Motion analysis for Medical and Bio-medical applications

Élodie Puybareau

Université Paris-Est
Laboratoire d'Informatique Gaspard Monge (LIGM)

***

Thesis submitted as requirements for the degree of doctorate in computer science of Université Paris-Est

***

Elodie Puybareau

*****

Motion analysis for Bio-medical applications

*****

Toru Tamaki	*Rapporteur*
Jesus Angulo	*Rapporteur*
Olivier Meste	*Examinateur, Président du Jury*
Caroline Chaux	*Examinateur*
Pierre-Régis Burgel	*Examinateur*
Hugues Talbot	*Examinateur*
Laurent Najman	*Directeur*
André Coste	*Co-Directeur*
Gabriel Pelle	*Invité*

Acknowledgements

I would like to warmly thank Jesus Angulo and Toru Tamaki who accepted to review this manuscript. I also want to thank Olivier Meste, Caroline Chaux and Pierre-Régis Burgel who accepted to be the examiners of my thesis.

Un manuscrit de thèse, c'est l'aboutissement d'une parenthèse de vie. Une sorte de marathon, lien entre les études et le monde du travail, la découverte du monde de la recherche dans lequel on plonge sans trop savoir ce qu'on va y trouver. Je termine mon manuscrit par les remerciements, dernière page que j'écrirai, et la plus dure. On pourrait croire que les remerciements ne sont qu'une formalité. Ce n'est pas le cas. Comment transmettre 3 ans de larmes et de rires, de prises de tête et de consensus, de conversations animées et de calme bienfaiteur, qui ont permis de faire naître de nouvelles idées et d'afûter mon esprit (et, soyons honnêtes, de survivre à ces 3 années nerveusement pas toujours simple) ?
Comme à la fin d'un bon roman, on dit au revoir aux protagonistes, on referme le livre pour passer à autre chose. Bien sûr, l'histoire continue, mais le décor est différent, c'est un nouveau tome qui commence dans lequel les rôles ont évolué.
Cher lecteur, comme lors d'une cérémonie des oscars, je vais citer de nombreuses personnes (même le chat!), que vous ne connaissez probablement pas, mais comme à cette cérémonie, c'est le passage obligé (vous pouvez aussi tourner la page et continuer votre lecture).
Par où commencer ? Par le plus récent peut-être, en remerciant mon jury, mes rapporteurs Jesus Angulo et Toru Tamaki qui ont accepté de rapporter mon manuscrit ainsi que mes examinateurs, Olivier Meste, Caroline Chaux et Pierre-Régis Burgel. Vous m'avez tous permis d'avoir des discussions pertinentes et ainsi donné des pistes d'amélioration. Merci pour le temps que vous m'avez consacré.
J'aimerais aussi remercier toute l'équipe de Créteil qui m'a adoptée pendant ces 3 ans. Merci à Estelle Escudier, Bruno Louis, Emilie Bequignon, Gabriel Pelle, Mathieu Bottier, Jean-François Papon, André Coste pour leurs conseils et leur aide, pour le temps que chacun a pu passer avec moi que ce soit pour les manips ou les discussions. Merci à toi André, d'avoir été mon directeur de thèse, d'avoir esayé de rentrer dans mon monde et de m'avoir ouvert la porte du tien. J'ai un remerciement tout particulier à faire aussi à Gabriel. Merci d'avoir cru en moi, et de

m'avoir accompagnée depuis avant l'ISBS jusqu'à la fin de ma thèse, soit pendant 6 ans. Merci de m'avoir offert ces opportunités, que ce soit à l'ISBS, au service d'explorations fonctionnelles ou avec la thèse.
Merci à Alexandra de Tassigny, à la fois à l'ISBS quand j'étais étudiante mais aussi "enseignante" (et de m'avoir fait confiance dans ce rôle là), et de m'avoir aidée pendant la thèse à réaliser une expérience des plus importante (décrite dans le manuscrit, vous la découvrirez plus tard...).
Merci aux personnes "administratives", en particulier Sylvie Cach et Caroline Farmouza, qui ont été là tout au long de la thèse pour nous aider avec les différentes démarches et problèmes administratifs...
Une thèse, c'est un peu un nouvel apprentissage de la vie, comme si l'on retournait en enfance. On commence à avancer, des fois on tombe, on se relève, on apprend, on fait nos premiers pas dans le monde de la recherche comme un jeune enfant apprend à marcher et à se construire. J'ai eu la chance d'avoir deux super "parents" de substitution, Hugues et Laurent. Je vous remercie tous deux pour tout ce que vous avez fait. Vous m'avez fait évoluer, grandir même. Même si on a pas toujours été d'accord sur tout, on a toujours réussi à avancer ensemble. Merci, tout ça n'aurait pas été possible sans vous.
Passons maintenant à la fine équipe des stagiaires et doctorants et membres de l'équipe. Merci Laurent, Mathieu, Geoffrey, Thibault, Ali, Diane, Julie, Clara, Ketan, Stephane, Bruno, Kacper, Vincent, pour tous ces jeux qui nous faisaient relâcher la pression (en particulier le loup-garou qui nous a permis de développer de superbes stratégies pour démasquer les loups garous!). Merci aussi à toute l'équipe A3SI, vous tous.

Eloïse, Odyssée, je ne sais pas par où commencer. Vous avez été les meilleures "co-bureautières" du monde. Vos "chansons" me manquent (pardon, je veux dire Odyssée qui essaye de chanter en yaourt et Eloïse qui corrige les paroles), nos discussions complètement futiles à propos de rouges à lèvres, de Disney, notre chasse aux figurines des oeufs Kinder, les conférences ensemble... Vous me manquez tellement. On a rit, on a douté, on a connu des victoires et des défaites, on a fait des essais, on a pleuré, on a stressé, on est allées au-delà de nos limites, on a grandi, et maintenant on est Docteures...et tout ça on l'a fait ensemble. Merci. Merci d'avoir été là, d'avoir été vous, de m'avoir aidée et soutenue, et pour tout ce qu'on a fait ensemble. Merci pour nos rires, nos soirées, nos vacances... Même si on se parle encore quotidiennement, c'est quand même pas pareil. L'image que j'ai de ma thèse, c'est nous trois dans le bureau, nous trois en Islande, nous trois à Disneyland... Bref, c'est nous trois. Je suis tellement fière de vous avoir, et du lien que nous avons créé, qui va bien plus loin que de l'amitié. Je vous aime les filles <3.

Viens maintenant la partie sur les parents, le chat et les poissons d'argent (à défaut de poissons rouge, ça prend moins de place...on fait ce qu'on peut dans un appartement parisien!).
Merci à mon mari (qui ne l'était pas à l'époque) de m'avoir aidée et soutenue dans tous ces moments, d'avoir géré les choses du quotidien quand je n'avais plus le temps de rien, de m'avoir rassurée quand j'avais des doutes... Merci pour tout.
Merci aussi à mes beaux parents, qui m'ont "adoptée" et soutenue dans les moments difficile. Merci d'avoir fait le déplacement pour venir me voir soutenir, le jour de l'anniversaire de ma belle-mère en plus (c'était un signe) ! Je suis heureuse de vous avoir dans ma vie !
A mes parents, et mes grands parents, merci. Merci de m'avoir soutenue, encouragée sur la voie des études longues, et de m'avoir fait confiance. Ca n'a pas été de tout repos, surtout avec la distance en km entre nous, ça a même été très difficile par moment, mais vous avez toujours été là pour moi. Je n'aurais pas imaginé la soutenance sans vous, que vous soyez là pour assister à l'aboutissement de ces années d'étude était indispensable ! Merci à vous pour tout ça, j'espère pouvoir continuer à vous rendre fiers de moi !

A vous tous, je vous dis merci d'avoir été mes compagnons de route durant cette partie de ma vie. Je referme ce premier tome de mes aventures, avec le même sentiment qu'à la fin d'un roman, un mélange de vide et d'accomplissement, et je vous dis à très bientôt.

Abstract

Motion analysis, or the analysis of image sequences, is a natural extension of image analysis to time series of images. Many methods for motion analysis have been developed in the context of computer vision, including feature tracking, optical flow, keypoint analysis, image registration, and so on. In this work, we propose a toolbox of motion analysis techniques suitable for biomedical image sequence analysis. We particularly study ciliated cells. These cells are covered with beating cilia. They are present in humans in areas where fluid motion is necessary. In the lungs and the upper respiratory tract, Cilia perform the clearance task, which means cleaning the lungs of dust and other airborne contaminants. Ciliated cells are subject to genetic or acquired diseases that can compromise clearance, and in turn cause problems in their hosts. These diseases can be characterized by studying the motion of cilia under a microscope and at high temporal resolution. We propose a number of novel tools and techniques to perform such analyses automatically and with high precision, both ex-vivo on biopsies, and in-vivo. We also illustrate our techniques in the context of eco-toxicity by analysing the beating pattern of the heart of fish embryo.

L'analyse du mouvement, ou l'analyse d'une séquence d'images, est l'extension naturelle de l'analyse d'images à l'analyse de séries temporelles d'images. De nombreuses méthodes d'analyse de mouvement ont été développées dans le contexe de la vision par ordinateur, incluant le suivi de caracteristiques, le flot optique, l'analyse de points-clef, le recalage d'image, etc. Dans ce manuscrit, nous proposons une boite a outils de techniques d'analyse de mouvement adaptées à l'analyse de séquences biomédicales. Nous avons en particulier travaillé sur les cellules ciliées qui sont couvertes de cils qui battent. Elles sont présentes chez l'homme dans les zones nécessitant des mouvements de fluide. Dans les poumons et les voies respiratoires supérieures, les cils sont responsables de l'épuration muco-ciliaire, qui permet d'évacuer des poumons la poussière et autres impuretés inhalées. Les altérations de l'épuration mucociliaire peuvent être liées à des maladies touchant les cils, pouvant être génétiques ou acquises et peuvent être handicapantes. Ces maladies peuvent être caractérisées par l'analyse du mouvement des cils sous un microscope avec une résolution temporelle importante. Nous avons développé

plusieurs outils et techniques pour réaliser ces analyses de manière automatiques et avec une haute précision, à la fois sur des biopsies et in-vivo. Nous avons aussi illustré nos techniques dans le contexte d'éco-toxicité en analysant le rythme cardiaque d'embryons de poissons.

Contents

List of Figures

Part I

Introduction

1 Introduction

Motion analysis is a well studied topic in computer vision. In this thesis, our objective is to develop motion analysis tools for medical and bio-medical applications. To this end, we adapt a large variety of known image processing, image analysis and computer vision techniques, such as image denoising, image segmentation, optical flow, registration and many others, to medical and biomedical image time series. We do this in order to provide novel, automated tools for physiological parameters estimation. We develop new methodologies and techniques, which we validate on real data in vivo and ex-vivo samples.

Contents

1.1 General context

Motion analysis concerns a large range of applications. They range from compartmental analysis to asteroid trajectories calculation, via insect behavior analysis. Each moving objects is different, has its own particularities and characteristics.
Many phenomena in medical and biomedical imaging involve moving objects: beating hearts, ventilation, muscular motion, etc. They can be either an artifact or an object of interest. Cardiac imaging needs gating to achieve consistent imaging, and so has to be taken into account. Ventilation causes artefacts in CT reconstructions and methodologies to extrapolate and register images was developed to remove these artefacts. In some cases, the motion is by itself an important cue. For instance calculating an ejection fraction in cardiac CT necessitates to obtain information linked to the motion of the heart. The diagnosis of thrombosis is directly linked to the blood flow. For these purposes, an analysis of the motion in the sequence is mandatory. However, most of the time the components of interest are not the only appearing in the sequences and they can be hard to distinguish. Motion analysis requires much more information than static image analysis, and is not easy to quantify. Indeed, most phenomena implying motion occur in 3D+t, but usually data is only available as 2D+t sequences. Hence the use of video sequence analyses in medicine or in bio-medical procedures remains difficult.

We focused on 2D microscopy applications, especially the high-speed bright field microscopy, which represent by far the most common techniques. These kinds of applications raise several issues: the 2D/3D problem mentioned above, illumination artifact that can be confused as motion, motion of the organisms studied induced by vibrations, fluid layers (which may induce deformations, may amplify vibrations etc.), living organism etc. These make the development of a pre-processing procedure mandatory in order to denoise, stabilize and prepare the sequence for analysis. Because of noise and experimental protocols, we consider a component is "non-moving" when the grey-level intensity remains sufficiently stable over time, i.e. when the variations induced by motion are of the same order as to those due to noise.
After preprocessing, the necessity of obtaining a segmentation of the motion component remains important. Indeed, even if pre-processing helps to analyze the sequence, most of the time the components of interests, the components we want to analyze, are not present in the entire image but in a small part. There can be some other motion areas in the sequences due to the objects studied themselves, for example intestinal motion can occur when studying blood flow in the tail section of a fish embryo. Hence, it is important to first segment the areas of interest before going further in the analysis. For this reason, we describe a method that segments regions according to features, i.e., we aim at obtaining feature-invariant regions. In this work, we propose a method using the power spectrum of Fourier transform wavelets-like decomposed as feature to obtain regions with homogeneous time variations.
Once the segmentation is performed correctly, parameters can be extracted. In this work for example, we develop a time-wise texture-like analysis allowing us to

evaluate frequencies but also to characterize the motion; we rely here on analyzing the pattern of the trajectory of the object.

