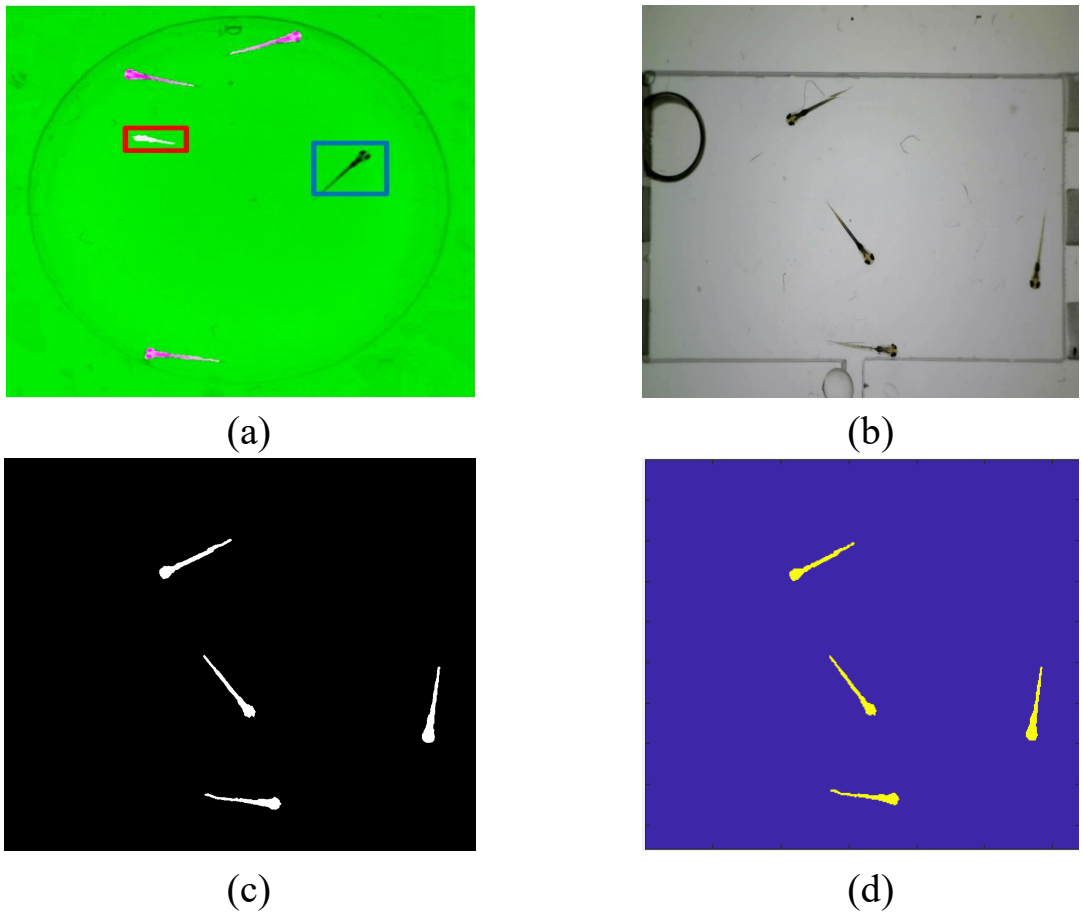
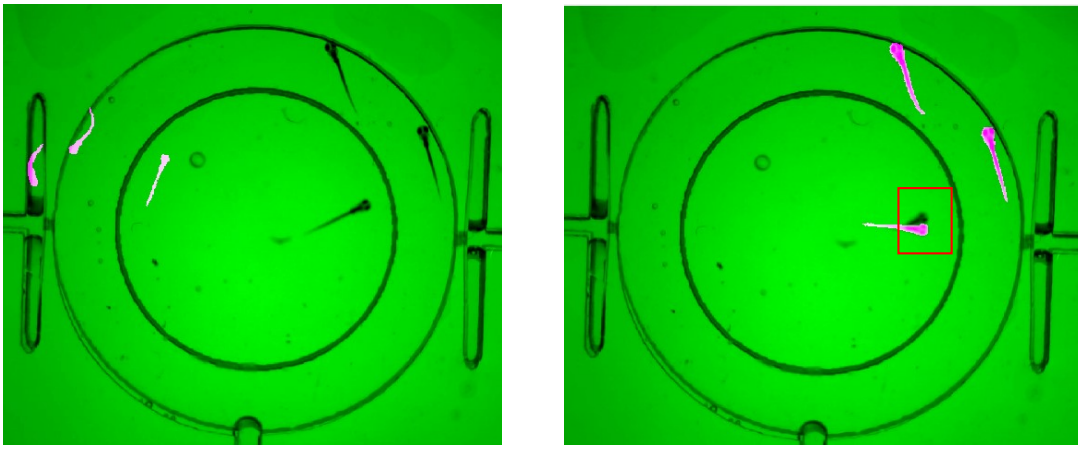


Figure 3.5: Visual examples of segmentation

(a) Visual assessment example; (b) example zebrafish larvae video frame 66 of seq. 4; (c) the crowdsourced segmentation ground-truth of (b); (d) the mask image of comparing (b) and (c) for quantitative assessment.

First-pass: Visual assessment

The received segmentation results are firstly visually assessed by the author to check for visually perceptible errors, such as whether the frame size has been changed by the worker, missing object/s or position deviation. To perform this visual assessment, a joint image (shown in Figure 3.5a) is created for each frame by overlaying each segmentation result from the worker over the corresponding original frame, where object pixels are highlighted in high contrast green-magenta colours for easy visual inspection for errors. Magenta denotes the segmented result for objects in the same position as the original frame, light pink (as highlighted with the red rectangle in Figure 3.5a) is a region for an object segmented in the wrong position, and black



(a) (b)

Figure 3.6: Typical visual perceptible errors.
(a) Error type I: resized frame ; (b) Error type II: position deviation.

highlighted with a blue rectangle is a missed object. If the segmentation deviation (as shown by the red and blue rectangles) of a frame is visually perceptible, then the 'ground-truth' for that frame is discarded and re-generated.

Figure 3.6 shows the two typical errors seen during the worker selection process from the crowdsourced platform, and to be detected by the fast initial visual assessment procedure. The frame was resized (with the segmentation 'ground-truth' generated by the worker not corresponding to its original frame), as shown in Figure 3.6a. The other typical error is shown by Figure 3.6b that the manually generated 'ground-truth' (shown by magenta colour) did not include the whole region of a target object (e.g. the head region of a zebrafish was missed and highlighted in black (within the red rectangle in Figure 3.6b) by the visual assessment software).

The retrieval of a segmentation result and its corresponding original frame from files, and the generation of their joint image in high contrast colours annotation is implemented automatically frame-by-frame using the visual assessment code programmed in Matlab, which is freely available online. No restart is needed for a whole video sequence. Since the initial verification is a fast inspection of visually perceptible errors, automatically looping a whole video with the number of frames generally included in segmentation datasets [23, 106, 157] is not overly large (no more than

10,000 frames with segmentation ground-truth annotated), this initial manual visual verification process will not limit the scalability of the proposed crowdsourced segmentation dataset generation approach. For example, the 10 zebrafish videos in the segmentation dataset generated required only 10 minutes to perform this initial manual visual assessment for 3056 frames.

Second-pass: Quantitative assessment

If no visually perceptible errors are detected, further under or over-segmentation errors from the workers' results are quantified for objective validation, and the performance consistency is calculated by 95% confidence error intervals. A mask image $J(x,y) = j(x,y)|x \in [1,N], y \in [1,M]$ (as shown in Figure 3.5d) of the same size $N \times M$ as the original frame $F(x,y)$ is created based on comparing the intensity values of the video frame (as shown in Figure 3.5b) with the binary 'ground-truth' $G(x,y)$ (as shown in Figure 3.5c) generated by the worker. Pixel values are initialized to 0 and updated by Equation 3.1.

$$A = \begin{cases} 1, & g(x,y) = 0 \quad \& \quad f(x,y) \in (p_{thr}, 255) \\ 2, & g(x,y) = 1 \quad \& \quad f(x,y) \in (p_{thr}, 255) \\ 3, & g(x,y) = 0 \quad \& \quad f(x,y) \in [0, p_{thr}) \end{cases} \tag{3.1}$$

Where p_{thr} is the intensity threshold value calculated using the Otsu algorithm [179] (with two classes per original frame). Then, the left pixels $j(x,y) = 0$ denotes an object in both the generated segmentation and the video frame. The over and under-segmentation error is estimated from the mask image $J(x,y)$:

$$D_{over} = \frac{|J(x,y|j(x,y) = 2)|}{|J(x,y|j(x,y) = 0 \quad \text{or} \quad 3)|} \tag{3.2}$$

$$D_{under} = \frac{|J(x,y|j(x,y) = 3)|}{|J(x,y|j(x,y) = 0 \quad \text{or} \quad 3)|} \tag{3.3}$$

3.5 Segmentation Results and Discussion

3.5.1 The generated zebrafish larvae segmentation dataset

Six workers were employed to generate the dataset, where the manual visual first-pass verification as described in Section 3.4.3 was conducted on the segmentation results generated. Figure 3.7 shows four pairs of video frames with their correspondence segmentation ground-truth annotated in the presented dataset. These example video frames are selected from different sequences with varying numbers of zebrafish (1, 3, 4 and 5 objects, respectively), round (as shown in Figure 3.7a, b, d) or square (as shown in Figure 3.7c, which is also the only adult zebrafish video) containers, and containers with (as shown in Figure 3.7b) or without human drawn labels. The annotated zebrafish larvae video segmentation dataset is freely available online, licensed under the Creative Commons Attribution-ShareAlike license [180].

Table 3.2: Verification of the generated ground-truth for the zebrafish segmentation dataset

Seq. No.	No. of objects	Length [frames]	D_{under} [%]	CI [%]
1	5	201	0.28	[0.27 0.29]
2	4	280	4.43	[4.40 4.47]
3	1	406	0.31	[0.30 0.32]
4	1	151	0.27	[0.26 0.28]
5	3	166	1.57	[1.55 1.59]
6	4	461	4.06	[4.04 4.08]
7	3	759	1.53	[1.52 1.53]
8	3	221	1.41	[1.40 1.43]
9	4	301	2.54	[2.52 2.57]
10	4	110	3.89	[3.74 4.04]

The second-pass quantified objective validation (as described in Section 3.4.3) on the generated ground-truth by the crowdsourced workers is summarised in Table 3.2.

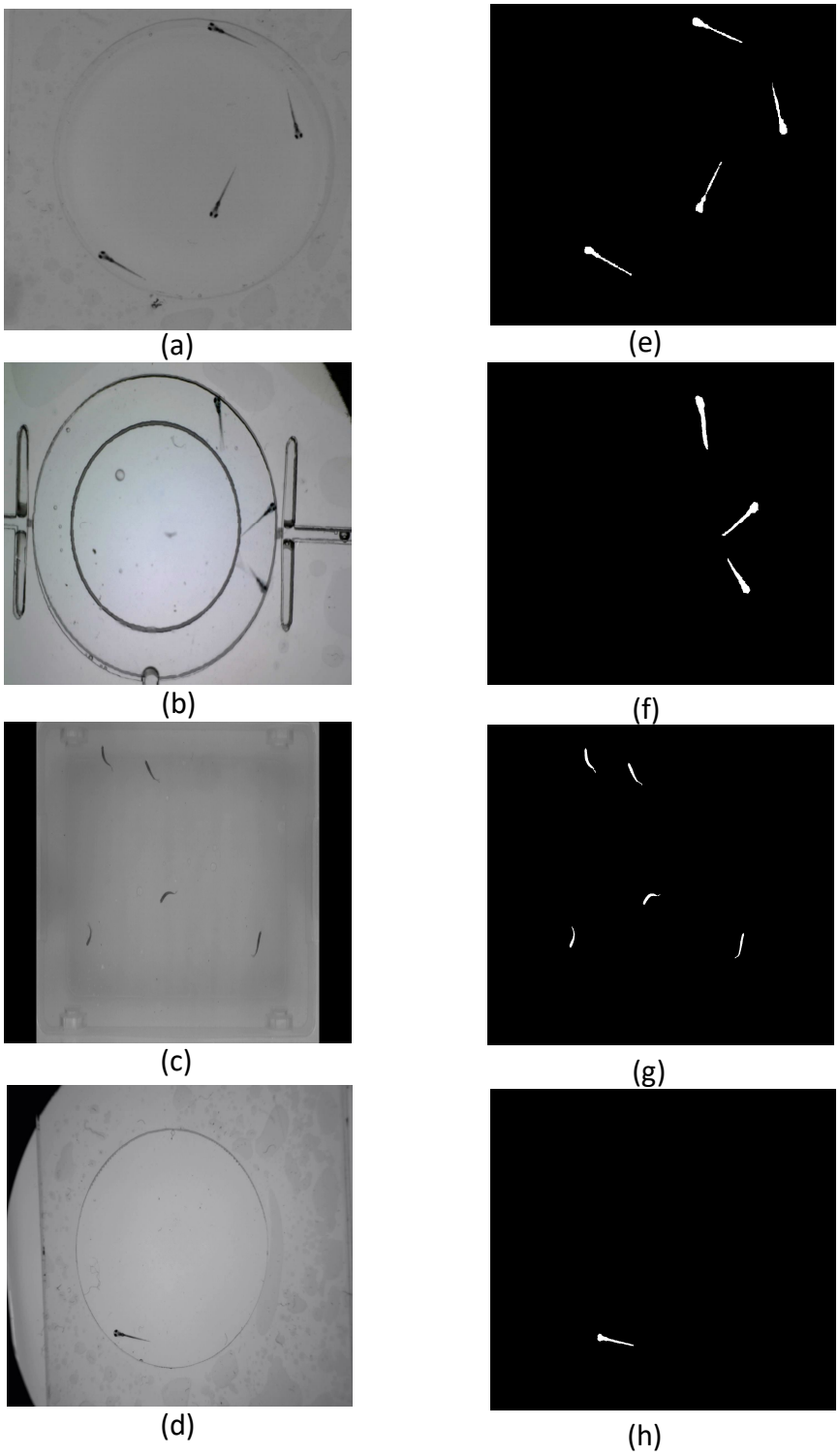


Figure 3.7: Visual example of the obtained segmentation ground-truth selected from the zebrafish segmentation dataset.

(a) Frame 76 of seq. 1; (b) Frame 36 of seq. 2; (c) Frame 49 of seq. 3; (d) Frame 43 of seq. 6; (e)-(h) Segmentation results for (a)-(d), respectively.

Because the transparent body peripheral of the zebrafish larvae causes low intensity contrast, no over-segmentation exists in the generated 'ground-truth' by workers. The overall error score of each sequence by the mean under-segmentation value and the 95% confidence interval (CI) is shown in Table 3.2. It can be seen that the overall under-segmentation error is within 5% for each sequence and indicates a high level of work segmentation consistency with the 95% confidence intervals for the under-segmentation error within 0.02%-0.3%.

3.5.2 Segmentation accuracy evaluation metrics

To enable the practical usage of the proposed dataset generation approach to test and evaluate segmentation algorithms, three standard metrics are presented and a novel metric is proposed to quantitatively evaluate the segmentation performance from videos by calculating the dissimilarity with the generated ground-truth.

Let S and $\bar{S}$ denote the segmentation ground-truth and the result of a segmentation algorithm for image $I = i(x,y)$ of size $N \times M$ pixels. Then, denote a pixel $i(x,y)$ in detected object region $C(S,i(x,y))$ and $C(\bar{S},i(x,y))$, in the ground-truth and segmentation algorithm result, respectively. The segmentation accuracy of an image is thus defined by the three standard metrics [181] as Equations 3.4 - 3.6:

$$Recall = \frac{|C(S,i(x,y)) \bigcap C(\bar{S},i(x,y))|}{C(S,i(x,y))} \tag{3.4}$$

$$Precision = \frac{|C(S,i(x,y)) \bigcap C(\bar{S},i(x,y))|}{C(\bar{S},i(x,y))} \tag{3.5}$$

$$F_{measure} = \frac{2 * Recall * Precision}{Recall + Precision} \tag{3.6}$$

The recall and precision metrics estimate under-segmentation and over-segmentation, respectively. The $F_{measure}$ is a weighted calculation of the precision and recall. Low recall with a very high precision value indicates under-segmentation, whilst low precision with a high recall suggests over-segmentation performance. Therefore, only comparing precision or recall independently does not meaningfully quantify segmentation evaluation, and the $F_{measure}$ is a weighted calculation of the precision and recall (with a perfect score at one).

The proposed Similarity Index (SI) metric in Equation 3.7 accounts for the number of correctly segmented objects, by penalizing missing objects or object occlusion, which has not been considered in the above standard metrics.

$$SI = F_{measure} - \frac{Num_{miss}}{2 * Num_{GT}} \tag{3.7}$$

where Num_{miss} and Num_{GT} are the number of objects missed, and objects detected in the ground-truth, respectively.

3.5.3 Evaluation of the segmentation dataset

To illustrate the utility of the generated zebrafish segmentation dataset and the segmentation evaluation metrics of this work as presented in Section 3.5.2, five leading segmentation algorithms were applied to this dataset with results measured using the proposed evaluation metrics. The segmentation algorithms applied to the zebrafish larvae dataset include motion feature based optical flow (denoted OptFlow in Figures) [182] and dense SIFT flow (denoted DenseSIFT in Figures)[154] methods, intensity feature based Discrete Region Competition (denoted DiscreteRegion in Figures) [183] and Squassh [184], and the well-known zebrafish tracking system, idTracker [31]. Figures 3.8 - 3.11 show the average precision, recall, $F_{measure}$, and similarity index scores presented with the 95% confidence intervals for each of the 10 video sequences in the dataset.

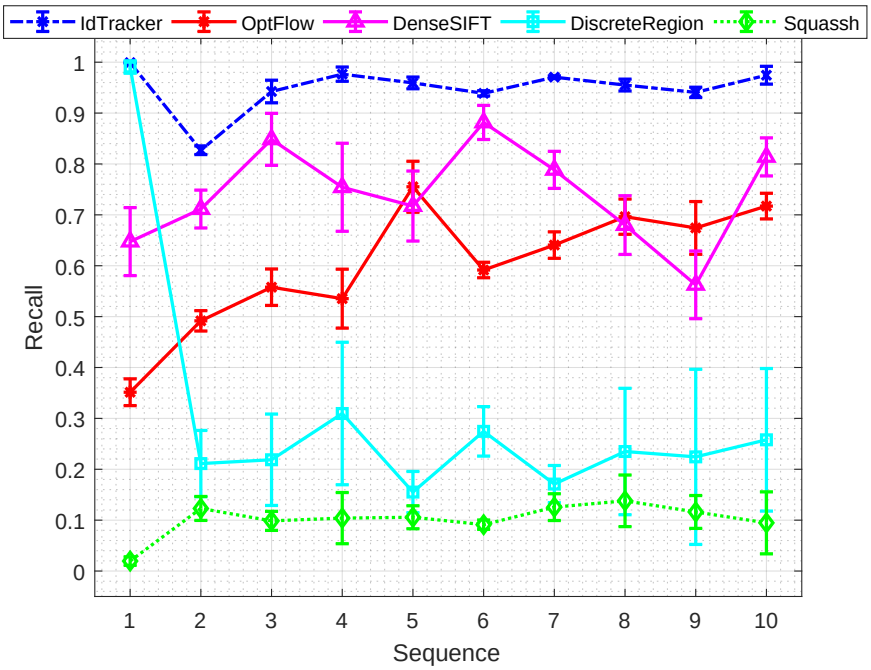


Figure 3.8: Average segmentation evaluation recall scores for each video sequence.

Recall. Figure 3.8 shows that the Discrete Region and Squassh methods obtained very low recall scores over the sequences, because many zebrafish pixels are classified as background. The Discrete Region method requires objects to have homogenous intensities, as demonstrated by only sequence 1 exhibiting a high recall due to adult zebrafish being in the sequence (where zebrafish larvae have inhomogeneous intensities). Comparing with the optical flow method, the dense SIFT flow method detects fewer irrelevant segments, which results in a higher recall performance. In contrast, the recall score of idTracker is consistently high. This is caused by the universal over-segmentation of idTracker: as the background pixels with similar intensity values with the zebrafish larvae are classified as objects.

Precision. In Figure 3.9, low precision scores for Dense SIFT, Discrete Region and Squassh illustrate the over-segmentation for zebrafish larvae of these methods. The precision score of idTracker is not consistently high under complex imaging conditions, as the zebrafish larvae body peripheral is transparent, the intensity contrast is therefore too low against the background to be separated. In contrast, the precision

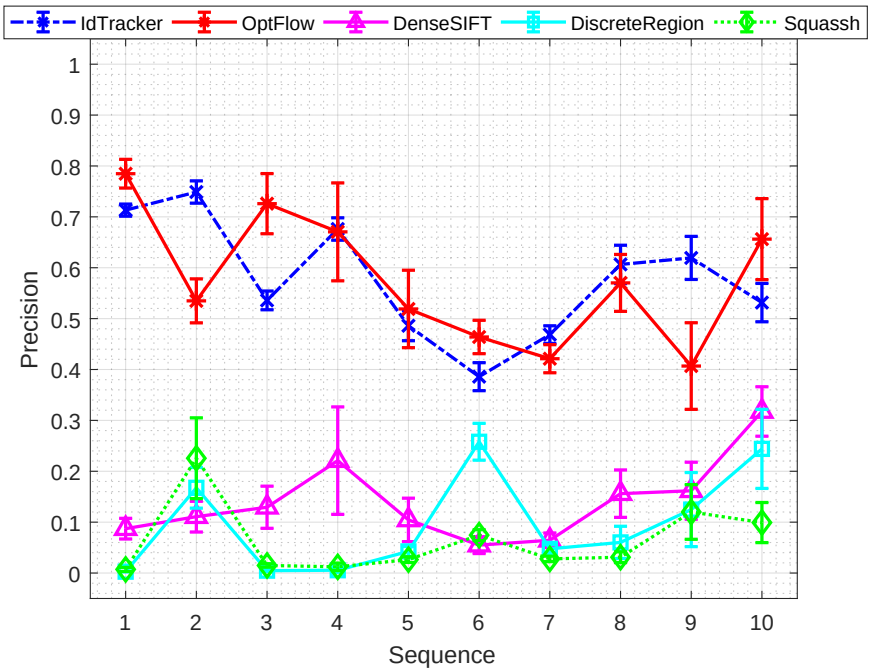


Figure 3.9: Average segmentation evaluation precision scores for each video sequence.

performance of the optical flow method steadily outperforms the other methods when using the high frame rate sequences 6 and 10 amongst the dataset, due to the short movement distance at high frame rates. In comparison, due to dense SIFT flow being very sensitive to movement, small moving water impurities and the projection of water surface ripples on the container edges are incorrectly detected as objects. This diminishes the portion of true zebrafish regions segmented to, therefore, result in a consistently low precision score, as seen from Figure 3.9.

$F_{measure}$ and similarity index. The $F_{measure}$ and similarity index in Figures 3.10 and 3.11 show the overall weighted segmentation accuracy combining Recall and Precision. The overall segmentation accuracy of idTracker has an average of 15.18%, and 48.57% higher $F_{measure}$ than optical flow and SIFT flow, respectively. The pure intensity contrast based Discrete Region and Squassh techniques have the poorest overall performance. The idTracker system also exhibits a 12.55% and 70.57% higher similarity index compared to optical flow and SIFT flow, respectively, indicating improved performance in relation to missing or occluded objects. In particular, the

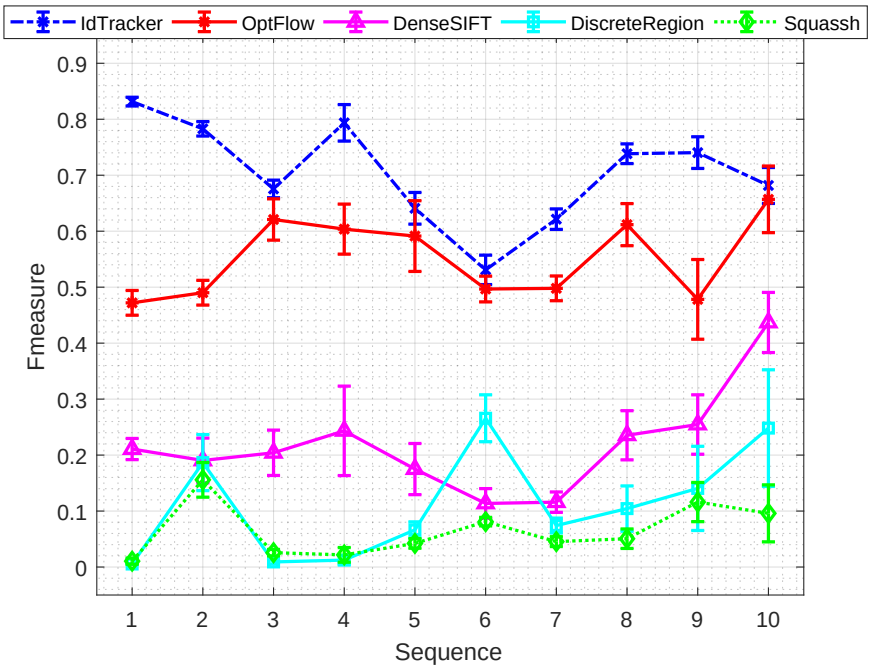


Figure 3.10: Average segmentation evaluation $F_{measure}$ scores for each video sequence.

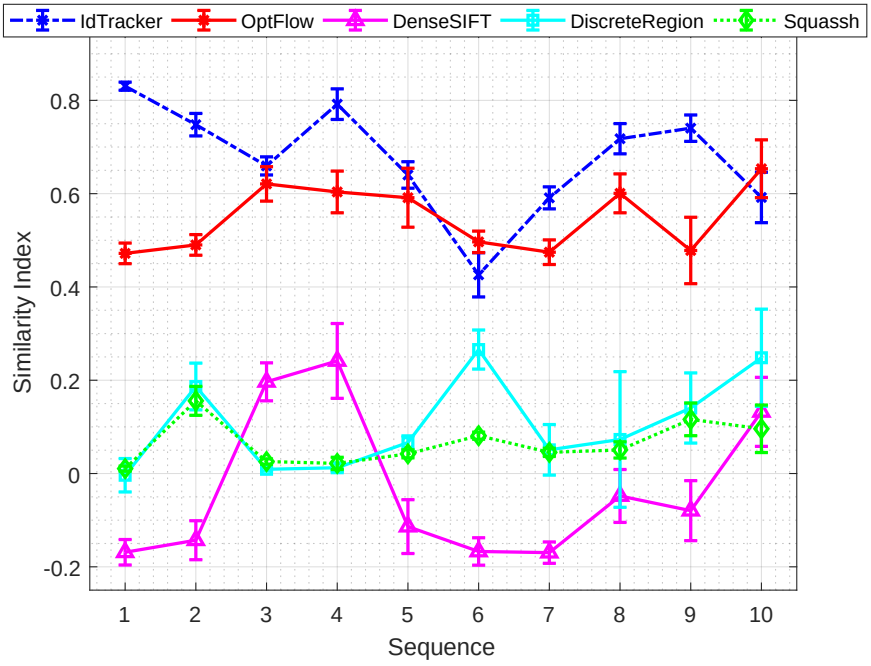


Figure 3.11: Average segmentation evaluation SI scores for each video sequence.

optical flow method is more robust against challenging background environments, such as unclear zebrafish well containers with labels (as illustrated by sequence 5) and high frame rate videos (as illustrated by sequences 6 and 10), demonstrated by the 1.82% and 1.14% smaller $F_{measure}$ and similarity index variance, respectively.

3.6 Tracking Ground-truth Generation

To extend upon the generation of segmentation ground-truth annotations and measure the subsequent object tracking performance, for the datasets used in this work the centroid positions of all objects in each frame are manually selected as the tracking ground-truth. The proposed approach and software for this tracking ground-truth generation is presented in Section 3.6.1, with the annotated tracking datasets generated presented in Section 3.6.2.

3.6.1 Tracking ground-truth generation software

To improve the efficiency of the tracking ground-truth generation, the centroid positions are generated frame-by-frame for one object using the GUI shown in Figure 3.12. Figure 3.13 shows the flow chart of the tracking ground-truth generation software proposed. The software first loads the video, then allows the user to select one object centroid, advances to the next frame automatically, and stores the annotated object centroid per frame in a matrix and writes file every 10 frames as shown in Figure 3.13. After selecting all the positions of one object from the whole sequence, the position ground-truth of another object from the first video frame is generated without having to restart the software.

For videos of multiple objects, especially for long duration microscopic videos, it is not uncommon that the ground-truth selection process cannot be finished at once,

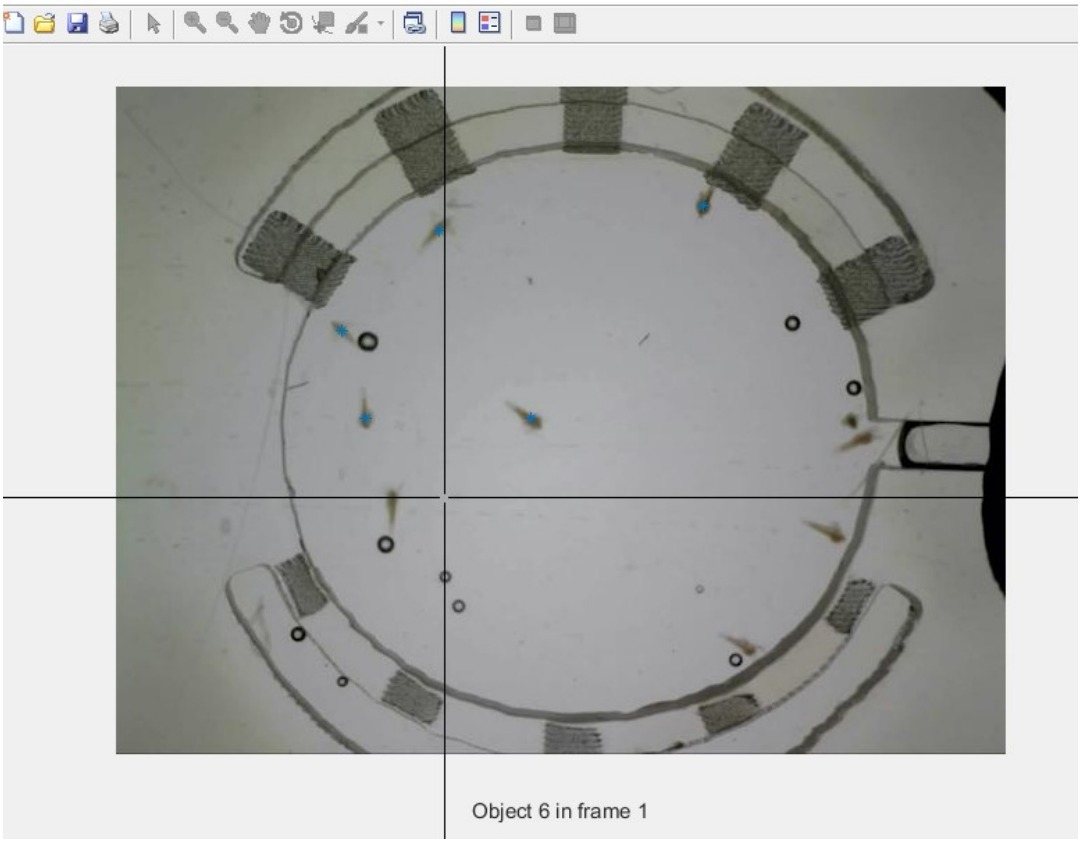


Figure 3.12: Interface of the proposed tracking ground-truth generation software to select the centroid positions of target objects in microscopic videos

where the track of the target object may be lost due to multiple instances of running the software. To avoid requiring re-generation of the tracking ground-truth, the generated ground-truth results for a target are stored every 10 frames. Then, prior to setting up a new matrix for storing the tracking ground-truth of a video, the software will search whether a ground-truth file already exists, and re-loads the existing ground-truth file to add new ground-truth data based on the existing file if a time-lapse sequence has already been partially generated (as indicated by the existence of a ground-truth file for the video sequence). Further, the positions of targets with tracking ground-truth generated are displayed on the current video frame using blue asterisks as shown in Figure 3.12, so that the track of the current target (to be selected by the centre of the cross in Figure 3.12) is visually obvious to the annotator. This visual feedback for generated ground-truth assists to distinguish the target object from its neighbour objects, especially in instances where two or more objects overlap and

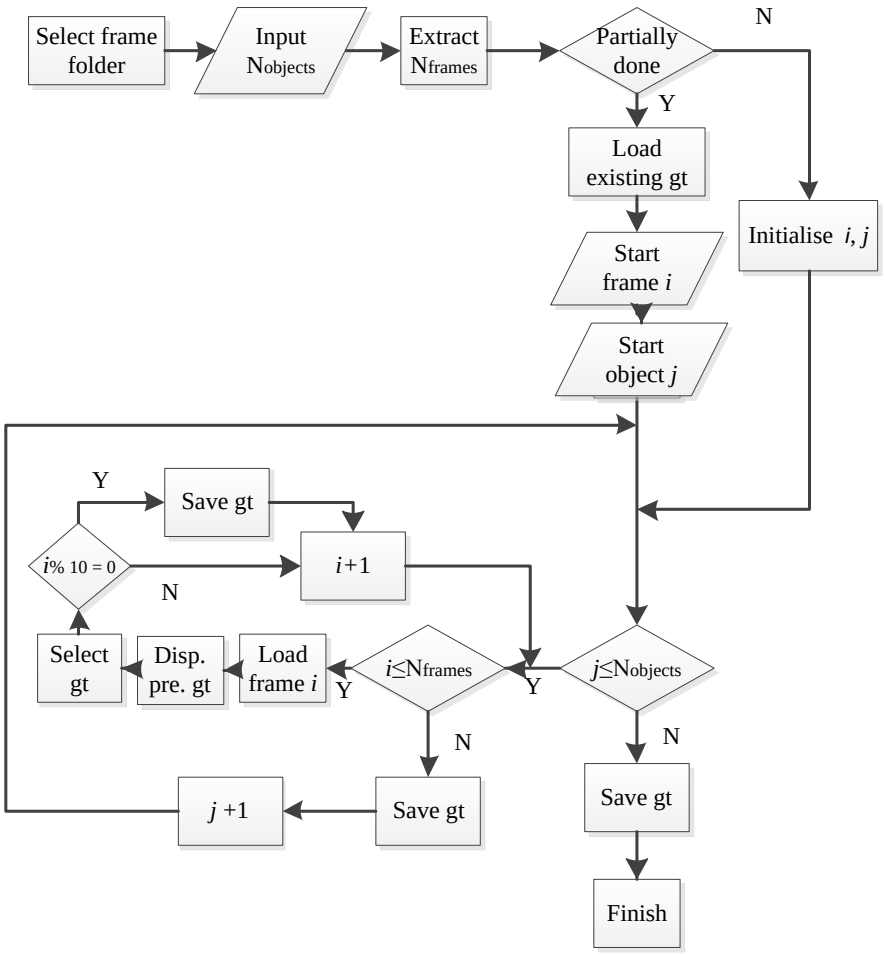


Figure 3.13: Flow chart of tracking ground-truth selection for whole video sequences.

separate again.