1.2 Contribution

We propose tools to help provide an answer to the problematics exposed above. We use computer vision tools to analyze biomedical data, as they are suitable to our applications. We adapt some existing tools for this purpose. The main tools we develop are motion analysis methods.
For biomedical purpose, we develop a method that automatically extract beating frequencies from sequences. We also develop a promising tool, the "curvescan", which is able to extract many other useful measurements. Finally, we investigate the possibility of analyzing cilia beating directly in vivo. We also fully describe two eco-toxicity applications, a mortality assessment tool which is now used in the industry, and a heart rate estimation tool currently at the feasibility stage.

1.3 Structure of this manuscript

- Chapter 2 and 3 present the context of our studies.

- Chapter 4 summarizes the contributions of Chapters 5, 6 and 7. Because this manuscrit is intended to be read also by medical practitioners and biologists, chapter 4 describes the essential elements presented in these chapters.

- Chapter 5 is a presentation of the image processing, image analysis and computer vision tools and methods we use in the manuscript.

- Chapter 6 details the procedures developed for the analysis of simple motion. Chapter 7 proposes a methodology for the analysis of more complex motions.

- Chapter 8 presents the pipelines of our cilia analysis procedures. The work presented in this chapter was published at ISBI 2015 [1] and selected for oral presentation at ICIP 2016[2].

- Chapter 9 is a feasability study for assessing cilia motility in vivo. It was published at the medical conference ERS 2016 [3].

- Chapter 10 refers to our fish embryo mortality assay, and was published at ISMM 2015 [4] and in the journal CBM [5]

- Chapter 11 presents the pipeline of our heart rate estimation on fish embryo's heart and tail. It was presented at an oral session at IPTA 2016 [6].

- Chapter 12 concludes and proposes avenues for further work resulting from this thesis.

As outlined in the next section, all the applicative chapters have been published in some fashion.

1.4 Publications associated with this manuscript

International Journal Papers

[1] E. Puybareau, D. Genest, E.Barbeau, M. Léonard, and H. Talbot. An automated assay for the evaluation of mortality in fish embryo. *In print in Computers in Bio-Medicine*

International Conferences

[1] E. Puybareau, H. Talbot, G. Pelle, B. Louis, J-F. Papon, A. Coste, and L. Najman. Automating the measurement of physiological parameters: a case study in the image analysis of cilia motion. In Image Processing (ICIP), IEEE International Conference on, Phoenix, September 2016

[2] E. Puybareau, H. Talbot, and M. Leonard. Automated heart rate estimation in fish embryo. In Image Processing Theory, Tools and Applications (IPTA), International Conference on, pages 379–384, Orleans, November 2015

[3] E. Puybareau, M. Léonard, and H. Talbot. An automated assay for the evaluation of mortality in fish embryo. In Mathematical Morphology and Its Applications to Signal and Image Processing, volume 9082 of Lecture Notes in Computer Science, pages 110–121. Springer, Reykjavik, May 2015

[4] E. Puybareau, H. Talbot, G. Pelle, B. Louis, J-F. Papon, A. Coste, and L. Najman. A regionalized automated measurement of ciliary beating frequency. In Biomedical Imaging (ISBI), IEEE 12th International Symposium on, pages 528–531, New-York, April 2015

Medical Conferences Abstracts

[1] E. Puybareau, E. Bequignon, M. Bottier, G. Pelle, B. Louis, E. Escudier, J.-F. Papon, L. Najman, H. Talbot, and A. Coste. Frequency-based region identification and delimitation for cilia beating pattern analysis. In Cilia 2016, Amsterdam,

October 2016

[2] E. Puybareau, E. Bequignon, M. Bottier, G. Pelle, B. Louis, E. Escudier, J.-F. Papon, L. Najman, H. Talbot, and A. Coste. Towards the in-vivo automated assessment of cilia motility. In Congress of the European Rhinologic Society, Stockholm, July 2016

[3] E. Puybareau, H. Talbot, G. Pelle, B. Louis, L. Najman, and A. Coste. Automatic detection of beating cilia with frequencies estimations. Cilia, 4(Suppl 1):P85, 2015

2

Ciliated cells analysis

This chapter describes methods developed for the analysis of cilated cells. This work is the result of a collaboration with INSERM UMR 955, Institut Mondor de Recherche Biomédicale, Cellular and Respiratory Biomechanics Laboratory. It constitutes one of the principal motivations of the research work reported in this manuscript. In this chapter, we set the context, the motivation and we review existing methods.

Contents

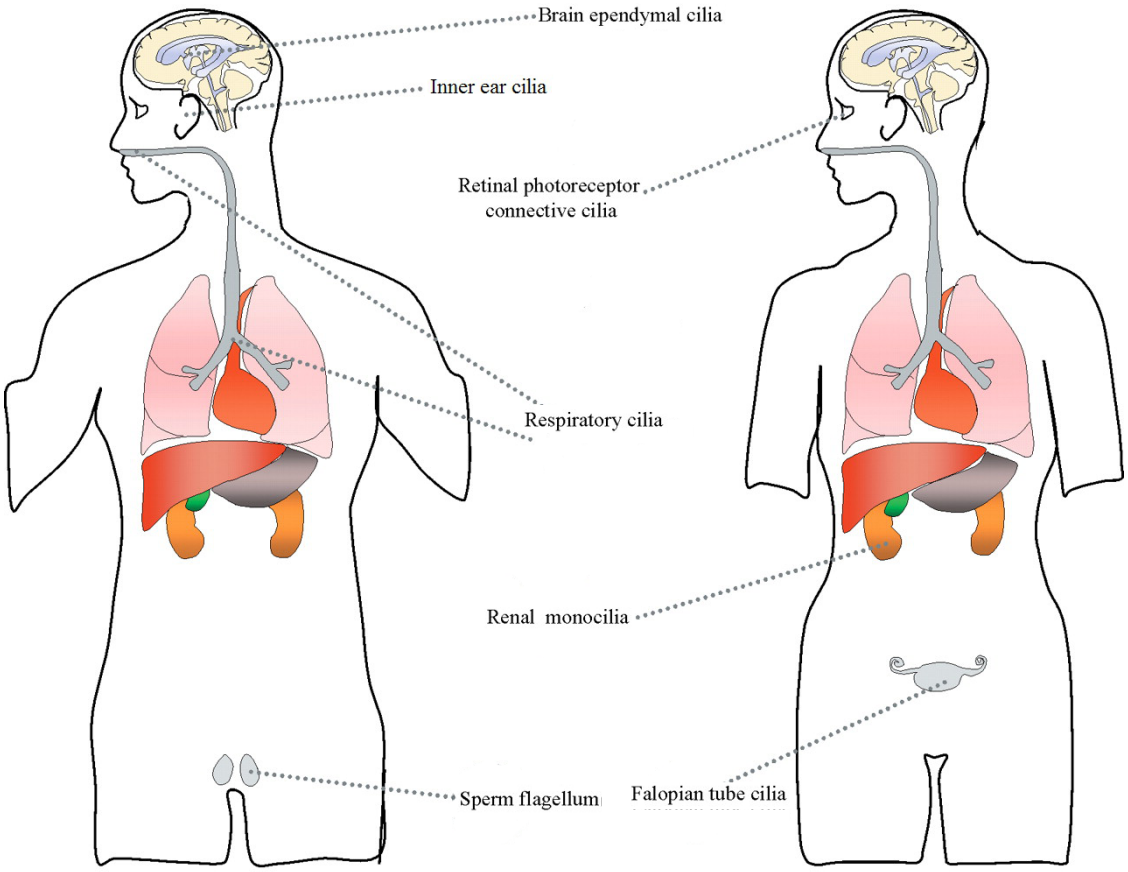


Figure 2.1: Ciliated cells in the body

2.1 Ciliated cells

Ciliated cells are specific cells present in many parts of the body (see Fig 2.1). These cells are covered by cilia, tiny hair-like structures that are motile. In contrast, nearly all cells in the human body have at least one *primary* cilium, which is not motile.

Respiratory cilia are composed of 9 external doublet of microtubules and a central pair (see Fig 2.2). The motility of microtubules is enabled by dynein.

Some genetic impairment can alter the motility of cilia. A loss of dynein or microtubule disrupts the beating of cilia. Each particular loss implies a different outcome in the beating. Some are illustrated in Fig. 2.3. Genetic cilia beating impairment may have profound effects. It may cause chronic disorders such as Primary Ciliary Dyskinesia (PCD), a disease with many life-altering symptoms, typically causing progressive damage to the respiratory system, and affecting 1 in 15,000-30,000 people worldwide [1]. A defect in ciliated cells in the renal tubes may cause Polycystic Kidney Disease (PKD) [7]. Lack of functional cilia in the fallopian tubes may cause ectopic pregnancies. Flagelum of human spermatozoid has an

[1] http://www.pcdfoundation.org/

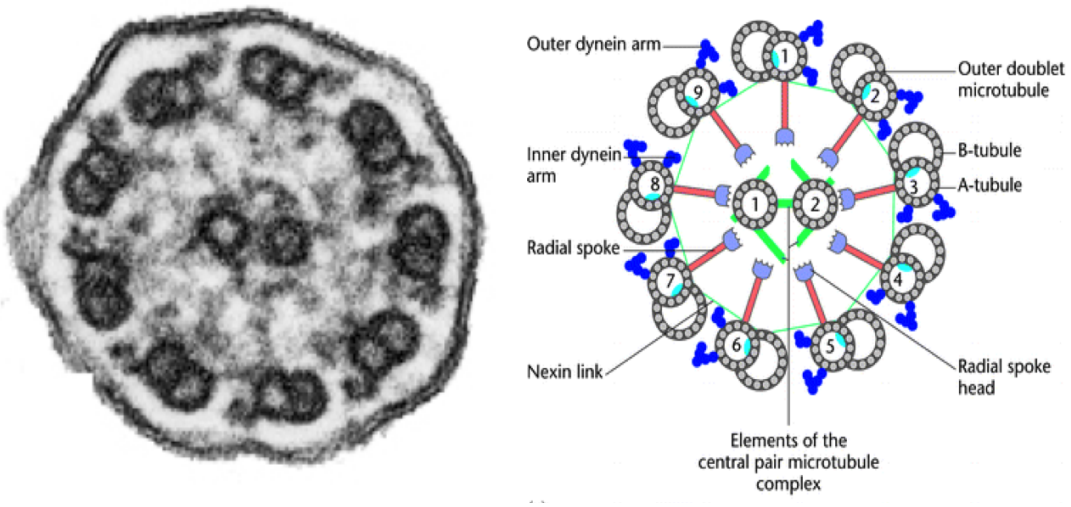


Figure 2.2: Normal axial structure of cilia. (a) Is the axial microscopic view of a cilia. (b) Is its schematization.

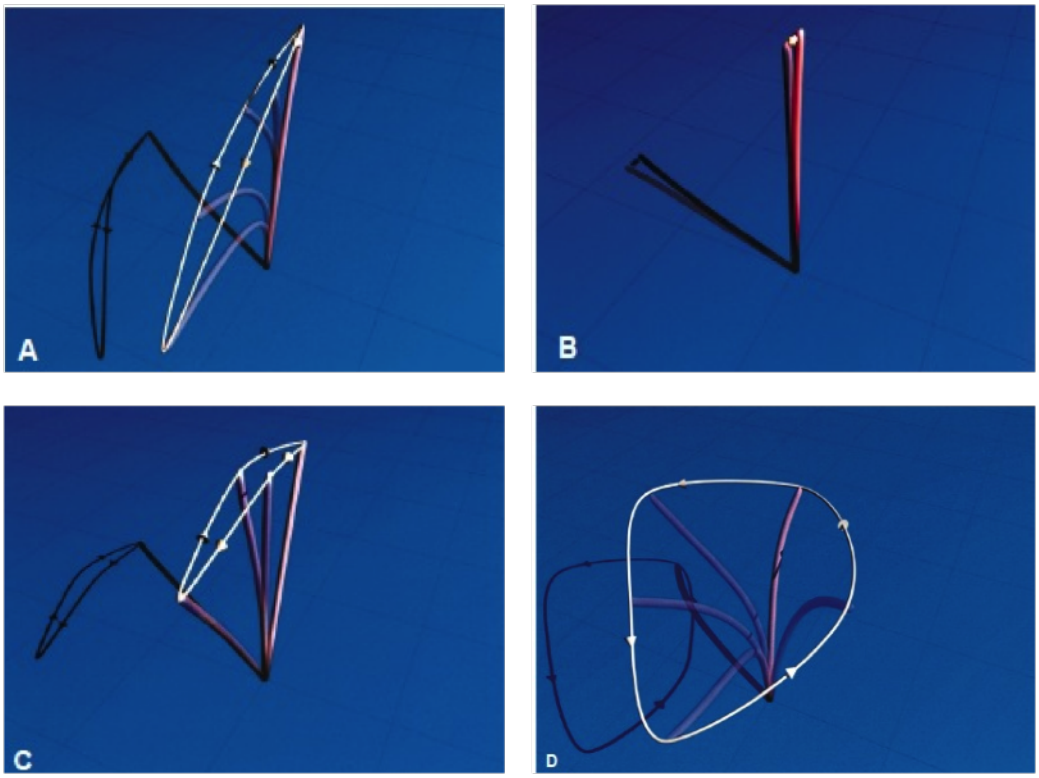


Figure 2.3: Illustration of beating issue according to the loss.(a) Normal beating. (b) Loss of external dynein. (c) Loss of intern dynein. (d)Loss of central pair

ultrastructure close to respiratory cilia, and thus infertility in males patients with PCD is frequent. Cilia dysfunction may cause male infertility. It has been shown that proper cilial function is responsible for the normal left-right asymmetry in mammals [8].