3.6.2 Generated tracking datasets

Since the tracking ground-truth generation requires less time and resources, three tracking datasets annotated with manually generated position ground-truth were generated in this work and provided for public usage. These annotated tracking datasets also enable the evaluation of small biological tracking systems, proposed and presented in Chapter 5 tested on zebrafish larvae and other small organisms.

Dataset 1: Zebrafish larvae tracking

To accompany the generated segmentation ground-truth as presented in Section 3.5.1, the centroid coordinate positions of all zebrafish larvae in each frame across the 10 video sequences were also manually selected using the GUI software for tracking ground-truth of this work, presented in Section 3.6.1.

Dataset 2: *Artemia* tracking

The *Artemia franciscana* dataset consists of 5 video sequences with 4802 frames in total. *Artemia franciscana* microscopic videos containing 5 organisms and artifacts (e.g., bubbles of different sizes) are provided in the dataset[1].

Dataset 3: *Daphnia* tracking

The *Daphnia magna* dataset[2] consists of 5 video sequences with 4804 frames in total. *Daphnia magna* microscopic videos contained 10 organisms and artefacts (e.g., bubbles and impurities of different sizes).

3.7 Summary

This Chapter proposed a crowdsourced approach with interactive segmentation methods to generate video segmentation datasets annotated with ground-truth to facilitate the evaluation of image and video segmentation algorithms, particularly for biological organisms due to the distinct lack of annotated evaluation datasets currently available. To verify the crowdsourced segmentation results, a two-pass verification process was proposed: firstly, the generated 'ground-truth' was visually assessed manually by creating a joint image with the original frame, using high contrast colors

[1] https://github.com/Xiao-ying/Artemia-tracking-dataset
[2] https://github.com/Xiao-ying/Daphnia-Video-Tracking-Dataset

to differentiate the dissimilarity; secondly, numerical quantification of the segmentation accuracy and reliability were obtained by comparing the result with frames thresholded with Otsu's algorithm. In addition, a set of segmentation evaluation metrics were presented to enable the evaluation of segmentation algorithms using such annotated datasets.

To illustrate the proposed crowdsourced segmentation dataset generation approach, a zebrafish larvae segmentation dataset composed of 10 sequences in various imaging conditions was generated for zebrafish segmentation applications. The segmentation performance of five leading segmentation approaches was evaluated based on the generated segmentation dataset and then quantified applying the proposed evaluation metrics.

For evaluation of the subsequent object tracking process after segmentation, manual tracking ground-truth annotation software for whole video sequences is also presented, with annotated tracking datasets for three types of small organisms generated as part of this work.

All software and annotated datasets presented in this Chapter are freely available online for repeatable research usage.

Chapter 4

Improved Cell Segmentation

4.1 Introduction

The automatic tracking of biological cell movement provides an efficient approach to investigate their dynamic behaviour characteristics in biomedical research and applications [105], compared with traditional manual visual inspection approaches. Because automatic cell tracking techniques require accurate cell segmentation from the video background, and the detection-based automatic tracking methods require cell segmentation results in each frame [28], improving cell segmentation accuracy has attracted significant research attention in recent years. However, current segmentation techniques are susceptible to errors due to non-ideal but realistic microscopic cell video conditions, including low contrast typical of cell microscopic images. Thus, this Chapter proposes a novel image pre-processing technique to enhance the low greyscale image contrast for improved cell segmentation accuracy.

Gaussian, uniform and exponential models are generally used histogram specification models for intensity contrast enhancement [185]. Using uniform models, grey levels of the output images are equally spread out over the entire greyscale range, such that the grey levels are not concentrated around certain regions of intensity values [124]. The method is predictable and simple to implement. However, an uniform Probability Distribution Function (PDF) model would not be a suitable option when

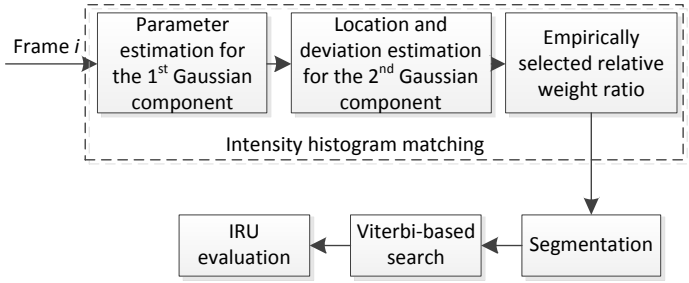


Figure 4.1: Flow chart of the proposed pre-processing method.

an image is dominated by large, dark intensity pixels with numerous pixels at higher grey levels (which is the case for cell microscopic images), because the transform model will rapidly increase the concentration of pixels from low grey levels to high grey level intensity [124]. The exponential model was comparatively less studied and proposed by Gyorgy in [186] to restore over- and under-exposure images, but this model was only tested on artificially distorted images. On the basis of the fact that each homogeneous area in natural images has a Gaussian-shaped intensity histogram [187], an adaptive bi-Gaussian mixture model is investigated. This model stretches the narrow histogram of low-contrast cell microscopic images to a larger greyscale range, with the highest similarity to the original intensity histogram whilst adapting to background changes. Moreover, the mean values of Gaussian components represent the average intensity values of major regions [187], so the contrast between the major regions can be precisely enhanced by spreading these calculated dominant intensity values.

This Chapter investigates the intensity PDF characteristics of time-lapse cell videos taken by microscopes. Figure 4.1 shows the flow chart of the proposed pre-processing method, subsequent processing (as per existing approaches [188, 189]) and the segmentation result evaluation. A primary dominant characteristic feature of the intensity PDF is found to be low intensity bins capturing the most intensity pixels in the intensity histogram (i.e., the first dominant feature of the cell image intensity

histogram, with further details as presented in Section 4.2). Accordingly, a shifted bi-Gaussian mixture model is proposed to match the image intensity histogram to increase the contrast differentiation between the target cells and the video background. A bi-Gaussian mixture model is selected based on Gonzalez et al [190] and details as presented in Section 4.2. The model is shifted because of the first dominant feature of the cell image intensity histogram, as detailed in Section 4.2.2. The model parameters of the first Gaussian component are estimated based on this dominant feature, presented in Section 4.2.2.

After further investigation of the intensity PDF, an additional dominant feature in the intensity histogram of cell video frames was found. This new dominant feature is used to determine the location and deviation parameters of the second Gaussian component to form a bi-Gaussian model, presented in Section 4.2.4.

Rather than using a model with fixed parameters across an entire video sequence, this Chapter presents an adaptive derivation of the mixture model parameters for each video frame to address changes in the video background. Changes in the video background are due to cell activities such as mitosis and apoptotic events, and have been documented as part of the dataset [106]. By deriving the model parameters for both bi-Gaussian components directly from the original intensity histograms, this work removes the need for human intervention in mathematical model construction to represent the video background. Also, direct derivation of the model parameters from the micrograph intensity histogram enables automatic intensity contrast enhancement, and optimally maintains the dominant features of the original intensity histogram.

For evaluation, the proposed cell micrograph segmentation pre-processing approach is applied to an existing Viterbi-based cell search algorithm [191], and compared with existing contrast enhancement methods, using a stem cell dataset (presented in Section 3.2.1 in Chapter 3) without using any (fluorescent) labels on the target cells. This stem cell dataset was used as an example cell database presenting various

challenges for cell segmentation as discussed in Section 2.2.

In the remainder of this Chapter, Section 4.2 presents the technical methodology of the proposed micrograph intensity contrast enhancement. Section 4.2.2 and Section 4.2.4 explain the model parameter estimation for the first and second Gaussian components from the primary and secondary dominant features of the cell micrograph studied, respectively. Section 4.3 details the experimental evaluation conducted and discusses the results obtained, while Section 4.4 concludes the Chapter.

4.2 Proposed Cell Micrograph Pre-processing

Based on the fact that each homogeneous area in natural images has a Gaussian-shaped histogram [187], a Gaussian-shaped model was initially chosen as a candidate model. In addition, there are few impurities or other particles in the nourishing environment, thus the cell images can mainly be categorised into cell foreground and image background regions. Thus, a bi-Gaussian distribution model was studied and applied in this work to match the cell microscopic image intensity histogram. Moreover, the intensity contrast between the individual major regions of the cell microscopic images is increased by spreading out the dominant grey levels, which further improves the global contrast of the images.

This Section presents the proposed image pre-processing approach for cell microscopic image contrast enhancement. Figure 4.2a shows an example but representative cell image from a microscopic time-lapse video sequence, with Figure 4.2b illustrating the corresponding (grey level) intensity histogram for the cell frame example. The cell image has 256 grey levels [0 255], with 0 denoting black and 255 indicating white colours (also pixels representing stem cell regions in the image). The low cell image contrast can be clearly seen from Figure 4.2a that the main area of the image is very dark in colour. It is also very hard to clearly distinguish the peripheral and boundary of cell bodies from the background. This low image contrast can be seen

from the (grey level) intensity histogram of the cell image as shown in Figure 4.2b. That is, the majority of the image intensity is populated within low intensity grey levels [0, 10] (shown in Figure 4.2b). This is the visually perceptible primary dominant feature of the original intensity histogram.

Gonzalez et al showed that images with intensity histograms that exhibit two distinct summits separated by a valley can be segmented well using a threshold-based segmentation method [190]. A mathematical function that exhibits these characteristics is a bi-Gaussian distribution, represented by Equation 4.1 [190].

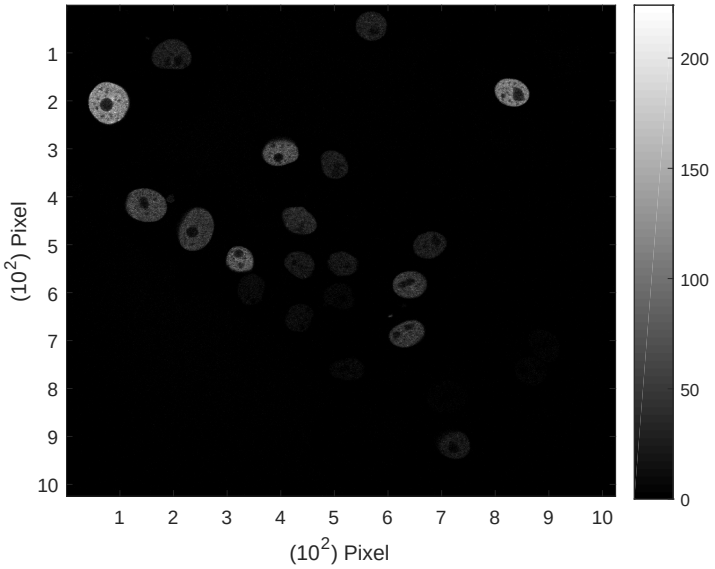
$$\begin{aligned} h(x) = C_1 \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-u_1)^2}{{\sigma_1}^2}\right) \\ + C_2 \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x-u_2)^2}{{\sigma_2}^2}\right) \end{aligned} \tag{4.1}$$

where C_1 and C_2 denote the weight of each Gaussian component, u_1 and u_2 are the distribution means, and σ_1 and σ_2 denote the standard deviations.

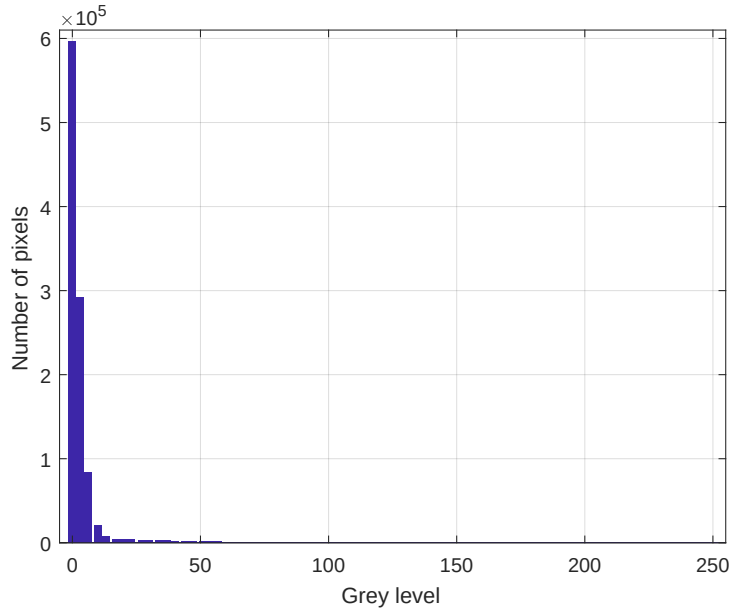
A bi-Gaussian mixture model example of this mathematical distribution represented by Equation 4.1 is presented in Figure 4.3 as a visual example.

4.2.1 First Gaussian component intensity histogram matching

This Chapter proposes an image pre-processing technique that matches the original cell image intensity histogram to a bi-Gaussian model, to introduce separated distribution summits based on the original intensity histogram to improve subsequent segmentation accuracy. To perform histogram matching for the greyscale cell images, let $p_r(r_k)$ and $p_z(z_k)$ denote the probability density distribution of the grey level intensity values in the original and desired post-processed images. The original probability density distribution $p_r(r_k)$ is estimated from the original image, and



(a) Microscopic stem cell video frame example



(b) Intensity histogram of (a)

Figure 4.2: Microscopic stem cell video frame example and its original intensity histogram

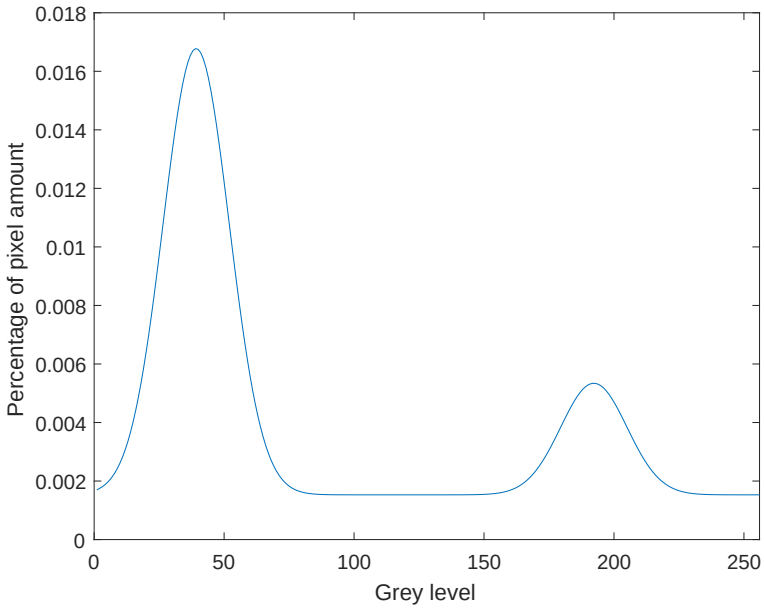


Figure 4.3: Bi-Gaussian mixture model example.

$p_z(z_k)$ is the probability density distribution desired (that is, the bi-Gaussian model [128]) for the output image (post-processed).

Let s_k be the random variable with the property [128]:

$$S_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n} k = 0, 1, 2, \cdots, L-1 \tag{4.2}$$

where n is the total number of pixels in the original image of $(L-1)$ grey levels and n_j is the number of pixels with the grey level r_j. Equation 4.2 presents the transformation function $T(r_k)$, obtained from the original image intensity histogram to obtain the required histogram.

Then, define z_k as a random variable with the property [128]:

$$v_k = G(z_k) = \sum_{i=0}^{k} p_r(r_i) = S_k \tag{4.3}$$

where v_k is the grey level value in the output image that has the specified bi-Gaussian probability density distribution $p_z(z_k)$; Equation 4.3 thus obtains the transformation

function $G(z_k)$ from the specified bi-Gaussian distribution.

Thus, by combining Equations 4.2 and 4.3 [128]:

$$G(z_k) = T(r_k) = S_k \tag{4.4}$$

Finally, z_k satisfies the following equation to calculate the grey levels for the shifted bi-Gaussian histogram [128]:

$$z_k = G^{-1}[T(r_k)] \tag{4.5}$$

4.2.2 Parameter estimation for the first Gaussian component

Applying this approach, a bi-Gaussian mixture model is derived and applied to match the original cell image intensity histogram. To maintain the primary and dominant feature of the intensity histogram as shown in Figure 4.2b, the applied bi-Gaussian model is proposed to shift towards the desired probability density distribution of bi-Gaussian shape, as shown in Figure 4.3.

To maintain the dominant feature of a small fraction of the histogram saturating between grey level range $[0, 10]$ as shown in Figure 4.2b (where Figure 4.2b is the representative of a typical cell image intensity histogram in the database studied):

- mean value u_1 of the first Gaussian mode is shifted to zero to ensure that the centroid location of the first mode is at the 'Y' coordinate axis; mean value u_2 of the second mode is initially shifted to grey level 200 to position the second summit at the original small convex location just below the grey level of 200, which corresponds to the small histogram fluctuations observed around this grey level across the database studied and ensures a valley between the two Gaussian modes.

- standard deviations σ_1 and σ_2 are initially fixed at 0.05 to obtain the same fluctuation width of the first Gaussian mode such that the first mode occupies grey level range [0, 50], consistent with the original histogram grey levels to maintain the shape of the original intensity histogram in the processed resultant image. Since the interval $u \pm 2\sigma$ covers 95% of the variation [192], σ_1 is obtained and updated by Equation 4.6.

- the weight of components C_1 and C_2 are derived according to the original image intensity histogram shape in Figure 4.2b: as the highest fraction of pixels cluster at the low grey levels, C_1 in the bi-Gaussian model is always greater than C_2.

To enable adaptive parameter derivation for u_2, σ_1 and σ_2, the secondary dominant characteristic of the intensity histogram of cell micrographs is discussed and presented in Section 4.2.4.

$$\sigma_1 = \frac{1}{2}\mathtt{arg}_j \left\{ \sum_{i=0}^{j} N_i = 0.95 \right\} \tag{4.6}$$

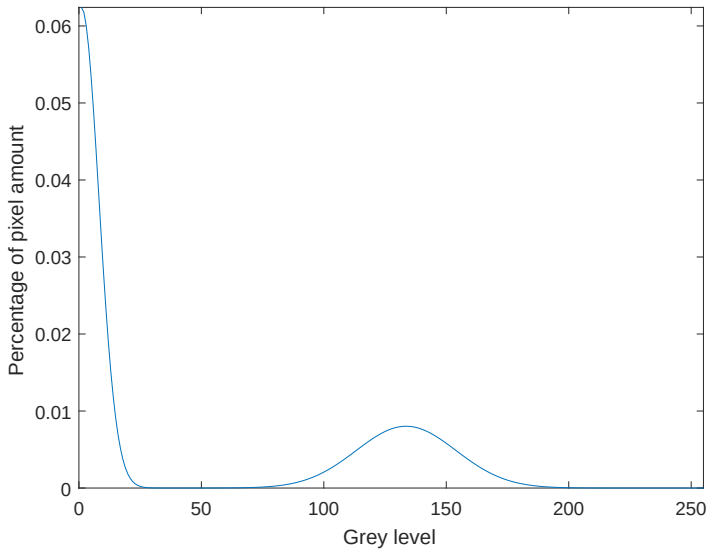


Figure 4.4: Shifted bi-Gaussian model.

Figure 4.4 illustrates an example resultant shifted bi-Gaussian histogram obtained from Equations 4.2 - 4.5, using the example cell image and intensity histogram shown in Figures 4.2a and 4.2b. Figure 4.4 shows that the two distribution summits are clearly exhibited with a valley separating the summits, but still exhibiting the original intensity histogram shape.

4.2.3 Second Gaussian component intensity histogram matching

The first bi-Gaussian component represents the primary dominant feature of original grey level intensity histogram exhibiting a majority (approx. 90% - 95%) of pixel intensities saturating at low grey levels. However, whilst the characteristics of the first bi-Gaussian component can be derived from the primary dominant histogram feature, the location and variation of the second Gaussian component (the parameters u_2 and σ_2) cannot be derived from the first dominant feature of the cell image intensity histogram as shown in Figure 4.2b.

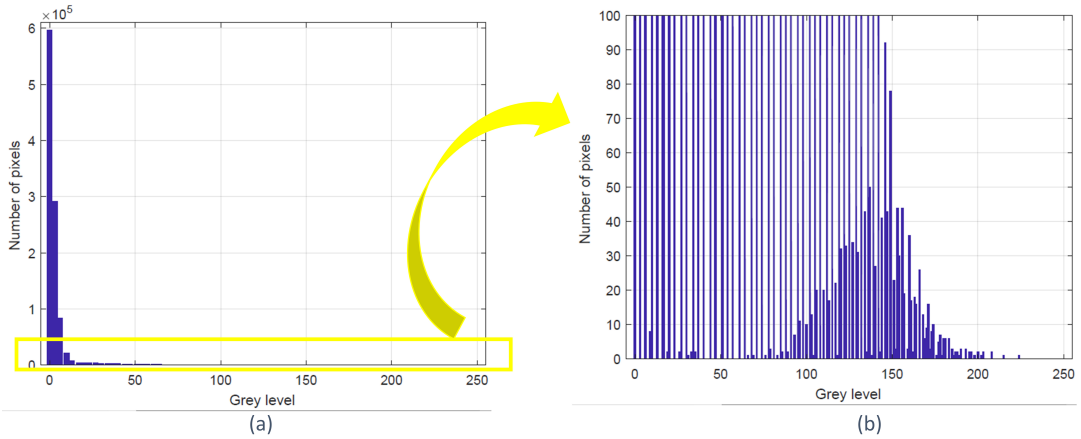


Figure 4.5: Location of the (hidden) secondary Gaussian shape.
(a) Intensity histogram of in Figure 4.2a; (b) Intensity histogram fragment of (a) within the range labelled by the yellow rectangle.

By further investigating the intensity histogram and inspecting the pixel number range [0, 100] (as labelled by the yellow rectangle in Figure 4.5a, and magnified for visibility in Figure 4.5b), the secondary dominant histogram feature can be seen.

That is, another Gaussian shape is exhibited in the intensity histogram, as shown between grey levels [60, 210] in Figure 4.5b. Figure 4.5 also displays the location of the second Gaussian shape, hidden in the low bins (the value of the bins are within [0, 100] against the dominant bin values on the order of 10^5 in scale as shown in Figure 4.2a) in the original intensity histogram. The second Gaussian shape only appears to be visually perceptible after considering the low bin range [0, 100], as per the example shown in Figure 4.5b. It should be clarified that the maximum number of pixels (i.e. the y axis value) in Figure 4.5b is 100, because Figure 4.5b is the zoomed-in region highlighted by the yellow box in Figure 4.5a. The difference between the number of pixels within the bin range [100, 200] is still very small as shown in Figure 4.5a.

By closer examination of the second Gaussian shape, as shown in Figure 4.6 (a magnified intensity histogram fragment), the grey level density within this pseudo-Gaussian shape is much greater than the remaining grey level distributions, to describe subtle intensity variation within cell regions.

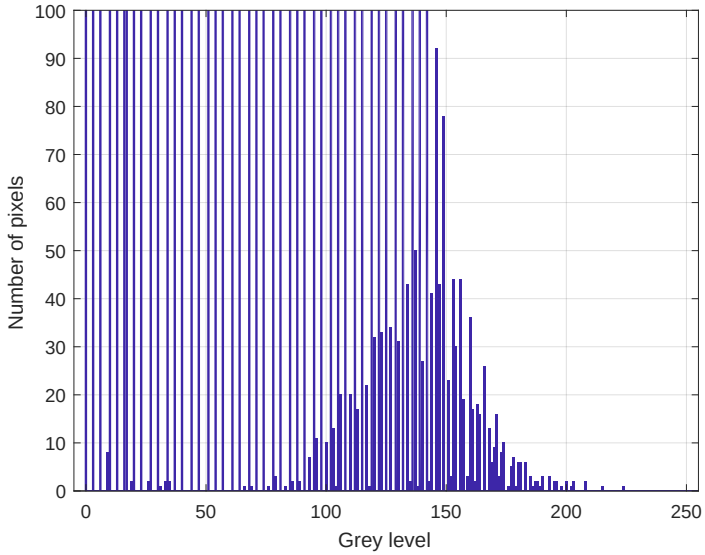


Figure 4.6: Close examination of the second Gaussian shape.

4.2.4 Parameter estimation for the second Gaussian component

The parameter estimation for the second Gaussian component presented in this Section extends the work in Section 4.2.2 [52], where the parameters of the second Gaussian component (indicating the cell pixels in the shifted bi-Gaussian mixture model) are empirically selected and fixed across the entire cell image sequence.

Here, the mean value u_2 and standard deviation σ_2 of the second component in the bi-Gaussian mixture model are derived from the secondary dominant histogram feature. To estimate and update the model parameters for the second Gaussian component, the number of pixels n_i of each grey level i in the intensity histogram are normalised using Equation 4.7. The number of pixels are normalised to ensure that all of the number of pixels are within the range of [0, 1].

$$N_i = \frac{n_i}{\sum_{i=0}^{255} n_i} \tag{4.7}$$

where $\sum_{i=0}^{255} N_i = 1$. N_i is the normalised number of pixels for grey level i with the range of N_i within [0, 1).

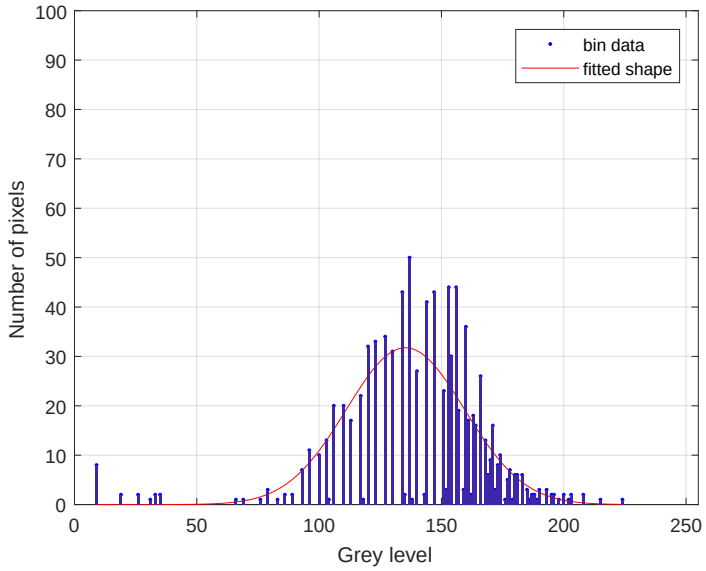


Figure 4.7: Fitted Gaussian model for the second Gaussian shape.

Firstly, the bin values within this secondary Gaussian shape $\mathbb{G}$ are extracted by eliminating zero and background bins from the original histogram (whose bin values are larger than the median bin value due to the first dominant feature in the intensity histograms). The median bin value is calculated from the overall normalised number of pixels i.e., $\text{median}(N_i)$. As a visual example, the resultant histogram bins after eliminating zero and background bins from the original intensity histogram are shown by the blue bars in Figure 4.7.

$$x = \frac{x - \text{mean}(X)}{\text{std}(X)} \tag{4.8}$$

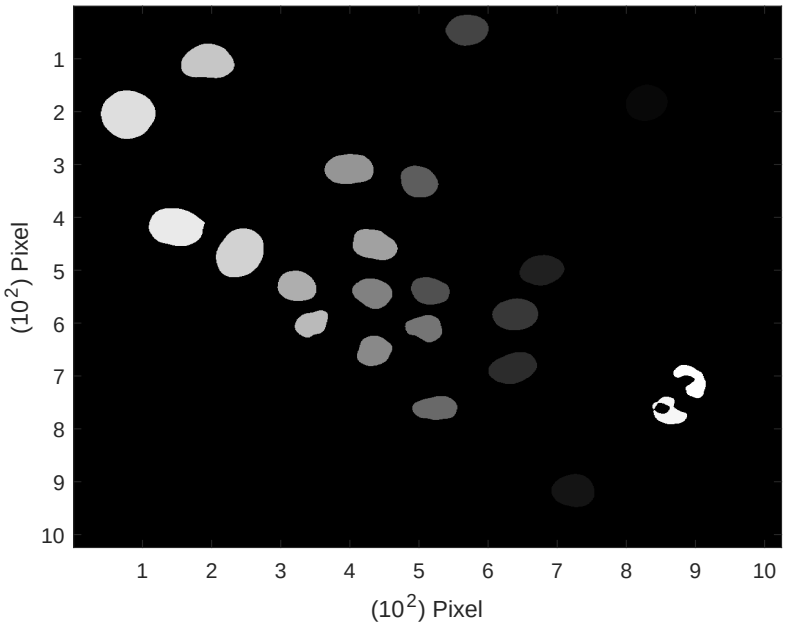
where X stores grey levels within $\mathbb{G}$.

$$y = coef(1) * \exp(-(\frac{x - u_2}{\sigma_2})^2)) \tag{4.9}$$

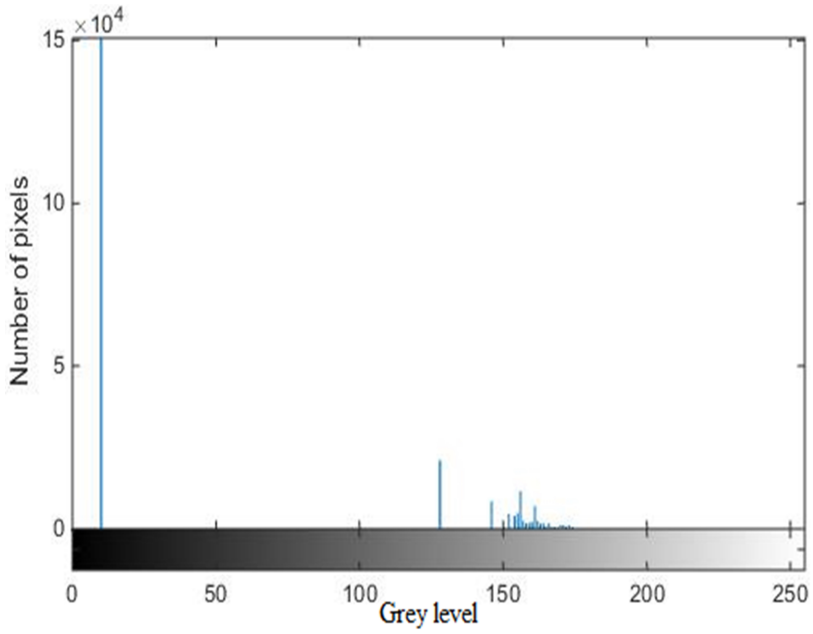
Then, the intensity values of $\mathbb{G}$ are updated using Equation 4.8. To numerically localise the resultant Gaussian shape and measure the deviation, the Gaussian model represented by Equation 4.9 is fitted with data interpolation using the Levenberg-Marquardt nonlinear squares fitting algorithm [193]. Therefore, the fitted curve is shown by the red line in Figure 4.7.

Following this proposed pre-processing, a morphological (opening) operation [15, 188, 194, 195] is then conducted to remove small fragmented cell fragments and separate connected cells as in [189] in the processed cell images prior to segmentation. Figure 4.8a shows the example resultant cell image after intensity histogram specification by the adaptively computed bi-Gaussian mixture model (as shown in Figure 4.4), and Figure 4.8b displays the intensity histogram of the example resultant cell image (shown in Figure 4.8a).

Figure 4.8b shows that two summits are exhibited in the processed cell image intensity histogram, with a valley separating the two summits (as described in [190])



(a) Resultant pre-processed cell image corresponding to Figure 4.2a;



(b) Intensity histogram of (a)

Figure 4.8: Resultant cell video frame example and its intensity histogram after processing by the proposed method.

whilst still maintaining the original intensity histogram shape.

Discovering the two Gaussian component features within the cell image intensity PDF addresses the model specification uncertainty problem to enable automation of previously semi-automatic pre-processing approaches that require human intervention to manually adjust the parameters of the second Gaussian component [52]. Thus, in the proposed work the parameters for both Gaussian components in the bi-Gaussian mixture model can be adaptively determined on a frame-by-frame basis in a fully automatic approach that does not require any human intervention. The proposed approach is also less computationally demanding compared with widely used mixture model parameter estimation approaches such as the Expectation and Maximisation (EM) algorithm [196–198] and Minimum Message Length (MML) algorithms [199–201], which iteratively search for parameters to minimise a log-likelihood function.

The parameters of the two Gaussian components are automatically derived from the first and second prominent features of the cell image intensity histogram. The only empirical parameter in identifying the bi-Gaussian distribution in the proposed work is the relative weight ratio between the two Gaussian components. The empirical selection of the relative weight ratio, as presented in Section 4.2.5, is the limitation of the proposed work.

4.2.5 Relative weight ratio between the two Gaussian components

The relative weight ratio $\lambda = C_2/C_1$ of the two Gaussian components is constrained to the range [0, 1) to maintain the primary dominant grey level intensity histogram feature. Thus, the summit of the first bi-Gaussian component is always greater than the summit of the second Gaussian component. However, the second Gaussian shape can be emphasised by increasing the value of the relative weight ratio, $\lambda = C_2/C_1$,

and then increase cell micrograph intensity contrast. The effect of the relative weight ratio between the Gaussian components is explored in Section 4.3.4.

4.3 Results and Discussion

A well-known iterative cell tracking method based on bandpass filtering and Viterbi search [191] is applied to evaluate the proposed cell segmentation pre-processing approach. The segmentation performance is compared to the original Viterbi search based system [191], the proposed approach and the commonly utilised HE applied to [191], and the Chan-Vese segmentation method [30]. Using the unsupervised segmentation accuracy evaluation method [202] based on Intra-Region Uniformity (IRU) theory [203] applied across a cell image database, a higher IRU indicates more accurate segmentation.