Cilia motility can also be altered by the environment, accident etc. For example,

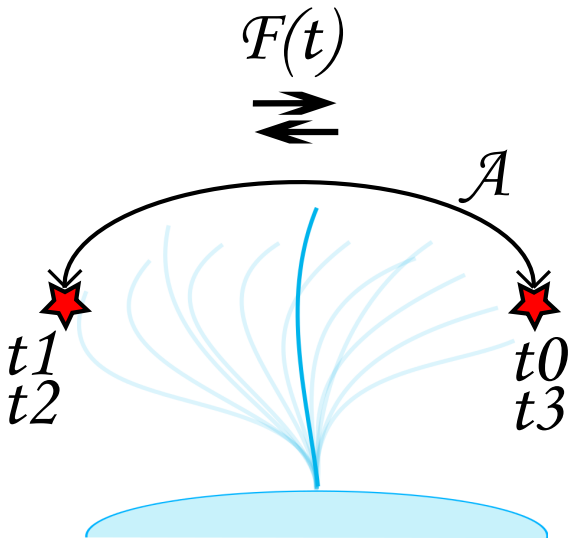


Figure 2.4: Measures needed by practitioners

prolonged or acute exposure to loud sounds can impair the ear's ciliated cells irremediably, without genetic influence. Example of such disease is chronic sinusitis, or chronic obstructive pulmonary disease. Chronic Obstructive Pulmonary Disease (COPD) is an acquired respiratory disease that affects cilia in the lungs. It is typically caused by pollution and smoking [9]. The affected cilia lower the lung clearance capacity and exacerbate pollutants effects in a downward spiral, eventually resulting in chronic bronchitis or emphysema [10, 11]. COPD affects more than 300 million people worldwide and caused 2.9 million deaths in 2013, 90% of which are in the developing world. The economic cost of this disease is in the order of 2 trillions annually and increasing. No cure currently exist for this disease, but it is important to understand cilia motion for diagnosis and associated care [12].

The ciliated cells we are interested in are the respiratory cells, in which cilia are 10μm in length, 0.1μm in diameter, and beat with an average beating frequency of 13Hz.

2.2 Context and state of the art

Muco-ciliary clearance is a crucial mechanism of defense against aerial environmental attacks such as micro-organisms or pollution. This clearance is achieved by the coordinated beating of the cilia covering the nasal epithelium. The most commonly used technique today for evaluating ciliary function in human being consists of collecting ciliated cells from nasal or tracheobronchial surface mucosa, to observe them under a microscope and to record their motion via high-speed video acquisition. Evaluation, via these records, of ciliary beating frequency and ciliary beating pattern was reported helpful in the diagnosis of primary ciliary dyskinesia [13, 14, 15]. Cilia beating frequency is the frequency of beating cilia, while the cilia beating pattern is the way cilia are beating.

It is of interest for practitioners to evaluate ciliary beating frequency easily, robustly and reliably. In clinical research, there exist several methods that estimate ciliary beating frequency. Cinematic analysis [16] counts the number of frames required to complete 10 ciliary beat cycles. The frequency is thus obtained as:

$$\text{CBF} = \frac{\text{frame rate}}{\text{number of frames for 10 beats}} \times 10$$

It is a time-consuming and user-dependent method, which has to be repeated several times to achieve a reliable result. Kymograph analysis [17] is a linescan-like method where the grey level of a line drawn by a human operator is analyzed. It is sensitive to illumination and vibrations, depends on the location of the line, and is thus also user-dependent.
Practitioners generally use both methods: with the kymography, they simply draw a line and compute a Fourier Transform on the image of the variation during time. This is illustrated in Fig. 2.5. This method is simple, fast and works on sequences that are not too noisy and regular enough. On the other hand, it is only capable of evaluating frequency and broadly assumes cilia arranged in a flat bed.

The cinematic analysis is even more time-consuming and user-dependent. The operator chooses a single cilium on a sequence, marks the base and the extremity of the cilium, and marks the extremity point frame by frame over as many frames as needed to amount to 10 cycles (see Fig. 2.5). This allows the analysis of the entire beating pattern: the frequency, but also the amplitude of beating, the speed and pauses the cilia makes (see Fig. 2.4). While this may sound relatively easy when following a single cilium in isolation, in practice a cilium is never isolated, only groups of cilia are visible, making the tracking of the cilia fairly complicated. Cilia are in reality beating in 3D. However only 2D sequences are available. Consequently cilia often come in and out of the focal plane and may become invisible. They can be missed or mistaken for a neighboring cilium. Moreover, it can be difficult to see the extremity of cilia. Analyzing sequences that way is typically time consuming and requires expertise to be carried out properly.

One of the challenge of this thesis is to provide practitioners with reliable tools to estimate automatically parameters that are reproducible and diagnostic, and to evaluate not only those parameters but also the efficiency of beating and by consequence the clearance efficiency.

2.2.1 Estimating cilia beating frequencies

The estimation of ciliary beating frequency has been a research topic since the middle of the 20th century. One of the first methods of reference for the measurement of cilary beating frequency was proposed in 1962 and used a photo-sensitive cell [18]. Stroboscopic methods have been replaced by more accurate techniques that use photomultiplier, photodiode and high-speed imaging. Those methods are described and compared in [19].

Some attempts to automate the measurement of CBF have been proposed in the literature. The SAVA System [20] estimates frequencies from small 4×4 pixels

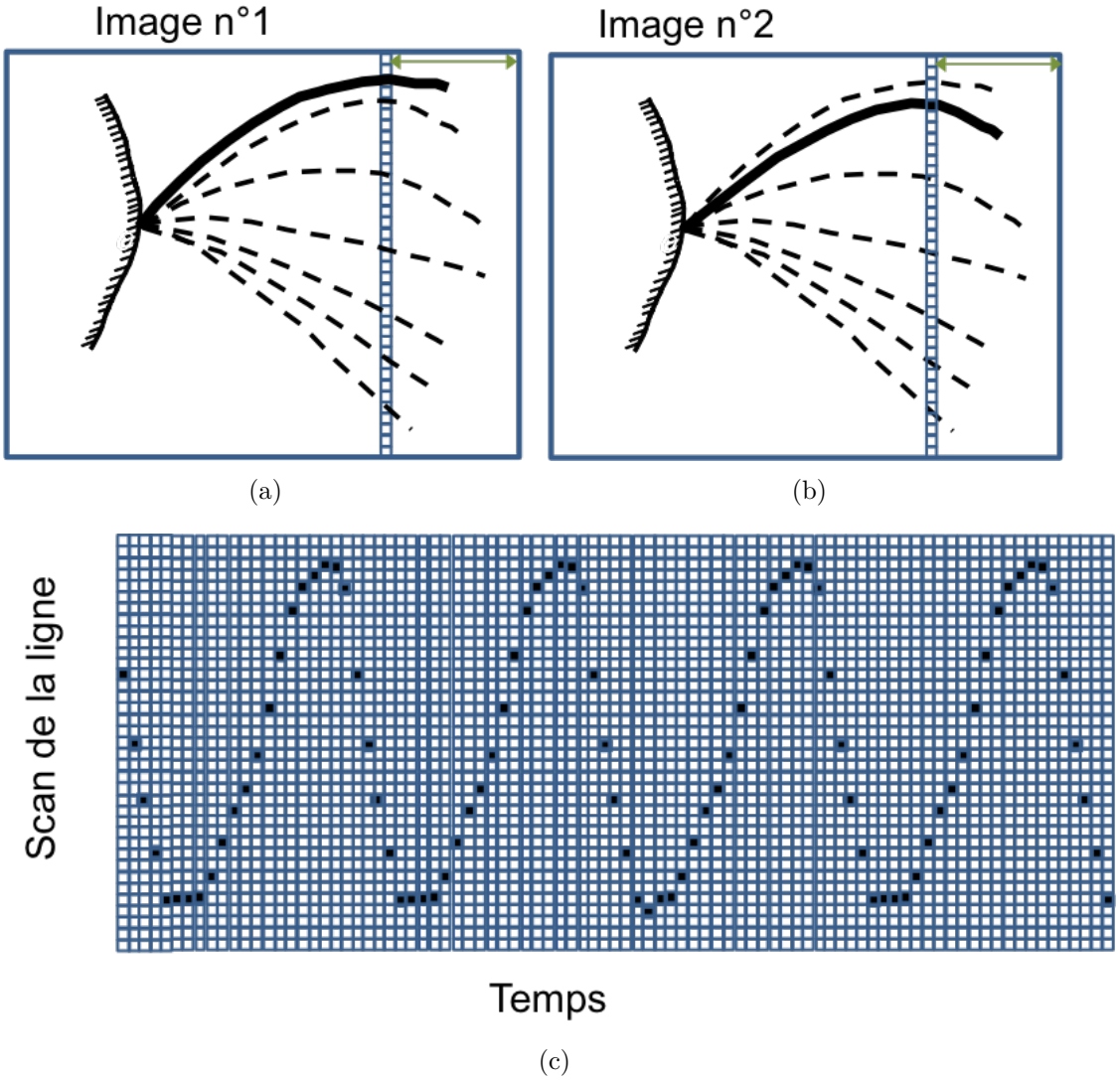


Figure 2.5: Illustration of a kymography procedure

windows. Whole frequency spectra can be simultaneously estimated. This method is based on grey-level intensity variation, which has shown some limitation if the contrast is not sufficient, rendering the reliability of the technique questionable [21]. CiliaFA [22] provides a frequency histogram of a large number of small regions of interest, assuming low noise and no cell proper motion. The method proposed in [23] uses a sparse optical flow to estimate a single frequency per image. Thus, it is not applicable when several different beating patterns are present in the sequence. Moreover, the method is very sensitive to noise and is easily perturbed by cells proper motion. A linescan-based technique is proposed in [24], coupled with the Fast Fourier Transform [25] (first developed in 1866 [26]), and is evaluated on slices on brain ciliated epithelium. It deals with acquisition problems: the removal of artefacts due to the camera sensor, and frame stabilization. However, the removal needs a blank acquisition sequence and thus access to the camera. More problematic

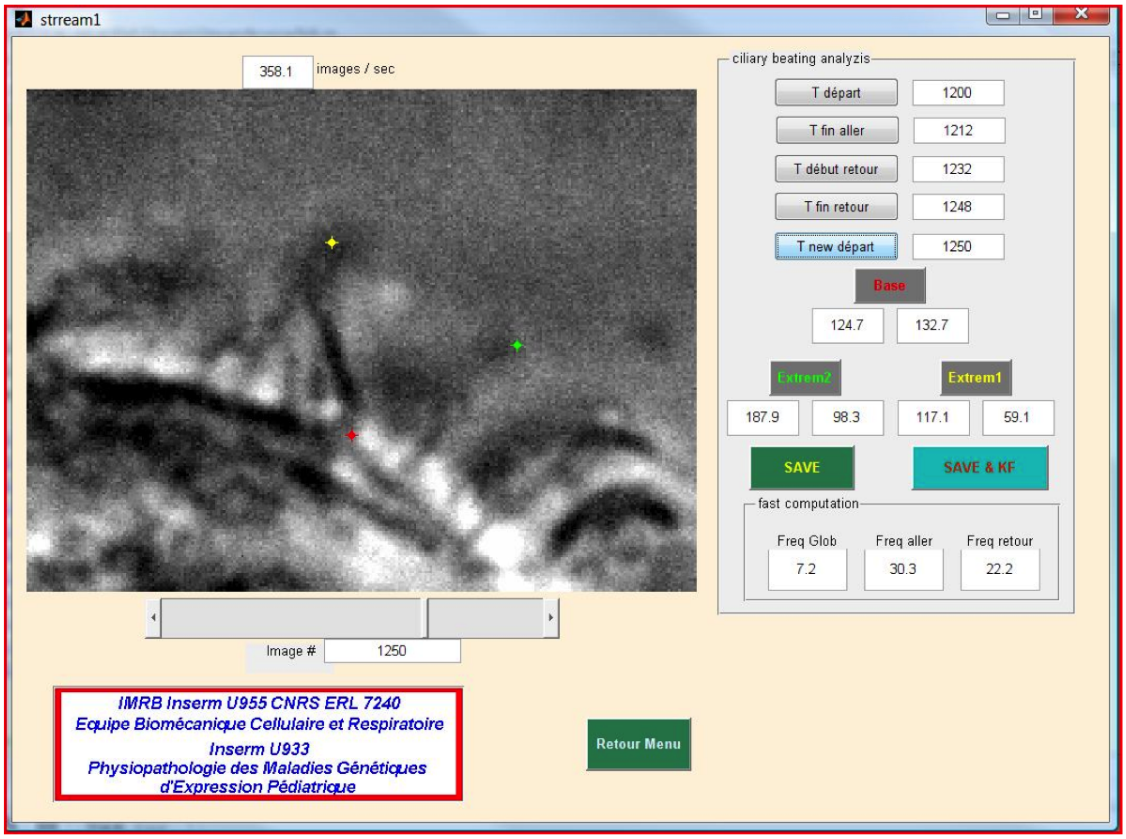


Figure 2.6: Screenshot of the software currently used for cinematic. The red dot is the base of the cilia, the green one is the first position of the extremity and the yellow is the current position of the extremity. analysis

for our application, the straight linescan technique needs a straight border of cells, something not always possible with harvested cells. In the field of view, multiple cell groups are often visible, and cilia on a given cell can beat at different frequencies. As a result, many frequencies can be measured in a single field of view. Such frequencies provide information on cilia synchronization, and ultimately on the status of the cells under scrutiny. We seek to segment the field of view into regions that are consistent from the point of view of the beating pattern. Nasal brushing produces significant amounts of cells with beating cilia. The diversity in sequence appearances can be appreciated on Fig. 2.7. Brushings were all recorded under a microscope minutes after the biopsy, at 358 frames per second with a high speed camera. The spatial resolution is 0.13μm, and the resolution is 256×192 pixels. The sequence was recorded on the border of the groups.

2.2.2 Cilia beating characterization and diagnosis

Cilia beating characterization remains a significant field of study. The difficulty is to analyze not only a single cilium but a group of cilia to evaluate the efficiency of beating. Indeed, a cilium can beat at normal frequency, but if the beating is

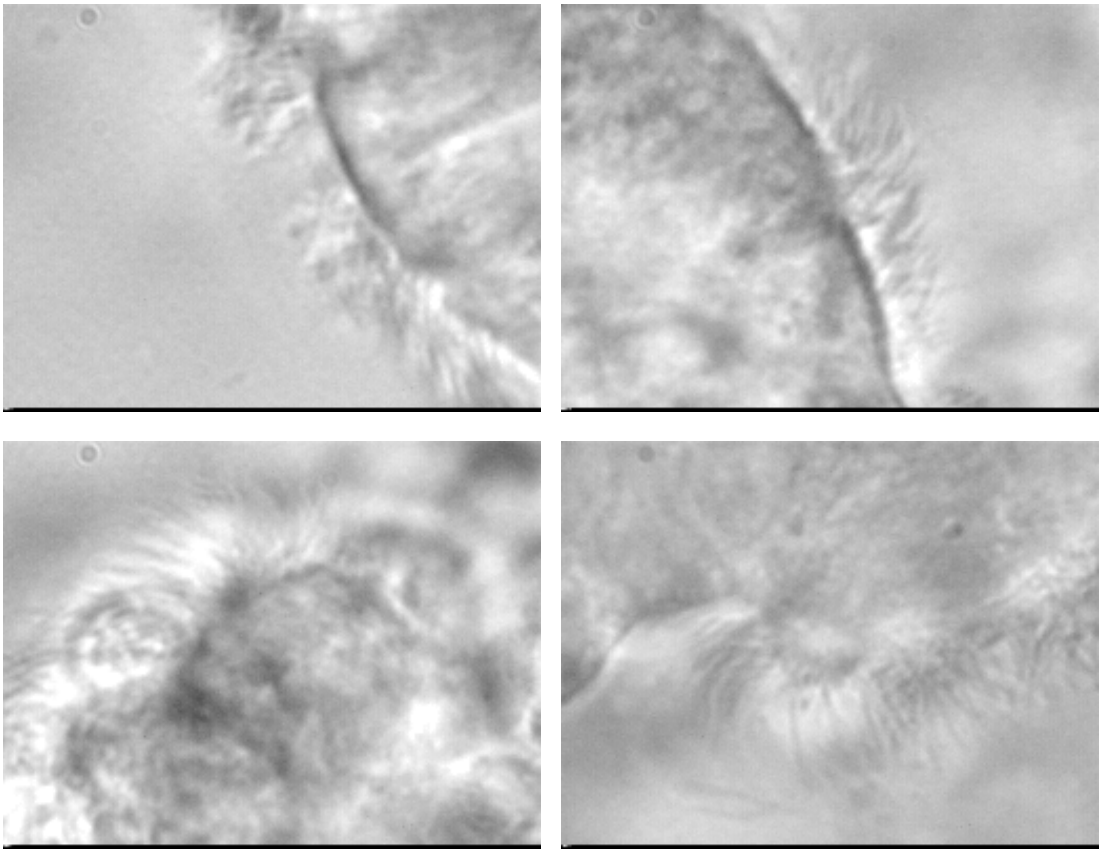


Figure 2.7: Four examples of ciliated cells, showing the large variability in our samples.

irregular or not synchronized with its neighbours, the clearance can be ineffective. Practitioners have to pair the measurements [27] to evaluate the characterization, which is difficult. Some attempts have been made to help the diagnosis of primary diskinesis patient such as in [28]. In that article, the procedure classifies the sequences in two categories (ill or healthy) with a machine learning procedure based on optical flow. In this study, authors show a significant distinction between classes but do not attempt to measure parameters interpretable by clinician, which would be useful for differential diagnosis.
Work in this field is ongoing but as yet not entirely conclusive.

2.2.3 Estimating cilia behaviour in vivo

The evaluation of mucociliary clearance in vivo has always been a challenge. The most reliable measurement consists of a saccharin protocol described by Andersen *et al.* in 1974. This method consists of applying particles of saccharin on the nasal mucous membrane and to measure the time spent before the patient detect a sweet taste. This method depends of the sensitivity of patients and is not sufficient to diagnose diseases. The only way to evaluate reliably the mucociliary clearance for diagnosis is the ex-vivo methods described above. They remain invasive for patients

because of the need for a biopsy, and the analysis is tedious to carry out, and so results can be slow to obtain. Moreover, cells can be injured during the biopsy leading to unreliable results. Finally, there remain questions about the behavior of cells in-situ and ex-situ.
For all these reasons, we sought to develop a procedure for observing and analysing cilia in vivo. This is extremely challenging because cilia are small (10μm in length, 0.1μm in diameter) and beat at relatively high frequency, from 0 to 30Hz. The results of the experiments related to these developments are reported in chapter 9.

In this chapter we presented the motivation for ciliated cells analysis and the context of the study. We now describe the eco-toxicity, begining with the reasons for using fish embryos.

3

Fish embryos and eco-toxicity

Fish embryo models are used increasingly for human disease modeling, chemical toxicology screening, drug discovery and environmental toxicology studies. These studies are devoted to the analysis of a wide spectrum of physiological parameters, such as mortality ratio or heart rate frequency.

Contents

3.1 Context

In recent years, it has become mandatory for the cosmetics industry to warrant the harmlessness of the products and compositions that they manufacture [29]. This includes the by-products and refuse of the manufacturing process as well as the degradation results of the products once released into the environment [30]. However, the cosmetics industry may no longer use animals for any study [31]. Fortunately, some living organisms are still permitted for toxicology assays. Among those, some varieties of fish embryos, particularly those that are easily available and do not require difficult husbandry techniques, are particularly suitable. Fish embryo are defined here as the stage of development following spawning but such that the yolk sac is still attached. At this stage, these organisms feed autonomously and passively from yolk nutrients and are not considered animals. Their size is small (a few mm in length) but do not require high-power microscopes or sophisticated image acquisition devices: a simple camera with a macro lens is enough. They are transparent, and so analyses can be carried out without using invasive techniques. Their circulatory system, in particular, including the heart and major blood vessels, is easily visible. Since they live in water they are ideal for studying waterways pollution [32]. These organisms also belong to the vertebrata subphylum, like humans. For all these reasons, they have been used as model organisms in toxicology studies for quite some time [33]. In the cosmetics industry, large numbers of compounds and formulation need to be tested for health-related effects. L'Oréal for example, use a range of five concentrations per chemical compound. For replication this represents one 24-well plate per concentration plus one plate with no compound present for comparison, for a total amount of 6 plate i.e. 144 wells per test. Because many chemical compounds need to be tested (several tens of thousands in the context of the cosmetics industry alone), this represents several millions of measurements. Manual or semi-automated procedures are no longer sufficient for this task [34].

3.2 The fish embryo model

The fish embryo model has become a classical model in research, and so in toxicology assays. The two most popular organisms used as models are Zebrafish and Medaka. Zebrafish (*Danio rerio*) is a small fish member of genus Danio, and Medaka (*Oryzia latipes*) is a small fish member of genus Oryzias.
Both have fast development and are cheap husbandry. Moreover, both are tranparent (see Fig. 3.1(a)) during the embryo phase, ensuring that their organs are visible and so readily analyzable. They are important model organisms for in vivo studies of vertebrates in biology, toxicology, health, and fundamental research [35, 36, 37, 38]. They are also used in new research fields such as nano-toxicity [39] (see Fig 3.1(b)). Their genetic material is well studied and bears some resemblance to that of humans: humans and zebrafish share 85% of their genome [40], making it a good model for toxicity studies. Moreover, their genome can be modified to express susceptibility to some diseases, or to imbue some of their organs with auto-fluorescence. We recall that while they are still embryos, they are not considered as lab animals and so can

be used for studies without special authorizations.

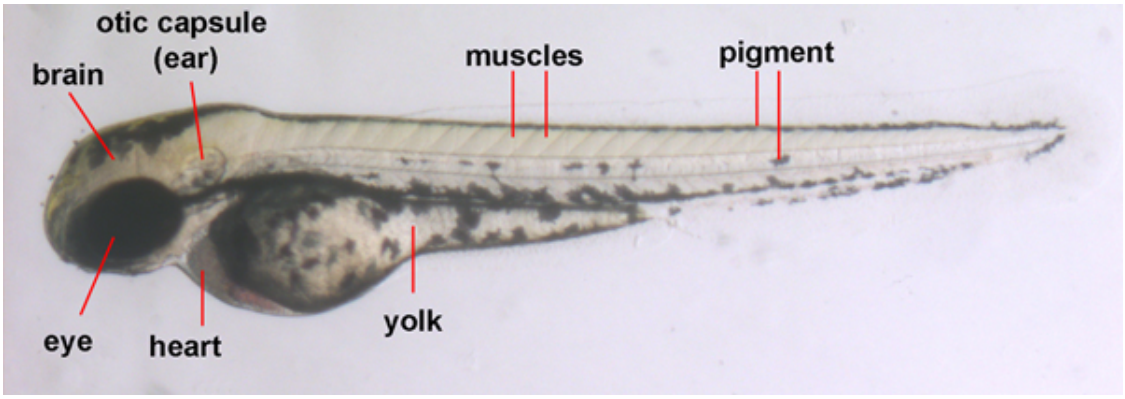
Zebrafish is somewhat smaller than Medaka and is more pigmented. In our studies transparency is essential, so they are not used for the same studies. We mainly studied the Medakas embryos. In our study, embryo were procured from Amagen (UMS 3504 CNRS / UMS 1364 INRA). They are placed as eggs in a neutral environment with Methylene blue to detect dead eggs, and incubated for 9 days at a temperature of 28°C. Embryos do not feed, so no nutrients are supplied. They were anesthetized with tricaine (0.1 g/L) immediately before study.

The protocol for fish analysis is always the same regardless the species or the technique. This protocol is illustrated in Fig. 3.2.

3.3 Image processing and fish studies

Image processing has been widely used in conjunction with fish studies, for instance for sizing [42], aging analysis [43], species recognition [44], automated counting [45], behaviour assessment [46, 47], and more recently for supervising micro-injections in fish embryo [48]. Indeed, fully automated image analysis can offer an interesting alternative to manual procedures, capable of rising to the challenge of offering fast and reliable vital signs assessments in fish embryos [49, 50].
In [41], Mikut and al. present a survey on the existing automated processing for fish embryo image analysis, and here for Zebrafish.
One of the most common vital sign is the pattern of a beating heart, and one of the strongest effect of toxic molecules is to induce death. It is hence meaningful to study the presence of a beating heart to evaluate the lethality of compounds.
In our study, we focused on the analyze of cardiac characteristics of Medaka embryos. To comply with industrial demands, the procedures have to be entirely automated. Speed and reproducibility favor reducing the action of humans during the entire protocol (see workflow on Fig. 3.2). Some tools already exists for automating all the non-imaging related protocol steps, such as the Hamilton robots[1] that cleanse and handle fish embryos and various molecules in the plate wells without any human intervention. For the acquisitions, the plates have to be removed from the apparatus, and brought to the acquisition system. This system can be manual (microscopy for example) or automated (the VAST system for Zebrafish [2] can take time-lapse images of a Zebrafish in a capillary glass vessel. Image analysis system vendors such as FEI [3] with Visilog, or Definiens can propose automated plaforms for taking images and videos of twell plates without requiring a microscope or a fast camera). These automated tools have some limitations. The VAST system is currently only available for Zebrafish, in which they are a very tight fit. Consequently they are not suitable for Medaka or Zebrafish with serious malformations. Visilog tool is not particularly well suited to analyze sequences, but is capable of detailed 2D studies, for instance including embryo malformations.