In Section 4.3.2, example segmentation results from cell images representative of the two image sequences of the N2DH-GOWT1 stem cell database are presented for a visual comparison. In particular, cells detected by the proposed approach (partial or complete) that manual ground-truth and Viterbi-based segmentation (with and without pre-processing) identify poorly or fail to identify are highlighted by a red rectangle and numerically labelled in Figures 4.9 - 4.16, corresponding to analyses in Section 4.3.2. Section 4.3.3 then presents the overall segmentation performance accuracy obtained using the full N2DH-GOWT1 stem cell image database [106].

4.3.1 Cell database

The proposed cell image contrast enhancement approach was applied to the group N2DH-GOWT1 of the training dataset [106], which is described in Section 3.2 in Chapter 3. The group consists of two cell video sequences of 182 images, which are

2D images of GOWT1 mouse embryonic stem cells, taken by a Leica TCS SP5 laser scanning confocal microscope.

4.3.2 Visual experimental results

Segmentation results from the first cell image sequence show that the proposed approach is able to segment cells that fail to be detected by existing approaches [191]. Example segmentation results are described below to illustrate this improvement.

Visual result example 1: 5^{th} frame of the first cell image sequence

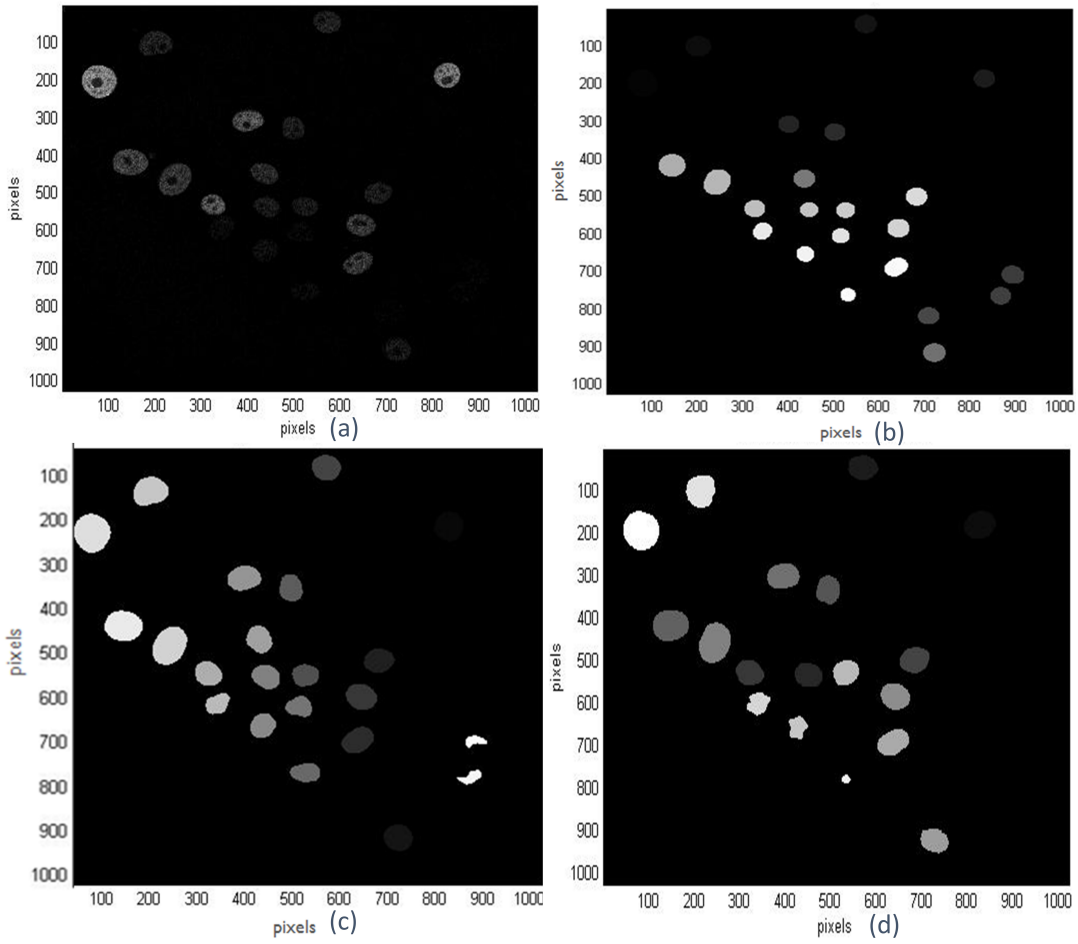


Figure 4.9: Visual example 1 of a segmentation result.

Example segmentation results from the 5^{th} frame of the first cell image sequence: (a) Original frame example; (b) Manual segmentation; (c) Segmentation result using [191]; (d) HE applied to [191].

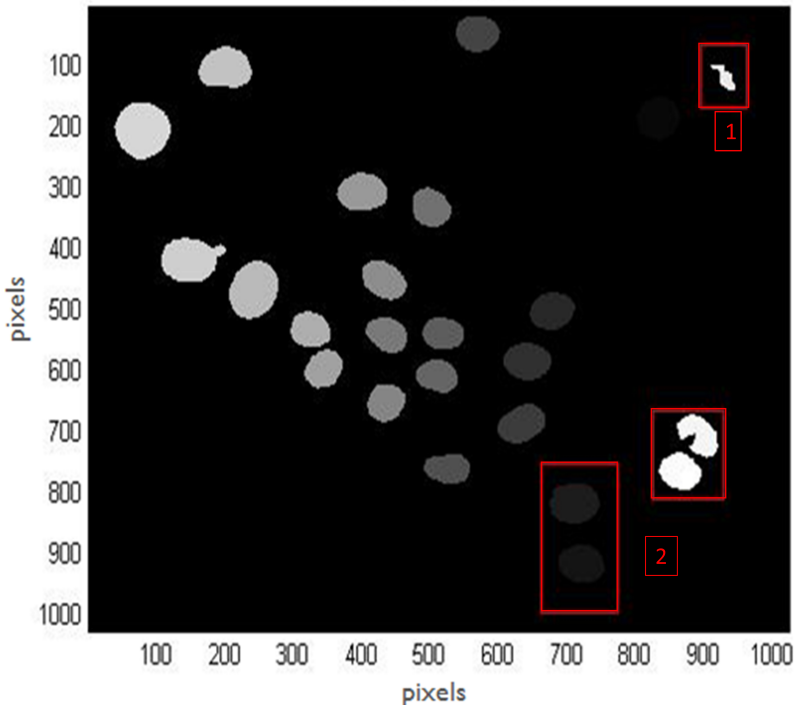


Figure 4.10: Segmentation result of the 5^{th} frame from the first cell video sequence using the proposed approach.

Figures 4.9 and 4.10 show the cell segmentation results obtained from the 5th frame of the first cell image sequence. The ground truth as shown in Figure 4.9b and the generated segmentation results as shown in Figure 4.9c and Figure 4.9d are presented using multi-grey level scaling as visual examples. The pixel values in the original ground truth and the generated segmentation results are integers within [0, N], where N is the number of labeled cell objects. The cells, for example, as shown in the top left of Figure 4.9b have pixel values that are close to 0, which denotes the background, so these objects were very dark visually but not missing. Clearer images can be seen by a large or high resolution display screen and original segmentation ground truth can also be seen in the dataset [106].

In this example frame, two cell regions were poorly detected by the Viterbi-based segmentation [191] with and without HE (Figures 4.9c and 4.9d, respectively), but improved with the proposed image pre-processing (Figure 4.10): (1) partial cell in the top right-hand corner was very clearly (partially) detected by the proposed approach (faintly detected by the manually segmentation, which is also called segmentation

ground-truth shown in Figure 4.9b, and [191] shown in Figure 4.9c); (2) two pairs of complete cells in the bottom right-hand corner were detected by the ground-truth and proposed approach but only partially detected by [191] (Figures 4.9c and 4.9d).

Visual result example 2: 14^{th} **frame of the first cell image sequence**

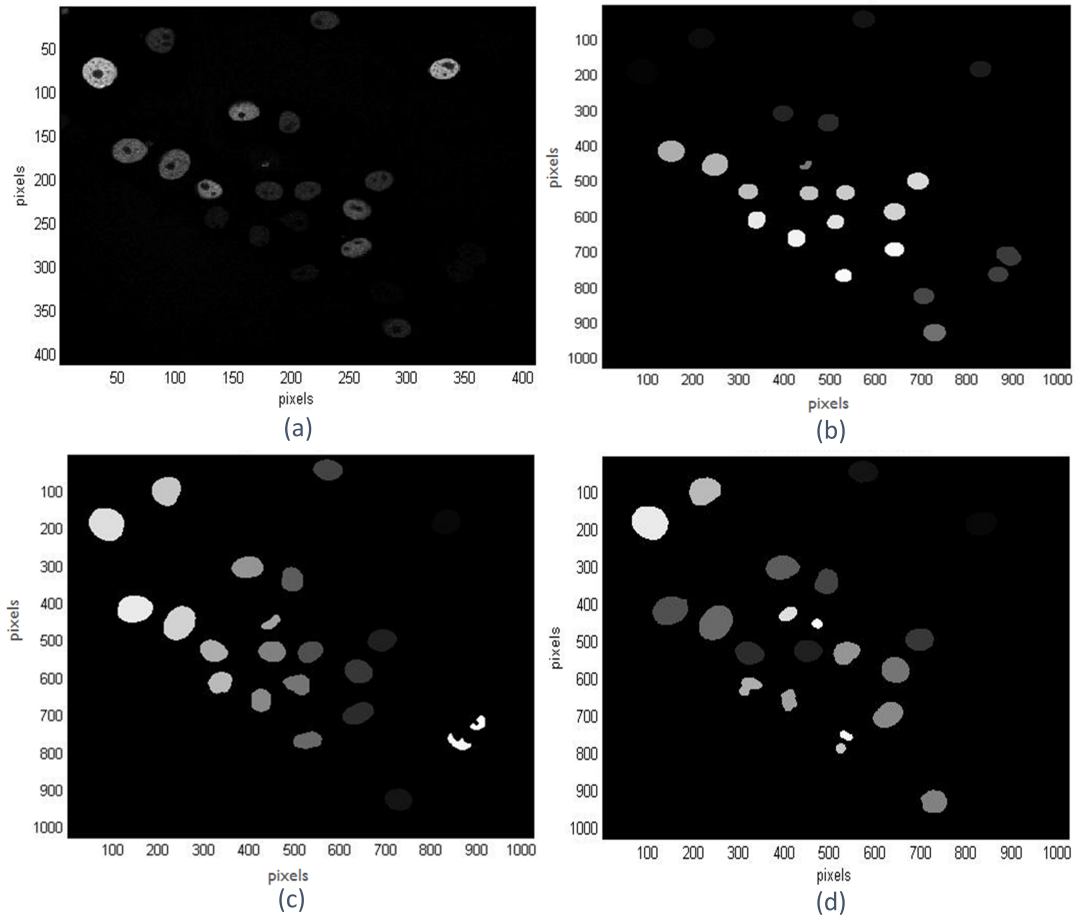


Figure 4.11: Visual example 2 of a segmentation result.

(a) Original frame example; (b) Manual segmentation; (c) Segmentation result using [191]; (d) HE applied to [191].

As another visual inspection example, segmentation results of the 14th frame of the first cell time-lapse video sequence are shown in Figures 4.11 and 4.12, illustrating the comparison of results between the proposed approach with existing approaches.

In the example cell image of Figure 4.11a, two cell regions (labeled as '1' in Figure 4.12b) were segmented by the Viterbi search approach [191] either with or without

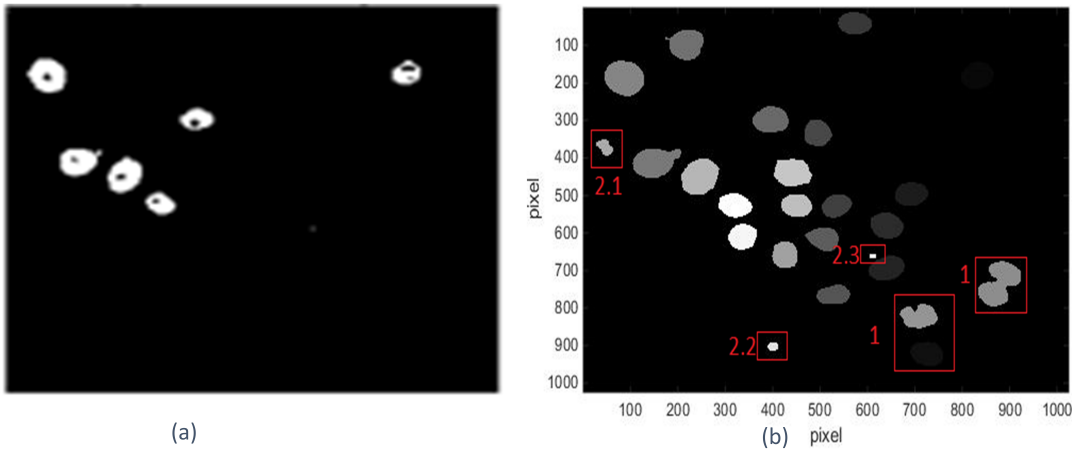


Figure 4.12: Segmentation result examples of the 14^{th} frame from the first cell video sequence.

(a) Chan-Vese Segmentation result [30]; (b) Segmentation result using proposed approach.

histogram equalization (Figure 4.11c and 4.11d, respectively), while all regions were detected in the ground-truth and by the proposed approach (4.12b). Cell regions labelled as '2.1-2.3' in Figure 4.12b are detected by the proposed approach but missed in the manually generated segmentation ground-truth due to the very low intensity contrast in the original cell image frame. Two pairs of complete cells in the bottom right-hand corner were detected by the ground-truth and proposed approach, but only partially detected by [191] (Figures 4.11c and 4.11d). The Chan-Vese segmentation approach (as shown in Figure 4.12a) can detect cells with clear intensity contrast against the background, but fails to detect cells with low intensity contrast. In comparison, the number of missed segmented objects using [30] depends on the intensity contrast of the microscopic video frames.

Visual result example 3: 14^{th} frame of the second cell image sequence

The segmentation results from the second cell image sequence further illustrated the ability of the proposed approach to detect cells that fail to be detected by the existing approach [191] and even the manually segmented result; examples are shown in Figures 4.13 - 4.16.

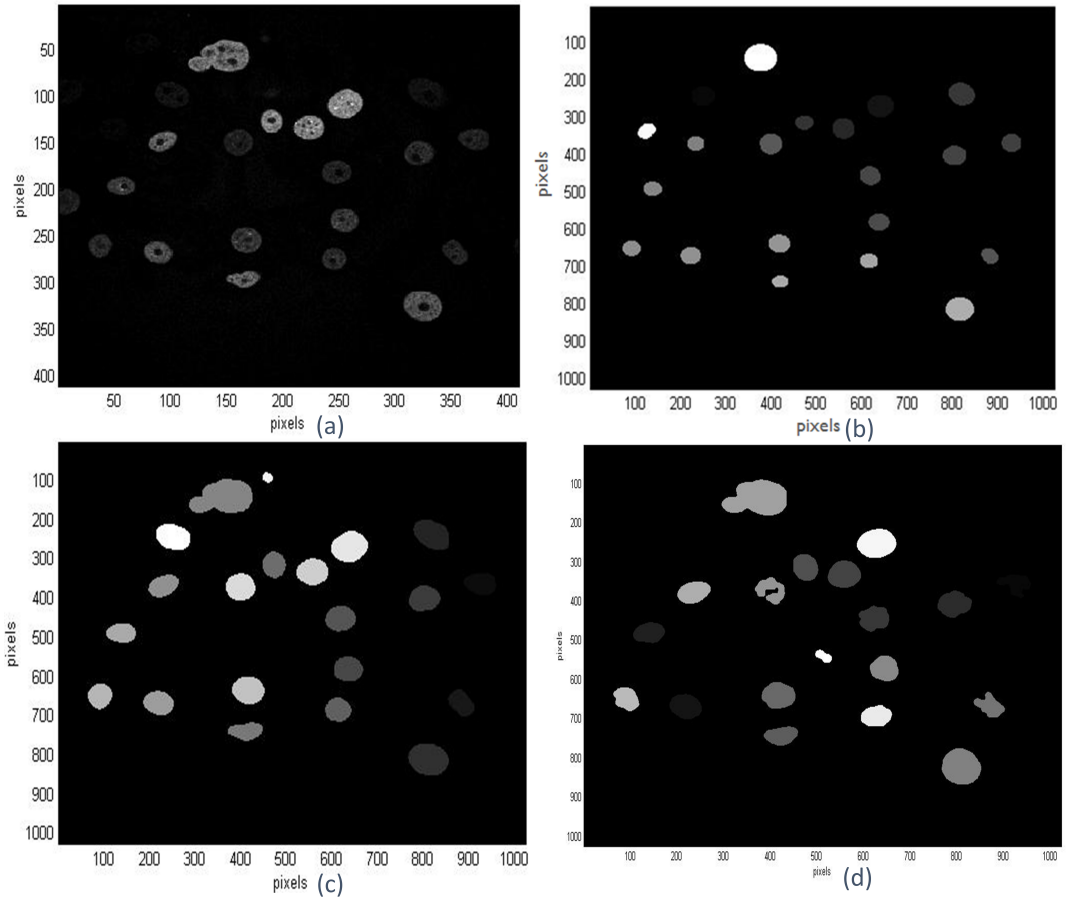


Figure 4.13: Visual example 3 of a segmentation result.

Example segmentation results from the 14^{th} frame of the second cell image sequence: (a) Original frame example; (b) Manual segmentation; (c) Segmentation result using [191]; (d) HE applied to [191].

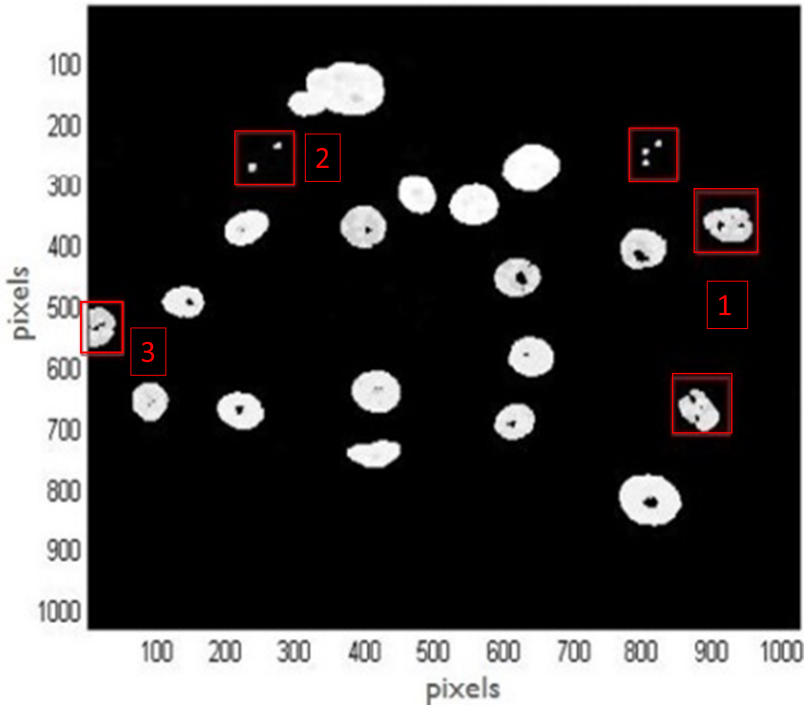


Figure 4.14: Segmentation result of the 14^{th} frame from the second cell video sequence using the proposed approach.

Figures 4.13 and 4.14 show the cell segmentation results obtained from the 14th frame of the second cell image sequence. Three cell regions poorly detected by [191] (Figures 4.13c and 4.13d) were improved upon with the proposed pre-processing (Figure 4.14): (1) three cells on the far right-hand side were detected by the ground-truth and proposed approach, but only faintly detected by [191]. Even where the proposed approach partially detected a cell in the top right, this partial segmentation indicates a cell presence identified by the proposed approach; (2) towards the image centre, a cell is partially segmented by the proposed approach and [191], but missed by the ground-truth and HE applied to [191]; (3) the proposed approach segmented the left-most cell cut-off in the frame, but was missed by both the ground-truth and [191] (Figures 4.13b-d).

Visual result example 4: 84^{th} **frame of the second cell image sequence**

Figures 4.15 and 4.16 show the cell segmentation obtained from the 84th frame of the second cell image sequence.

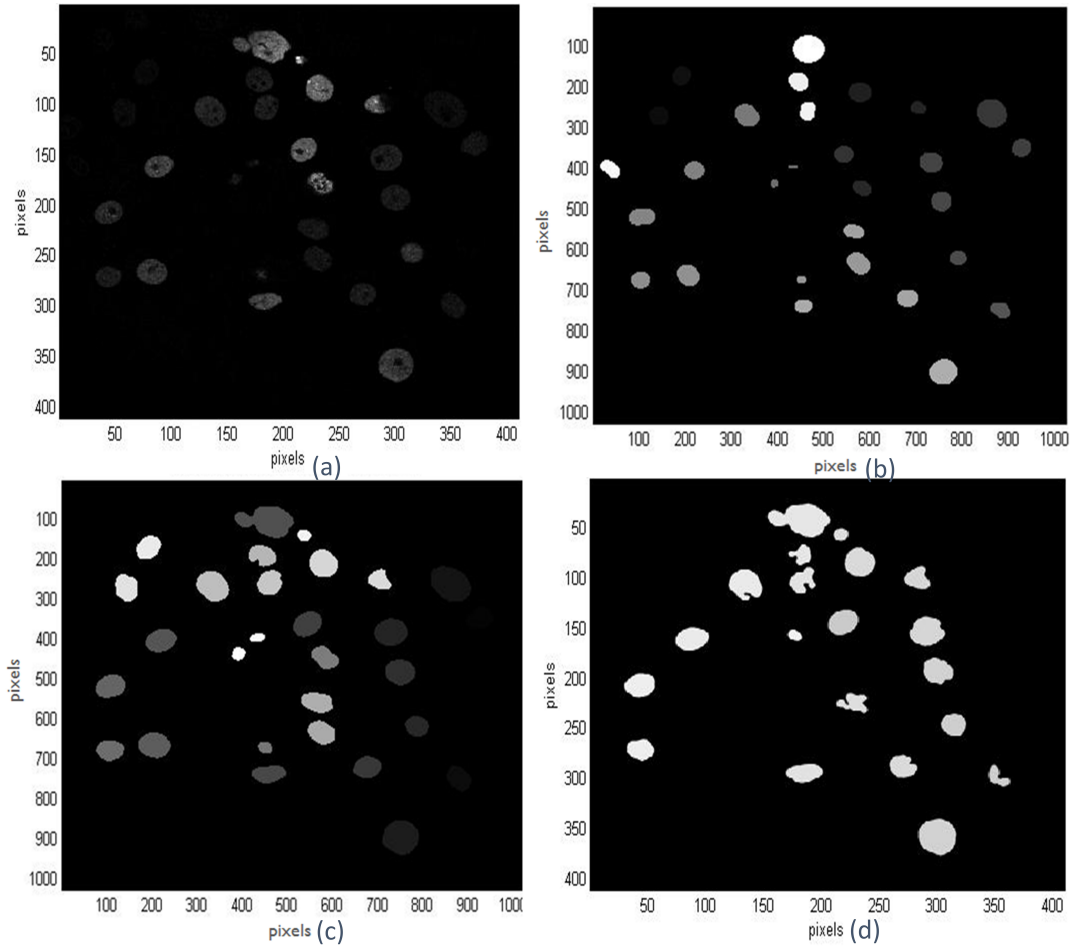


Figure 4.15: Visual example 4 of a segmentation result.

Example segmentation results from the 84^{th} frame of the second cell image sequence:(a)Original frame example; (b)Manual segmentation; (c)Segmentation result using [191]; (d) HE applied to [191].

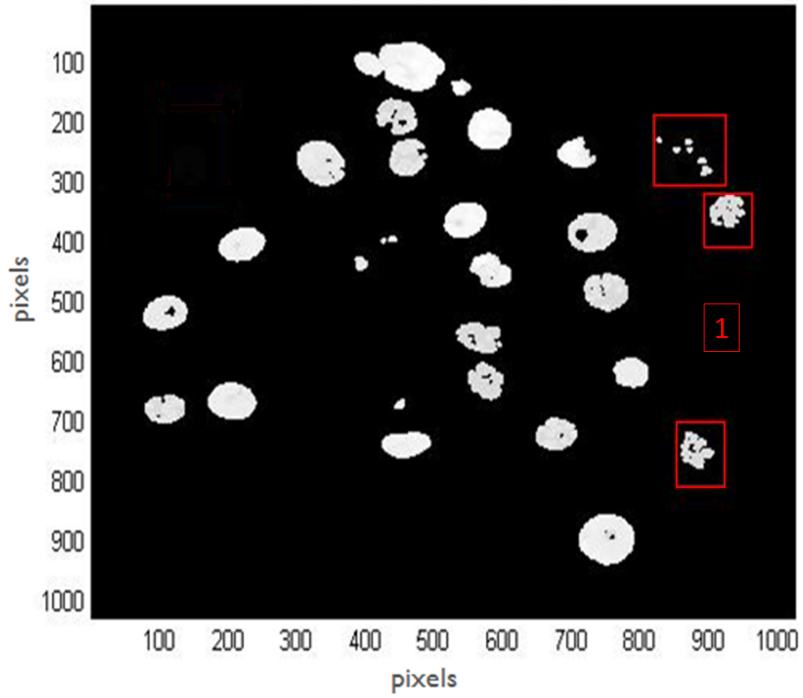


Figure 4.16: Segmentation result of the 84^{th} frame from the second cell video sequence using the proposed approach.

In the same right-hand side region as Figure 4.14, the three right-most cells are detected by the ground-truth (Figure 4.15b) and proposed approach (labeled as (1) in Figure 4.16), but only faintly detected by [191], though improved by HE pre-processing (Figures 4.15c and 4.15d).

4.3.3 Overall cell segmentation accuracy

Table 4.1 summarises the quantitative segmentation accuracy across the two image sequences in the database tested. Since the segmentation results by Chan-Vese segmentation algorithm [30] show the inability to detect cells with low intensity values in microscopic images, as illustrated by Figure 4.12a, the Chan-Vese algorithm is not included in the quantitative overall segmentation evaluation. As shown in Table 4.1, the proposed approach exhibits the highest IRU for both cell image sequences of the N2DH-GOWT1 database, improving the segmentation accuracy by up to 37% and 33%, respectively, compared to the existing Viterbi-based approach of [191] (with and without HE). The proposed method exhibits the recorded highest IRU values

and improves the segmentation accuracy of the work in [52] by increased mean IRU values of 3.3% and 6.7% over the two sequences, respectively.

Table 4.1: Mean cell segmentation accuracy across the cell image database based on IRU

	Seq. 1 (91 frames)	Seq. 2 (91 frames)
Viterbi method [191]	0.46	0.12
HE applied to Viterbi method [191]	0.58	0.13
Previous work [52] applied to [191]	0.61	0.15
Proposed method applied to [191]	**0.63**	**0.16**

4.3.4 Effect of relative weight on segmentation accuracy

The effect of the relative weight ratio λ is explored by a series of experiments to vary the λ value.

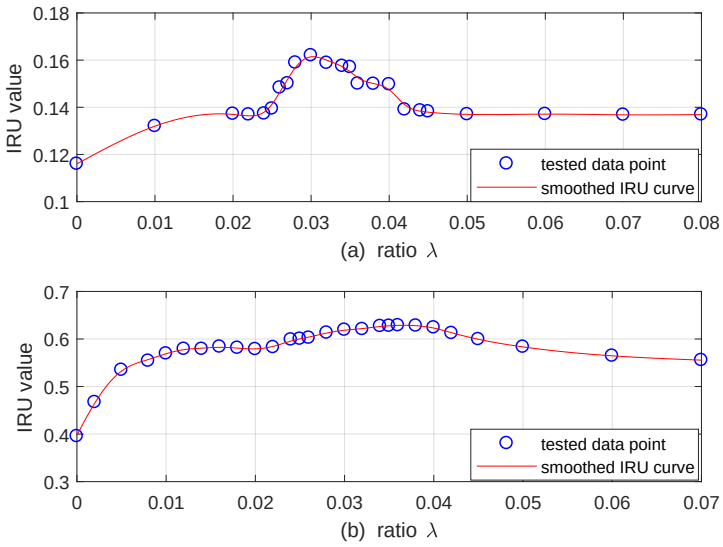


Figure 4.17: Relative modal amplitude effect on segmentation accuracy.

(a) IRU vs relative amplitude in the first cell image sequence; (b) IRU vs relative amplitude in the second cell image sequence.

Figure 4.17 shows the effect of varying the relative weight λ in the bi-Gaussian model on the overall segmentation accuracy across the studied cell image database. A similar consistency across the two cell image sequences is exhibited in Figure 4.17: the segmentation accuracy is maximal with a relative weight λ within the range $[0.03, 0.04]$. That is, the relative amplitude of the bi-Gaussian model can be fixed within $[0.03, 0.04]$ for the cell image database studied, which indicates that the first Gaussian component weight C_1 cannot overly suppress the second Gaussian component weight C_2 to maintain the original intensity histogram featuring saturation between low grey level range. In addition, the resultant small λ value range indicates that the relative distribution amplitude of cell regions is less than 5% of the background distribution amplitude. This demonstrates the necessity to maintain the shape characteristics of the original intensity histogram: consistently increasing the relative weight does not result in corresponding increases in segmentation accuracy, which can be attributed to why the histogram equalisation technique does not work well for the dataset studied.

4.4 Summary

The accurate detection and segmentation of cells from time-lapse microscopic video sequences provides a critical foundation for understanding dynamic cell behaviours and cell characteristics when using automatic cell tracking systems. However, cell tracking approaches typically assume accurate segmentation results.

To address the segmentation challenge of low image intensity contrast characteristic to cell images, this Chapter introduced a novel image pre-processing technique to enhance cell microscopic image intensity contrast prior to segmentation (and subsequent tracking). The novelty of the image pre-processing method is the proposed adaptive, shifted bi-Gaussian mixture model that is used to match the original cell image intensity histogram for greater differentiation between the cell foreground and

microscopic image background, while maintaining the original intensity histogram shape. The proposed pre-processing matched the intensity histogram of the original cell microscopic image to a shifted bi-Gaussian distribution. Image contrast is enhanced with two summits in the new intensity histogram for target cell segmentation, whilst the shape of the original histogram is maintained. The model parameters of the shifted bi-Gaussian mixture are automatically derived frame-by-frame from cell video sequences to adapt to background changes in the video sequence. Evaluation experiments on a cell image database showed more accurate cell image segmentation by up to 37% and 33% for the two image sequences tested, compared to Viterbi-based segmentation. The proposed approach obtained as well as or better than the (manual) ground-truth and Viterbi search approaches, with and without existing contrast enhancement techniques applied.

However, there is a limitation in the proposed pre-processing method in the empirical selection of the relative weight ratio between the two Gaussian components. This parameter was selected by trial-and-error and comparing a range of possible values. The relative weight ratio between the two Gaussian components may vary when applying to other microscopic cell video datasets. Thus, extra testing on a range of values is recommended.

Chapter 5

Single and Multiple Zebrafish Larvae Tracking

5.1 Introduction

The accurate tracking of zebrafish larvae movement is fundamental to research in many biomedical, pharmaceutical, and behavioural science applications [3, 7, 66, 73]. However, the microscopic video conditions of zebrafish larvae are more complex than those of cell microscopic videos, such as the existence of impurities, small particles, turned up water bubbles, surface projections etc. Thus, the pre-processing technique presented in Chapter 4 using two Gaussian components cannot fully represent the zebrafish larvae microscopic video conditions. Further, the locomotive characteristics of zebrafish larvae are significantly different from adult zebrafish, where existing adult zebrafish tracking systems cannot reliably track zebrafish larvae [149]. The detection of zebrafish larvae is often lost during tracking systems [5, 31, 34] based on movement features (i.e., pixel displacement thresholding method [35]) when the larvae stop moving. This erroneous detection of zebrafish larvae is not uncommon due to their 'bursty' movement characteristics, as explained in Section 2.3.1. Further, the far smaller size differentiation between larvae and the container render the detection of water impurities inevitable, which further affects the tracking

of multiple zebrafish larvae or require very strict video imaging conditions that typically result in unreliable tracking results for realistic experimental conditions. As a result, zebrafish larvae misdetection and overlapped movement trajectories especially restrict current automatic multiple zebrafish larvae tracking techniques.

In addition, constraints on the input imaging conditions are commonly required by existing automatic zebrafish tracking systems [5, 31, 33–35], which are largely due to poor zebrafish detection and segmentation results from the input videos. Thus, improving the segmentation method will remove the need for input imaging constraints, where it has been assumed that improving the segmentation accuracy can result in more reliable tracking performance [139]. However, this assumption of improving segmentation accuracy to enhance tracking performance has not yet been examined.

Thus, instead of using the developed pre-processing technique presented in Chapter 4, this Chapter investigates advanced background subtraction and segmentation techniques, including dense SIFT flow, SURF features, and Gaussian mixture models, to improve the segmentation accuracy. Dense SIFT flow has been shown to be robust for image registration with large appearance changes. Any arbitrary natural images can be assumed to consist of individual near-homogenous areas, which inherently have a Gaussian-shaped histogram [187]. The mean values and the variances of the Gaussian components indicate the average intensity values and the texture details of the individual areas, respectively. The Dirichlet prior and the Minimum Message Length criteria automatically select the number of Gaussian components and their parameters to optimally enhance the image contrast. This Chapter also investigates removing the input microscopic video imaging conditions when developing the automatic zebrafish larvae tracking system in Section 5.3.1. Section 5.2.2 also presents the investigation on other feature detectors (SURF, HOG, BRISK and SIFT). For comparison, the single (LSRtrack tracker and original SIFT flow tracker) and multiple (SimpleTracker, IdTracker and LoliTracker) organism trackers are presented in Section 5.2 and Section 5.3, respectively.

5.1.1 Single zebrafish larva tracking

This Chapter first investigates single zebrafish larva detection and tracking in Section 5.2, using the dense Scale-Invariant Feature Transform (SIFT) flow method [154], which has been shown to be robust for image registration with large appearance changes. The dense SIFT flow is proposed to detect sudden tail beating movements of zebrafish larva [147], where the approaches and experiments proposed are the first application of dense SIFT flow directly on microscopic videos to study the dynamic behaviour of zebrafish larvae.