[1] http://www.hamiltoncompany.com
[2] http://www.unionbio.com/vast/
[3] https://www.fei.com



(a)

(b)

Figure 3.1: Anatomy of a Zebrafish larva, from http://www.devbio.biology.gatech.edu/?page_id=399 and Medakas' malformation as a result of being exposed to nano particles of silver [39]

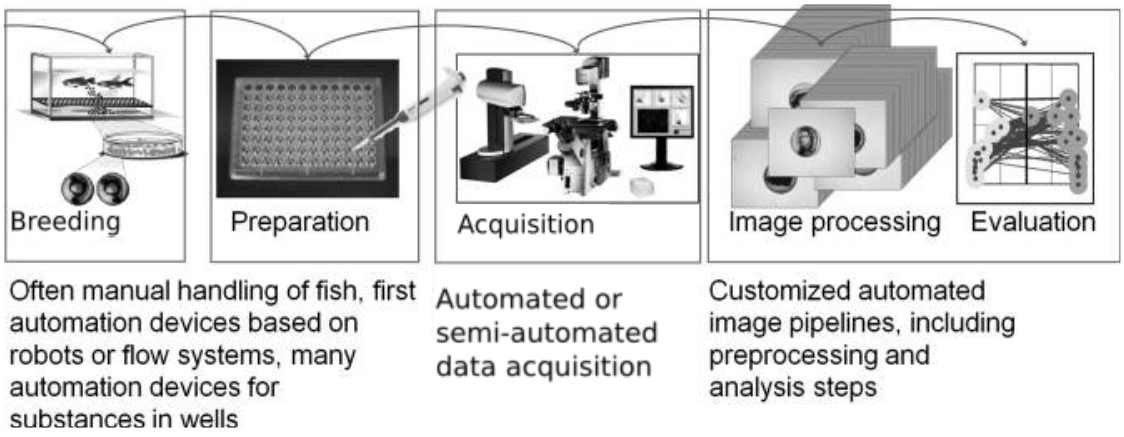


Figure 3.2: Workflow for fish analysis procedure (modified from [41]).

Some automated software packages for the analysis of various embryo characteristics do exist. A quantification of Zebrafish behaviour from low temporal resolution video sequences was proposed in [47]. A screening method to estimate heart frequency exists in the form of a patent [51], but is sensitive to noise and requires a high spatial resolution. The photocardiography method [52] is promising but requires specific tools. ViewPoint developed an automatic tool to measure heartbeat frequencies in Zebrafish via bloodflow analysis[4], but their method is not open-source and twe have not found clear documentation describing their methodology.
One challenge is to adapt existing tools to the available platform and industrial constraints. The methods has to be reliable, easy to use, robust and fast, and finally it has to be adapted to the demands of industrial applications.
We note that some analyses can be simplified by using stainings and other markers. For instance using methylene blue, dead eggs and embryos can be easy to detect as they become dark as the pigment concentrates in the dead organisms.

In most cases, however, a readily available staining does not exist and so a specific image analysis procedure must be developed.

In this chapter, we have presented our second field of application. We now present the image analysis tools we developed, with a first chapter summarizing the methodologies we proposed.

[4] http://www.viewpoint.fr/en/p/software/microzebralab-zebrafish-cardiology-and-physiology

Part II

Technical contributions

4

Methodology essentials

This chapter sumarizes in less technical terms all the methodologies presented in the following three chapters. the objective of this chapter is to communicate the essential developments that we propose while eschewing technical jargon, for a broader audience.

Contents

4.1 Tools

Many image processing, image analysis and computer vision techniques were used in this thesis. We summarize these in technical details in chapter 5. That chapter also contains two applications for removing artifacts, which we present here in a simpler context. These two applications are the “Sensor pattern removal" and the “Image stabilization procedure". We developed these applications because of our experimental constraints that induce strong artifacts on our sequences. The two main artifacts we faced were:

- **Sensor pattern.** Most of our image data were acquired under a microscope using high speed camera. Due to constraints in high-speed sensor design, the captured sequences look like they are taken behind a non-moving grid. This grid is the sensor pattern relative to the camera itself and must be removed before any further the analysis as it affects the quality of the sequences.

- **Undesirable motion.** As we are studying motion, undesirable motion artifacts can induce incorrect results. Because we are working on living organisms that have to be kept in a liquid medium even during the analysis, the probability of observing undesired motion such as vibrations, sliding etc. increases and these have to be corrected for.

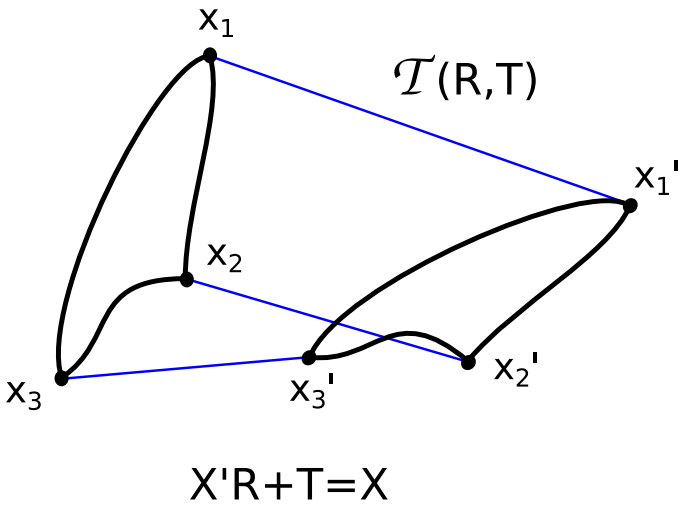
These two points have to be handled in a precise order: first the sensor pattern removal and then the motion correction. The reason of this order is that if we started the other way, the texture of the camera would no longer be fixed, and would induce a new type of artifact that would be more difficult to remove.

4.1.1 Sensor pattern removal

The sensor pattern is a fixed, texture-like grid. When we do not remove this texture, watching the sequence is like through a noisy grid. More importantly, this fixed grid artifact perturbs motion detection and induces further artifacts.
To remove it, we have first to identify it. By computing the average image of the sequence, only the immobile parts of the sequence including the grid remain. Indeed, since the time-wise average of a non-moving element is this element itself. But the average image also provides other meaningful information on the sequence, and not just the grid, so we have to separate the grid from the rest of the average image (for example something moving slowly). Because the grid is composed of fairly isolated pixels of different varying grey-level intensities (like a lattice), we can blur the average image with a small Gaussian filter: this erases only the small components including the grid, and slightly blur the other components.
The result is an average image with the grid and one without (see 5.14). The grid is identified by subtracting these 2 images. We can now remove the identified grid from all frames of the sequence without affecting the quality of the sequence.

4.1.2 Sequence stabilization

We need to stabilize the object of study. For this we are going to perform a so-called “registration" between the objects on the frames. Registration allows to superimpose objects between frames. Using a registration method permits to stabilize the sequence, as the object of study, once registered, appears in the same fixed position in all the images of the sequence.
We use a registration method based on “key points". A key point is a point located with high precision on the image, associated with a set of “descriptors". The descriptors describe the point itself and its environment. A keypoint is a point whose descriptors are sufficiently specific to be distinguished from other nearby points.
Registration uses a reference frame and a moving frame. The idea is to consider a point in the reference frame and to find its corresponding point in the moving frame. A set of 2 corresponding points forms a pair. With a large enough set of pair, we are able to find the transform between the coordinates of the points in the reference frame and in the moving one. The figure below illustrates this.



We have $X = R \times X' + T$ where R is the rotation matrix and T the translation (more detailed equations are given in the methodology paragraph). It means that to superimpose the points between the two frames, we need to apply a transform on a set of point. In our case, we consider 3 families of transforms: identity (i.e. no transformation at all), translation only and translation + rotation. These are called “rigid" transform because they do not change the scale or the shape of the objects. We selected the best transform by comparing the differences between each result with the reference frame. We consider these three cases in succession because if we look for a complex transform we always find parameters for it, but sometimes the parameters are not significant.
We transformed our sequence using this method in order to obtain a stabilized sequence.

4.2 Simple motion analysis

We present here tools that we used for the analysis of simple motion such as heart frequency estimation. These tools are independent of each other. They can be used alone or in combination with others, depending on the aim of the study. We recommend to apply these procedures at least after image stabilization procedure if the sequence shows undesirable motion, to ensure that the moving parts are components of interest.

4.2.1 Motion highlighting

This procedure erases all the fixed elements of a sequence, so that only the moving part are visible. Because we are working on motion analysis, it is useful to enhance the sequences by highlighting the moving components of this sequence. This procedure is intended to be used in a sequence where there is no longer any motion artifact present, so after sequence stabilisation is performed.
We used again the time-wise averaging method: the average of a stabilized sequence yields an image of the non moving parts of this sequence. We do not need to blur this image, as it is non-moving. We just subtract this average from all the frames of the sequence, yielding a sequence consisting only of the moving parts.

4.2.2 Motion segmentation by temporal gradient

We present here a simple segmentation (i.e. separation) between fixed and moving parts in a sequence. It is possible to visualize the motion between two frames when computing the difference between two frames. This allows us to detect what changes between the two frames (see for example fig. 11.4(a)). The asymmetric temporal gradient is a sequence of the difference between each frame and the one immediately after it. When we sum this temporal gradient, this shows all the area where motion occurred during the entire sequence, in a single image (see fig. 11.4(b)). It is not however of interest to us to consider the motion over a whole sequence. This could be difficult to interpret since slow motion components would begin to play a part. So we do integrate (we sum) the sequence only over a few frames. In these integrated images, the intensity represents the amount of change. The highest intensities refer to the largest motion. We filter the slow motion pixels with a threshold: all the pixels under the threshold value are set to 0. This value is determined with an adaptive thresholding method named “Otsu’s criterion". It automatically adapts this value to the content of the image. We then connect the pixels of this image with morphological tools (see Chap. 5.2.1 for the definitions and explanations of erosion and dilation). We obtain an image of the motion areas where each pixel belonging to an area have the same value (see fig. 11.4(d)).

4.2.3 Motion segmentation by temporal variance

We also developed a method to segment motion areas with a more general criterion called the temporal variance. For this we calculate the same temporal gradient as previously, but we do not perform a simple sum. Instead we compute a time-wise variance at each pixel location, i.e. we consider the set of values that a particular pixel takes over time. This hilights the areas where motion is present. This yields an image of where motion is present. To blend irrelevant variations, we perform a spatial Gaussian blur, which smooths the resulting image. Each resulting smooth blob corresponds to an area of motion. We then segmented each area with “Watershed procedure", described in chapter 5. In essential terms, the watershed procedure finds an optimal contour between sets of “markers", which are areas in the image that are characteristic of the inside of the objects of interest (inner markers) and of the outside of objects (outer markers). The markers of the watershed were obtained using morphological tools, the inner markers are the highest values of each blob, and the outer marker is the area where no motion is detected. We then obtain a segmentation in which each region is related to a local maximum motion intensity.

4.2.4 Spurious motion elimination

Cyclic motion can be distinguished from other kinds of motions by their own characteristics, and can be of interest (the heartbeat for example is a cyclic motion). Cyclic motion is a motion that has a repeated pattern over time. We developed a tool that distinguishes cyclic motion from other kinds of motion. If we split the sequence into sub-sequences long enough to contain one cycle, we will notice the same (or nearly the same) motion in each sequence, whereas spurious motion will not appear on all sub-sequence. By computing the variance of each sub-sequence, we highlight the main motion in the sub-sequences, which will always be present in the case of cyclic motion. The median of all sub-sequences yields the motion that is present on the majority of the sub-sequences, especially the cyclic motion.

4.2.5 Frequency estimation

The estimation of frequency is a parameter that is of interest for clinicians interested in diseases involving cilia, or for eco-toxicity studies. We developed several automated or semi-automated methods to estimate frequencies.

Semi-automatic grey-level intensity based frequency estimation

A semi-automatic way for the estimation of frequency is to measure the variations of grey-level intensities, for instance by computing the average pixels intensities in a window. Frequencies can be estimated by Fourier analysis on this signal.

Optical flow

Optical flow is the apparent motion between frames. Computing the optical flow means provides a local translation vector field between these frames. In our work, we developed an automatic optical flow-based method to evaluate frequency in

areas of interest. When we study the intensity of this vector field, which we term "displacement" in the following, we can observe variations: they correspond to speed variations, which can yield frequency information. For each motion region detected, we hence obtain a displacement vector which is the median of all the displacement vectors belonging to that region. The analysis of the variation of the magnitude with a Fourier transform yields the frequency in the region.

4.3 Complex motion analysis

Some parameters may be more difficult to analyze than others. Because we do not always know in advance the characteristics of the motion under study, we developed tools that are suitable for complex motion analysis.

4.3.1 Feature-based region segmentation

A powerful approach is to take into account beating patterns when segmenting areas of motion. Each region should correspond to a single motion pattern, and so moving elements in this region should be moving in a similar way. Because we are mostly interested in beating patterns, and these patterns are periodic, it makes sense to use Fourier analysis. Using Fourier analysis, we compute the power spectrogram of the luminosity at various scales in a moving square window. Each spectrogram forms a “signature", or a vector of features that can be compared with the feature vector of the neighboring regions. We use a clustering method to compare vectors and separate groups, each with sufficient beating pattern similarity. This allows us to identify markers, which are groups of pixel with the highest similarity-score and so beat in the same way. We then attribute the remaining pixels according to their similarity and proximity with the seeds to an area. For each pixel, our procedure yields a score of similarity with each seed, from which we derive region identifications. This procedure ensures that in each region, the behavior of its moving elements is coherent.

4.3.2 Curvescan

Assuming we achieve a reasonable segmentation of areas of motion, we are interested in analyzing the trajectories of the objects belonging to these regions. It is useful to visualize the path followed by the objects of interest in the sequence, in particular the trajectories of cilia. An idea is to outline the pattern they draw while beating. We can imagine that cilia have a pen at their extremity, and they are made to draw their path over time, like on seismographs. The curvescan is derived from this idea. To capture these trajectories, we use the contours of our segmentation. The segmentation surrounds cilia. The contour of this segmentation are typically located near the cilia extremities. By taking the first frame, we obtain this contour, and we "unroll" it: we obtain the position of each extremity of cilia in the first image. We paste it in a new image. Then we do the same thing in the second image, the third etc and we paste the new unrolled contour below the previous one like in

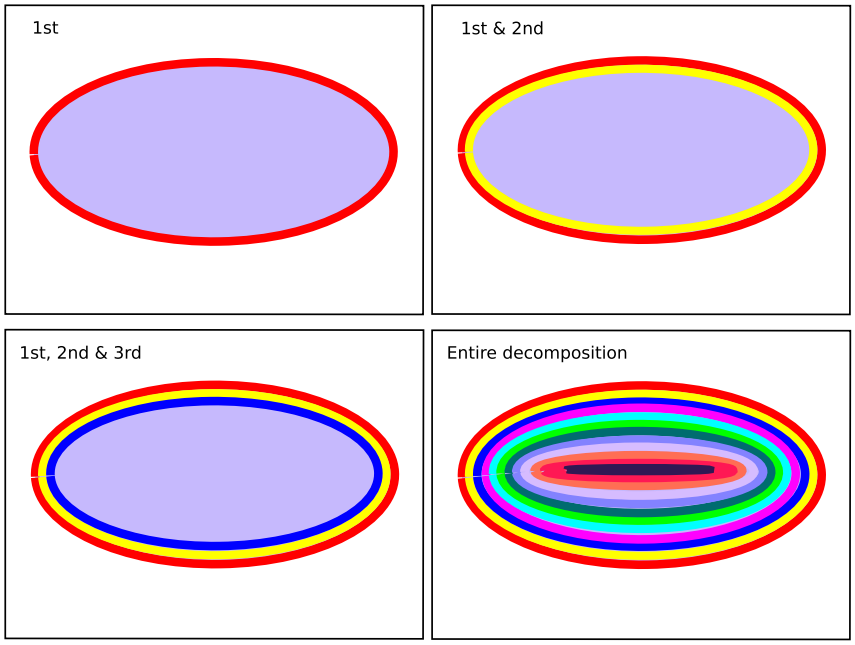


Figure 4.1: Distance map simplification: the red line is the line labeled 1 because it is necessary to only cross 1 pixel to reach the background, the yellow is labeled 2 because crossing two pixels is now necessary, etc.

the linescan: we obtain a follow-up of the position of cilia. The difference between linescan and curvescan is the shape of the studied pixels. To obtain our curves, we computed a "distance map" to our segmentation: it attributes to each pixel of the segmentation the value of its minimal distance to the background (see Fig. 4.1). We can then study all the curves present in our segmentation by doing this procedure on each line of the distance map.

In the next Chapter, we provide all the technical details for the tools that we have developed during the course of our research. Readers who are not interested in a high level of technical details regarding our image analysis contribution may directly skip to part III, describing the applications.

5

Tools for motion analysis

This chapter gives a presentation of the basic image processing tools we used in our study.

All the data we analyzed were acquired on devices such as bright field, fluorescence or stereo-microscopes, and were recorded with digital cameras. the organisms we studied, whether fish embryo or ex-vivo biopsy samples, were imaged in a fluid. The protocols we used are fraught with artifacts that need to be removed before motion analysis in the videos. These artifacts typically consist of noise or undesirable motion.

Contents

5.1 Definitions

5.1.1 Images

An image is an association of pixels, the size of an image is its total number of pixels. A 2D-image is an image with 2 dimensions: numbers of rows and columns. Let m be the number of rows, and n the number of columns, the total number of pixels of the image will be m×n. In a 3D-image, the third dimension is the depth, d, and the total size will be m×n×d.
A grey level image denotes images whose pixels value corresponds to a grey shade, from black to white. The lowest values will correspond to a darker shade whereas the highest values are lighter.

5.1.2 2D+t sequences

A 2D+t sequence is a collection of 2D images acquired (generally with a camera) over a duration t. Each image in the sequence is called a frame. The total number of frames depends on both t and the temporal resolution of the camera: a sequence acquired with a resolution of 25 frames per second will contain fewer frames than the same duration acquired with a resolution of 300 frames per second.

5.2 Basic tools

5.2.1 Mathematical morphology

Let I, J be grey level images defined on a parallelepipedic support (i.e rectangular for 2D or box-shaped for 3D) Ω, taking 8-bit discrete values, i.e. $I, J : \Omega \to \mathbb{Z} \cap [0, 255]$. We denote the Boolean image $(I)_{\geq\theta} = \{\forall x \in \Omega, I(x) \geq \theta\}$, i.e. the thresholded image of I at value θ.

Connexity

In 2D on a square grid, one pixel is said to be 4 or 8 connected if it belongs to a specific neighborhood of that pixel. The 4-connexity neighborhood is defined as pixels sharing an edge. In other words, the neighborhood consists of the 4 nearest neighbours: the 4 pixels of top, bottom, left and right. The 8-connexity is composed of the pixels sharing an edge or a corner. In other word, it is formed of the 4 pixel that are 4-connected plus the 4 pixels in the corners. These are illustrated in fig. 5.1.

Structuring element

A structuring element is a set of points plus an origin point. They can be of various sizes and shapes, and their associated with may or may not be one of the considered points. The shape, size and origin depend on the application they are used for. However, some classical structuring elements are widely used, such as discrete approximations of Euclidean balls. Their symmetric, or transpose, are the

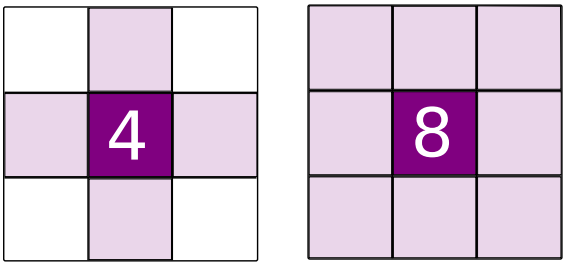


Figure 5.1: Illustration of connexity. The dark pixel is the "reference" pixel, and the lights are its 4/8 neighborhood

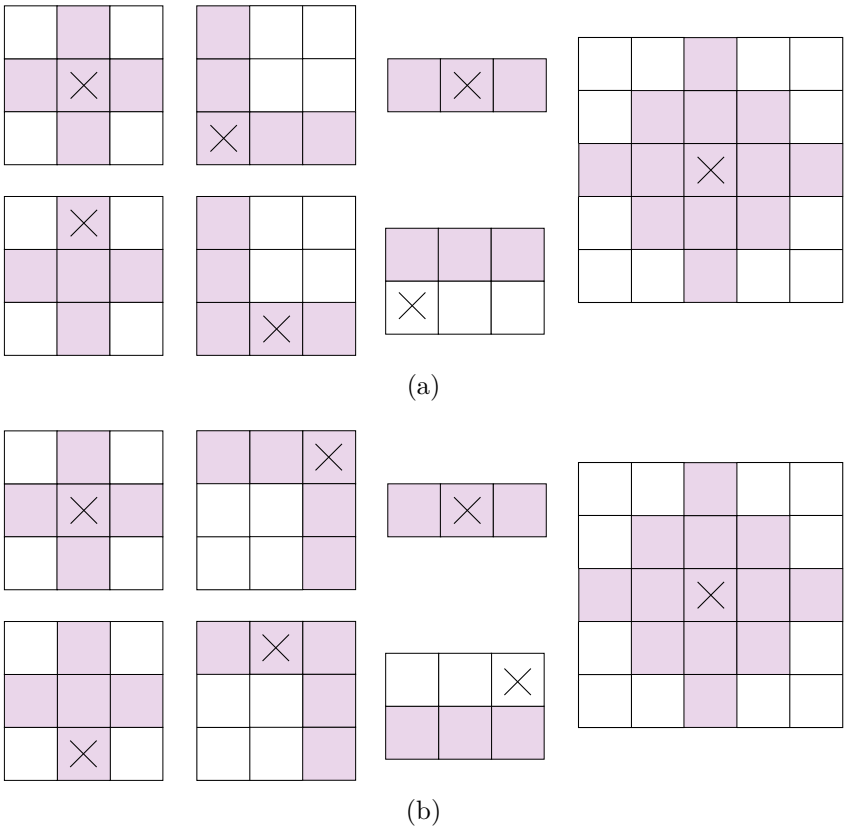
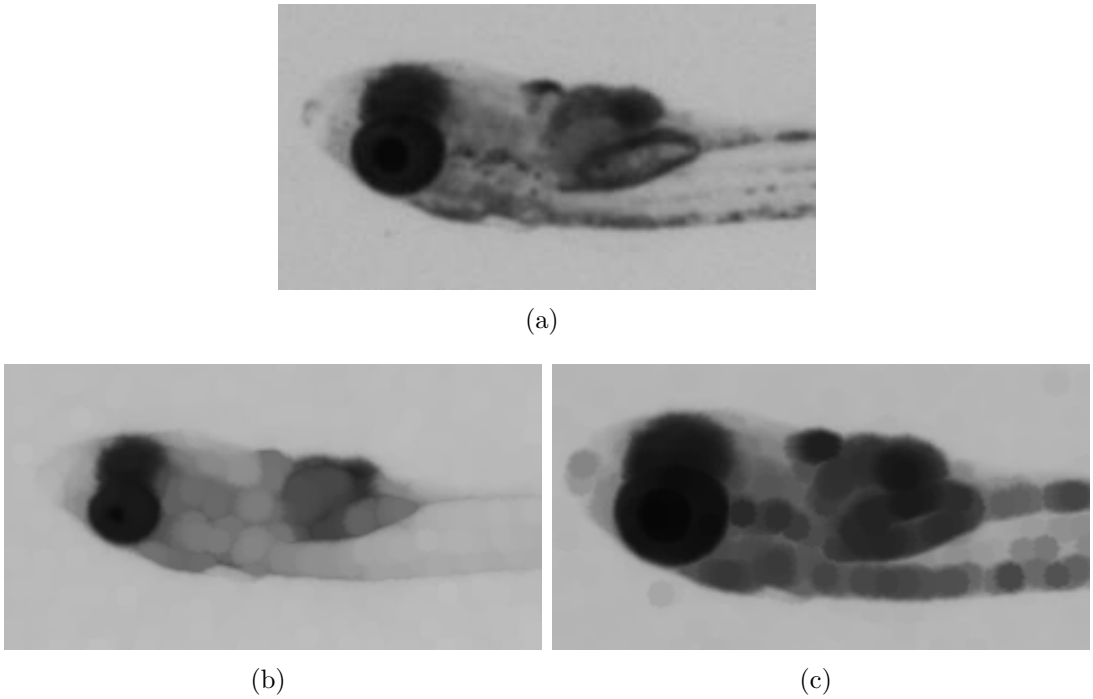


Figure 5.2: Structuring elements and their "transposed" or symmetric with respect to their origin. The cross indicates the origin. The same shape of structuring element can have different origins, yielding different results while used with morphological operators. (b) are the symmetrics of (a)

180-degree rotation of the structuring elements about their origin. We note $\mathcal{B}$ the structuring element, and $\check{\mathcal{B}}$ its symmetric, we classically note $\mathcal{B}_r$ the structuring element associated with a size r, for example a discrete Euclidean ball of radius r.

Dilation and Erosion

$\delta_{\mathcal{B}}(I)$ is the dilation of I by the structuring element $\mathcal{B}$, $\varepsilon_{\mathcal{B}}(I)$ is the adjunct erosion. These are defined as follow:



(a)

(b) (c)

Figure 5.3: Erosion and Dilation: (a) is the original image. (b) and (c) are respectively the dilation and the erosion of (a) with a ball of radius 5

$$[\delta_{\mathcal{B}}(I)](x) = \max\{I(y), y \in \check{\mathcal{B}}(x)\} \tag{5.1}$$

$$[\varepsilon_{\mathcal{B}}(I)](x) = \min\{I(y), y \in \mathcal{B}(x)\} \tag{5.2}$$

Fig. 5.3 illustrates the erosion and dilation.

The erosion and dilation with the same structuring element form an adjunction. This means that :

$$J \subseteq \varepsilon_{\mathcal{B}}(I) \Leftrightarrow I \subseteq \delta_{\mathcal{B}}(J) \tag{5.3}$$

This property allows us to define composed operators, as follows:

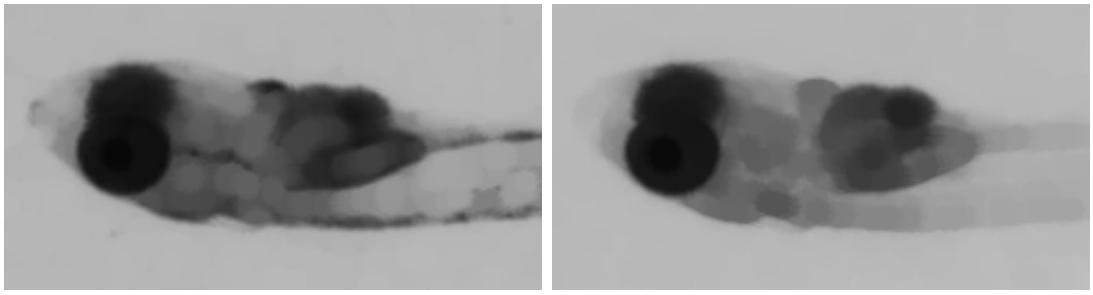
Opening and Closing

$\gamma_{\mathcal{B}}(I)$ and $\varphi_{\mathcal{B}}(I)$ are the morphological opening and closing [53]. These operators are composed from the erosion and dilation. The opening corresponds to the dilation of the eroded image, while the closing corresponds to an erosion of the dilated image.

$$\gamma_{\mathcal{B}}(I) = \delta_{\mathcal{B}}(\varepsilon_{\mathcal{B}}(I)) \tag{5.4}$$

$$\varphi_{\mathcal{B}}(I) = \varepsilon_{\mathcal{B}}(\delta_{\mathcal{B}}(I)) \tag{5.5}$$

Opening and closings can also be defined from their algebraic properties, without reference to a structuring element [54], these are called algebraic openings and closings. For instance, we also define, for a binary image X, $\gamma^{\lambda}(X)$, the area opening of X with parameter λ [55]. This opening removes all the components with an area smaller than a parameter λ. This opening can be defined for grey-level images, in



(a) (b)

Figure 5.4: Opening and Closing: (a) and (b) are respectively the opening and the closing of Fig 5.3(a) with a ball of radius 5

which case it arases peaks until they have an area larger than λ. Fig. 5.4 illustrates the erosion and dilation.