Further, the proposed approach considers the specific movement characteristic of zebrafish larvae, with sudden swimming locomotion interspersed with substantially stationary periods of little movement [204]. Specifically, to address the lost detection of zebrafish larva when there is no movement, the Speeded Up Robust Features (SURF) approach [205] is applied to detect non-moving zebrafish larva. This combination of dense SIFT and SURF features is therefore proposed and used for the accurate and reliable detection and segmentation of zebrafish larvae.

In experimental validations of the proposed single zebrafish larva tracking system, the determined positions of larval zebrafish are evaluated by the Average Centroid Location Error (ACLE), and the obtained results in the proposed approach are compared to that of the widely used single zebrafish larva tracking system, LSRtrack system [5].

5.1.2 Multiple zebrafish larvae tracking

Section 5.3 extends upon the single zebrafish larva tracking system to track multiple zebrafish larvae, consisting of segmentation and association modules. To achieve this purpose, this work proposes the adaptation of Gaussian Mixture Model (GMM) based video background subtraction and object segmentation techniques to detect and

segment zebrafish larvae from the microscopic videos under realistic experimental conditions, without imaging constraints.

Segmentation errors caused by zebrafish misdetection and overlapping (which further result in overlapped trajectories) are additional common issues for zebrafish larvae tracking that currently restrict its application. Thus, Section 5.3.3 investigates multiple object association method, robust against non-ideal zebrafish detection and segmentation, based on the Kuhn-Munkres association algorithm [206]. In the proposed multiple object association approach, the positions of mis-detected and overlapped zebrafish larvae are calculated using inter-frame knowledge.

The proposed automatic multiple zebrafish larvae tracking system is evaluated on a set of single and multiple adult and larvae zebrafish videos across a wide variety of (complex) video conditions, including shadowing, labels, water bubbles and background artefacts. To facilitate this evaluation on multiple zebrafish segmentation and tracking research, the datasets with annotated ground-truth, presented in Sections 3.5.1 and 3.6.2 in Chapter 3, are used in this Chapter. The datasets and software are made publicly accessible[1].

5.2 Single Zebrafish Larva Tracking

Due to the reliable detection of moving zebrafish larvae using the Scale-Invariant Feature Transform (SIFT) flow algorithm [154], this work proposes the application of SIFT flow to extract the zebrafish larva from the image background in each microscopic video frame, followed by position detection indicated by the centroid location of the detected object (same as for the LSRtrack approach [5]). Figure 5.1 shows the workflow of the proposed single zebrafish larva tracking system.

SIFT flow has not previously been applied to zebrafish larvae tracking due to computational complexity requirements. With an entire moving object extracted, the

[1] https://github.com/Xiao-ying/moving-zebrafish-larvae-segmentation-and-tracking-dataset-

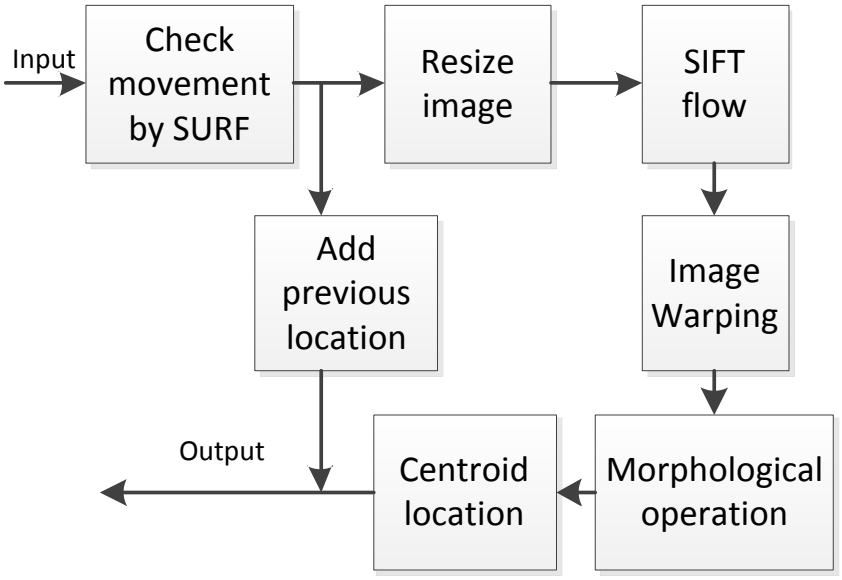


Figure 5.1: Flow chart of the proposed method.

subsequent object tracking process is, therefore, simpler and more reliable. However, SIFT flow fails once no movement exists in the successive frames; thus, SURF feature [205] detection is proposed to apply to detect the zebrafish without movement between two connective frames prior to the SIFT flow calculation. If no movement exists, the position of the zebrafish in the secondary frame is equal to the previous location and the SIFT flow in the current frame does not need to be calculated, thus the computational time is furthermore decreased.

5.2.1 SURF features

Speeded Up Robust Features (SURF) is a widely used state-of-the-art scale and rotation invariant sparse image feature representation. The SURF descriptor is a 64 dimensional blob detector, where blob response maps are calculated to convolve an image with the Hessian matrix [205]. The SURF algorithm utilises a fast interest point detection and description scheme, which outperforms existing state-of-the-art

feature descriptions, such as the Difference of Gaussian (DoG) detector [207], the Harris corner descriptor [208] and Hessian-Laplace descriptor [209], both in speed and accuracy [205].

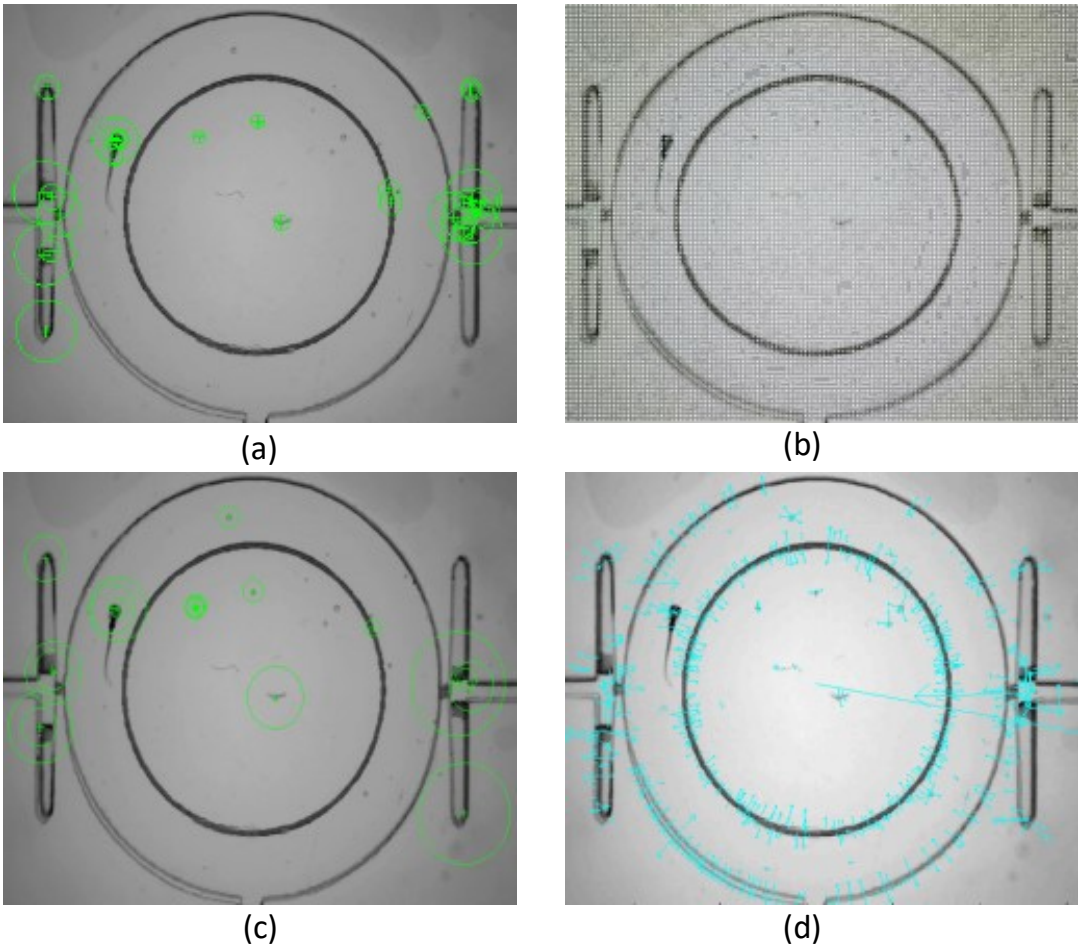


Figure 5.2: Sparse feature points from various feature descriptors:
(a) SURF, (b) HOG, (c) BRISK (d) SIFT.

The strongest SURF feature points, which are generally located at the head of the zebrafish, are calculated for all frames across the studied single zebrafish video sequences. Then, the distances of the SURF points are computed between successive frames to determine the object displacement.

5.2.2 Dense SIFT flow

In the proposed approach, the differentiation of the zebrafish larvae from the image background is excluded through the dense SIFT flow algorithm [154]. Figure 5.3 illustrates each step in the proposed system (as shown in Figure 5.1) as applied to an example frame pair from a zebrafish larva input video. To reduce the computational complexity required for calculating the dense SIFT image and SIFT flow (without loss of tracking accuracy), the input video frame is downsampled to reduce the number of SIFT features calculated, and various downsampling scale factors are investigated as part of the experimental results in Section 5.2.6. The application of frame downsampling was also used in the original SIFT feature literature to reduce the number of selected SIFT feature points [210].

The dense SIFT flow algorithm [154] is designed to match similar images from different scenes. Unlike the extraction of sparse feature points employed in previous methods [128, 210], SIFT flow [154] uses dense SIFT, extracting a SIFT histogram at a scale for all pixels with overlapping patches [211]. Sparse image feature representations, such as SIFT [212, 213], SURF [205], Binary Robust Invariant Scale Key-points (BRISK) [214], and Histogram of Oriented Gradients (HOG) [215] consist of feature extraction and detection.

Here, this work only uses the feature extraction component of dense SIFT [154]. Dense SIFT vectors more accurately preserve the intrinsic uncertainty of the processed image compared to sparse feature approaches [205, 213–215], so that more robust decisions can be made in proceeding tracking stages [211]. Dense SIFT flow can, therefore, extract an entire moving object from the video background if there is movement displacement between two successive frames [154]. In contrast, the state-of-art feature point detection methods [205, 213–215] only detect feature points based on the difference between pixels and their neighbours such that feature points located in the background as well as on the object are selected (as shown in Figure

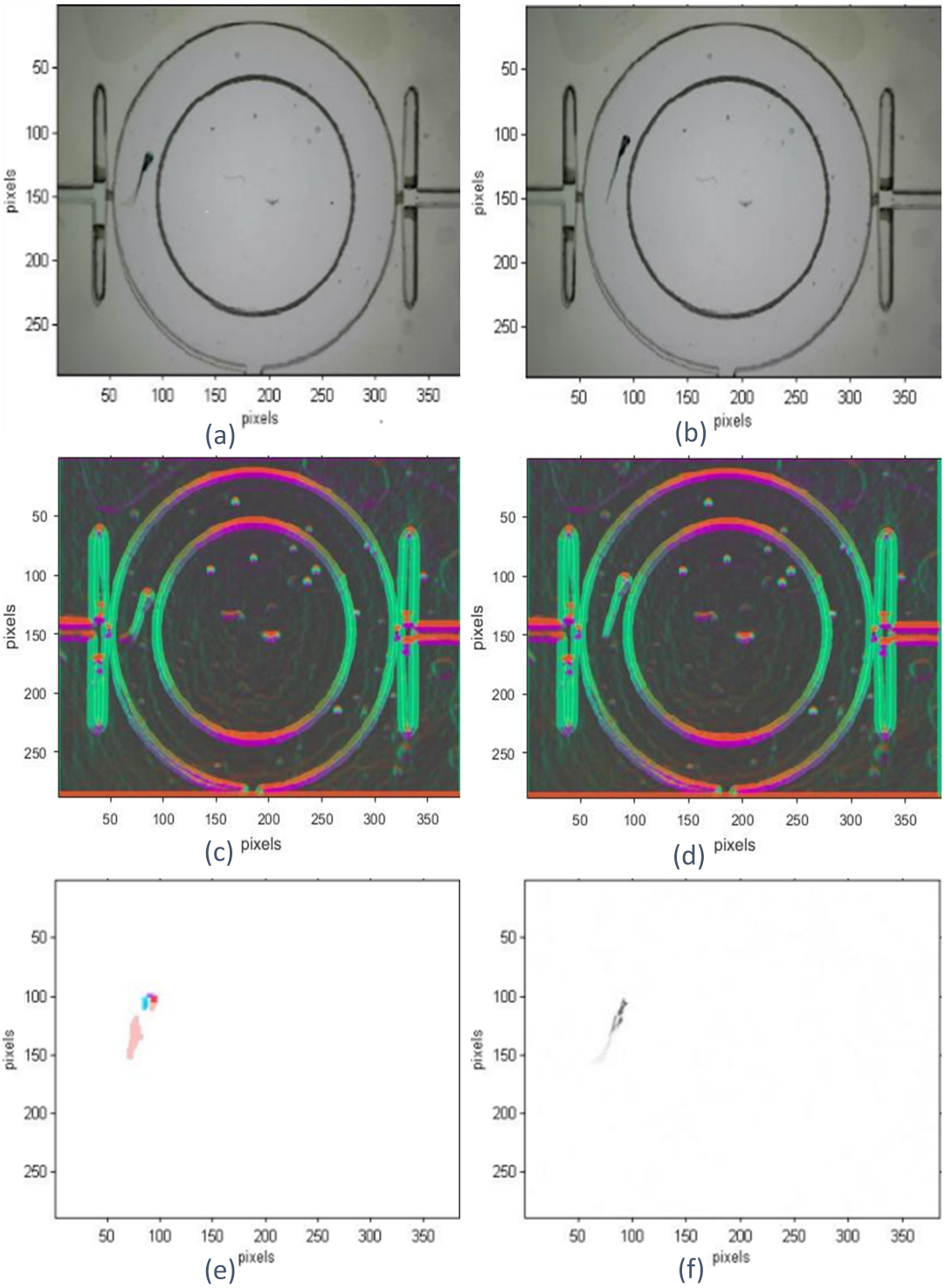


Figure 5.3: Visual demonstration of results of each step by the proposed single zebrafish larva tracking system.

(a) Original frame $I(t)$; (b) The following frame of (a), $I(t+1)$; (c) Visual illustration of the SIFT image of (a); (d) Visual illustration of the SIFT image of (b); (e) SIFT flow $I(t,t+1)_{SIFT-flow}$ from (c) to (d); (f) Warped image $I(t)_{Warp}$.

5.2 applied to zebrafish larvae).

In SIFT flow, the dense SIFT algorithm [154] creates a SIFT image, $I(t)_{SIFT}$, from each original frame $I(t)$ in zebrafish videos. Each pixel in the SIFT image, $I(t)_{SIFT}$, is replaced by the SIFT descriptor describing its neighbour region of size $PS \times PS$, which is the window size of the normalised Gaussian kernel required to produce the SIFT descriptor [154]. The generated SIFT images, $I(t)_{SIFT}$ and $I(t+1)_{SIFT}$(as shown by Figures 5.3c and 5.3d), are then used to estimate the moving region between the two successive frames. This moving region is determined by SIFT flow, $I(t,t+1)_{SIFT-flow}$ (as shown in Figure 5.3e) by estimating the dense correspondences between these two SIFT images in a 2D flow field.

5.2.3 Image warping

Finally, the detection of the moving object (i.e., the zebrafish larva) is conducted based on the flows in each pixel using the pixel warping function in Equation 5.1. The original frame $I(t)$ is seen as the query, applying a pixel warping function to align the SIFT flow, $I(i,i+1)_{SIFT-flow}$ to the query, and displaying the warped image, $I(t)_{Warp}$ (as shown by Figure 5.3f) with respect to the dense correspondence.

Image warping is the act of distorting a source image to a destination image based on the correspondence between source and destination spaces [216]. Image warping is primarily used for correcting image geometric distortions introduced by imperfect imaging systems [216], and primarily in experimental conditions though [217] provides a mathematical expression for optical flow application; segmentation based on optical flow was further compared and presented in Section 3.5, denoted as OptFlow, and Section 5.3. This work uses image warping to determine the zebrafish larva displacement between two connected video frames, where Equation 5.1 represents how SIFT flow is combined with image warping [217].

$$I(t)_{Warp} = I(t)(x + dx, y + dy) \tag{5.1}$$

where (dx, dy) is the SIFT flow vector estimated by the dense SIFT flow algorithm [154]. Every pixel of image $I(t)$ is warped separately according to horizontal and vertical displacements (dx, dy) [154].

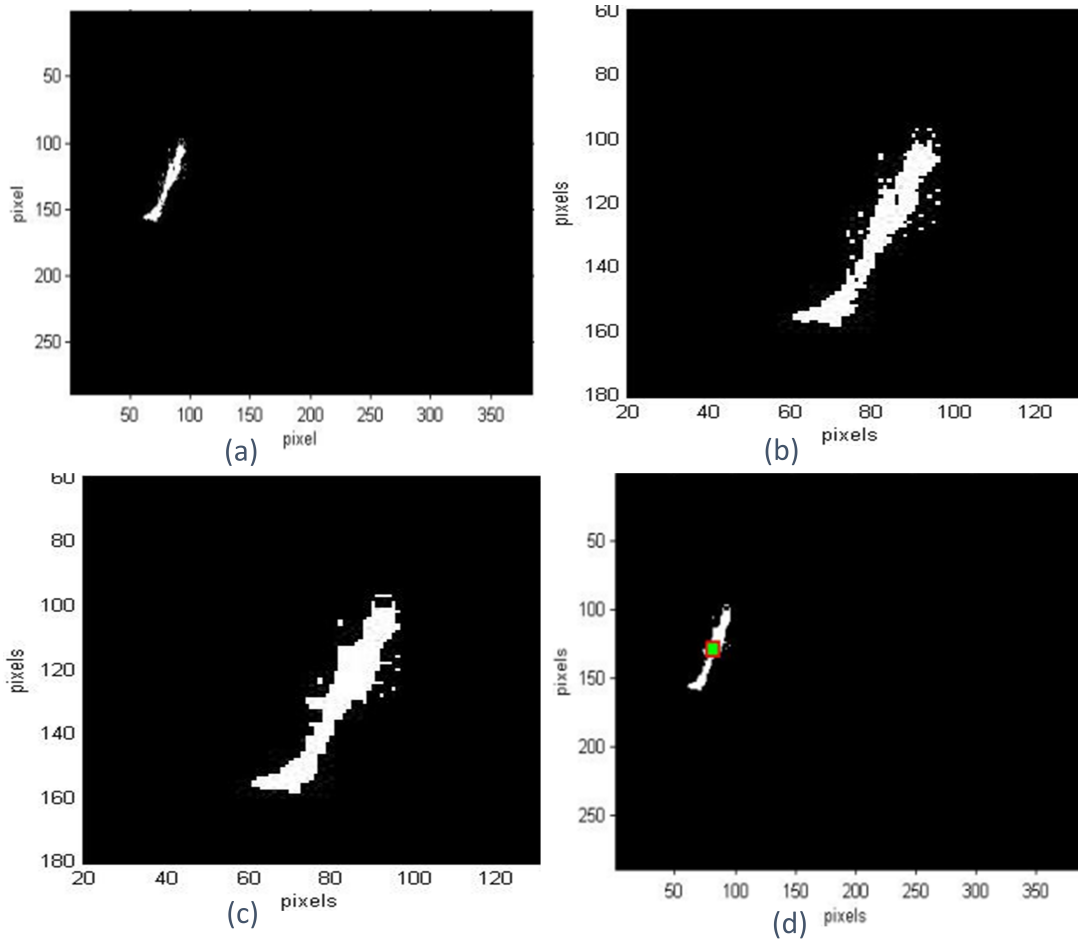


Figure 5.4: Visual demonstration of results of each step by the proposed single zebrafish larva tracking system.

(a) Image $I(i)_{Warp-bw}$; (b) Zoomed-in image of a zebrafish region after image binarisation on (a); (c) Close operation on (b); (d) Calculated zebrafish centroid position.

Since dense SIFT flow extracts the moving object from the background, the resultant grey level image $I(t)_{Warp}$ as illustrated in Figure 5.3f is primarily white space. Thus, binarisation (with the threshold selected according to the Otsu thresholding method

[179]) is applied to the warped image $I(t)_{Warp}$, obtaining a binary image as shown in Figure 5.4a.

To enhance the tracking accuracy of the zebrafish larva movement, the binarised image exhibits 'noise' (sparse, small image artefacts and holes, where an example detected object region is shown in Figure 5.4b), which results in disconnected zebrafish larva objects (as illustrated in Figure 5.4b). Thus, a morphological 'close' operation is applied to the binary image before calculating the zebrafish larva centroid position, in order to join narrow breaks, and fill in long thin gulfs and holes [218].

5.2.4 Estimation of zebrafish object centroid

Comparing the centroid of a moving object for two successive frames is a computationally simple and efficient method for estimating the distance that the object has moved. Generally, an image can be seen as a matrix I of intensities of both an object and a background. Equations 5.2 and 5.3 represent the calculation of the centroid for both axes [219]:

$$C_x = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i \cdot I_{ij}) \Bigg/ \sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij} \tag{5.2}$$

$$C_y = \sum_{i=1}^{n}\sum_{j=1}^{m}(y_i \cdot I_{ij}) \Bigg/ \sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij} \tag{5.3}$$

where x_i and y_i are the horizontal and vertical coordinates of a pixel, respectively, I_{ij} is the correspondence intensity value of this pixel, and n and m are the dimensions of image I. The distance that an object moves is therefore measured by the Euclidean distance between the calculated centroid point (C_x, C_y) of a frame with that of the subsequent frame.

The correct computation of the centroid using Equations 5.2 and 5.3 assumes that the object intensity differs compared to the background. Vital to the centroid calculation is the exclusion of as much of the image background as possible, lest it strongly bias the centroid calculation to the centre of the image [219].

Though the existing LSRtrack system also uses the zebrafish centroid for tracking, the background in the LSRtrack system [5] is removed by setting an intensity threshold. Thus, pixel values must exceed this threshold to be classified as an object region. However, this background subtraction method of LSRtrack is susceptible to impurities inside water, glares on water surface [153] etc. due to the use of a fixed and empirically selected threshold value. Consequently, zebrafish larva can be mis-detected in LSRtrack due to incorrect threshold selection.

5.2.5 Improving the tracking accuracy

In the proposed approach, two main measures are taken to enhance the accuracy of tracking the zebrafish larva movement:

1. **Morphological Operation**. A morphological close operation is applied to the binary images $I(i)_{Warp-bw}$ before tracking to remove 'noise'. An example result of applying this morphological close operation to Figure 5.4b is shown in Figure 5.4c. Figure 5.4d shows the final calculated centroid position, labelled by the red square with green infill.

2. **Object position correction by SURF**. The dense SIFT flow algorithm detects the movement of zebrafish larva; however, the method fails to track the position of the object if there is no movement between successive frames. In this situation, the position of the zebrafish in the previous frame will be assigned as the updated position in the current frame. An example tracking result is illustrated in Figure 5.5, which shows that 91% of the frames can detect a zebrafish larva position using the SIFT flow algorithm (i.e. 9 out of the 100 example

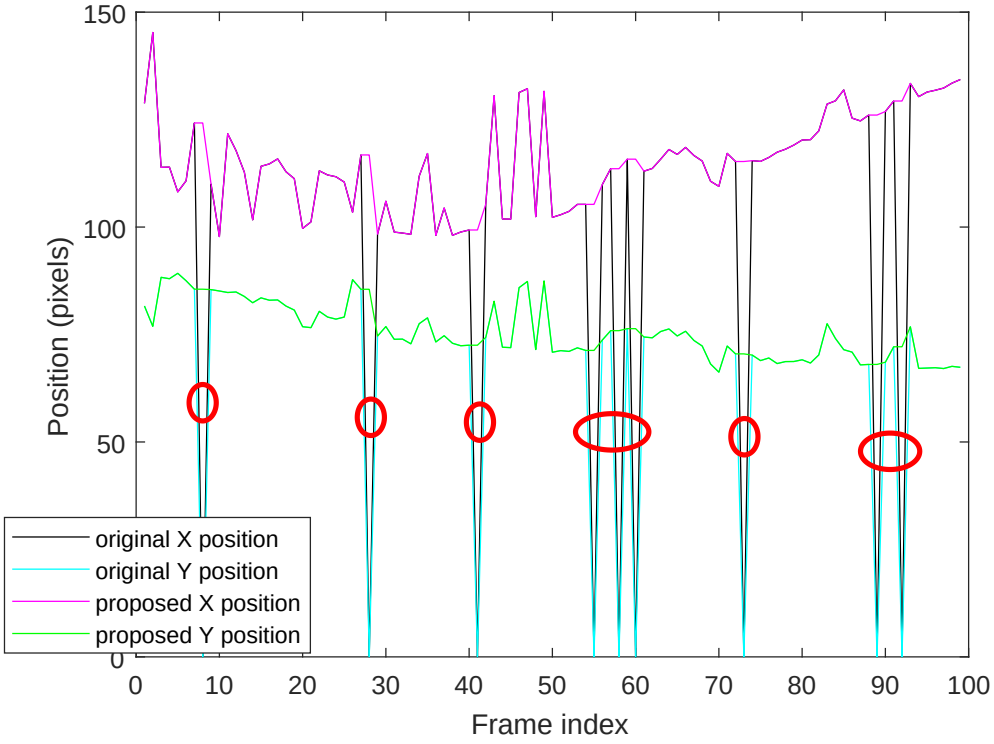


Figure 5.5: Zebrafish position correction when no movement is detected.

frames evaluated fail because of no displacement of the zebrafish between successive frames). Thus, the frames without displacement indicated by the zero values of X and Y coordinates (black and aqua lines in Figure 5.5) are assigned the zebrafish position of the previous frame (green and magenta lines in Figure 5.5). It can be seen from Figure 5.5 that the position detection failures (highlighted by the red circles) in the original system are corrected in the position detection results by the proposed method (green and magenta lines in Figure 5.5).

5.2.6 Results and discussion

The proposed single zebrafish larva tracking system based on SIFT flow and image warping is evaluated on a microscopic video of 100 video frames. The video of a zebrafish larva (housed in a flat round well, as shown in Figure 5.3a and 5.3b) was taken

at a frame rate of 15 fps by a Dino-Lite premier AD7013MT digital microscope, as presented in Section 3.3.

To illustrate the tracking accuracy of the proposed system, a comparison is performed with the LSRtrack method [5], original SIFT flow [154], and the proposed image processing approach using SIFT flow with frame downsampling and morphological operations applied to LSRtrack. To evaluate the tracking accuracy, the estimated zebrafish centroid obtained by each approach is compared to the manual ground-truth to estimate the Average Centroid Location Error (ACLE).

Tracking accuracy evaluation

Table 5.1: Tracking accuracy performance comparison.

	ACLE (pixels)	NDFP (pixels)
LSRtrack [5]	29.09	$3.51 * 10^3$
LSRtrack [5] (30% downsampling)	28.62	$1.38 * 10^2$
Original SIFT Flow [154]	12.50	$1.23 * 10^6$
Proposed (50% downsampling)	7.72	$3.07 * 10^5$
Proposed (30% downsampling)	**7.39**	$1.11 * 10^5$
Proposed (10% downsampling)	27.17	$1.23 * 10^4$

*ACLE = average centroid location error

*NDFP = Number of Detected Feature Points

As shown in Table 5.1, the LSRtrack system [5] exhibits poor tracking accuracy, with the highest ACLE value exhibited. Whilst LSRtrack is widely utilised in single zebrafish tracking due to the low computational complexity, it can be seen from Table 5.1 that the intensity thresholding-based LSRtrack system [5] is not robust for tracking zebrafish larva movement with or without image downsampling and morphological operation processing. In contrast, dense SIFT flow [154] with intensity thresholding exhibits a general 16.59 pixels smaller ACLE error than LSRtrack [5]. However, the proposed approach has the highest tracking accuracy, demonstrated by the decreased ACLE value by 21.7 pixels and by 5.11 pixels than LSRtrack [5] and dense SIFT flow with intensity thresholding [154], respectively.

Effect of varying downsampling rate on tracking accuracy

In the proposed method, to reduce the calculation of SIFT vectors in each video frame, the original zebrafish video frame is downsampled. To investigate the effect of the frame downsampling rate on tracking accuracy, the frame resolution is reduced at different scale factors: 50%, 30%, and 10% of the original video frame resolution. As shown in Table 5.1, the proposed method with frame downsampling scale factors at 30% and 50% improved the tracking accuracy by 40.88% and 38.24%, respectively, compared to the original SIFT flow [154], respectively. This improvement is due to decreasing the effect of background noise and water impurities on the moving zebrafish larvae detection. However, further downsampling the video frame results in decreased tracking accuracy for the 10% scale factor with a higher ACLE value of 19.78 pixels compared with that from the 30% downsampling scale factor. This increased error is due to the position displacement of the zebrafish in successive frames not being fully detected by the SIFT flow method due to too few feature points being detected.

Conversely, the downsampled resolution at a 50% scale factor does not further increase tracking accuracy for the proposed approach compared with a 30% scale factor, as the higher resolution images result in a larger-sized object (in terms of the number of pixels); thus, larger deviations can result when compared with the ground-truth centroid position.

The effect of frame downsampling on tracking accuracy is also tested for the LSRtrack system [5]. Table 5.1 shows that the tracking accuracy at 30% frame resolution reduction is comparable to the original LSRtrack (with 1.62% improvement in the ACLE tracking accuracy). Frame resolution downsampling therefore does not significantly influence LSRtrack tracking accuracy.

Thus, the results shown in Table 5.1 show that the proposed method applied with downsampling at a 30% scale factor returns the highest tracking accuracy (i.e., the

lowest ACLE value), improving the tracking accuracy by average 21.7 pixels compared to the original LSRtrack system [5].

5.3 Multiple Zebrafish Larvae Tracking

Extending upon the single zebrafish tracking system proposed in this work as presented in Section 5.2, this Section presents an accurate and reliable multiple zebrafish larvae tracking system. The proposed system consists of three stages: background subtraction, zebrafish larvae segmentation, and association or matching of the larvae between successive frames. Based on the novel adaptation of advanced computer vision techniques (as explained below) using the Kuhn-Munkres multiple object association algorithm and the proposed theoretical object position calculation using inter-frame knowledge, the proposed system is robust against non-ideal zebrafish larvae detection and segmentation.

In this Section, the proposed background subtraction and segmentation approach for multiple zebrafish larvae is firstly presented in Section 5.3.1, and evaluated in Section 5.3.2. Then, an improved association approach is proposed in Section 5.3.3 to enable the estimation of individual zebrafish larvae movement trajectories. Individual zebrafish movement parameters can therefore be estimated from these individual trajectories as presented in Section 5.3.4, and the overall proposed multiple zebrafish larvae tracking system is evaluated and motion parameters analysed in Section 5.3.5.

5.3.1 Zebrafish segmentation in unconstrained microscopic video conditions

Figure 5.6 outlines the proposed automatic multiple zebrafish larvae tracking system, which consists of three stages: background subtraction, zebrafish larvae segmentation, and association or matching of the larvae between successive frames. The

robustness of the proposed segmentation method is achieved by the efficient estimation and removal of the video frame background, further median filter processing and morphological greyscale erosion. The detection and segmentation of zebrafish larvae then are not affected by the input video imaging conditions. The performance of the proposed segmentation method is evaluated and compared with three state-of-the-art segmentation algorithms, as detailed in Section 5.3.2. To examine and verify the assumption of improving segmentation accuracy to enhance tracking performance, the existing association algorithm, the Munkres' implementation of the Hungarian algorithm [206] is applied after the segmentation stage. Then, this association procedure is extended to form the novel tracking system, proposed in this work, as detailed in Section 5.3.3.

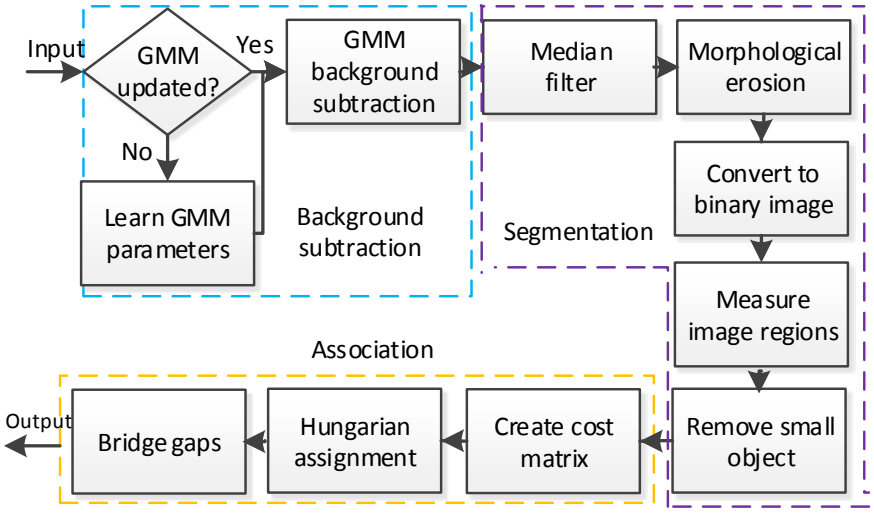


Figure 5.6: Overview of the proposed zebrafish larvae segmentation method with an existing tracking algorithm for further evaluation.

Background subtraction

The efficient estimation and removal of the video frame background is essential for zebrafish larvae tracking, to address water impurities and 'bursty' movement characteristics. The water particles, ripples and bubbles (as shown in Figure 2.5) when stirred up by the larvae tail beating produce motion data, which means that larvae detection methods based on motion (such as optical flow [182] and SIFT flow [154]) and frame-to-frame differences (such as segmentation methods used in idTracker [211] and LoliTrack [34]) cannot be applied [220]. As the movement of water impurities is relatively temporary compared to the larvae movement, the proposed system applies an improved adaptive Gaussian Mixture Model (GMM) [221], as detailed in the following.

Building GMM training samples. The time interval of training samples, T, is the length of samples and is determined as per the work in [199], where the first 500 frames of videos are used to estimate the GMM model parameters so as to obtain a consistent background model. In practice, however, there are many short microscopic zebrafish larvae videos where the number of frames is less than the required time interval, T. In this work, it is proposed that duplicate video frames are added at the beginning of short videos to allow GMM model background estimation.

For a short video where the length is less than the training sample required, duplicate video frames are added to the beginning of the video sequence to estimate the background model. Figure 5.7 illustrates the process of constructing the GMM training samples from a short video.

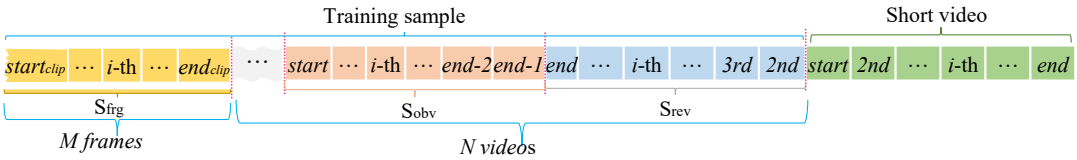


Figure 5.7: Building GMM training samples for a short video.

The green frame series in Figure 5.7 indicate the original time-lapse zebrafish larvae microscopic short video $\mathbb{V}$. The frames in front of this video are its training sample $\mathbb{S}$, used to estimate the background model parameters of $\mathbb{V}$. The set of training samples $\mathbb{S}$ consists of: S_{obv}, frames from the original short video $\mathbb{V}$ in obverse order from $[start, end-1]$ frames, shown by the pink frame series; S_{rev}, frames from the original short video $\mathbb{V}$ in reverse order from $[end, 2nd]$ frames, shown as the light blue; and, S_{frg}, the video fragment shown in orange. The obverse order frame series S_{obv} and reverse order frame series S_{rev} are alternately linked to each other to construct the training sample, and the last frame to connect to the original video is the reverse order frame to ensure a smooth background transition between the last frame in the reverse order frame series to the first frame of the original microscopic short video.

To construct a training set with length of L_{sample} images, a video fragment, S_{frg}, taken from the short video studied will be added to the training set for the set length requirement when the S_{obv} and S_{rev} series do not have exact L_{sample} images.

The required number of videos, N, and number of frames, M, are calculated using Equation 5.4 and Equation 5.5.

$$N = L_{sample} / N_{frames} \tag{5.4}$$

where the division is integer division with fractional quotients being rounded toward negative infinity to the nearest integer.

$$M = L_{sample} - L_{sample} / N_{frames} \tag{5.5}$$

Based on the required number of videos N (in obverse order and reverse order), the video fragment S_{frg} is constructed according to Equation 5.6.

$$S_{frg} = \begin{cases} \{S_{rev}(i) \mid i = M+1\,,\,\ldots\,,\,3,\,2\}, & \frac{N}{2}\ is\ even \\ \{S_{obv}(i) \mid i = 2\,,\,\ldots\,,\,3,\,M+1\}, & \frac{N}{2}\ is\ odd \end{cases} \tag{5.6}$$

Background GMM model estimation. The distribution probability density model parameters for the GMM are calculated according to the mixing weight and Minimum Message Length (MML) criterion [221], and the improved adaptive GMM [221] introduces an exponential decay envelope shaped by the constant factor α to adapt to background changes. The decay envelope factor α strongly weights the pixel samples representing the temporary movement of these water impurities and illumination changes, to minimise the influence of this temporary movement and enable fast adaptation to background changes. Therefore, the zebrafish larvae regions in the video frames are distinguished as moving objects and segmented by subtracting the calculated background model from the original video frames. That is, in the proposed system there are no requirements on the input video images as per existing zebrafish tracking approaches to have a clear background, use transparent containers without edge shadows, or have high intensity contrast between the zebrafish larvae and the background.

To address the 'bursty' movement specific to zebrafish larvae, with sudden swimming locomotion interspersed with substantially stationary periods of little movement [204], the applied GMM model [221] in the proposed system adds flexibility to the description of the background by adaptively and recursively selecting the number of Gaussian components used to represent each pixel compared with traditional GMM models with one or a fixed number of components to model each pixel. To determine the number of Gaussian components, the Dirichlet prior [116] and the MML criteria [221] are used to select the number of components on the basis of the final value of the component mixing weights. The same approach and initialization settings used in [222] are applied to recursively update the mixing weights for each

new sample. Whilst the initialisation stage is based on existing work [222], a novel contribution to the background subtraction is the proposed GMM training sample building method.

For the initialisation stage, the approach and number of randomly generated components are taken from [222], and the Dirichlet prior is applied. After each update, components holding a negative weight will be discarded by the MML criteria, with the components remaining taken as the number used in the model. The background model enables the removal of stationary background regions such as the zebrafish container and labels drawn on the petri dish; hence, unlike existing techniques, the proposed system is able to process larvae videos under practical experimental conditions.

However, if a larva in the video becomes static for some time, its body pixels will start to generate an additional stable cluster of pixels. But with the previously calculated background being occluded, the starting weight of the new stable cluster is very small. The cluster will only be classified to the background model when its weight is larger than a threshold (referred to as c_f in [221]) when the larvae remains static for long enough, which will consistently increase the weight of the newly generated cluster. Thus, the detection period of larvae with no movement is extended for approximately $\log(1-c_f)/\log(1-\alpha)$ frames as calculated in [221].

Zebrafish larvae segmentation

After background subtraction, the resultant image generally still contains distortion due to a scattering of small noise fragments that are detected by the GMM model as moving objects. Thus, to remove these noise fragments the proposed system applies a median filter with a 3×3 square moving window (the smallest available window size for removing small fragments) [15] and morphological greyscale erosion [189].

Water ripples evoked by zebrafish larvae movement are also often detected by the GMM model. Further, the ripples cannot be completely removed by the median filter and mathematical morphological operation because their relative region size is typically larger than the noise. For these distortions, a binary image is firstly obtained from the greyscale image based on the global normalized threshold calculated using Otsu's method [179]. Then, the system calculates the number of pixels from each connected component in the binary image, and estimates the average size across the regions in the image. Regions which are less than 20% of the average larvae size (based on the typical relative size of ripples evoked by the larvae) are then removed from the segmentation image. This 20% threshold is taken from the experiments in [31], and empirical tests were conducted to verify this threshold as appropriate to the proposed system.

5.3.2 Multiple zebrafish segmentation performance evaluation

The proposed multiple zebrafish larvae tracking system consists of three stages: background subtraction, segmentation and association. As the assumption of improving segmentation accuracy to enhance the overall tracking performance has not yet been examined for multiple zebrafish larvae, this Section evaluates the segmentation approach in the proposed multiple zebrafish tracking system, compared with the segmentation method within idTracker [31], and the well-known motion feature based optical flow [182] and SIFT flow [154] methods.

To evaluate the tracking performance due to the proposed improved segmentation approach, the existing tracking approach utilised in SimpleTracker [223] is applied based on the segmentation results of the proposed segmentation approach. SimpleTracker [223] uses the Kuhn-Munkres tracking algorithm for the initial association and the nearest neighbour to directly connect trajectory fragments, without considering missed objects. The overall tracking accuracy of the proposed system (as shown in Figure 5.6) is then compared with idTracker [31], and the widely used commercial

LoliTrack system [34]. All evaluation experiments are performed using the zebrafish larvae segmentation and tracking dataset presented in Sections 3.5.1 and 3.6.2 in Chapter 3, annotated with manually generated segmentation and tracking ground-truth.

Zebrafish larvae association between frames

To enable evaluation of tracking accuracy due to proposed improved segmentation, the Munkres' implementation of the Hungarian association algorithm is applied to the proposed segmentation result to form a tracking system (and an improved association algorithm is presented in Section 5.3.3). That is, after zebrafish larvae segmentation the moving larvae objects are associated or matched between successive frames to obtain individual larvae tracking trajectories. An $n \times m$ matrix $\mathbb{D}$ is created to annotate the cost of associating source objects $O = O_1, O_2, \cdots, O_n$ in the frame t to the target objects $T = T_1, T_2, \cdots, T_m$ in the frame $t+1$:

$$D(O,T) = \begin{pmatrix} d_{O_1,T_1} & d_{O_1,T_2} & \cdots & d_{O_1,T_m} \\ d_{O_2,T_1} & d_{O_2,T_2} & \cdots & d_{O_2,T_m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{O_n,T_1} & d_{O_n,T_2} & \cdots & d_{O_n,T_m} \end{pmatrix} \tag{5.7}$$

where n is the detected number of zebrafish larvae in the frame t, and m is the number of zebrafish larvae segmented in the successive frame $t+1$. The element d_{O_i,T_j} in the matrix denotes the cost to connect the i-th object in the frame t, to the j-th object in the frame $t+1$. The value of d_{O_i,T_j} is calculated as the Euclidean distance from the source object to the target object based on the centroids of the segmented regions in Cartesian coordinates $O_i(x_i, y_i)$ and $T_j(x_j, y_j)$, as given by:

$$d_{O_i,T_j} = (x_j - x_i)^2 + (y_j - y_i)^2 \tag{5.8}$$

The frame-to-frame zebrafish larvae assignment based on the cost matrix is performed using the Muncres' implementation of the Hungarian algorithm [206], which searches for unique assignments i.e., assigns source object i to only one target object j in the secondary frame. The assignment is based on the global minimum of the smallest sum of squares distance amongst all of the possible associations, and allows $n \neq m$ in case of zebrafish larvae detection failure or larvae occlusion. Water impurities and well edge shadows may still remain in the binary bitmap, and to avoid these remaining noise fragments the maximum value $\max(dist_{GT})$ of organism displacement extracted from the tracking ground-truth is defined as the distance threshold [206]. In the cases where d_{O_i,T_j} is greater than the distance threshold, the value of d_{O_i,T_j} in the cost matrix $\mathbb{D}$ is set to ∞ before mapping association.

In the cases where zebrafish larvae fail to be detected or segmented in one frame but reappear in subsequent video frames, a 'gap' in the moving trajectory of this object will appear at the frame where the zebrafish larvae detection initially failed, with a resulting new trajectory created from the frame where the zebrafish reappears. Scenarios of multiple object occlusion [32] and misdetection of long-term stationary larvae objects can generate such trajectory 'gaps'. Thus, in the proposed tracking system a 'gap bridging' stage is performed using the nearest neighbour algorithm [224] to connect trajectory fragments and improve the inter-frame larvae object association. However, the trajectory gap will not be connected if the squared distance calculated between the two frames of trajectory fragments is greater than the distance condition calculated as $1.2 * (\max(dist_{GT}))^2$. The ratio of 1.2 is extended by 20% beyond unity to set a margin for rebound, similar to the threshold ξ in the gap filling stage of [32], with the trajectory at that frame recorded as an error.

Results and discussion

Segmentation evaluation metrics. This work uses three standard metrics and a proposed metric as Equations 3.4 - 3.7 (presented in Section 3.5.2) to numerically

quantify the segmentation performance. The metric Similarity Index (SI) accounts for the number of correctly segmented objects by penalizing missing objects or object occlusion. The recall and precision metrics estimate under-segmentation and over-segmentation, respectively. The $F_{measure}$ is a weighted calculation of the precision and recall. More detailed information about $F_{measure}$ and SI are presented in Section 3.5.2.

Tracking evaluation metrics. To enable the objective evaluation of tracking performance on the database presented in 3.6.2, this work employs the widely utilized standard Multiple Object Tracking (MOT) metric: Classification of Events, Activities and Relationships (CLEAR MOT) [225].

CLEAR MOT consists of two metrics: Multiple Object Tracking Precision (MOTP), which estimates the location precision of all detected objects compared to that of the manually labelled zebrafish larvae positions in each frame (known as ground-truth); and, Multiple Object Tracking Accuracy (MOTA), which measures the accuracy in tracking object trajectories (producing exactly one trajectory per object), and the ability to consistently label objects over time. Mathematically, the MOTP and MOTA metrics [225] are represented as:

$$MOTP = \frac{\sum_{i,t} |D_{i,t} - GT_{i,t}|}{\sum_t N_t} \tag{5.9}$$

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \tag{5.10}$$

where $|D_{i,t} - GT_{i,t}|$ indicates the Euclidean distance between the pair-wise matched position of the i-th segmented object in the t-th frame $D_{i,t}$ and the position of this object in the ground-truth $(GT_{i,t})$, averaged by the total number of matches in the entire video sequence.

In the MOTA metric, m_t, fp_t, and mme_t for each frame t indicate the number of missed zebrafish detections, false positive segmentation (e.g., image noise fragment detected as zebrafish), and the swapping of identities for individual zebrafish larvae, respectively. g_t represents the total number of objects in frame t.

Segmentation performance

Figure 5.8 shows the average $F_{measure}$ and SI scores presented with the 95% confidence intervals for each of the 10 zebrafish video sequences in the dataset, presented in the order that the first sequence has the clearest background and the 10th (last) sequence has the most complex background. Details of the tested dataset, such as the number of frames for each sequence, are presented in Table 3.2 in Section 3.5.1.

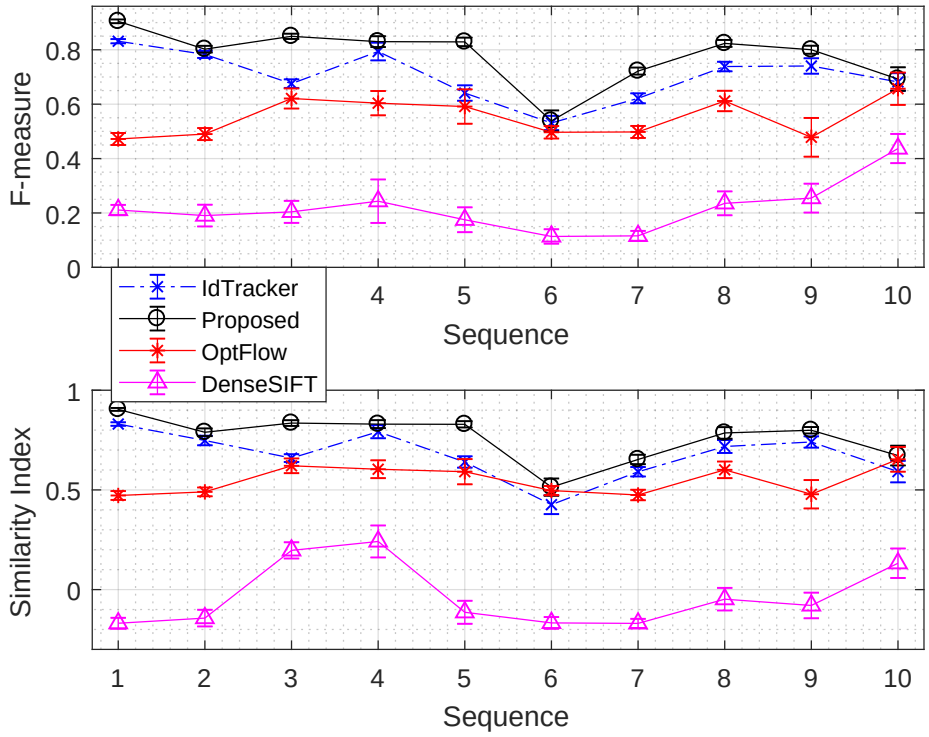


Figure 5.8: Segmentation accuracy over the 10 zebrafish video sequences.

The results as seen in Figure 5.8 show that the overall segmentation accuracy of the proposed method has an average 7.54%, 22.72% and 56.12% higher $F_{measure}$ than idTracker, optical flow and SIFT flow, respectively. The proposed approach also

exhibits an 8.74% and 21.30% higher similarity index compared to idTracker, and optical flow, respectively, indicating an improved performance in relation to missing or occluded objects. In particular, the proposed method is more robust against challenging background environments, such as unclear zebrafish well containers with labels (as illustrated by sequences 8 and 9). The robust segmentation accuracy as seen in Figure 5.8 across the 10 videos under variant background conditions evaluated with the proposed system further shows that the segmentation performance does not depend on video input tested.

The sensitivity of the segmentation accuracy due to the tuning factor α was examined in [220], where the range of α values evaluated showed a consistent and reliable segmentation performance. In turn, the robust segmentation accuracy seen in Figure 5.8, which illustrates the 10 videos under variant background conditions evaluated with the proposed system shows that the segmentation performance does not depend on the video input tested. Further, optical flow, SIFT flow and idTracker respectively exhibit 1.95%, 2.45% and 0.47% more variance in all of the evaluation metrics studied, which suggests that the segmentation results are less reliable across the complex video sequences evaluated.

Effect of improved segmentation on tracking accuracy

Figure 5.9 summarises the tracking accuracy using the MOTP and MOTA metrics [225] evaluated over the 10 video sequences. Seq. 1 has the clearest background, seq. 2-6 each have one type obstruction (well edge shadows, particles, particles, labels on well, and well edge shadows, respectively), and seq.7-8 each have two types of obstruction, and seq. 10 has the most complex container background conditions. Both the proposed tracking system and existing tracking approaches perform reliably when the videos have restricted background conditions, as shown by the MOTP and MOTA values for seqs. 1- 4 in Figure 5.9. However, when the video conditions increasingly degrade from seqs. 5-10, the proposed system is more reliable than the existing systems as illustrated by the consistent performance of the proposed system

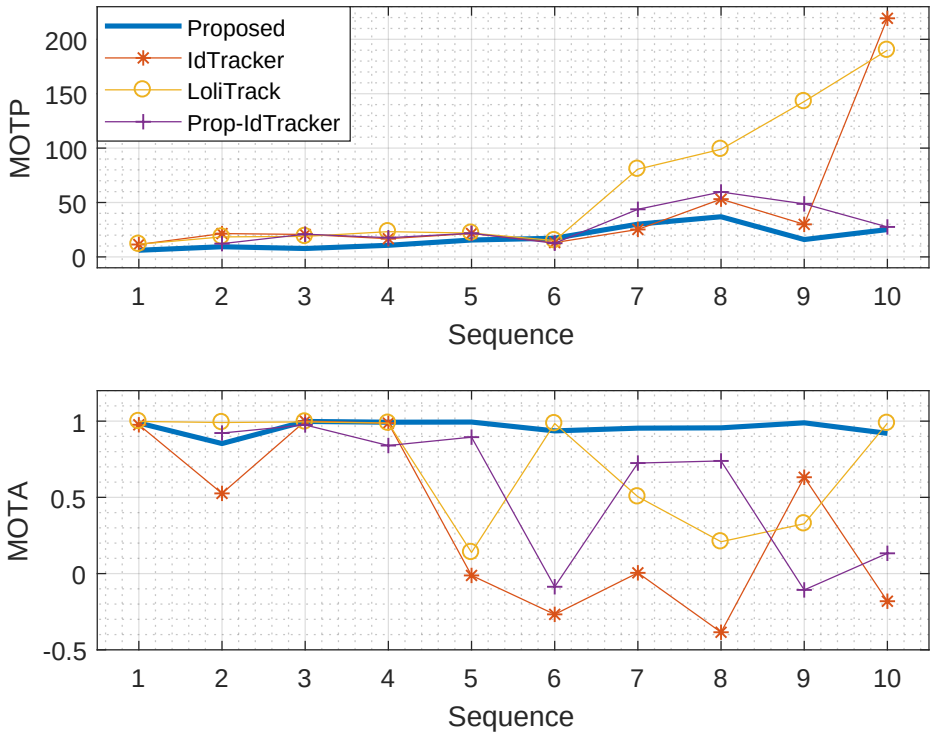


Figure 5.9: Tracking accuracy over the 10 zebrafish video sequences.

compared with the existing systems as measured by both metrics from seqs. 5-10 in Figure 5.9. Further, the proposed system exhibits the smallest position detection error, with a decreased overall error of 25.61 and 44.49 pixels using MOTP compared to idTracker and LoliTrack, respectively, and an increased accuracy of 31.57% and 27.2% using MOTA compared to idTracker and LoliTrack, respectively.

The proposed segmentation approach is also applied as pre-processing to the idTracker system to determine the effect of the proposed segmentation approach on the overall tracking accuracy. The result of both the MOTP and MOTA values improved by 32.00% and 22.91%, respectively, compared with the original idTracker system. The proposed background subtraction and segmentation processing also removes the need to constrain the input zebrafish larvae video imaging conditions, and enables the testing of videos under realistic experimental conditions using idTracker. That is, researchers who already use idTracker can apply the proposed segmentation method as pre-processing to obtain tracking results of higher accuracy using the

Table 5.2: Total number of swapped individual identities

No.	idTracker	Segment + idTracker	LoliTrack	Proposed
1	6	NaN*	2	0
2	2	3	4	3
3	0	0	0	0
4	4	0	0	0
5	3	2	6	0
6	1	1	1	2
7	2	1	3	1
8	11	41	24	14
9	4	17	3	0
10	4	8	0	8
Average	3.7	8.1	4.3	2.8

*Note: No valid data generated due to the running error when testing Seq. 1.

existing idTracker system, with video data in unconstrained imaging conditions.

Table 5.2 summarises the total number of individual identities swapped across each tested video in the dataset. The proposed system exhibits the smallest identity swapping rate, with 24.32% less identity swap than the idTracker system. In addition to the proposed segmentation method exhibiting a consistently higher accuracy segmentation than idTracker as shown in Figure 5.8, applying the proposed segmentation method to idTracker reduces the zebrafish larvae misdetection and false positive rates, as shown in Figure 5.9. However, the identity swapping rate is doubled as shown by Table 5.2, due to the generated binary foreground images providing limited intensity information for idTracker to generate the required fingerprint.

Summary of the proposed segmentation approach

Compared to the tracking of adult zebrafish in microscopic videos, the dynamic 'bursty' locomotive characteristics and complex video imaging conditions of zebrafish larvae due to their small size relative to background imaging artefacts poses many different challenges for the tracking of multiple zebrafish larvae. This Section proposed a zebrafish larvae segmentation approach for both single and multiple

zebrafish larvae under complex video conditions, applying an adaptive GMM probability density model, median filter and morphological operations to segment larvae objects from the background. Comparisons with the segmentation performance of existing state-of-the-art biological small organism tracking systems illustrated the accuracy of the proposed method, where the proposed segmentation approach also removes the strict limitations on input video imaging conditions to enable the testing of unconstrained experimental videos.

Enhanced overall tracking performance due to the improved segmentation accuracy was also illustrated by augmenting the proposed segmentation approach with the Hungarian assignment algorithm as a tracking system. Further, the proposed background subtraction and segmentation approaches applied alone as pre-processing to existing tracking systems (such as idTracker) improved the multiple zebrafish tracking accuracy by up to 32%. This is due to decreased zebrafish larvae misdetection and false positive rates; however, the identity swapping rate may increase if the identity is generated using intensity variance information, such as the approach used in idTracker. The proposed segmentation will also work for adult zebrafish as illustrated by the adult zebrafish sequence (seq. 1). Together with the increased size and intensity contrast, the continuous swimming movements of adult zebrafish thus provide consistent motion features that can be easily captured by the adaptive GMM model and subsequent object tracking.

5.3.3 Improved zebrafish larvae association across video frames

This Section presents the proposed tracking system, which improves upon the association stage in the system as shown in Figure 5.6, to improve robustness against non-ideal zebrafish larvae detection and segmentation results obtained using microscopic time-lapse videos taken under practical laboratory experimental conditions.

Figure 5.10 outlines the overall workflow of the proposed multiple zebrafish larvae tracking system. The improved assignment algorithm within the proposed system takes advantage of inter-frame information to construct the individual tracking trajectories by linking the detected zebrafish in each video frame to correspondences in the next successive frame. Then, the positions of misdetections or occlusions are calculated, and the calculated organisms are re-assigned to their correct trajectories by adjusting initial assignment results.

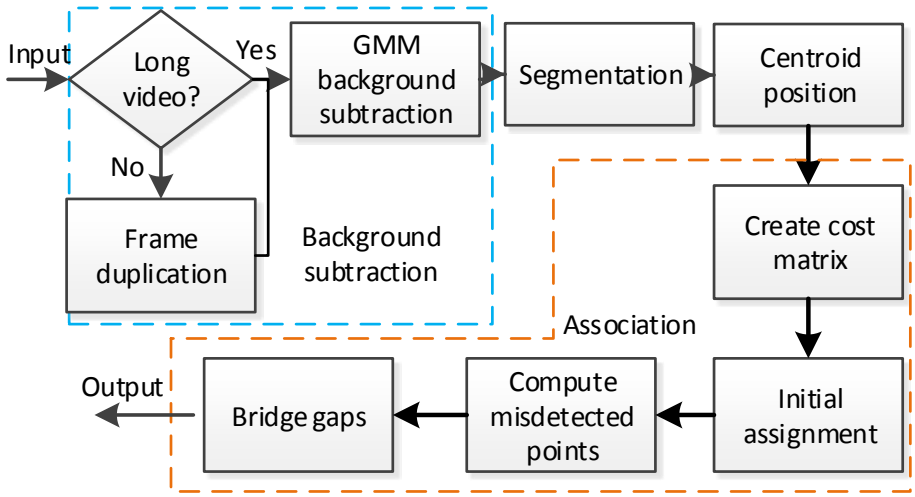


Figure 5.10: Overall multiple zebrafish larvae tracking system.

The proposed system applies background subtraction-based computer vision segmentation to detect and segment the zebrafish larvae from each video frame as presented in Section 5.3.1. The Kuhn-Munkres algorithm [223], which guarantees one to one association, is then used for target association or mapping between successive video frames to generate individual tracking trajectories based on the (non-ideal) segmentation results. The positions of misdetected and overlapped zebrafish larvae are calculated through knowledge of their neighbours' locations within the individual tracking trajectories. This step plays an essential role in maintaining consistent identities for detected individual zebrafish larvae over time. The individual tracking

trajectories then enable further movement analysis, where this work estimates the velocity, acceleration and movement direction, presented in Section 5.3.4.

Representing detected zebrafish larvae

To represent the positions of the detected zebrafish larvae in each video frame, the centroid locations of segmented foreground regions in Cartesian coordinates are used and stored in a vertical cell array matrix, as shown by the parallelogram series (indicating video frames) and $Points\{t\}$ matrix in Figure 5.11. In the $Points\{t\}$ cell array, the first column stores the temporary identity, numbered from 1 to the number of detected zebrafish larvae in each frame, where $n_{obj}(t)$ indicates the number of detected zebrafish in frame t. The second and third column stores the horizontal and vertical positions of each detected organism in X and Y coordinates, respectively. This cell array and matrix representation allow for varying element length to indicate the number of detected foreground zebrafish, which can change frame-to-frame due to detection and segmentation errors.

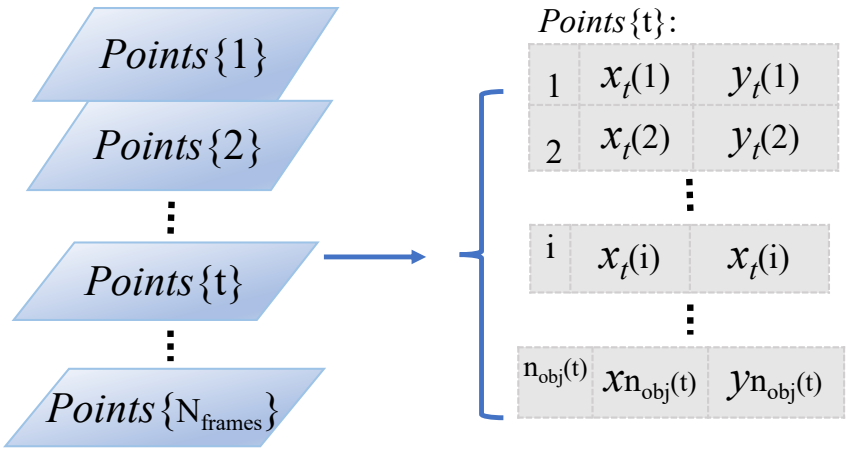


Figure 5.11: Storage structure of detected zebrafish larvae in a video sequence.

The centroid positions of detected zebrafish larvae are obtained in the segmentation process and represented frame-by-frame using a cell array for a video sequence as described in this Section; however, the individual identities in each frame are still

unknown. That is, which zebrafish larvae in the current frame correspond to which target in the following frame has not been mapped. In addition, there are still some remaining misdetected zebrafish larvae that have been classified into background clusters or overlapped with other detected zebrafish larvae.

Cost matrix and initial assignment

The initial assignment of detected points in each video frame is a partial assignment using an extension of the Kuhn-Munkres algorithm to frame-by-frame processing for rectangular arrays [223]. Extending the original Hungarian algorithm [226] from solving the assignment problem with an equal number of workers and tasks represented in an $n \times n$ matrix, the number of workers and tasks can be unequal and represented in a rectangular matrix. This approach can then be applied to zebrafish larvae tracking, where the number of detected zebrafish can change due to non-ideal segmentation resulting from zebrafish misdetection and occlusion.

In the initial assignment process in this extended Kuhn-Munkres approach [223], all the detected zebrafish $Points\{t\}$ in frame t are taken as source points, and the segmented zebrafish $Points\{t+1\}$ in the following frame $t+1$ are seen as the target points between connected frames. The target points in $Points\{t+1\}$ are mapped to the source points in $Points\{t\}$ frame-by-frame across a video sequence.

The construction of the cost matrix $\mathbb{D}$ (details as presented under "Zebrafish larvae association between frames" in Section 5.3.2) of size $n \times m$ for associating source zebrafish larvae $O = \{O_1, O_2, \cdots, O_n\}$ in the frame t to the target zebrafish larvae $T = \{T_1, T_2, \cdots, T_m\}$ in the frame $t+1$ is given by Equation 5.7. The element d_{O_i,T_j} in the cost matrix $\mathbb{D}$ denotes the cost to connect the i-th point in the frame t, to the j-th point in the frame $t+1$, and is calculated by the Euclidean distance between the source point to the target point using Equation 5.8.

The frame-to-frame target initial assignment searches for unique mappings to assign source point i to only one target point j in the successive frame. The final sum of the resultant complete assignment between $Points\{t+1\}$ and $Points\{t\}$ is a global optimal cost, which is the lowest summed distance amongst all of the possible assignments within two successive frames. The matched target points propagate the identities of their matched source points; thus, after obtaining the final assignment map for the whole video, connecting the points with the same identities over all frames of a video results in the individual organism tracking trajectories.

To eliminate false positive points from the segmentation results when building the individual trajectories, a distance constraint is set as a threshold in the source-target cost matrix when applying the initial frame-to-frame assignment. The threshold is calculated by $\delta * \text{median}(dis_{i,j})$ [32]. When the minimum value of the i-th row in the source-target cost matrix is larger than the threshold value of this video sequence, which indicates that the distance between the source point i to all of the points in the successive frame exceeds the threshold value, the source point i with its corresponding position information will be removed from the point matrix and considered as a segmentation noise fragment to not be further assigned to a target point.

When a zebrafish larva disappears in a frame t due to misdetection or occlusion (which further results in overlapped trajectories as presented in Section 5.1.2), a source point in frame $t-1$ therefore cannot be assigned to a target. A gap will then occur in the tracking trajectory where the target fails to be detected, and a new tracking trajectory will start from the frame when the zebrafish larvae is correctly detected again. This source point in frame $t-1$ without a mapped target is saved in an unmatched source matrix.

When the zebrafish larvae is re-detected in the frame $t+n$ after being missed for n frames, there is one more point in $Points\{t+n\}$ compared to $Points\{t+n-1\}$. To map the points $Points\{t+n-1\}$ to $Points\{t+n\}$, a point in the frame $t+n$ cannot be assigned to a source point in the previous frame; this point in frame $t+n$

is saved in the unmatched target matrix.

The methods to calculate the theoretical positions between the unmatched source points and unmatched target points and adding these points to their correct tracking trajectories are explained in the following two sections, respectively.

Position estimation for misdetected and occluded zebrafish larvae

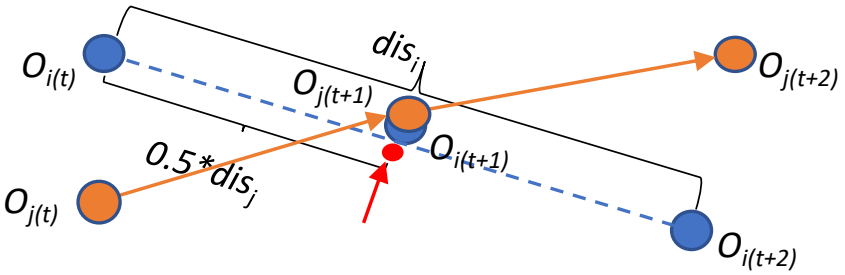


Figure 5.12: Point calculation for occluded organisms

Figure 5.12 illustrates the location computation of an overlapped zebrafish larva. The two points O_i (shown by the blue dot) and O_j (shown by the orange dot) overlap with each other in the frame $t+1$, and this overlapped point at time $t+1$ is assigned to the object O_j by the initial tracking process. Since the point O_i in the frame t cannot find a target in the frame $t+1$, and the point in the frame at time $t+2$ cannot be assigned to a source point in frame $t+1$, the point O_i is classified as an unmatched source point in the frame t, and unmatched target point in frame $t+2$ by the initial assignment approach described in the previous section, respectively.

The position of the missed point due to occlusion (as shown by the blue dot partially covered by the orange dot in Figure 5.12) or misdetection (for example, when the blue dot is totally covered by the orange dot in Figure 5.12) is calculated by the unmatched source point $O_i(t)$ and unmatched target point $O_i(t+2)$. For example, as shown in Figure 5.12, the location of the missed point O_i at frame $t+1$ is calculated by the median point between $O_i(t)$ and $O_i(t+2)$.

When there are multiple unmatched pairs, the mapping from unmatched target points to the unmatched source points is also based on the extended Hungarian assignment algorithm [223]. In searching for target points, the unmatched source points firstly search for possible correspondences in the following frame. If no assignment can be mapped, the search extends to the unmatched targets in the following 3rd frame. It was shown in [10] that a trajectory fragment can be connected to its subsequent fragment track so long as the frame separation is less than 6 video frames. Thus, the default search range for the proposed multiple zebrafish larvae tracking system is from the second to the sixth frames following the frame when the misdetections and occlusions originally occur. The positions of missed organisms (x_c, y_c) are calculated by Equations 5.11 and 5.12 using the matched point pair from an unmatched source point (x_s, y_s) and an unmatched target point (x_t, y_t).

$$x_c = x_s + \frac{1}{j} * (x_t - x_s) \tag{5.11}$$

$$y_c = y_s + \frac{1}{j} * (y_t - y_s) \tag{5.12}$$

where j indicates the following j-th frame from the unmatched source point.