Connected component and geodesic reconstruction in binary images

A connected component $\mathcal{C}$ is a set of pixels that are connected (i.e. there exist a path consisting of neighboring pixels between any two points of C), and is separated from the other components by the background. $\gamma^r(I, J)$ is the geodesic binary reconstruction by dilation of I under J (which is an algebraic opening). In other words, it allows the reconstruction of objects from seeds.

$$\gamma^r(I, J) = \{\mathcal{C}_i \in I \backslash \mathcal{C}_i \cap J \neq \emptyset\} \tag{5.6}$$

Fig. 5.5 illustrates this operator on a map of Australia's mainland and the island of Tasmania (composed of two different connected components), where the seed only intersects the Australian mainland, and so Tasmania disappears after the reconstruction.

H-Maxima

The h-maxima operator finds all the regional maxima in grey-level images with an height greater than a parameter h [56]. Regional maxima are connected components (groups of neighbors pixels with the same intensity value), and that are surrounding by pixels that all have a lower value. The h-maxima operator considers all the local maxima of an image and arases them by a height of h, then computes the difference with the original image, which forms the result (see Fig 5.6). We note $H_h^{\mathbf{max}}(I)$ the height maxima of I with parameter h. The geodesic reconstruction and h-maxima operator are examples of connected operators. To implement connected operators [57, 58], an efficient max-tree/min-tree framework is used [59].

Watershed

The Watershed procedure was proposed for the first time for segmentation purpose in 1978 by Lantuejoul and Digabel [60]. The first publication on Watershed for contour detection appeared in 1979 [61].

(a) (b)

Figure 5.5: Geodesic reconstruction. (a) is the initial image on which an opening has been applied to obtain a seed. The inner parts of the Australia are the seeds for the reconstruction. (b) is the reconstruction: note that the island of Tasmania was not reconstructed.

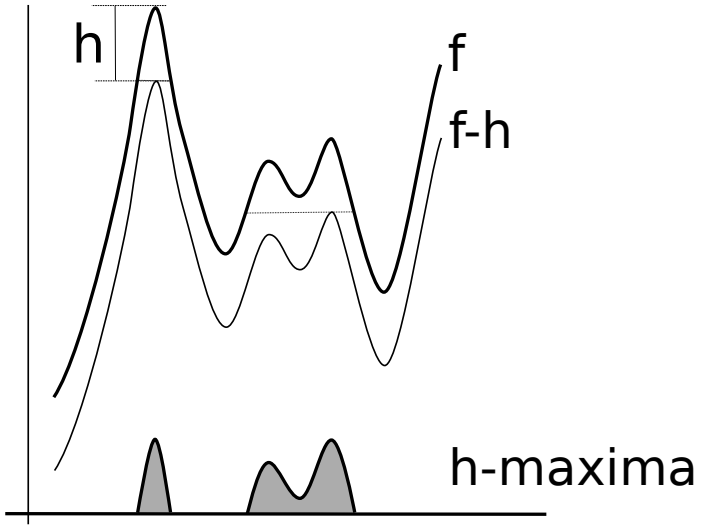


Figure 5.6: Illustration of the h-maxima operator.

The classical representation is to consider a terrain or "relief". This relief may be the magnitude of the gradient of an image, with high values near the contour of objects. The markers are representative sets of points that belong to each object, or to the background. The grey level values of an image can be considered in 3D, the high values are the top of "mountains", and the bottom are the "valleys". The watershed line is obtained by growing the markers from the valleys to the top of the mountains, as if simulating an inundation coming from the markers.. When a growing region

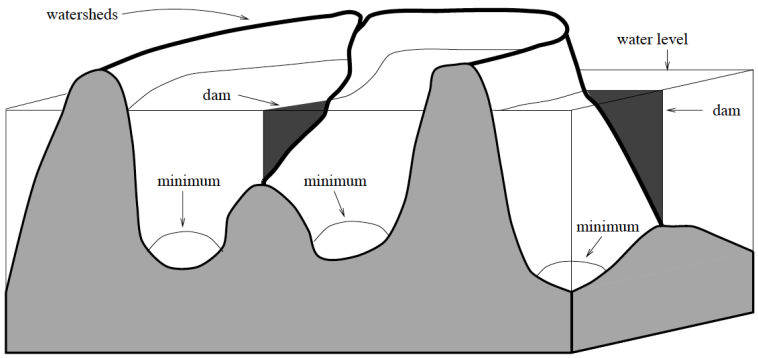


Figure 5.7: Representation of the watershed procedure in 3D

encounters an other one, the border separating these regions is a watershed line and is kept in the result. When the image has been entirely flooded, the result is an image of all the watershed lines. Fig. 5.7 illustrates the mountain/valley concept. Meyer presented a short history of the watershed and its variants in [62].

However, while this classical view of the watershed transform is intuitive, it does not tell the whole story, indeed the watershed transform can be interpreted in other ways: a topological transform [63], a graph transform leading to fast algorithms [64], the result of an optimisation procedure [65], and links with other segmentation algorithms [66]. For this it is useful to turn to a graph-based representation.

5.2.2 Graph-based optimisation model

The more modern interpretation of the watershed is to consider it as a graph transform, specifically a minimum spanning forest [67]. Recently, the watershed and other similar transforms were shown to correspond to a unified graph-based framework, described in the next section.

Definition

Let $G(I) = (V(I), E(I))$ a non-directed graph such as $V(I) = I$ are vertices and $E(I)$ are the edges.
Each point $p \in V(I)$ is a pixel of the image I = a vertex. Each couple $(p, q) \in E(I)$ is an undirected couple of neighbors (p,q) = (q,p), for example representing the 4-connectivity.
Let X be the unknown expected result, and Y the image of markers. Markers are labeled pixels that constitute data fidelity, e.g. representative areas in a segmentation.
We write the problem of the segmentation as an optimization problem where we minimize the energy of:

$$F(X) = \sum_{p \in V(I)} W^k(p)(x_p - y_p)^l + \sum_{(p,q) \in E(I)} W^k{}_D(p,q)(x_p - x_q)^l \qquad (5.7)$$

l \ k	0	1	2	...	$\rightarrow +\infty$
0	V	...	...	...	...
1	...	GC	GC	...	WC
2	...	RW	RW	...	PW
...	...	...	...	...	...
$\rightarrow +\infty$	...	MF	...	...	...

Table 5.1: V = Voronoï , GC = Graph Cut [68], RW = Random Walker [69], WC = WatershedCut [70], PW = PowerWatershed [68], MF = MaxForest [71]

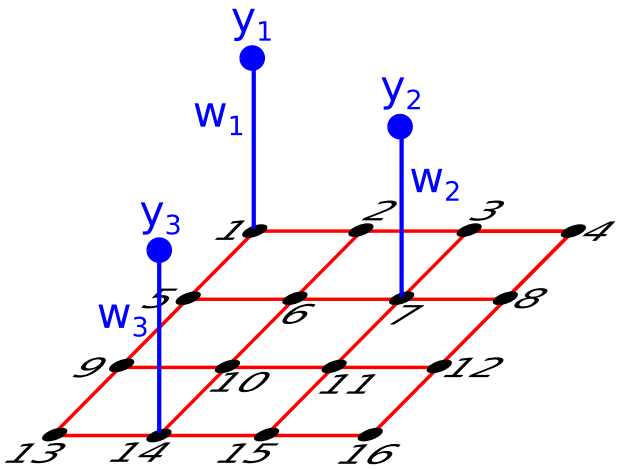


Figure 5.8: Illustration of 3 markers on the image

$$X^* = \arg\min_X(F(X)) \tag{5.8}$$

Where for $p \in I$ $(p = (i, j))$, we write $x_p = X(p)$, $y_p = Y(p)$.
The power (k,l) will determine the model used for resolution. We present here the corresponding models, without specifying which ones we use in the remainder. Table 5.1 presents the models according to the values of k and l.

Each edge has a weight represented in W including the edges w between markers and image (see Fig 5.8).

Data similarity: comparison between X and Y

$$\sum_{p \in V(I)} W^k(p)(x_p - y_p)^l \tag{5.9}$$

Initially The only known ouput data in X are the markers. The first term compares $X(p)$ and $Y(p)$ when $Y(p) \neq$ N/A. This is implemented with $W^k(p)$:

$$W(p) = \begin{cases} 0 & \text{if } y_p = \text{N/A} \\ \lambda & \text{if } y_p = \text{class 1 or 2} \end{cases}$$

If Y p is not a marker (and so is N/A), $W^k(p)(x_p - y_p)^l$ is equal to 0, whereas if it is a marker, to minimize this term, the optimization procedure will tend to make x_p and y_p tend toward each other: it will select the output X in which x_p is close to y_p. The sum will be equal to 0 if all the x_p corresponding to markers Y are well classified. This influence will vary in importance according to the value of λ.

Neighbor similarity and regularization of X

$$\sum_{(p,q)\in E(I)} {W^k}_D(p,q)(x_p - x_q)^l \tag{5.10}$$

When p and q are neighbors. x_p and x_q are their respective values in X. This term helps to ensure label similarity in a region. The similarity is determined by ${W^k}_D(p,q)$. If W is high, it means p and q in Im should be similar according to D, and the resolution will select X in which x_p and x_q tend to be equal. If W is small, x_p and x_q are more likely to be different.
The definition of D is largely up to the user, but is generally a functional that has a low value if the difference between pixel is high, and vice versa. For example, for a grey-level intensity segmentation, $W_D(p,q)$ can be given by:

$$W_D(p,q) = \exp(-\beta(I(p) - I(q))^2) \tag{5.11}$$

with $\beta > 0$, and $I(p)$ and $I(q)$ the original grey-level intensities. If they are equal, the value of $W_D(p,q)$ will be equal to 1, whereas the more p and q differ, the more the corresponding $W_d(p,q)$ tends to zero. This behaviour will tend to place transitions in the values of X in areas where p and q differ, which is the desired outcome.
The choice of D and W are directly linked to the problem considered.

The combination of the two terms allows us to select the best X that respects the markers and the coherence of the regions according to the descriptors.

Solution

The resolution of Eq. (5.7) depends very much on l. If $l = 1$, the equation can be interpreted as solving a linear programming problem, more specifically a maximum-flow/minimum cut transport problem [72]. This can be solved with fast algorithms, but these are only exact in the binary case (two markers). With $l = 2$, a fast exact solution exists for any number of labels, with the Random Walker algorithm [69]. In this case, Eq. (5.7) is a strictly convex quadratic functional that can be optimized by classical differentiation, yielding a linear system.

Random walker Let X be n rows and m columns. $v_x = nm$ is the number of vertices in X and $e_x = 2nm - n - m$ the number of edges in X. Let $v_y = e_y$ be the number of markers. The total number of vertices is $v = v_x + v_y$ and the total number of edges is $e = e_x + e_y$

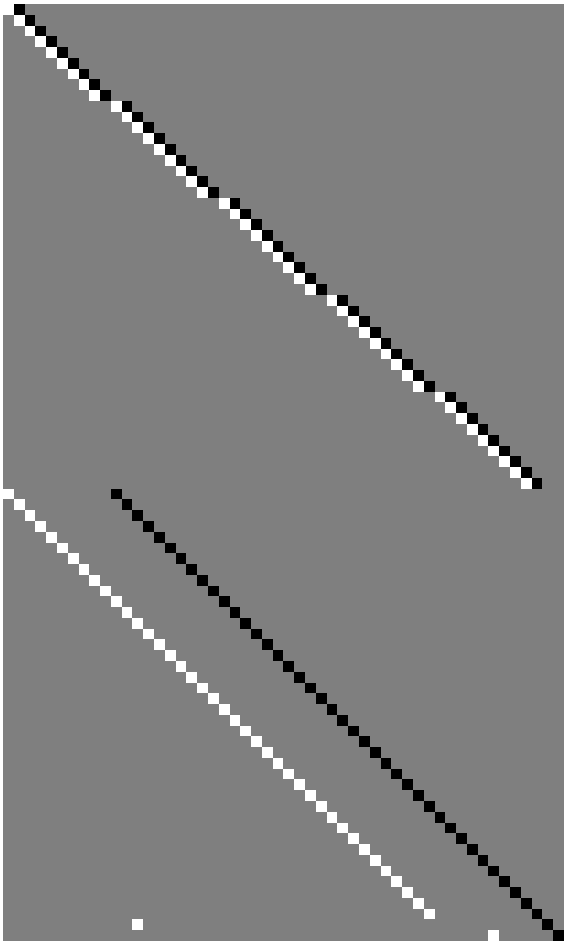


Figure 5.9: Illustration of the adjacence matrix with 2 markers. The grey pixels are 0, white +1 and black -1. It is the adjacence matrix of a 5×10 image, with markers on pixels [1,2] and [4,5]

A is the adjacence matrix: it represents the edges from a vertex with a +1 and to a vertex with a -1. It contains the markers and their edges in the last rows.

$$\mathrm{A} = \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & \dots & v \end{array} \\ \begin{pmatrix} +1 & -1 & 0 & 0 & 0 & \dots & 0 \\ 0 & +1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 0 & +1 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix} \end{array} \begin{array}{c} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_e \end{array}$$

The upper part of A are the horizontal edges, and the middle are the vertical ones and the e_y last are the lines of the markers. An illustration of this can be found on Fig 5.9.