Bridging trajectory gaps for misdetected and occluded zebrafish larvae

Individual tracking trajectories for each zebrafish larvae are obtained through connecting the matched points with the same identities after the initial assignment process frame-by-frame over a video sequence. However, the tracking trajectories obtained from the initial assignment process are usually trajectory fragments, separated when zebrafish larvae are misdetected or overlapped to result in segmentation errors. In the proposed system, these trajectory gaps are bridged by adding the location

points of these misdetected or overlapped zebrafish larvae as estimated and described in the previous section.

To connect trajectory fragments, the points stored in the unmatched source matrix are mapped to the points in the unmatched target matrix, and the positions of the missed points between the newly unmatched source to an unmatched target are also calculated during this search process. For example, as shown in Figure 5.12, the unmatched source $O_i(t)$ as the end point of its trajectory fragment is connected to the unmatched target point $O_i(t+2)$, which is the start point of its trajectory fragment, and their middle point shown by the red dot is added between points $O_i(t)$ and $O_i(t+2)$ as the theoretical position of the overlapped point $O_i(t+1)$.

5.3.4 Locomotion characteristic analysis

After obtaining the individual tracking trajectories for each zebrafish larvae in video sequences, the movement characteristics of individual zebrafish can thus be analysed based on the individual tracking trajectories. This work presents the calculation of basic locomotive parameters to analyse the zebrafish movement characteristics after obtaining individual tracking trajectories.

The calculation of three movement parameters are presented: movement velocity, acceleration and direction as represented by Equations 5.13 - 5.15, respectively. These calculated locomotive parameters provide an unbiased approach to dynamic behaviour analysis [8].

$$velocity = \frac{\sqrt{(x_{t+1}(i) - x_t(i))^2 + (y_{t+1}(i) - y_t(i))^2}}{dt} \tag{5.13}$$

where $dt = 1/fs$, fs is the video frame rate, and $x_t(i)$ and $y_t(i)$ are the Cartesian coordinates of organism i (i is the zebrafish larva identity as assigned by the association algorithm) in the frame t. The $velocity$ described by Equation 5.13 is calculated

based on the Cartesian position changes of individual zebrafish larvae between the successive frames [227]. The resultant *velocity* by this method enables the analysis of zebrafish larvae movement speed and direction over time, which provides quantitative measurement on the dynamic behaviour of the studied organisms.

$$acceleration = \frac{d(velocity)}{dt} \tag{5.14}$$

The *acceleration* parameter is calculated by quantifying the *velocity* changes over time on microscopic image pixels based on using the same Cartesian coordinates as movement velocity. The algebraic sign of the *acceleration* value indicates the velocity change of the zebrafish larvae: a positive *acceleration* value denotes that the studied individual larva is increasing in velocity, whilst a negative *acceleration* value indicates a decreasing velocity movement pattern. By examining the algebraic sign of the *acceleration* parameter over a period of time, the resultant *acceleration* parameter can provide insight into the status of the zebrafish larvae nervous system e.g., being nervous or relaxed [8, 228, 229]. Further, the larvae accelerations and decelerations can be used as an equivalent way to analyse momentum in fluid aquatic locomotion [227, 230].

$$direction = \text{atan2}\,\frac{y_{t+1}(i) - y_t(i)}{x_{t+1}(i) - x_t(i)} \tag{5.15}$$

The arc tangent in Equation 5.15 expressed in radians is used to compute the movement direction of individual zebrafish larvae from the positions in two successive frames. The resultant direction value is confined within $(-\pi, \pi]$. The signs of the calculated arc tangent are used to determine the quadrant of the calculated movement direction in Cartesian coordinates.

5.3.5 Results and discussion

To evaluate the proposed system, zebrafish larvae microscopic videos presented in Section 3.6.2 are used. No chemical stimuli are tested on the studied zebrafish larvae models, so their behaviour response does not correspond to any specific chemical stimuli. In addition to the tracking accuracy evaluation, the natural locomotive characteristics for individual zebrafish larva trajectories described by movement velocity, acceleration and direction as presented in Section 5.3.4 are also analysed on the video dataset to test the dynamic behaviour analysis capabilities of the proposed tracking system.

Tracking accuracy evaluation

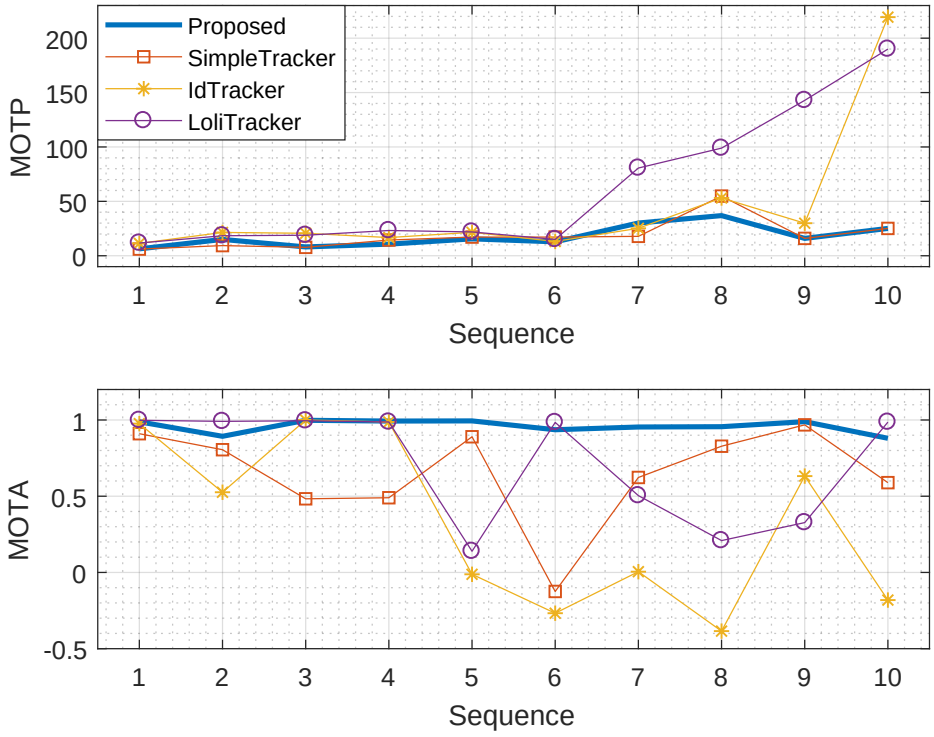


Figure 5.13: Tracking results evaluation among the methods compared using zebrafish larvae videos.

To evaluate the proposed tracking system, the overall tracking accuracy over a zebrafish video sequence is compared with well-known multiple object tracking approach idTracker [31], SimpleTracker [223] (as applied to examine the effect of improved segmentation on the overall tracking performance in Section 5.3.2), and the widely used commercial LoliTrack system [34]. Figure 5.13 shows the tracking accuracy results measured by the MOTP and MOTA tracking evaluation metrics (described in Section 5.3.2). The number of miss-detections was accounted for calculating the *SI* (*Similarity Index*, which penalises missing objects or object occlusion). The resultant *SI* was presented in Figure 5.8 in Section 5.3.2. Details of the tested dataset, such as the number of zebrafish tracked for each sequence, are presented in Table 3.2 in Section 3.5.1.

It can be seen from Figure 5.13 that the positions of detected organisms of the proposed system consistently exhibit the smallest distance differences with the manually labelled ground-truth positions amongst all the methods tested. The proposed system resulted in overall smaller MOTP values of 0.92, 25.59, and 44.48 pixels compared to SimpleTracker, idTracker, and LoliTrack, respectively. The generally improved zebrafish larvae position detection results demonstrated the effectiveness of the theoretical position estimation based on the zebrafish larvae detection and segmentation results using the proposed tracking system.

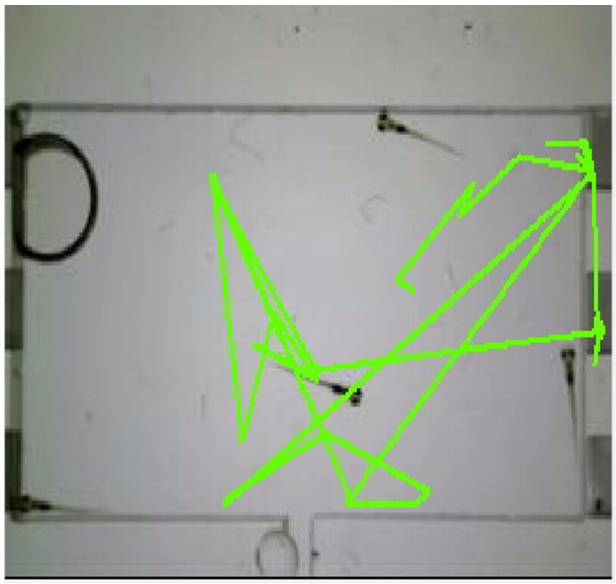
All of the methods performed well for position accuracy of the detected zebrafish when the videos had a clear background as shown by sequences 1-6. However, the location errors of the proposed system measured by MOTP for the detected organisms in the zebrafish videos compared to the ground-truth did not decrease as dramatically as LoliTrack or idTracker with increasingly complex video backgrounds, as shown by the MOTP values in sequences 7-10 in Figure 5.13. Accordingly, the proposed system still out-performed the existing approaches when taking into account misdetection, false positive segmentation and identity swapping, with 31.20%,

63.01%, and 24.61% higher MOTA values than SimpleTracker, idTracker, and LoliTrack, respectively. This was mainly achieved by the ability to estimate the positions of the misdetected or overlapped targets using knowledge of neighbour positioning from the segmentation results in the proposed system.

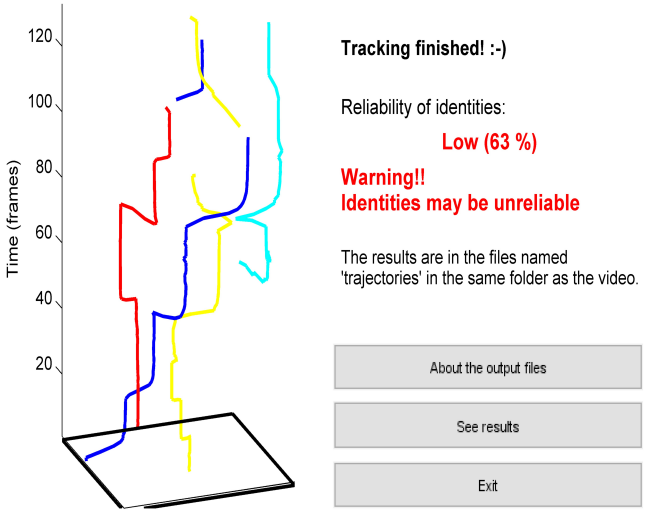
In addition, the approach to bridge trajectory fragments in the proposed system using the extended Hungarian assignment algorithm [223] decreased the possibility of individual identity swaps. This is especially the case when there are multiple unmatched trajectory fragment pairs based on unmatched source-target points. In contrast, SimpleTracker only uses a distance metric based on the nearest neighbour algorithm, and can additionally cause identity swapping during the gap bridging process.

Tracking trajectory example

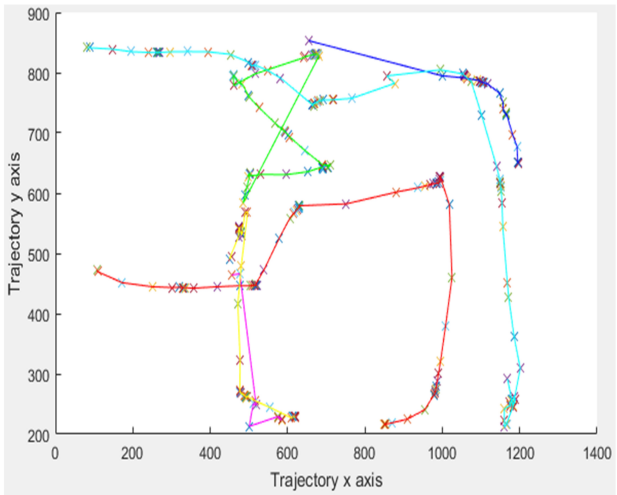
Figure 5.14 is a visual example of the tracking trajectory obtained for video sequence 4 by LoliTrack, idTracker and the proposed tracking system. Detailed quantitative evaluation on segmentation and tracking results for the zebrafish dataset compared with the ground truth can be referred to in previous sections of this Chapter. For example, miss-detection for each zebrafish microscopic video was accounted when calculating the *SI*. The resultant *SI* for each video was presented in Figure 5.8 in Section 5.3.2, and the quantitative tracking trajectory accuracy against ground truth over the tested dataset was presented in Figure 5.13 in Section 5.3.5. It can be seen that the proposed tracking system exhibits the most complete tracking trajectories estimated from realistic experimental video conditions. In contrast, LoliTrack (Figure 5.14a) detects the well edge shadow as zebrafish due to their similar intensity values, whilst idTracker system (Figure 5.14b) produces many trajectory gaps primarily caused by the false detection (as shown by the light blue line) of larvae objects due to their 'bursty' locomotive characteristics and small size differentiation with impurities inside the water. The resulting identity estimated from idTracker is therefore



(a)



(b)



(c)

Figure 5.14: Visual example comparing tracking trajectories.

also not reliable, with an estimated identity reliability of 63%, calculated according to the trajectory analysis of idTracker as shown in Figure 5.14b.

Analysis of zebrafish larvae movement characteristics

To explore the capability of the proposed multiple zebrafish larvae tracking system applied to target movement characteristic analysis due to the improved tracking trajectory accuracy, the movement parameters velocity, acceleration and movement direction were calculated for the zebrafish dataset as shown in Figures 5.15 - 5.17.

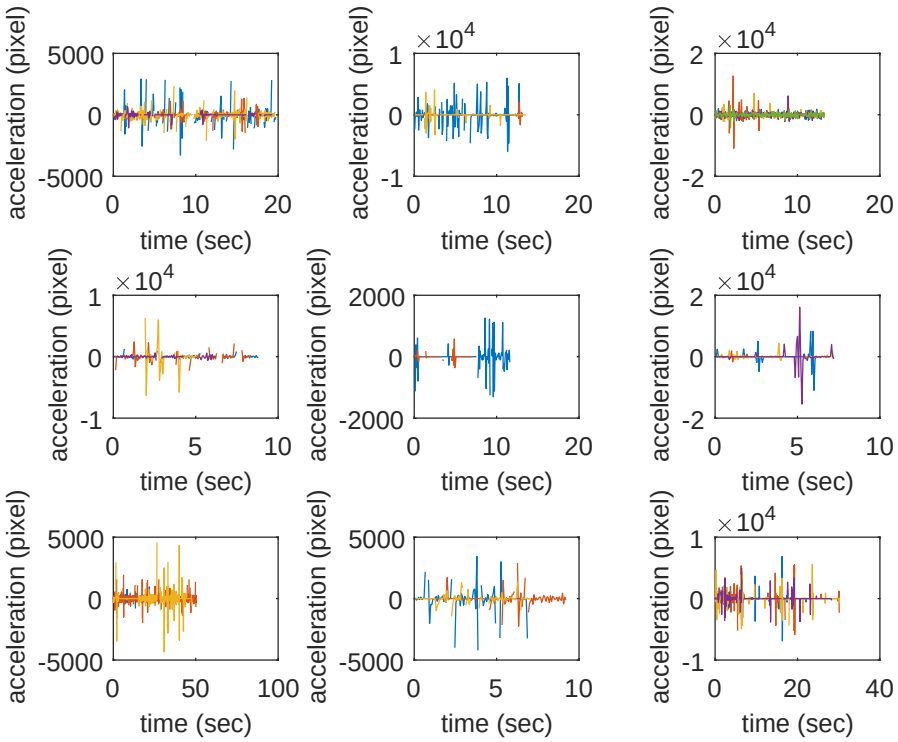


Figure 5.15: Zebrafish larvae movement acceleration.

Calculated acceleration for zebrafish individuals (denoted by different colours in sub-figures, which represent seqs.) using the 10 tested zebrafish videos presented in Section 3.3. Objects in seqs. 5 and 6 (one zebrafish each) are presented in the central sub-figure.

Figure 5.15 shows the movement acceleration analysis for each zebrafish video using the resultant tracking trajectories generated by the proposed tracking system (as presented in Figure 5.10). It was found in [231] that the interaction and movement of

zebrafish larvae were very close to zero by 7 dpf (days post fertilization). As shown in Figure 5.15, in general the variation in the zebrafish acceleration that is obviously visually perceptible to the human eye occurred only before approximately 10 seconds in the 10 adult and larvae zebrafish videos tested, due to the anxius response from the zebrafish when the imaging camera is introduced [232]). The zebrafish speed stabilises from this time onwards, which is consistent with known zebrafish movement characteristics [232] [231].

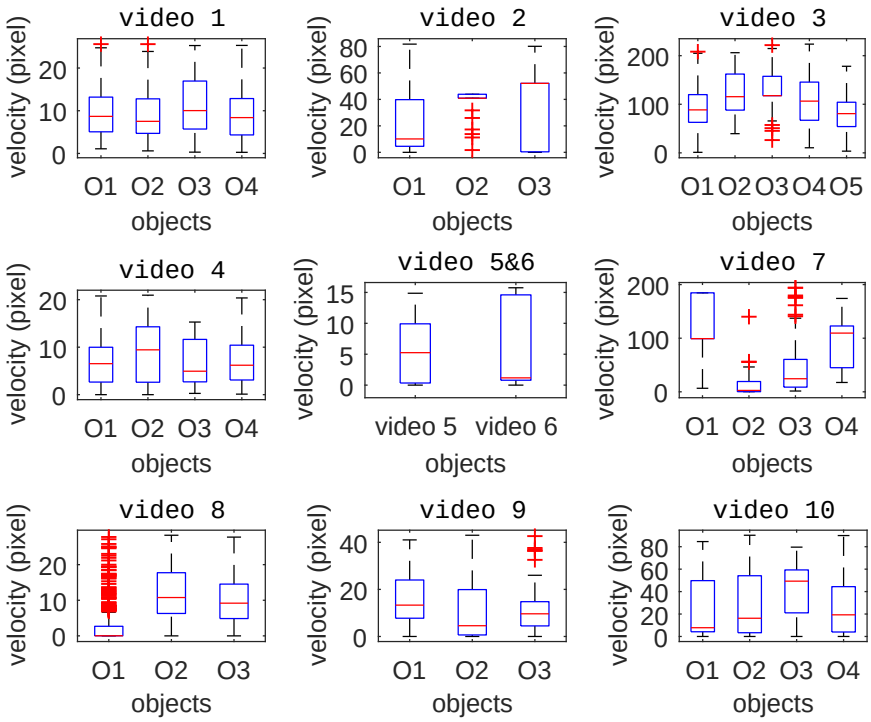


Figure 5.16: Zebrafish larvae movement velocity.

Calculated velocity for zebrafish individuals (denoted by boxes in sub-figures for each seq.) using the 10 tested zebrafish videos presented in Section 3.3. Objects in seqs. 5 and 6 (one zebrafish each) are presented in the central sub-figure.

As zebrafish organism movement speed and changes in movement direction provide insight into the individual dynamic behaviour and interaction rules [231], Figures 5.16 and 5.17 show examples of the movement speed and direction analysis results for each zebrafish in the tested videos, respectively. The number of boxes for each microscopic video show that every zebrafish larvae was successfully assigned with

one identity (which also illustrates the one-to-one zebrafish larvae mapping in the association method of the proposed system as presented in Section 5.3). The median velocity (labelled by the red line inside each box) of each larva indicates the consistent movement characteristic within the same housing well. If required, the mean, minimum and maximum velocity values and quartiles can be easily obtained and analysed through the visual box plot representations shown in Figures 5.16 and 5.17.

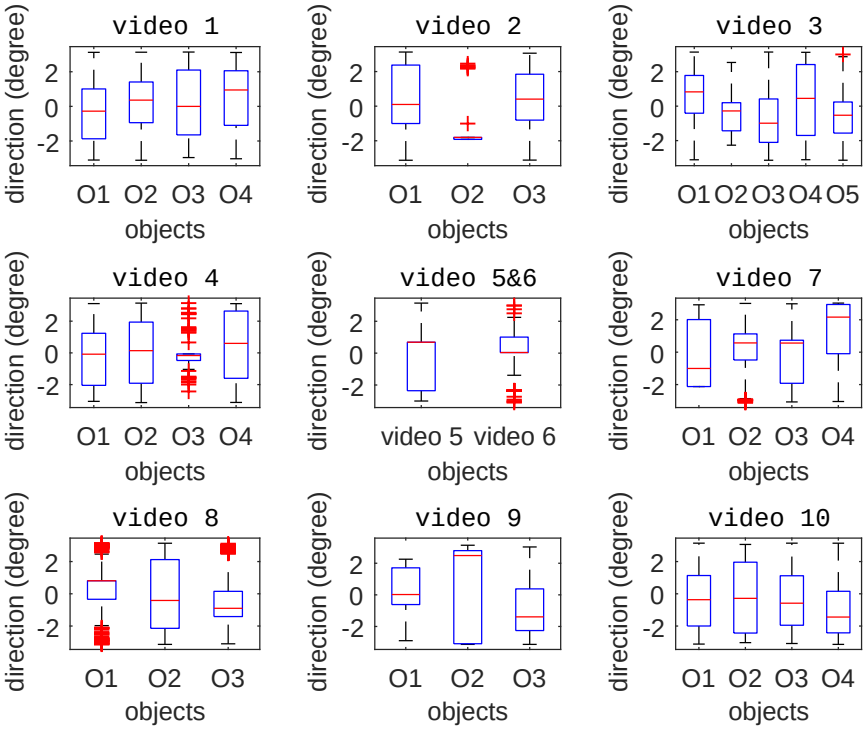


Figure 5.17: Zebrafish larvae movement direction analysis.

Calculated movement direction for zebrafish individuals, presented the same way as in Figure 5.16.

5.4 Summary

Single and multiple zebrafish larvae, whilst fundamental to many biological, ecotoxicity and medical applications, are typically manually tracked, which is both subjective and time-consuming. This Chapter presented and proposed single and multiple

zebrafish larvae tracking systems designed to automatically segment and track the 'bursty' movement characteristics specific to zebrafish larvae.

This work first proposed an automatic single zebrafish larva tracking method based on dense SIFT descriptors in a flow field combined with image warping to determine the moving zebrafish region in microscopic video frames. The proposed approach exploits the intrinsic advantage of dense SIFT flow to accurately represent the zebrafish larvae displacement, combined with image downsampling to reduce the computational complexity without affecting the zebrafish larvae detection accuracy, and a morphological operation to improve the single zebrafish tracking accuracy due to the'bursty' movement characteristics. The experimental results presented in Section 5.2.6, demonstrated the improved tracking performance of the proposed single zebrafish larva tracking method compared to the LSRtrack and the original SIFT flow method applied to pixel intensity thresholding. Further, experiments showed that downsampling of the frame resolution reduced the SIFT flow computational complexity without influencing tracking accuracy.

The computational complexity increases when applying dense SIFT flow to multiple zebrafish larvae tracking due to the increased number of feature descriptors required to represent the multiple targets. Thus, this work further investigated the novel adaptation of advanced computer vision techniques and multiple object tracking algorithms to develop an automatic and accurate multiple zebrafish larvae tracking system, that does not require any constraints on the input microscopic video imaging conditions as per existing tracking approaches.

However, zebrafish larvae misdetection and occlusions are inevitable problems when detecting and segmenting these small zebrafish larvae from time-lapse microscopic videos, which in turn affects the subsequent zebrafish larvae tracking processes. To improve the multiple zebrafish tracking accuracy based on non-ideal zebrafish larvae detection and segmentation results, this work combined an extended Kuhn-Munkres

algorithm for multiple object association to link detected zebrafish larvae frame-by-frame. To address segmentation errors due to misdetected or occluded zebrafish larvae, the proposed approach estimated the positions of organisms in interim frames using corresponding points in neighbouring frames. Finally, the calculated points are applied to connect and adjust the tracking trajectory fragments from the initial object associations. For the multiple zebrafish tracking system, the resultant tracking trajectories generated by the proposed system can then be used for further study, including the analysis of larvae movement characteristics.

The performance of the proposed multiple zebrafish larvae tracking system is evaluated based on segmentation and tracking accuracy of the zebrafish larvae dataset presented in Section 5.3, and outperforms three existing and well-known multiple zebrafish tracking systems. Moreover, the proposed system also provides the capability for locomotive characteristic analysis, using the individual tracking trajectories generated by the tracking to facilitate zebrafish larvae behaviour analysis research. Behavioural rules and new medicine or chemical effects on the dynamic behaviour of zebrafish larvae can therefore be further analysed in unbiased ways using the provided movement analysis approaches to facilitate dynamic behaviour related experiments.

Chapter 6

Generalised Multiple Small Biological Organism Tracking System (MSBOTS)

6.1 Introduction

In recent years, small biological organisms such as zebrafish larvae, *Artemia franciscana*, and *Daphnia magna* have become the most prominent vertebrate models used to facilitate genetic studies [67], studies of neural activity [233] and behavioural measurements [68]. Tracking techniques are vital for understanding the biology and ecology underlying the movements of these organisms [6, 20, 136, 234, 235]. The proposed automatic multiple zebrafish larvae tracking system presented in Section 5.3 of Chapter 5 was developed based on the specific 'bursty' movement characteristics of zebrafish larvae.

Due to the applied GMM model-based background subtraction and moving target segmentation techniques in the proposed tracking system, the automatic multiple zebrafish larvae tracking system has the potential to be generalised to detect and segment general moving small organisms from microscopic videos without any imaging

conditions. In addition, the tracking approach in the proposed system searches correspondences between successive video frames for each detected moving targets based on object centroid points. That is, the tracking algorithm does not 'know' which organism it is working on to create linking mappings after calculating organism region centroids from the segmentation results; the moving organisms are simply represented as points or nodes in the subsequent tracking procedure. Thus, after obtaining the centroid positions for each organism in video frames from the segmentation process, the type of organism in the microscopic videos does not affect the tracking results in the proposed system.

To evaluate and verify whether the proposed automatic multiple zebrafish larvae tracking system presented in Section 5.3 can be generalised to biological small organisms, this Chapter tests the system presented in Section 5.3 with two other types of small living organisms: *Artemia franciscana*, and *Daphnia magna*. And thus the system is denoted as the Multiple Small Biological Organism Tracking System (MSBOTS) in this Chapter. Five typical microscopic videos for each organism obtained under realistic experimental conditions (as described in Section 3.6.2) are tested using MSBOTS.

6.2 Applying MSBOTS to other small organisms

The overall procedure of the proposed MSBOTS mainly consists of organism differentiation from the video background (i.e. organism segmentation) and source-target assignment among the detected organisms (i.e. detected organism tracking). The accurate differentiation of organisms from the video background and image artefacts (e.g., water impurities) in each video frame is the critical foundation for multiple organism tracking systems. In the proposed MSBOTS, the video background is estimated by an adaptive GMM model [221]] as presented in Section 5.3.1. Organisms in every video frame are segmented after background subtraction; this segmentation

approach based on background subtraction is summarised and reported in Section 5.3.1. The following source-target assignment procedure is based on the computed centroid point locations of the segmented regions. This step plays an essential role in maintaining consistent identities of individual detected organisms over time. To find the corresponding organisms between frames, the Kuhn-Munkres algorithm [223] is applied in the association stage, as presented in Section 5.3.3.

This Section evaluates the proposed MSBOTS with *Artemia franciscana* and *Daphnia magna* to investigate the extended application of the tracking system to other small organisms with movement characteristics different to zebrafish larvae. The *Artemia franciscana* and *Daphnia magna* datasets are presented in Section 3.6.2 consisting of 10 video sequences. Again, no chemical stimuli were tested on the studied small organism models, and thus their behavioural responses do not correspond to any specific chemical stimuli.

6.2.1 Movement characteristics of tested organisms

The morphologies and locomotory systems of active *Artemia* larvae are different than those of adults [227]. The Reynolds numbers in their fluid regimes increases from 2 to nearly 40 during the whole larval stage, and the drag scale differs at low and high Reynolds numbers during development. These differences render *Artemia* an ideal model for locomotion comparison studies that include successive ontogenetic stages with a broad range of size or shape changes [227]. The locomotive pattern of *Artemia* changes gradually and average swimming speed increases as they grow, where the jerky swimming of *Artemia* larvae becomes smooth and continuous gliding through the water as seen with adult *Artemia*. In addition, the movement of *Artemia* larvae is influenced by changes in the surrounding fluid [227].

The dynamic behaviour parameters of the freshwater crustacean *Daphnia magna* are

affected upon exposure to sublethal concentrations of toxins such as dimethyl sulfoxide (DMSO)(a universally used aprotic solvent in ecotoxicology tests [155, 236]). The induction of significant behavioural abnormalities of *Daphnia magna* by toxicants is currently emerging in behavioural biotests in ecotoxicology [91, 155].

6.2.2 Evaluated microscopic videos

Figure 6.1 shows an example of a microscopic video frame of *Artemia franciscana*. In the amplified microscope view, the sizes of the water particles and labels on the container (as shown by the black circles in Figure 6.1) are comparable with the size of *Artemia*, and thus accurate differentiation of the *Artemia franciscana* from the microfluidic impurities is essential for the overall tracking performance.

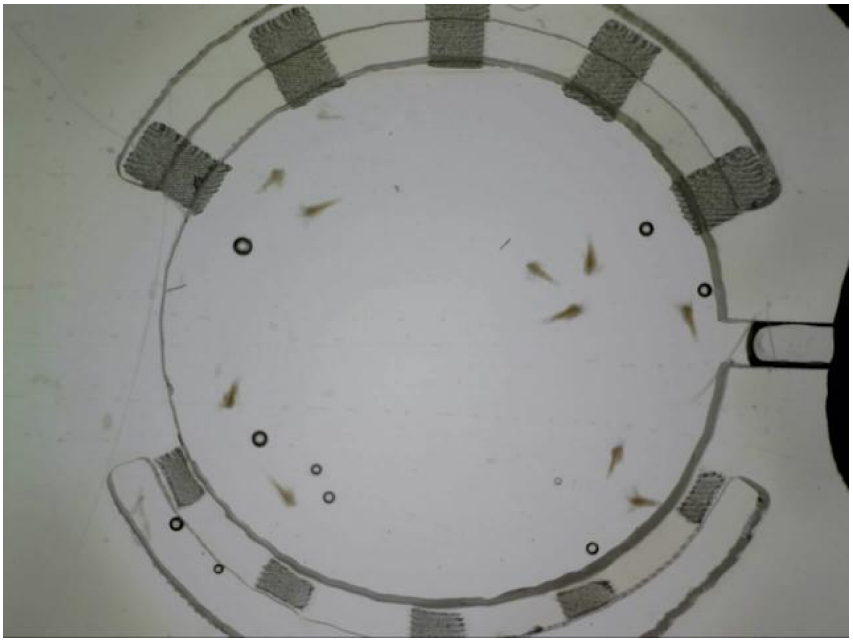


Figure 6.1: Microscopic video frame examples of *Artemia franciscana*.

Figure 6.2 shows an example microscopic video frame of *Daphnia magna*. Only the area containing *Daphnia* is of interest in the video frame, and thus the surrounding

area of the container was protected from illumination (as shown by the black regions on the micrograph margin in Figure 6.2).

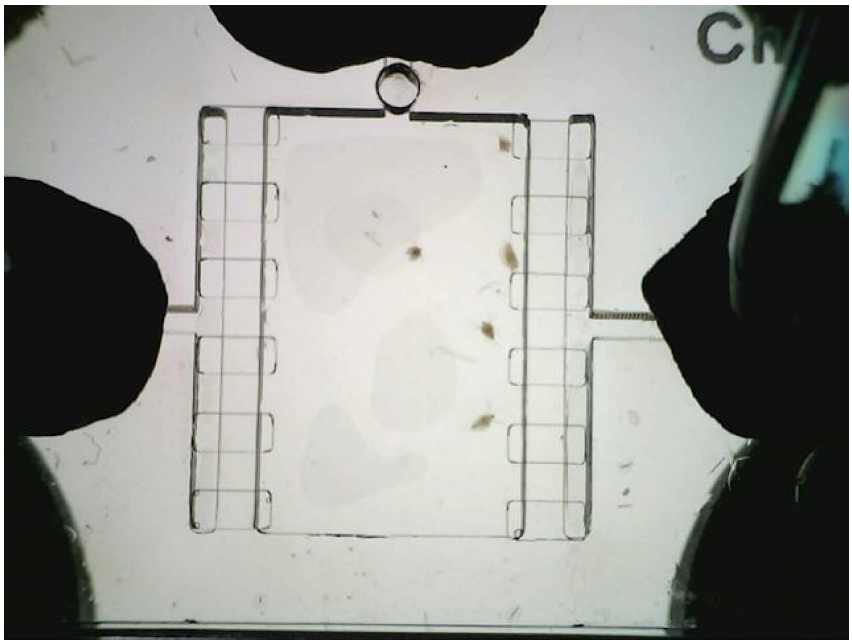


Figure 6.2: Microscopic video frame examples of *Daphnia magna*.

6.3 Results and Discussion

To enable the objective and quantitative evaluation of the tracking performance of the proposed MSBOTS approach, this Section employs the same Multiple Object Tracking (MOT) metrics used for the zebrafish larvae tracking performance evaluation as described in Section 5.3.2: MOTP and MOTA. The MOTP and MOTA tracking performance measurement metrics estimate the location precision of organism detection compared to manual labelling of organism positions in each frame, the accuracy in tracking object trajectories (producing exactly one trajectory per object), and the ability to consistently label organisms over time.

To evaluate the overall tracking performance of the proposed MSBOTS, the overall tracking accuracy over video sequences is also compared with those of three existing multiple organism tracking platforms: the well-known multiple object tracking approach idTracker [31], which has the ability to distinguish visually indistinguishable organism individual identities; SimpleTracker [223] (which combines the Kuhn-Munkres tracking algorithm for initial association and the nearest neighbourhood to connect trajectory fragments, without considering missed organisms from non-ideal segmentation results); and the widely used commercial LoliTrack system [34]. Five videos are tested for each organism type and the results are presented in Figure 6.3 and Figure 6.4, respectively.

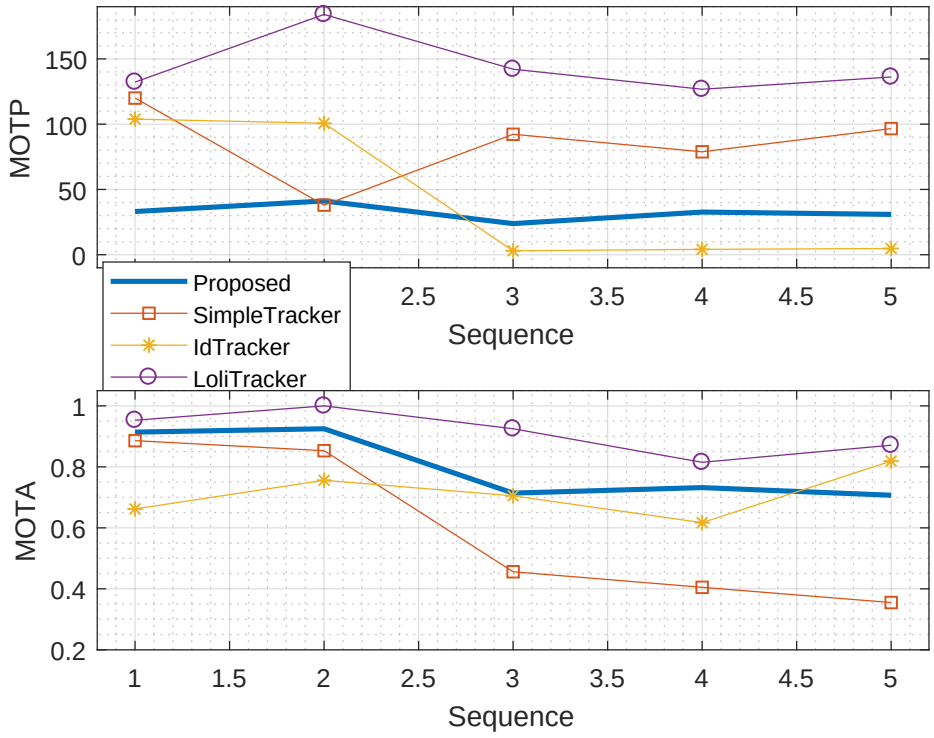


Figure 6.3: Tracking results evaluation among the tracking methods compared using *Artemia franciscana* microscopic videos.

Figure 6.3 and Figure 6.4 show the tracking accuracy evaluation using *Artemia franciscana* and freshwater *Daphnia magna* microscopic videos, respectively, to test the application of the proposed MSBOTS on small organisms with movement characteristics different to those of zebrafish larvae. In addition, *Artemia* exhibits flexible

movements that vary depending on the surrounding fluidics [227, 237], and *Daphnia* exhibit short, jerky hopping movements in water [238]. The proposed MSBOTS exhibits consistent overall tracking accuracy performance with the tested videos for these two organism types, whereas the tracking accuracy measured by the MOTP and MOTA values varies among these videos.

6.3.1 Overall tracking performance for *Artemia* videos

As evidenced by the MOTP values for the *Artemia* videos in Figure 6.3, the proposed method exhibits a 47.68 pixel smaller standard deviation than the idTracker system, which illustrates the reliability of the proposed system for *Artemia microscopic* videos and improved tracking accuracy. Although idTracker produces smaller organism position estimation errors (MOTP), as shown by sequences 3-5 in Figure 6.3, the mean MOTA value is 7.07% and 6.44% lower than the values for the proposed method and LoliTrack, respectively. These values indicate a potential organism detection problem similar to that for zebrafish larvae detection (e.g., detection of the organism as background and of impurities as the organism due to their small size differences and similar movement characteristics, especially when the organism stops moving or water impurities are stirred up by organism movement, which causes further identity confusion due to detection errors).

6.3.2 Overall tracking performance for *Daphnia* videos

Figure 6.4 shows the tracking accuracy results measured by MOTP and MOTA for the tested *Daphnia magna* microscopic videos. The overall tracking accuracy shows that the proposed MSBOTS and idTracker consistently outperform the other tracking systems compared. The mean MOTA value for the proposed MSBOTS is 5.48% higher than that of idTracker. Similar to the proposed MSBOTS, SimpleTracker employs the Kuhn-Munkres tracking algorithm for the initial association, whereas the

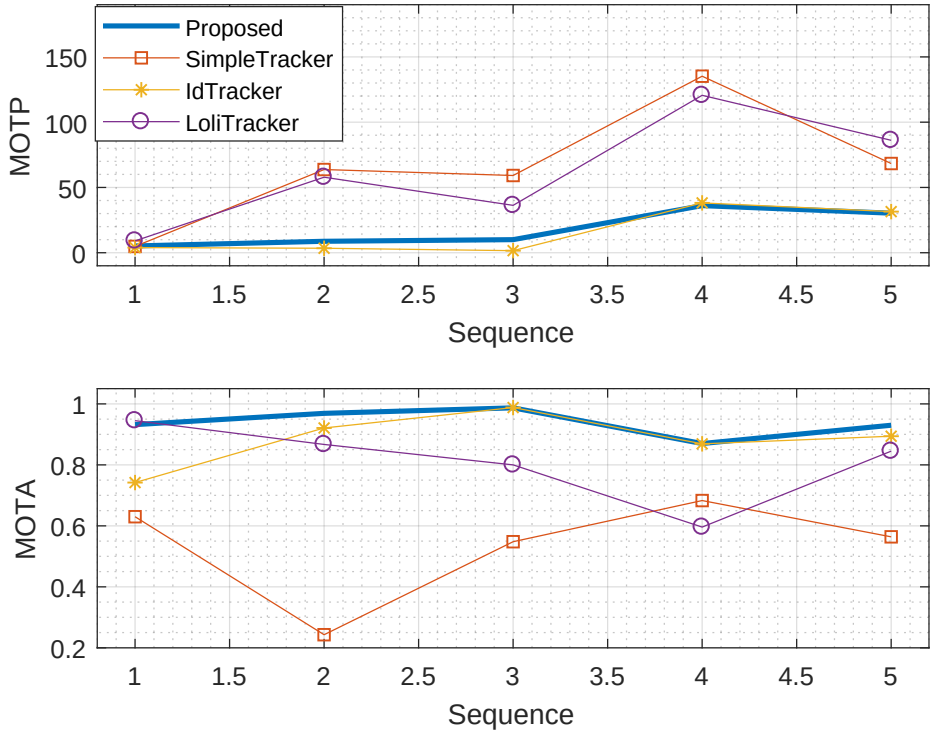


Figure 6.4: Tracking results evaluation among the comparison methods using *Daphnia magna* microscopic videos.

association method differs in the following tracking trajectory fragments bridging step. SimpleTracker uses the nearest neighbourhood algorithm, and the proposed MSBOTS combines the position calculation for misdetected and overlapping *Daphnia* with the unmapped points from the initial assignment for trajectory bridging. The 40.4% higher MOTA value for the proposed MSBOTS compared with SimpleTracker illustrates the improved individual identity maintenance due to the gap bridging method after the initial association by the proposed MSBOTS. This tracking reliability and accuracy of the proposed MSBOTS thus illustrates its applicability to *Daphnia*, which feature short, jerky hopping movements, compared with existing tracking systems.

6.4 Summary

Accurate automatic tracking of multiple small biological organisms provides an efficient approach for many biomedical and ecotoxicity applications. Thus, this Chapter presents evaluations of the proposed multiple zebrafish larvae tracking system presented in Section 5.3 to generalise and track other small organisms. Two types of small organisms are evaluated, *Artemia franciscana* and *Daphnia magna*, with movement characteristics different to zebrafish.

The MSBOTS background subtraction approach enables the removal of stationary backgrounds, such as the organism container and labels drawn on the containers; hence, in contrast to existing techniques, the MSBOTS system can process videos under practical experimental imaging conditions. The Kuhn-Munkres algorithm, which guarantees one-to-one association, is then used for object association or mapping between successive video frames to generate individual tracking trajectories based on the (non-ideal) segmentation results. The positions of misdetected and overlapped objects are calculated based on knowledge of their neighbours' locations within the individual tracking trajectories. The proposed system was evaluated based on tracking accuracy and outperformed existing multiple organism tracking systems.

With a total of 10 microscopic videos of five videos for each organism tested, results show that the tracking accuracy of the proposed system outperforms those of three existing tracking systems. Thus, the proposed system, denoted as MSBOTS in this Chapter, can be applied to track different organisms and is robust against non-ideal object detection results obtained using existing video segmentation techniques applied to microscopic time-lapse videos taken under practical laboratory experimental conditions.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

The analysis of dynamic behaviour of cells and small biological organisms provides essential information for cell studies, neurobiology, pharmacology and behavioural sciences and further understanding biological processes. The traditional manual observation of the locomotive movement of these biological objects using microscopic videos are time-consuming, tedious and subjective to human errors. In addition, the visual inspection in current days is nearly impossible due to the rapidly increasing experimental data obtained. Accurate automatic tracking presented in this work provided an efficient approach for analysing the dynamic behaviour of multiple cells and small biological organisms, followed by quantitative estimation of the kinematic living behaviour characteristics. However, there is a limitation of the proposed work and that is the empirical selection of the relative weight ratio between the two Gaussian components (as detailed in Section 4.2.5).

The list below summarises the main technical contributions of this work within the associated Chapters:

- Chapter 3

  - Proposed a scalable crowdsourced approach to generate video segmentation datasets annotated with ground-truth (published in the conference paper [50]).
  - Presented an annotated zebrafish larvae video segmentation dataset.
  - Proposed a two-pass verification process to evaluate the manually generated segmentation ground-truth.

- Chapter 4
  - Proposed a shifted bi-Gaussian mixture model to enhance the low intensity contrast of cell micrographs (published in the conference papers [51, 52]).

- Chapter 5
  - Exploited the intrinsic advantages of dense SIFT flow for single zebrafish larvae tracking.
  - Developed an accurate, reliable and efficient automatic multiple small biological organism tracking system (published in the conference paper [55] and presented in the submitted manuscript [54]).
  - Developed a segmentation method to address the unresolved misdetection and occlusion problem (published in the journal paper [53]).
  - Improved association method robust against non-ideal detection and segmentation (published in the conference paper [55]).

The lack of a dataset with ground-truth segmentation and suitable evaluation metrics is a common obstacle, particularly for biomedical image processing research. This thesis developed and proposed an efficient and scalable crowdsourced approach to generate video segmentation ground-truth to facilitate database generation for general biological organism segmentation and tracking algorithm evaluation. To illustrate the proposed approach, the methodology was applied to generate an annotated

zebrafish larvae video segmentation dataset comprised of 10 video sequences. To further facilitate the evaluation of segmentation algorithms using such an annotated dataset, a set of segmentation evaluation metrics was also presented. Finally, tracking datasets for three types of organisms annotated with centroid positions for each organism in 20 microscopic videos were presented for the evaluation of overall tracking performance in this work. Both the datasets generated and Matlab code for the evaluation metrics have been made freely available online to the research community.

As an example biological organism and given cell time-lapse microscopic videos, this work investigated a novel image pre-processing technique to enhance the low greyscale image intensity contrast for improved cell image segmentation accuracy. A novel image pre-processing technique for salient region object detection was proposed to match an adaptive, shifted bi-Gaussian mixture model to the intensity histogram of the original (stem) cell image frame to obtain greater contrast differentiation between the target cell and microscopic video background and noise (while maintaining the original intensity histogram shape) and further improve segmentation accuracy. Rather than using a model with fixed parameters across an entire video sequence, this research proposed the adaptive derivation of the mixture model parameters to match the intensity histogram for each video frame to adaptively address changes in the video background. Experiments conducted on a stem cell time-lapse video dataset indicated up to 37% improved segmentation accuracy, and in some frames (partially or completely), cells that the manual ground-truth and/or existing segmentation approaches failed to identify were detected.

The thesis then extended to segment and track more complex organisms: single and multiple zebrafish larvae. Zebrafish larvae misdetection and occlusion are inevitable problems when detecting and segmenting these small biological organisms from time-lapse microscopic videos. The level of difficulty of detection and segmentation increases compared with general objects when tracking small biological aquatic organisms, and these problems affect the subsequent zebrafish larvae tracking

processes.

This work first exploited the intrinsic advantages of dense SIFT flow (due to its potential for detailed kinematic analysis of individual zebrafish larval movements such as trunk curvature and tail-beating frequency, etc.) combined with position updates to the previous frame and a morphological operation to develop a single zebrafish larva tracking system. The experimental results demonstrated the improved tracking accuracy of the proposed method compared to the intensity thresholding-based single zebrafish larva tracking system LSRtrack. Furthermore, experiments showed that downsampling of the frame resolution reduced the dense SIFT flow computational complexity without influencing tracking accuracy.

Due to the increased calculation complexity of the dense SIFT flow feature points, this work further explored more efficient multiple zebrafish larvae tracking methods. The development of multiple zebrafish larvae tracking systems can be divided into two parts: zebrafish larvae segmentation from the microscopic video background and the association of the detected zebrafish larvae between successive video frames over time.

In this work, a GMM model based on background subtraction was investigated for improving the segmentation accuracy; both the number of Gaussian components and component parameters adapted to changes in the video background. After GMM background subtraction, however, the resultant image is typically distorted by a scattering of small noise fragments. In addition, the imaging environment is usually non-ideal; e.g., for zebrafish larvae, small particles or impurities in the water are stirred up by the strong tail beating movements of the swimming larvae. Thus, a segmentation approach addressing these image degradations was proposed using a median filter and morphological erosion techniques to remove noise fragments and distortion in the resultant segmentation image.

To further improve the tracking accuracy for non-ideal zebrafish larvae detection

and segmentation results, this thesis presented the proposed multiple zebrafish larvae tracking system, which denoted as Multiple Small Biological Organism Tracking System (MSBOTS) after being generalised. MSBOTS combined the multiple object association algorithm for linking detected objects frame-by-frame and tracking trajectory adjustment techniques. To address segmentation errors due to misdetected or occluded zebrafish larvae, the proposed MSBOTS estimated the positions of zebrafish larvae in interim frames using corresponding points in neighbouring frames. Finally, the calculated points were applied to connect and adjust the tracking trajectory fragments from the initial association by an extended Kuhn-Munkres algorithm. Although the state-of-the-art idTracker system can reasonably maintain object identity, the identity maintenance is highly reliant on the object tracking accuracy. In the proposed MSBOTS tracking approach, to differentiate the individual trajectory/tracks of independent objects when detecting and analysing the movement behaviour of overlapping biological organisms, the tracking trajectory created by the extended Kuhn-Munkres algorithm (which obtains a global optimal assignment by minimising the sum of the Euclidean distance between frame pairs) and the following trajectory adjustment process generated less identity-swapping events than idTracker. The proposed MSBOTS was evaluated on three different types of small organisms with different movement characteristics using a total of 20 videos (10 zebrafish larvae microscopic videos and 10 microscopic videos of *Artemia* and *Daphnia* for generalised application of the proposed MSBOTS).

The proposed system exhibits a decrease in the overall Multiple Object Tracking Precision (MOTP) error of up to 44.48 pixels compared to the commercial LoliTrack system, and an increase in Multiple Object Tracking Accuracy (MOTA) of up to 63.01% compared with the state-of-the-art idTracker system. In addition, the proposed system decreases the standard deviation by up to 47.68 pixels compared with idTracker. The individual tracking trajectory results offer an additional advantage for locomotive characteristic analysis and calculations of velocity, acceleration and

movement direction. Moreover, MSBOTS provided locomotive characteristic analysis capability using the generated individual tracking trajectories to facilitate small organism behavioural analysis research. Example movement parameters of velocity, acceleration and direction were calculated using the proposed approach for the three organisms tested. Behavioural rules and the effects of new medicine or chemicals on the dynamic behaviour of organisms can be studied in future using the provided or extending the movement analysis to facilitate specific experiments.

7.2 Future Work

7.2.1 Applying the proposed MSBOTS in behavioural ecotoxicology

Reports of behavioural responses as highly sensitive, 'early warning' indicators of toxicant stress on aquatic ecosystems have increased rapidly in the literature [91]. Small living organisms such as zebrafish larvae (*Danio rerio*), *Artemia franciscana*, and *Daphnia magna* are widely used in behavioural ecotoxicology research [91, 155]. However, some dynamic behaviour parameters are still obtained manually, such as counting the number of dead organisms based on the absence of a physical response in some behavioural bioassays [155]. In addition, the systematic study of behavioural effects and comparisons among different effect modes is hampered by the limited number of chemicals examined, the exposure conditions, or locomotive parameters [91].

Applying the proposed MSBOTS in behavioural biotests can increase efficiency and reduce the required manual labour and time by automatically and accurately detecting and tracking the locomotive movement of small living organisms using microscopic video recordings. Commonly used behaviour parameters such as movement

speed (velocity), acceleration, deceleration and movement directions are generated after obtaining the individual tracking trajectories using the proposed MSBOTS.

Thus, researchers in ecotoxicology can focus on designing biotest procedures and comparing behavioural responses to different chemicals, toxicants and exposure conditions. The previous entirely manual and labour-intensive process of movement tracking and behavioural parameter calculation can now be conducted automatically by using the proposed MSBOTS.

7.2.2 Examination of detailed organism body movement

The centre of a biological organism is usually utilised as the centroid of the body mass in existing zebrafish tracking systems [5, 31, 75]. This work also estimated organism centroid positions to indicate the studied organisms' locations. However, the centroid point alone cannot provide information on detailed postures, such as wing positions (for *Drosophila*), body curvature (for zebrafish), and tail beating frequency. Detailed kinematic posture information can provide biologists and ecologists with highly quantitative estimated behaviour data enabling the exploration of general mechanisms and principles of individual behaviour and between individual organisms as correlated from the acquired microscopic video data [36, 37, 239].

This research investigated the application of dense SIFT flow to single zebrafish larva tracking in Section 5.2 due to the algorithm's potential to permit the estimation of detailed zebrafish postures by extracting an entire moving object from the background when there is displacement of the object in successive video frames. However, the SIFT flow algorithm does not apply to multiple zebrafish larvae tracking due to computational complexity requirements. Future work could investigate the estimation and analysis after obtaining the locomotive parameters from automatically tracking the small biological organisms using the proposed MSBOTS.

The gradient magnitude technique [135] could be an option to calculate the organism body outline based on the segmentation results using the proposed MSBOTS. The calculated body outlines can be used as the basis of detailed kinematic evaluation in an organism video frame from the proposed datasets.

7.2.3 Applying machine learning techniques to organism tracking

Studies of machine learning (and deep learning, which requires a larger dataset than machine learning) in artificial intelligence have expanded greatly in the last decade, driving advances in many different fields. Machine learning has enabled breakthrough outcomes in computer vision tasks such as automatic object detection in images (an important component of driver-assisted and self-driving cars), object recognition using images, and knowledge discovery in biomedical image research (to improve the understanding of complex diseases) [240, 241].

Future work could investigate the application of machine learning techniques [242], such as TensorFlow [241], Support Vector Machine (SVM) [243], neural networks [32] and decision trees [244, 245], to the detection and tracking of small biological organisms (e.g., zebrafish larvae, *Artemia franciscana*, and *Daphnia magna*) from microscopic videos. Machine learning has been reported to achieve better and more reliable outcomes than computer vision in object recognition and classification [240, 246].

References

[1] A. Sacan, H. Ferhatosmanoglu, and H. Coskun. “CellTrack: an open-source software for cell tracking and motility analysis”. In: *Bioinformatics* 24.14 (2008), pp. 1647–1649.

[2] M. Möller, M. Burger, P. Dieterich, et al. “A framework for automated cell tracking in phase contrast microscopic videos based on normal velocities”. In: *Journal of Visual Communication and Image Representation* 25.2 (2014), pp. 396–409.

[3] R. Gerlai, M. Lahav, S. Guo, et al. “Drinks like a fish: zebra fish (Danio rerio) as a behavior genetic model to study alcohol effects”. In: *Pharmacology Biochemistry and Behavior* 67.4 (2000), pp. 773–782.

[4] S.R. Kirchner, M. Fedoruk, T. Lohmüller, et al. “Analyzing the movement of the nauplius’ artemia salina’by optical tracking of plasmonic nanoparticles”. In: *Journal of Visualized Experiments: JoVE* 89 (2014).

[5] Y.Z. Zhou, R.T. Cattley, C.L. Cario, et al. “Quantification of larval zebrafish motor function in multi-well plates using open-source MATLAB® applications”. In: *Nature Protocols* 9.7 (2014), pp. 15–33.

[6] P.R. Martineau and P. Mourrain. “Tracking zebrafish larvae in group–Status and perspectives”. In: *Methods* 62.3 (2013), pp. 292–303.

[7] A.V. Kalueff, M. Gebhardt, A.M. Stewart, et al. “Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond”. In: *Zebrafish* 10.1 (2013), pp. 70–86.

[8] M.B. Orger, E. Gahtan, A. Muto, et al. “Behavioral screening assays in zebrafish”. In: *Methods in Cell Biology*. Vol. 77. Elsevier, 2004, pp. 53–68.

[9] R.S. Zou and C. Tomasi. “Deformable graph model for tracking epithelial cell sheets in fluorescence microscopy”. In: *IEEE Transactions on Medical Imaging* 35.7 (2016), pp. 1625–1635.

[10] C. Noss, A. Lorke, and E. Niehaus. “Three-dimensional tracking of multiple aquatic organisms with a two camera system”. In: *Limnology and Oceanography: Methods* 11.3 (2013), pp. 139–150.

[11] M.A.A. Dewan, M.O. Ahmad, and M.N.S. Swamy. “Tracking biological cells in time-lapse microscopy: an adaptive technique combining motion and topological features”. In: *IEEE Transactions on Biomedical Engineering* 58.6 (2011), pp. 1637–1647.

[12] N.K. Satija, V.K. Singh, Y.K. Verma, et al. “Mesenchymal stem cell-based therapy: a new paradigm in regenerative medicine”. In: *Journal of Cellular and Molecular Medicine* 13.11-12 (2009), pp. 4385–4402.

[13] T. Schroeder. “Imaging stem-cell-driven regeneration in mammals”. In: *Nature* 453.7193 (2008), pp. 345–351.

[14] J.V. Frangioni and R.J. Hajjar. “In vivo tracking of stem cells for clinical trials in cardiovascular disease”. In: *Circulation* 110.21 (2004), pp. 3378–3383.

[15] E. Puybareau, H. Talbot, N. Gaber, et al. “Morphological analysis of Brownian motion for physical measurements”. In: *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer. 2017, pp. 486–497.

[16] F. Cagnacci, S. Focardi, M. Heurich, et al. “Partial migration in roe deer: migratory and resident tactics are end points of a behavioural gradient determined by ecological factors”. In: *Oikos* 120.12 (2011), pp. 1790–1802.

[17] J. Brodersen, P.A. Nilsson, L.A. Hansson, et al. “Condition-dependent individual decision-making determines cyprinid partial migration”. In: *Ecology* 89.5 (2008), pp. 1195–1200.

[18] B.B. Chapman, K. Hulthén, D.R. Blomqvist, et al. "To boldly go: individual differences in boldness influence migratory tendency". In: *Ecology Letters* 14.9 (2011), pp. 871–876.

[19] J.P. Marechal, C. Hellio, M. Sebire, et al. "Settlement behaviour of marine invertebrate larvae measured by EthoVision 3.0". In: *Biofouling* 20.4-5 (2004), pp. 211–217.

[20] M.T. Ekvall, G. Bianco, S. Linse, et al. "Three-dimensional tracking of small aquatic organisms using fluorescent nanoparticles". In: *PloS One* 8.11 (2013), e78498.

[21] G. Dur, S. Souissi, F. Schmitt, et al. "Effects of animal density, volume, and the use of 2D/3D recording on behavioral studies of copepods". In: *Hydrobiologia* 666.1 (2011), pp. 197–214.

[22] D. Martin, C. Fowlkes, D. Tal, et al. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics". In: *Proceedings of Computer Vision ICCV 2001. Eighth IEEE International Conference on*. Vol. 2. IEEE. 2001, pp. 416–423.

[23] F. Galasso, N.N. Shankar, C. T. Jimenez, et al. "A unified video segmentation benchmark: Annotation, metrics and analysis". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013, pp. 3527–3534.

[24] B. Thomas and M. Jitendra. "Object segmentation by long term analysis of point trajectories". In: *Computer Vision–ECCV 2010* (2010), pp. 282–295.

[25] P. Sundberg, T. Brox, M. Maire, et al. "Occlusion boundary detection and figure/ground assignment from optical flow". In: *Proceedings of Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE. 2011, pp. 2233–2240.

[26] J.K. Udupa, V.R. LeBlanc, Y. Zhuge, et al. "A framework for evaluating image segmentation algorithms". In: *Computerized Medical Imaging and Graphics* 30.2 (2006), pp. 75–87.

[27] A. Yilmaz, O. Javed, and M. Shah. “Object tracking: A survey”. In: *Acm Computing Surveys (CSUR)* 38.4 (2006), pp. 1–45.

[28] T. Kanade, Z.Z. Yin, R. Bise, et al. “Cell image analysis: Algorithms, system and applications”. In: *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*. IEEE. 2011, pp. 374–381.

[29] O. Dzyubachyk, W.A. van Cappellen, J. Essers, et al. “Advanced level-set-based cell tracking in time-lapse fluorescence microscopy”. In: *IEEE Transactions on Medical Imaging* 29.3 (2010), pp. 852–867.

[30] P. Getreuer. “Chan-vese segmentation”. In: *Image Processing On Line* 2 (2012), pp. 214–224.

[31] A. Pérez-Escudero, J. Vicente-Page, R.C. Hinz, et al. “idTracker: tracking individuals in a group by automatic identification of unmarked animals”. In: *Nature Methods* 11.7 (2014), pp. 743–748.

[32] Z.P. Xu and X.E. Cheng. “Zebrafish tracking using convolutional neural networks”. In: *Scientific Reports* 7 (2017).

[33] E.E. Conklin, K.L. Lee, S.A. Schlabach, et al. “VideoHacking: automated tracking and quantification of locomotor behavior with open source software and off-the-Shelf video equipment”. In: *Journal of Undergraduate Neuroscience Education* 13.3 (2015), A120.

[34] L. Závorka, B. Koeck, J. Cucherousset, et al. “Co-existence with non-native brook trout breaks down the integration of phenotypic traits in brown trout parr”. In: *Functional Ecology* 31.8 (2017), pp. 1582–1591. ISSN: 1365-2435.

[35] J. Rihel, D.A. Prober, A. Arvanites, et al. “Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation”. In: *Science* 327.5963 (2010), pp. 348–351.

[36] A.I. Dell, J.A. Bender, K. Branson, et al. “Automated image-based tracking and its application in ecology”. In: *Trends in Ecology & Evolution* 29.7 (2014), pp. 417–428.

[37] E.A. Dam, J.E. Harst, C.J. Braak, et al. “An automated system for the recognition of various specific rat behaviours”. In: *Journal of Neuroscience Methods* 218.2 (2013), pp. 214–224.

[38] H. Dankert, L. Wang, E.D. Hoopfer, et al. “Automated monitoring and analysis of social behavior in Drosophila”. In: *Nature Methods* 6.4 (2009), pp. 297 –303.

[39] B.L. Xu and M.L. Lu. “An ant-based stochastic searching behavior parameter estimate algorithm for multiple cells tracking”. In: *Engineering Applications of Artificial Intelligence* 30 (2014), pp. 155–167.

[40] K.E. Magnusson and J. Jaldén. “A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages”. In: *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*. IEEE. 2012, pp. 382–385.

[41] M. Kabra, A.A. Robie, M. Rivera-Alba, et al. “JAABA: interactive machine learning for automatic annotation of animal behavior”. In: *Nature Methods* 10.1 (2013), pp. 64 –67.

[42] W.J. Godinez, M. Lampe, S. Wörz, et al. “Deterministic and probabilistic approaches for tracking virus particles in time-lapse fluorescence microscopy image sequences”. In: *Medical Image Analysis* 13.2 (2009), pp. 325–342.

[43] S.H. Rezatofighi, S. Gould, B.N. Vo, et al. “A multiple model probability hypothesis density tracker for time-lapse cell microscopy sequences”. In: *International Conference on Information Processing in Medical Imaging*. Springer. 2013, pp. 110–122.

[44] I. Smal, K. Draegestein, N. Galjart, et al. “Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis”. In: *IEEE Transactions on Medical Imaging* 27.6 (2008), pp. 789–804.

[45] N. Chenouard, I. Bloch, and J.C. Olivo-Marin. “Multiple hypothesis tracking for cluttered biological image sequences”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.11 (2013), pp. 2736–3750.

[46] L. Liang, H.Y. Shen, C.P. Camilli, et al. “A novel multiple hypothesis based particle tracking method for clathrin mediated endocytosis analysis using fluorescence microscopy”. In: *IEEE Transactions on Image Processing* 23.4 (2014), pp. 1844–1857.

[47] Karl K. Rohr, William J W.J. Godinez, Nathalie N. Harder, et al. “Tracking and quantitative analysis of dynamic movements of cells and particles”. In: *Cold Spring Harbor Protocols* 2010.6 (2010), pdb.top80.

[48] S.M. Rezatofighi, S. Gould, B.T. Vo, et al. “Multi-target tracking with time-varying clutter rate and detection profile: application to time-lapse cell microscopy sequences”. In: *IEEE Transactions on Medical Imaging* 34.6 (2015), pp. 1336–1348.

[49] k. Branson. “Distinguishing seemingly indistinguishable animals with computer vision”. In: *Nature Methods* 11.7 (2014), pp. 721 –722.

[50] X.Y. Wang, E. Cheng, I.S. Burnett, et al. “Crowdsourced generation of annotated video datasets: a zebrafish larvae dataset for video segmentation and tracking evaluation”. In: *Proceedings of Life Sciences (IEEE LSC 2017), IEEE International Conference on*. IEEE. 2017, pp. 87–90.

[51] X.Y. Wang, E. Cheng, and I.S. Burnett. “Improved cell segmentation with adaptive bi-Gaussian mixture models for image contrast enhancement preprocessing”. In: *Proceedings of Life Sciences (IEEE LSC 2017), IEEE International Conference on*. IEEE. 2017, pp. 274–277.

[52] X.Y. Wang, E. Cheng, and I.S. Burnett. “Improved (STEM) cell segmentation with histogram matching image contrast enhancement”. In: *Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on*. IEEE. 2015, pp. 816–820.

[53] X.Y. Wang, E. Cheng, I.S. Burnett, et al. “Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions”. In: *Scientific Reports* 7.8 (2017), 17596(1–8).

[54] X.Y. Wang, E. Cheng, I.S. Burnett, et al. “MSBOTS: a Multiple Small Biological Organism Tracking System robust against non-ideal detection and segmentation”. In: *Submitted to Nature Methods [under review]* (2018).

[55] X.Y. Wang, E. Cheng, I.S. Burnett, et al. “Automatic tracking of multiple zebrafish larvae with resilience against segmentation errors”. In: *Proceedings of Biomedical Imaging (ISBI 2018), IEEE International Symposium [to be presented in April 4-7, 2018, Washington D.C., U.S.]* IEEE. 2018.

[56] R. Mikut, T. Dickmeis, W. Driever, et al. “Automated processing of zebrafish imaging data: a survey”. In: *Zebrafish* 10.3 (2013), pp. 401–421.

[57] Y.S. Huang, A. Nigam, O. Campana, et al. “Miniaturized video-microscopy system for near real-time water quality biomonitoring using microfluidic chip-based devices”. In: *SPIE BioPhotonics Australasia*. Vol. 10013. International Society for Optics and Photonics. 2016, 100131R.

[58] T.W. Nattkemper. “Automatic segmentation of digital micrographs: a survey”. In: *Medinfo*. 2004, pp. 847–851.

[59] T.W. Nattkemper, T. Twellmann, H. Ritter, et al. “Human vs. machine: evaluation of fluorescence micrographs”. In: *Computers in Biology and Medicine* 33.1 (2003), pp. 31–43.

[60] B.K. Rath and J. Frank. “Fast automatic particle picking from cryo-electron micrographs using a locally normalized cross-correlation function: a case study”. In: *Journal of Structural Biology* 145.1 (2004), pp. 84–90.

[61] M.U. Maška, D. Svoboda, P. Matula, et al. “A benchmark for comparison of cell tracking algorithms”. In: *Bioinformatics* 30.11 (2014), pp. 1609–1617.

[62] M. Riffle and T.N. Davis. “The yeast resource center public image repository: a large database of fluorescence microscopy images”. In: *BMC Bioinformatics* 11.1 (2010), pp. 263–272.

[63] T.W. Nattkemper, H. Wersing, W. Schubert, et al. “A neural network architecture for automatic segmentation of fluorescence micrographs”. In: *Neurocomputing* 48.1 (2002), pp. 357–367.

[64] G.P. Gupta and J. Massagué. “Cancer metastasis: building a framework”. In: *Cell* 127.4 (2006), pp. 679–695.

[65] P. Friedl and K. Wolf. “Tumour-cell invasion and migration: diversity and escape mechanisms”. In: *Nature Reviews Cancer* 3.5 (2003), pp. 362–374.

[66] W.B. Barbazuk, I. Korf, C. Kadavi, et al. “The syntenic relationship of the zebrafish and human genomes”. In: *Genome Research* 10.9 (2000), pp. 1351–1358.

[67] R. Dahm and R. Geisler. “Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species”. In: *Marine Biotechnology* 8.4 (2006), pp. 329–345.

[68] M.B. Orger and G.G. de Polavieja. “Zebrafish behavior: opportunities and challenges”. In: *Annual Review of Neuroscience* 0 (2017).

[69] A.K. Canger, M. Passini, W.S. Asch, et al. “Restricted expression of the neuronal intermediate filament protein plasticin during zebrafish development”. In: *Journal of Comparative Neurology* 399.4 (1998), pp. 561–572.

[70] D. Leake, W.S. Asch, A.K. Canger, et al. “Gefiltin in zebrafish embryos: sequential gene expression of two neurofilament proteins in retinal ganglion cells”. In: *Differentiation* 65.4 (1999), pp. 181–189.

[71] S. Higashijima, H. Okamoto, N. Ueno, et al. “High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin”. In: *Developmental Biology* 192.2 (1997), pp. 289–299.

[72] S. Isogai, M. Horiguchi, and B.M. Weinstein. “The vascular anatomy of the developing zebrafish: an atlas of embryonic and early larval development”. In: *Developmental Biology* 230.2 (2001), pp. 278–301.

[73] W. Katherine and S.E. Fraser. “Order and coherence in the fate map of the zebrafish nervous system”. In: *Development* 121.8 (1995), pp. 2595–2609.

[74] C.A. Dlugos and R.A. Rabin. “Ethanol effects on three strains of zebrafish: model system for genetic investigations”. In: *Pharmacology Biochemistry and Behavior* 74.2 (2003), pp. 471–480.

[75] R. Creton. “Automated analysis of behavior in zebrafish larvae”. In: *Behavioural Brain Research* 203.1 (2009), pp. 127–136.

[76] A. Kurta and B.G. Palestis. “Effects of ethanol on the shoaling behavior of zebrafish (Danio rerio)”. In: *Dose Response* 8.4 (2010), pp. 527–533.

[77] R.S. Ryback. “The use of fish, especially goldfish, in alcohol research.” In: *Quarterly Journal of Studies on Alcohol* 31.1 (1970), pp. 162–166.

[78] K. Howe, M.D. Clark, C.F. Torroja, et al. “The zebrafish reference genome sequence and its relationship to the human genome”. In: *Nature* 496.7446 (2013), pp. 498–503.

[79] S. Guo. “Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish?” In: *Genes, Brain and Behavior* 3.2 (2004), pp. 63–74.

[80] D.A. Prober, J. Rihel, A.A. Onah, et al. “Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish”. In: *Journal of Neuroscience* 26.51 (2006), pp. 13400–13410.

[81] P. Panula. “Hypocretin/orexin in fish physiology with emphasis on zebrafish”. In: *Acta Physiologica* 198.3 (2010), pp. 381–386.

[82] I.V. Zhdanova, S.Y. Wang, O.U. Leclair, et al. “Melatonin promotes sleep-like state in zebrafish”. In: *Brain Research* 903.1 (2001), pp. 263–268.

[83] T.C. Farrell, C.L. Cario, C. Milanese, et al. “Evaluation of spontaneous propulsive movement as a screening tool to detect rescue of Parkinsonism phenotypes in zebrafish models”. In: *Neurobiology of Disease* 44.1 (2011), pp. 9–18.

[84] H.A. Burgess and M. Granato. “Modulation of locomotor activity in larval zebrafish during light adaptation”. In: *Journal of Experimental Biology* 210.14 (2007), pp. 2526–2539.

[85] T.C. Amiard, J.C. Amiard, and P.S. Rainbow. *Ecological biomarkers: indicators of ecotoxicological effects*. CRC Press, 2012.

[86] M.T. Rashid, M. Frasca, A.A. Ali, et al. “Artemia swarm dynamics and path tracking”. In: *Nonlinear Dynamics* 68.4 (2012), pp. 555–563.

[87] E. Klavins and R.M. Murray. “Distributed algorithms for cooperative control”. In: *IEEE Pervasive Computing* 3.1 (2004), pp. 56–65.

[88] J. Hellou. “Behavioural ecotoxicology, an “early warning” signal to assess environmental quality”. In: *Environmental Science and Pollution Research* 18.1 (2011), pp. 1–11.

[89] E. Jonczyk and G. Gilron. “Acute and chronic toxicity testing with daphnia sp.” In: *Small-scale Freshwater Toxicity Investigations*. Springer, 2005, pp. 337–393.

[90] G.M. Gajardo and J.A. Beardmore. “The brine shrimp artemia: adapted to critical life conditions”. In: *Frontiers in Physiology* 3 (2012), p. 185.

[91] J. Chevalier, E. Harscoët, M. Keller, et al. “Exploration of Daphnia behavioral effect profiles induced by a broad range of toxicants with different modes of action”. In: *Environmental Toxicology and Chemistry* 34.8 (2015), pp. 1760–1769.

[92] E.J. Sutton, T.D. Henning, B.J. Pichler, et al. “Cell tracking with optical imaging”. In: *European Radiology* 18.10 (2008), pp. 2021–2032.

[93] M. Lard, J. Backman, M. Yakovleva, et al. “Tracking the small with the smallest: using nanotechnology in tracking zooplankton”. In: *PloS One* 5.10 (2010), e13516.

[94] M. Chalfie. “Green fluorescent protein”. In: *Photochemistry and Photobiology* 62.4 (1995), pp. 651–656.

[95] D.A. Persons, J.A. Allay, J.M. Riberdy, et al. “Use of the green fluorescent protein as a marker to identify and track genetically modified hematopoietic cells”. In: *Nature Medicine* 4.10 (1998), pp. 1201–1205.

[96] K. Hatta, H. Tsujii, and T. Omura. “Cell tracking using a photoconvertible fluorescent protein”. In: *Nature Protocols* 1.2 (2006), pp. 960–967.

[97] N.C. Shaner, P.A. Steinbach, and R.Y. Tsien. “A guide to choosing fluorescent proteins”. In: *Nature Methods* 2.12 (2005), pp. 905 –909.

[98] D.M. Chudakov, S. Lukyanov, and K.A. Lukyanov. “Fluorescent proteins as a toolkit for in vivo imaging”. In: *Trends in Biotechnology* 23.12 (2005), pp. 605–613.

[99] R. Ando, H. Hama, M. Yamamoto-Hino, et al. “An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein”. In: *Proceedings of the National Academy of Sciences* 99.20 (2002), pp. 12651–12656.

[100] V. Ntziachristos, C. Bremer, and R. Weissleder. “Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging”. In: *European Radiology* 13.1 (2003), pp. 195–208.

[101] N.C. Shaner, R.E. Campbell, P.A. Steinbach, et al. “Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein”. In: *Nature Biotechnology* 22.12 (2004), pp. 1567–1572.

[102] D.K. Chatterjee, A.J. Rufaihah, and Y. Zhang. “Upconversion fluorescence imaging of cells and small animals using lanthanide doped nanocrystals”. In: *Biomaterials* 29.7 (2008), pp. 937–943.

[103] J. Ma, L.N. Wu, Z. Hou, et al. “Visualizing the endocytosis of phenylephrine in living cells by quantum dot-based tracking”. In: *Biomaterials* 35.25 (2014), pp. 7042–7049.

[104] S.W. Hell. “Toward fluorescence nanoscopy”. In: *Nature Biotechnology* 21.11 (2003), pp. 1347–1355.

[105] G.T.S. Jasmin, R.A. Louzada, P.H. Rosado de Castro, et al. “Tracking stem cells with superparamagnetic iron oxide nanoparticles: perspectives and considerations”. In: *International Journal of Nanomedicine* 12 (2017), pp. 779–793.

[106] “Cell segmentaion dataset”. In: 10th IEEE International Symposium on Biomedical Imaging (ISBI 2013). 2013. URL: http://www.celltrackingchallenge.net/datasets.html.

[107] D.P. Mukherjee, N. Ray, and S.T. Acton. “Level set analysis for leukocyte detection and tracking”. In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 562–572.

[108] G. Bishop and G. Welch. “An introduction to the kalman filter”. In: *Proc of SIGGRAPH, Course* 8.41 (2001), pp. 27599–23175.

[109] L.D. Krista and P.T. Gallagher. “Automated coronal hole detection using local intensity thresholding techniques”. In: *Solar Physics* 256.1-2 (2009), pp. 87–100.

[110] H.S. Wu, J. Berba, and J. Gil. “Iterative thresholding for segmentation of cells from noisy images”. In: *Journal of Microscopy* 197.3 (2000), pp. 296–304.

[111] N. Paragios and R. Deriche. “Geodesic active contours and level sets for the detection and tracking of moving objects”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.3 (2000), pp. 266–280.

[112] A. Blake and M. Isard. *Active contours: the application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion*. Springer Science & Business Media, 2012.

[113] G. Li, T.M. Liu, J. Nie, et al. “Segmentation of touching cell nuclei using gradient flow tracking”. In: *Journal of Microscopy* 231.1 (2008), pp. 47–58.

[114] A.A. Aly, S.B. Deris, and N. Zaki. “Intelligent algorithms for cell tracking and image segmentation”. In: *International Journal of Computer Science & Information Technology* 6.5 (2014), p. 21.

[115] K. Miura. “Tracking movement in cell biology”. In: *Microscopy Techniques*. Springer, 2005, pp. 267–295.

[116] G. Andrew, J.B. Carlin, H.S. Stern, et al. “Bayesian data analysis”. In: *Chapman and Hall* (1995).

[117] D. Anoraganingrum. “Cell segmentation with median filter and mathematical morphology operation”. In: *Image Analysis and Processing. International Conference on*. IEEE. 1999, pp. 1043–1046.

[118] L. Kang, E.D. Miller, M. Chen, et al. “Cell population tracking and lineage construction with spatiotemporal context”. In: *Medical Image Analysis* 12.5 (2008), pp. 546–566.

[119] X.W. Chen, X.B. Zhou, and S. Wong. “Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy”. In: *IEEE Transactions on Biomedical Engineering* 53.4 (2006), pp. 762–766.

[120] S. Leila, P. Maria, and K. Josef. “Automatic watershed segmentation of randomly textured color images”. In: *IEEE transactions on Image Processing* 6.11 (1997), pp. 1530–1544.

[121] K. Jiang, Q.M. Liao, S.Y. Dai, et al. “A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering”. In: *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics*. IEEE. 2003, pp. 2820–2825.

[122] J.S. Beis and D.G. Lowe. “Shape indexing using approximate nearest-neighbour search in high-dimensional spaces”. In: *CVPR*. IEEE. 1997, p. 1000.

[123] A.J. Hand, T. Sun, D.C. Barber, et al. “Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration”. In: *Journal of Microscopy* 234.1 (2009), pp. 62–79.

[124] R.C. Gonzalez and E. Richard. *Digital image processing (3rd ed.)* Harlow: Pearson/Prentice Hall, 2008.

[125] P.S. Hiremath, P. Bannigidad, and S. Geeta. “Automated identification and classification of white blood cells (leukocytes) in digital microscopic images”. In: *IJCA Special Issue on Recent Trends in Image Processing and Pattern Recognition (RTIPPR)* (2010), pp. 59–63.

[126] Y.C. Chang and C.M. Chang. “A simple histogram modification scheme for contrast enhancement”. In: *IEEE Transactions on Consumer Electronics* 56.2 (2010), pp. 737–742.

[127] R. Aedla, G.S. Dwarakish, and D.V. Reddy. “Satellite image contrast enhancement algorithm based on plateau histogram equalization”. In: *Region 10 Symposium, 2014 IEEE*. IEEE. 2014, pp. 213–218.

[128] W. Gonzalez and R.E. Woods. “Digital image processing using MATLAB”. In: *Third New Jersey: Prentice Hall* (2004).

[129] A.A. Salihah, M.Y. Mashor, and N.H. Harun. “Improving colour image segmentation on acute myelogenous leukaemia images using contrast enhancement techniques”. In: *Biomedical Engineering and Sciences, IEEE EMBS Conference on*. IEEE. 2010, pp. 246–251.

[130] A.A. Salihah, M.Y. Mashor, N.H. Harun, et al. “Colour image enhancement techniques for acute leukaemia blood cell morphological features”. In: *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*. IEEE. 2010, pp. 3677–3682.

[131] Z. Salleh, M.Y. Mashor, N.R. Noor, et al. “Colour contrast enhancement based on bright and dark stretching for ziehl-neelsen slide images”. In: *Intelligent Information Hiding and Multimedia Signal Processing.* Vol. 2. IEEE. 2007, pp. 205–208.

[132] W. Marie, L. John, W. Birgit, et al. “Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer”. In: *Radiotherapy and Oncology* 98.1 (2011), pp. 117–125.

[133] M.M. Fraz, S.A. Barman, P. Remagnino, et al. “An approach to localize the retinal blood vessels using bit planes and centerline detection”. In: *Computer Methods and Programs in Biomedicine* 108.2 (2012), pp. 600–616.

[134] W.M. Wells. “Efficient synthesis of Gaussian filters by cascaded uniform filters”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2 (1986), pp. 234–239.

[135] W.F. Xue, X.Q. Mou, and L. Zhang. “Decoupled marginal distribution of gradient magnitude and Laplacian of Gaussian for texture classification”. In: *CCF Chinese Conference on Computer Vision*. Springer. 2015, pp. 418–428.

[136] R.M. Colwill and R. Creton. “Locomotor behaviors in zebrafish (Danio rerio) larvae”. In: *Behavioural Processes* 86.2 (2011), pp. 222–229.

[137] J. Akagi, C.J. Hall, K.E. Crosier, et al. “Immobilization of zebrafish larvae on a chip-based device for environmental scanning electron microscopy (ESEM) imaging”. In: *SPIE Micro+ Nano Materials, Devices, and Applications*. Vol. 8923. 46. International Society for Optics and Photonics. 2013, pp. 1–6.

[138] J. Akagi, F. Zhu, C.J. Hall, et al. “Integrated chip-based physiometer for automated fish embryo toxicity biotests in pharmaceutical screening and ecotoxicology”. In: *Cytometry Part A* 85.6 (2014), pp. 537–547.

[139] K.N. Wu, D. Gauthier, and M.D. Levine. “Live cell image segmentation”. In: *IEEE Transactions on Biomedical Engineering* 42.1 (1995), pp. 1–12.

[140] A. Bhalerao, L. Pase, G.J. Lieschke, et al. “Local affine texture tracking for serial registration of zebrafish images”. In: *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*. IEEE. 2012, pp. 434–437.

[141] S. Bhattacharya, S. Gupta, and V.K. Subramanian. “Localized image enhancement.” In: *Communications (NCC), Twentieth National Conference on*. 2014, pp. 1–6.

[142] C.P. Loizou, M. Pantziaris, I. Seimenis, et al. “Brain MR image normalization in texture analysis of multiple sclerosis”. In: *Information Technology and*

Applications in Biomedicine (ITAB). 9th International Conference on. IEEE. 2009, pp. 1–5.

[143] J. Altmann. "Observational study of behavior: sampling methods". In: *Behaviour* 49.3 (1974), pp. 227–266.

[144] O. Mirat, J.R. Sternberg, K.E. Severi, et al. "ZebraZoom: an automated program for high-throughput behavioral analysis and categorization". In: *Frontiers in Neural Circuits* 7 (2013), pp. 1–12.

[145] W.J. Godinez and K. Rohr. "Tracking multiple particles in fluorescence time-lapse microscopy images via probabilistic data association". In: *IEEE Transactions on Medical Imaging* 34.2 (2015), pp. 415–432.

[146] E. Wang, R.F. Tungaraza, F. Rosalia, et al. "Multi-subject connectivity-based parcellation of the human IPL using Gaussian mixture models and hidden Markov random fields". In: *Biomedical Imaging (ISBI), IEEE 10th International Symposium on*. IEEE. 2013, pp. 520–523.

[147] Y.W. Liu, P. Ma, P.A. Cassidy, et al. "Statistical analysis of zebrafish locomotor behaviour by generalized linear mixed models". In: *Scientific Reports* 7 (2017).

[148] H. Cong, M.Z. Sun, D.Y. Zhou, et al. "Multi-target tracking of zebrafish based on particle filter". In: *Proceddings of CCC 2016, 35th Chinese Control Conference on*. IEEE. 2016, pp. 10308–10313.

[149] Y. Liu, P. Ma, P.A. Cassidy, et al. "Statistica analysis of zebrafish locomotor behaviour by generalized linear mixed models". In: *Scientific Reports* 7 (2017), pp. 29–37.

[150] M. Usami. "An ultra-small RFID chip:/spl mu/-chip". In: *Advanced System Integrated Circuits. Proceedings of 2004 IEEE Asia-Pacific Conference on*. IEEE. 2004, pp. 2–5.

[151] Y. Habibi, D.R. Sulistyaningrum, and B. Setiyono. "A new algorithm for small object tracking based on super-resolution technique". In: *AIP Conference Proceedings*. Vol. 1867. 1. AIP Publishing. 2017, pp. 20–24.

[152] S.C. Baraban, W.T. Dinday, P.A. Castro, et al. “A large-scale mutagenesis screen to identify seizure-resistant zebrafish”. In: *Epilepsia* 48.6 (2007), pp. 1151–1157.

[153] F. Emran, J. Rihel, and J.E. Dowling. “A behavioral assay to measure responsiveness of zebrafish to changes in light intensities”. In: *Journal of Visualized Experiments: JoVE* 20 (2008), pp. 1–6.

[154] C. Liu, J. Yuen, and A. Torralba. “Sift flow: dense correspondence across scenes and its applications”. In: *Dense Image Correspondences for Computer Vision*. Springer, 2016, pp. 15–49.

[155] Y.S. Huang, R. Cartlidge, M. Walpitagama, et al. “Unsuitable use of DMSO for assessing behavioral endpoints in aquatic model species”. In: *Science of the Total Environment* 615 (2018), pp. 107–114.

[156] H. Zhang, J.E. Fritts, and S.A. Goldman. “Image segmentation evaluation: a survey of unsupervised methods”. In: *Computer Vision and Image Understanding* 110.2 (2008), pp. 260–280.

[157] C. Held, R. Palmisano, L. Häberle, et al. “Comparison of parameter-adapted segmentation methods for fluorescence micrographs”. In: *Cytometry Part A* 79.11 (2011), pp. 933–945.

[158] F. Ge, S. Wang, and T.C. Liu. “New benchmark for image segmentation evaluation”. In: *Journal of Electronic Imaging* 16.3 (2007), pp. 1–16.

[159] X.Y. Jiang, C. Marti, C. Irniger, et al. “Distance measures for image segmentation evaluation”. In: *EURASIP Journal on Applied Signal Processing* 2006 (2006), pp. 209–218.

[160] K. McGuinness and N.E. O’connor. “A comparative evaluation of interactive segmentation algorithms”. In: *Pattern Recognition* 43.2 (2010), pp. 434–444.

[161] O. Peter, M. Jitendra, and B. Thomas. “Segmentation of moving objects by long term video analysis”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.6 (2014), pp. 1187–1200.

[162] G. Habeler, K. Natter, G.G. Thallinger, et al. “YPL. db: the yeast protein localization database”. In: *Nucleic Acids Research* 30.1 (2002), pp. 80–83.

[163] M. Kals, K. Natter, G.G. Thallinger, et al. “YPL. db2: the yeast protein localization database, version 2.0”. In: *Yeast* 22.3 (2005), pp. 213–218.

[164] M. Frenkel-Morgenstern, A.A. Cohen, N. Geva-Zatorsky, et al. “Dynamic proteomics: a database for dynamics and localizations of endogenous fluorescently-tagged proteins in living human cells”. In: *Nucleic Acids Research* 38.1 (2009), pp. 508–512.

[165] A.A. Cohen, N. Geva-Zatorsky, E. Eden, et al. “Dynamic proteomics of individual cancer cells in response to a drug”. In: *Science* 322.5907 (2008), pp. 1511–1516.

[166] D.H. Rapoport, T. Becker, A.M. Mamlouk, et al. “A novel validation algorithm allows for automated cell tracking and the extraction of biologically meaningful parameters”. In: *PloS One* 6.11 (2011), e27315.

[167] Microsoft direct developer center. *DirectX SDK*. 2009. URL: https://www.microsoft.com/en-us/download/details.aspx?id=23549 (visited on 12/12/2017).

[168] MicroBioTests Inc. *Artemia toxicity screening test for estuarine and marine waters*. URL: http://www.microbiotests.be/SOPs/Artoxkit%20M%20SOP%20-%20A5.pdf (visited on 12/15/2017).

[169] MicroBioTests Inc. *Daphtoxkit f magna: crustacean toxicity screening test for freshwater*. URL: http://www.microbiotests.be/SOPs/Daphtoxkit%20magna%20F%20SOP%20-%20A5.pdf (visited on 12/15/2017).

[170] J.A. Solis, Y.S. Huang, D. Wlodkovic, et al. “Microfluidic environment and tracking analysis for the observation of Artemia Franciscana”. In: *Proceedings of sthe 26th British Machine Vision Conference* (2015).

[171] H. Irshad, L. Montaser-Kouhsari, G. Waltz, et al. “Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology:

evaluating experts, automated methods, and the crowd." In: *Proceedings of Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*. NIH Public Access. 2015, pp. 294–305.

[172] M. Burlick, O. Koteoglou, L. Karydas, et al. "Leveraging crowdsourced data for creating temporal segmentation ground truths of subjective tasks". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2013, pp. 743–750.

[173] R. Adams and L. Bischof. "Seeded region growing". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.6 (1994), pp. 641–647.

[174] W.D. Wright. "Color science: concepts and methods, quantitative data and formulas". In: *Physics Bulletin* 18.10 (1967), p. 353.

[175] Y.Y. Boykov and M.P. Jolly. "Interactive graph cuts for optimal boundary & region segmentation of objects in ND images". In: *Proceedings of Computer Vision ICCV 2001. 8th IEEE International Conference on*. Vol. 1. IEEE. 2001, pp. 105–112.

[176] D.M. Greig, B.T. Porteous, and A.H. Seheult. "Exact maximum a posteriori estimation for binary images". In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1989), pp. 271–279.

[177] G. Friedland, K. Jantz, and R. Rojas. "Siox: Simple interactive object extraction in still images". In: *Multimedia, 7th IEEE International Symposium on*. IEEE. 2005, pp. 7–13.

[178] O.J. Morris, M.J. Lee, and A.G. Constantinides. "Graph theory for image analysis: an approach based on the shortest spanning tree". In: *IEE Proceedings (Communications, Radar and Signal Processing on)*. Vol. 133. 2. IET. 1986, pp. 146–152.

[179] N. Otsu. "A threshold selection method from gray-level histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66.

[180] Creative Commons S.A. Attribution. *Creative commons*. URL: https://creativecommons.org/licenses/by-sa/3.0/deed.en_GB..

[181] D.M. Powers. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation". In: (2011).

[182] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. "Performance of optical flow techniques". In: *International Journal of Computer Vision* 12.1 (1994), pp. 43–77.

[183] J. Cardinale, G. Paul, and I.F. Sbalzarini. "Discrete region competition for unknown numbers of connected regions". In: *IEEE Transactions on Image Processing* 21.8 (2012), pp. 3531–3545.

[184] A. Rizk, G. Paul, P. Incardona, et al. "Segmentation and quantification of subcellular structures in fluorescence microscopy images using Squassh". In: *Nature Protocols* 9.3 (2014), pp. 586–596.

[185] D.N. Coltuc, P. Bolon, and J.M. Chassery. "Exact histogram specification". In: *IEEE Transactions on Image Processing* 15.5 (2006), pp. 1143–1152.

[186] K. Gyorgy, F. Attila, and H. Andras. "Exponential contrast maximization of intensity images". In: *Image and Signal Processing and Analysis (ISPA), 7th International Symposium on*. IEEE. 2011, pp. 139–142.

[187] A. Mohsen, S. Hossein, G. Mohammad, et al. "Gaussian mixture model-based contrast enhancement". In: *IET image processing* 9.7 (2015), pp. 569–577.

[188] C.W. Chen, J. Luo, and K.J. Parker. "Image segmentation via adaptive K-mean clustering and knowledge-based morphological operations with biomedical applications". In: *IEEE Transactions on Image Processing* 7.12 (1998), pp. 1673–1683.

[189] A. Pnevmatikakis and L. Polymenakos. "Robust estimation of background for fixed cameras". In: *Proceddings of Computing CIC'06. 15th International Conference on*. IEEE. 2006, pp. 37–42.

[190] H. Papasaika-Hanusch. “Digital image processing using matlab”. In: *Institute of Geodesy and Photogrammetry, ETH Zurich* 63 (1967).

[191] K.E.G. Magnusson, J. Jaldén, P.M. Gilbert, et al. “Global linking of cell tracks using the Viterbi algorithm”. In: *IEEE Transactions on Medical Imaging* 34.4 (2015), pp. 911–929.

[192] E. Limpert and W.A. Stahel. “Problems with using the normal distribution–and ways to improve quality and efficiency of data analysis”. In: *PLoS One* 6.7 (2011), e21403.

[193] J.J. Moré. “The Levenberg-Marquardt algorithm: entation and theory”. In: *Numerical Analysis*. Springer, 1978, pp. 105–116.

[194] L. Koskinen, J. Astola, and Y. Neuvo. “Soft morphological filters”. In: *Proceedings of SPIE*. Vol. 1568. 1991, pp. 262–270.

[195] M.L. Comer and E.J. Delp. “Morphological operations for color image processing”. In: *Electronic Imaging (J.)* 8.3 (1999), pp. 279–289.

[196] A.P. Dempster, N.M. Laird, and D.B. Rubin. “Maximum likelihood from incomplete data via the EM algorithm”. In: *Journal of the Royal Statistical Society. Series B (methodological)* (1977), pp. 1–38.

[197] L.V. Maria, S.O. Gerardo, A.G. Elkington, et al. “Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm”. In: *Medical Image Analysis* 8.3 (2004), pp. 255–265.

[198] S. Belongie, C. Carson, H. Greenspan, et al. “Color- and texture-based image segmentation using EM and its application to content-based image retrieval”. In: *Computer Vision, 1998. Sixth International Conference on*. IEEE. 1998, pp. 675–682.

[199] Z. Zivkovic and F. Van Der Heijden. “Efficient adaptive density estimation per image pixel for the task of background subtraction”. In: *Pattern Recognition Letters* 27.7 (2006), pp. 773–780.

[200] C.S. Wallace. “Classification by minimum-message-length inference”. In: *International Conference on Computing and Information*. Springer. 1990, pp. 72–81.

[201] J. Oliver, Rohan R. Baxter, and C. Wallace. “Minimum message length segmentation”. In: *Research and Development in Knowledge Discovery and Data Mining* (1998), pp. 222–233.

[202] S. Chabrier, B. Emile, H. Laurent, et al. “Unsupervised evaluation of image segmentation application to multi-spectral images”. In: *Pattern Recognition. ICPR 2004. Proceedings of the 17th International Conference on*. Vol. 1. IEEE. 2004, pp. 576–579.

[203] Y.J. Zhang. “A survey on evaluation methods for image segmentation”. In: *Pattern Recognition* 29.8 (1996), pp. 1335–1346.

[204] C. Wyatt, E.M. Bartoszek, and E. Yaksi. “Methods for studying the zebrafish brain: past, present and future”. In: *European Journal of Neuroscience* 42.2 (2015), pp. 1746–1763. DOI: 10.1111/ejn.12932. URL: http://dx.doi.org/10.1111/ejn.12932.

[205] H. Bay, T. Tuytelaars, and G.L. Van. “Surf: speeded up robust features”. In: *Computer Vision ECCV* (2006), pp. 404–417.

[206] C. Yi. “Munkres’ assignment algorithm, modified for rectangular matrices”. Sept. 2011. URL: http://csclab.murraystate.edu/bob.pilgrim/445/munkres.html.

[207] D.M. Catarious, A.H. Baydush, and C.E. Floyd. “Characterization of difference of Gaussian filters in the detection of mammographic regions”. In: *Medical Physics* 33.11 (2006), pp. 4104–4114.

[208] C. Harris and M. Stephens. “A combined corner and edge detector.” In: *Alvey Vision Conference*. Vol. 15. 50. Manchester, UK. 1988, pp. 10–5244.

[209] K. Mikolajczyk, T. Tuytelaars, C. Schmid, et al. “A comparison of affine region detectors”. In: *International Journal of Computer Vision* 65.1-2 (2005), pp. 43–72.

[210] D.G. Lowe. “Distinctive image features from scale-invariant keypoints”. In: *International Journal of Computer Vision* 60.2 (2004), pp. 91–110.

[211] W.C. Qiu, X.G. Wang, X. Bai, et al. “Scale-space sift flow”. In: *Dense Image Correspondences for Computer Vision*. Springer, 2016, pp. 71–82.

[212] L. Bo and T. Whangbo. “A SIFT-color moments descriptor for object recognition”. In: *IT Convergence and Security (ICITCS), 2014 International Conference on*. IEEE. 2014, pp. 1–3.

[213] B. Wang, D.H. Lin, H.K. Xiong, et al. “Joint inference of objects and scenes with efficient learning of text-object-scene relations”. In: *IEEE Transactions on Multimedia* 18.3 (2016), pp. 507–520.

[214] S. Leutenegger, M. Chli, and R.Y. Siegwart. “BRISK: Binary robust invariant scalable keypoints”. In: *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE. 2011, pp. 2548–2555.

[215] K.V. Suresh. “HOG-PCA descriptor with optical flow based human detection and tracking”. In: *Communications and Signal Processing (ICCSP), 2014 International Conference on*. IEEE. 2014, pp. 900–904.

[216] P.S. Heckbert. “Fundamentals of texture mapping and image warping”. In: (1989).

[217] T. Brox, A. Bruhn, N. Papenberg, et al. “High accuracy optical flow estimation based on a theory for warping”. In: *Computer Vision ECCV 2004* (2004), pp. 25–36.

[218] D. Li and Y.H. Zhang. “A fast offset estimation approach for InSAR image subpixel registration”. In: *IEEE Geoscience and Remote Sensing Letters* 9.2 (2012), pp. 267–271.

[219] M.K. Cheezum, W.F. Walker, and W.H. Guilford. “Quantitative comparison of algorithms for tracking single fluorescent particles”. In: *Biophysical Journal* 81.4 (2001), pp. 2378–2388.

[220] N. Katsarakis, A. Pnevmatikakis, Z.H. Tan, et al. “Improved gaussian mixture models for adaptive foreground segmentation”. In: *Wireless Personal Communications* 87.3 (2016), pp. 629–643.

[221] Z. Zivkovic and F. Van Der Heijden. “Efficient adaptive density estimation per image pixel for the task of background subtraction”. In: *Pattern Recognition Letters* 27.7 (2006), pp. 773–780.

[222] Z. Zivkovic and F. van der Heijden. “Recursive unsupervised learning of finite mixture models”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.5 (2004), pp. 651–656.

[223] F. Bourgeois and J.C. Lassalle. “An extension of the Munkres algorithm for the assignment problem to rectangular matrices”. In: *Communications of the ACM* 14.12 (1971), pp. 802–804.

[224] J.W. Suurballe. “Disjoint paths in a network”. In: *Networks* 4.2 (1974), pp. 125–145.

[225] K. Bernardin and R. Stiefelhagen. “Evaluating multiple object tracking performance: the CLEAR MOT metrics”. In: *EURASIP Journal on Image and Video Processing* 2008.1 (2008), pp. 246309–246318. ISSN: 1687-5281. DOI: 10.1155/2008/246309. URL: https://doi.org/10.1155/2008/246309.

[226] W.K. Harold. “The Hungarian method for the assignment problem”. In: *Naval Research Logistics Quarterly* 2 (1955), pp. 83–97.

[227] T.A. Williams. “A model of rowing propulsion and the ontogeny of locomotion in Artemia larvae”. In: *The Biological Bulletin* 187.2 (1994), pp. 164–173.

[228] H.W. Detrich, M. Westerfield, and L.I. Zon. “Overview of the zebrafish system”. In: *Methods in Cell Biology*. Vol. 59. Elsevier, 1998, pp. 3–10.

[229] M.Q. Jin, X.F. Zhang, L.J. Wang, et al. “Developmental toxicity of bifenthrin in embryo-larval stages of zebrafish”. In: *Aquatic Toxicology* 95.4 (2009), pp. 347–354.

[230] T.L. Daniel. “Unsteady aspects of aquatic locomotion”. In: *American Zoologist* 24.1 (1984), pp. 121–134.

[231] R.C. Hinz and G.G. de Polavieja. “Ontogeny of collective behavior reveals a simple attraction rule”. In: *Proceedings of the National Academy of Sciences* 114.9 (2017), pp. 2295–2300.

[232] X.L. Peng, J. Lin, Y.D. Zhu, et al. “Anxiety-related behavioral responses of pentylenetetrazole-treated zebrafish larvae to light-dark transitions”. In: *Pharmacology Biochemistry and Behavior* 145 (2016), pp. 55–65.

[233] A. Comeche, M.V. Maria, P. Yolanda, et al. “Effect of methylparaben in artemia franciscana”. In: *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 199 (2017), pp. 98–105.

[234] N. Shubham, H. Whidul, B. Anamika, et al. “A novel method for automated tracking and quantification of adult zebrafish behaviour during anxiety”. In: *Journal of Neuroscience Methods* 271 (2016), pp. 65–75.

[235] H. Alyuruk, G.K. Demir, and L. Cavas. “A video tracking based improvement of acute toxicity test on Artemia salina”. In: *Marine and Freshwater Behaviour and Physiology* 46.4 (2013), pp. 251–266.

[236] J. Galvao, B. Davis, M. Tilley, et al. “Unexpected low-dose toxicity of the universal solvent DMSO”. In: *The FASEB Journal* 28.3 (2014), pp. 1317–1330.

[237] G.E. Tyson. “Ultrastructure of a spirochete found in tissues of the brine shrimp, Artemia salina”. In: *Archives of Microbiology* 99.1 (1974), pp. 281–294.

[238] R.W. Rottmann, J.S. Graves, C. Watson, et al. *Culture techniques of moina: the ideal Daphnia for feeding freshwater fish fry*. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, 1992.

[239] G.J. Berman, D.M. Choi, W. Bialek, et al. “Mapping the structure of drosophilid behavior”. In: *BioRxiv* (2014), pp. 1 –22.

[240] J. Watt, R. Borhani, and A. Katsaggelos. *Machine learning refined: foundations, algorithms, and applications*. Cambridge University Press, 2016.

[241] M. Abadi, P. Barham, J.M. Chen, et al. "TensorFlow: A System for large-scale machine learning." In: *OSDI*. Vol. 16. 2016, pp. 265–283.

[242] L. Wang, G.Y. Zhao, L. Cheng, et al. *Machine learning for vision-based motion analysis: theory and techniques*. Springer, 2010.

[243] A. Altaf and A. Raeisi. "Presenting an effective algorithm for tracking of moving object based on support vector machine". In: *Indian Journal of Science and Technology* 8.17 (2015), pp. 1–7.

[244] J.J. Xiao, R. Stolkin, and A. Leonardis. "Single target tracking using adaptive clustered decision trees and dynamic multi-level appearance models". In: *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*. IEEE. 2015, pp. 4978–4987.

[245] M. Smirnov and M.A. Pusateri. *Object recognition and tracking using a classifier comprising cascaded stages of multiple decision trees*. US Patent App. 14/212,312. 2015.

[246] B. Sidahmed. "Combining boosting machine learning and swarm intelligence for real time object detection and tracking: towards new meta-heuristics boosting classifiers". In: *International Journal of Intelligent Robotics and Applications* 1.4 (2017), pp. 410–428.