W is the matrix of the weights. It is a diagonal matrix of size $e \times e$ where the diagonal components are the weight on the edge e_i.

$$\mathrm{W} = \begin{pmatrix} w_{e_a} & 0 & \dots & \dots & 0 \\ 0 & w_{e_b} & \ddots & & \vdots \\ \vdots & \ddots & w_{e_c} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & w_{e_e} \end{pmatrix}$$

We define the Laplacian matrix as follows: $\mathrm{L} = \mathrm{A}^\top \mathrm{W} \mathrm{A}$. It is a v×v matrix.

Minimizing $F(X)$ when $l = 2$ and any arbitrary k, corresponds to the Random Walker algorithm. In this case the Euler-Lagrange of Eq. (5.7) reduces to a linear system:

Let X be the vector of x_i and Y the vector of y_i, we have:

$$\mathrm{V} = \mathrm{X} \cup \mathrm{Y} = \{x_1, ..., x_{v_x}\} \cup \{y_1, ..., y_{v_y}\} = \begin{bmatrix} \mathrm{X} \\ \mathrm{Y} \end{bmatrix} \tag{5.12}$$

Separating L taking X and Y into account yields:

$$\mathrm{L} = \begin{array}{c} \begin{matrix} \mathrm{X} & & \mathrm{Y} \end{matrix} \\ \begin{pmatrix} \mathrm{L}_X & \vdots & \mathrm{B} \\ \dots & \vdots & \dots \\ & \vdots & \\ \mathrm{B}^\top & \vdots & \mathrm{L}_Y \end{pmatrix} \end{array}$$

Expressing F in terms of V yields:

$$\mathrm{F}[\mathrm{V}] = \begin{bmatrix} \mathrm{Y}^\top & \mathrm{X}^\top \end{bmatrix} \begin{bmatrix} \mathrm{L}_X & \mathrm{B} \\ \mathrm{B}^\top & \mathrm{L}_Y \end{bmatrix} \begin{bmatrix} \mathrm{X} \\ \mathrm{Y} \end{bmatrix} = \frac{1}{2} \left[\mathrm{X}^\top \mathrm{L}_X \mathrm{X} + 2\mathrm{Y}^\top \mathrm{B}^\top \mathrm{X} + \mathrm{Y}^\top \mathrm{L}_Y \mathrm{Y} \right] \tag{5.13}$$

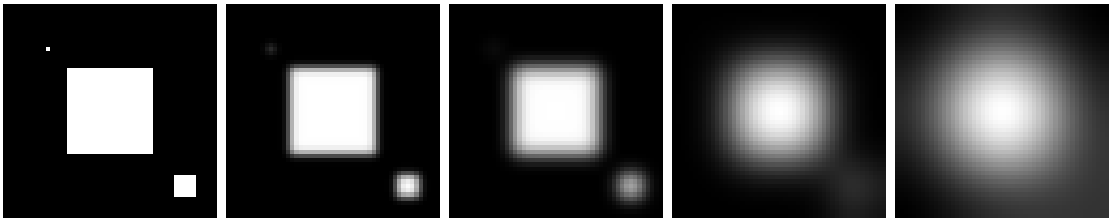
Deriving with respect to X and setting to 0 yields:

$$\frac{\partial \mathrm{D}}{\partial \mathrm{X}} = \frac{1}{2} \left[2\mathrm{B}^\top \mathrm{Y} + 2\mathrm{L}_X \mathrm{X} \right] = \mathrm{B}^\top \mathrm{Y} + \mathrm{L}_X \mathrm{X} = 0 \tag{5.14}$$

Finally X is given by solving:

$$\mathrm{L}_X \mathrm{X} = -\mathrm{B}^\top \mathrm{Y} \tag{5.15}$$

Watershed cuts In the case where $k \to +\infty$, the solution to Eq. 5.7 is also easily obtained: it is a watershed cut [67]. Watershed cuts at least as efficient as random walkers, and also exact for any number of labels. Watershed cut solutions are usually not unique, though the specific one obtained by letting k tend to infinity is indeed unique, and is termed the power watershed solution [68], for which an efficient algorithm exist. As for the general watershed-cut algorithm, it is derived from maximum spanning tree algorithms, for which quasi-linear algorithms exist [73, 74]. The Image Foresting Transform [75] is a related algorithm.



(a) (b) (c) (d) (e)

Figure 5.10: Effect of a Gaussian filter. (a) is the initial image, (b) the result of the Gaussian filter with $\sigma = 1$, (c) with $\sigma = 2$, (d) with $\sigma = 5$ and (e) with $\sigma = 10$.

5.2.3 Gaussian filter

A Gaussian filter smooths the image with a convolution of a Gaussian function centered in 0 (mean = 0) and with a standard deviation σ. The convolution is given by the following equation:

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(x-t)g(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} f(t)g(x-t)\,\mathrm{d}t \tag{5.16}$$

In 1D, the Gaussian function is written in this way:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{x^2}{2\sigma^2}) \tag{5.17}$$

The 2D filter is an isotropic combination isotropic of the 1D filter. The filter is a convolution of the image with a Gaussian kernel. The 2D Gaussian kernel is defined by:

$$g(x,y) = \frac{1}{2.\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\sigma^2}) \tag{5.18}$$

The result is a blurred image. The image is becomes increasingly blurred as σ increases. (see Fig 5.10).

5.2.4 Bilateral filter

The bilateral filter is a spatially variant convolution. It is an edge-preserving filter. It averages pixels based on their spatial closeness and similarity. The closeness is determined in a window and the similarity by a spatial σ. It is based on the Gaussian filters but takes into account the variation in intensity [76]. For each pixel pixel at coordinate (x,y,t) in a 2D+t image, for each pixel (k,l,m) in a window W around this pixel, we have:

$$g(i,j,t) = \frac{1}{\sum_{(k,l,m)\in W} w(i,j,t,k,l,m)} \sum_{(k,l,m)\in W} I(k,l,m)w(i,j,t,k,l,m), \tag{5.19}$$

where

$$w(i,j,t,k,l,m) = \exp\big(-\frac{(i-k)^2+(j-l)^2+(t-m)^2}{2\sigma_d^2} - \frac{(I(i,j,t)-I(k,l,m))^2}{2\sigma_r^2}\big). \tag{5.20}$$

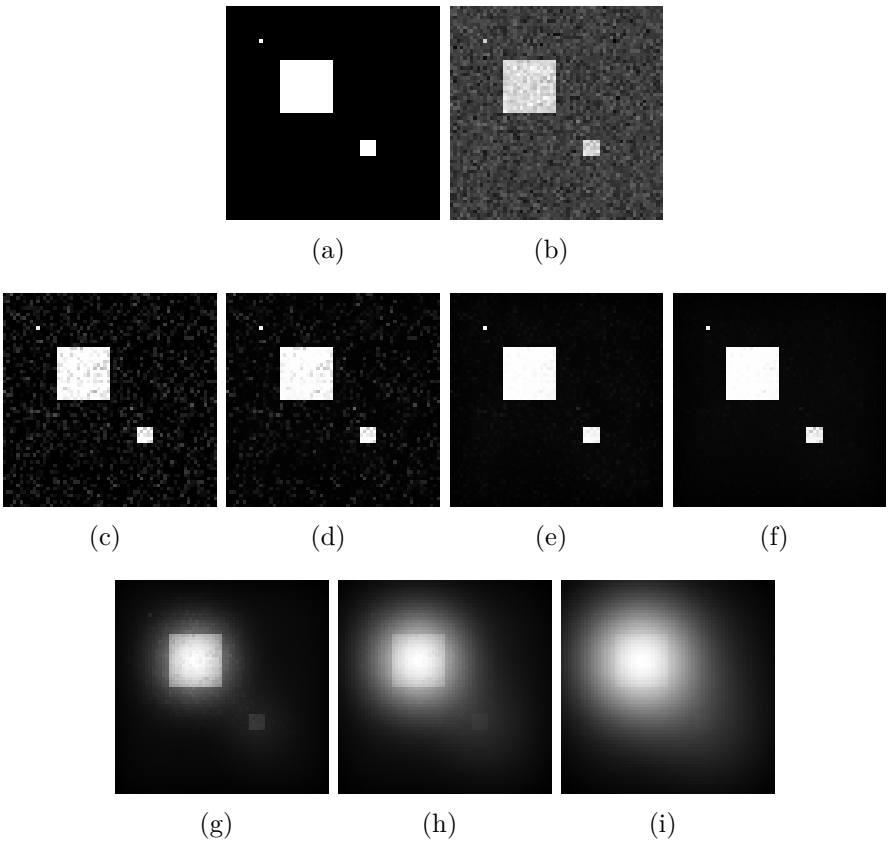


Figure 5.11: Example of Bilateral filter with various parameters. (a) is the initial image, and (b) the noisy one. (c) the result of the Bilateral filter with $\sigma_r = 0.01$ and $\sigma_d = 3$, (d) with $\sigma_r = 0.06$ and $\sigma_d = 3$, (e) with $\sigma_r = 0.15$ and $\sigma_d = 5$, (f) with $\sigma_r = 0.25$ and $\sigma_d = 7$, (g) with $\sigma_r = 1$ and $\sigma_d = 7$, (h) with $\sigma_r = 2$ and $\sigma_d = 9$ and (i) with $\sigma_r = 4$ and $\sigma_d = 11$.

Fig 5.11 illustrates the Bilateral filter on a noisy image with varying parameters σ_r and σ_d. We can notice that if parameters are too low, the filter has a limited filtering effect, whereas if the parameters are too high, the filter tends to blur the image like the Gaussian filter. However, for a range of parameters, the filters reduces the noise effectively while keeping contours intact.

5.2.5 Optical flow

Optical flow is the computation of the apparent oriented motion of objects between two consecutive frames. Optical flow is more sophisticated than a temporal gradient: it provides translation vectors between objects from one frame to the next. Taking a set of points or pixels in the first image, the algorithm locates the equivalent points in the second image. The vector of translation is the vector with which we can find the coordinates of the point in the previous or next image.
This vector provides the direction and the speed of translation obtained from the magnitude and the frame rate. Several methods exist to estimate the optical flow, that may be dense or sparse. Sparse optical flow works with the extraction of

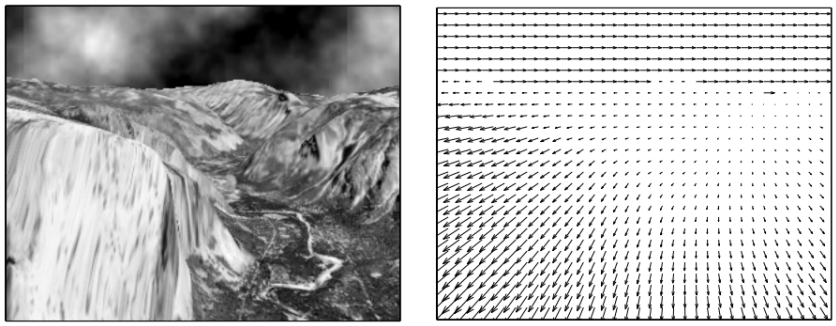


Figure 5.12: One frame of the Yosemite sequence and the corresponding true velocity field (subsampled), from [80]

features (such as edges). Dense optical flow will work on all the pixels or a subset of pixels. Sparse optical flow is generally faster but can miss some areas if the object are not selected as interest points. Dense optical flow is usually slower, but efficient methods do exist.
The Lucas–Kanade method [77] is a famous historical method. Proposed in 1981, it used a differential method for optical flow estimation. It is a local method as it will calculate the flow only in the neighborhood of the pixels selected from the features. It assumes that the flow is essentially constant in a local neighborhood of a selected pixel and uses least squares to solve the optical flow equations for all the pixels in that neighborhood.
The Horn-Schunck method [78] is often compared to the previous one. It is a global method that yields a high density of flow vectors. In a homogeneous object where there are no features, the flow information is extracted from the boundaries of the objects. It yields a plain map of motion for objects of interest but it is more sensitive to noise than local methods.

The Gunnar-Farneback [79] algorithm is a semi-local dense flow. It analyzes only a few points that are regularly spaced on a grid applied to the sequence. The Farneback method uses pyramidal calculations and polynomial approximations to estimate the flow. Each neighborhood of both frames is approximated by quadratic polynomials using the polynomial expansion transform.

$$f(x) \sim x^T Ax + b^T x + c \tag{5.21}$$

where A is a symmetric matrix, b a vector and c a scalar. The coefficients are estimated from a weighted least squares fit to the signal values in the neighborhood (cited from [80]) This algorithm does not uses the a-piori of the spatio-temporal consistency over several frames, which is in our case a good point. Fig. 5.12 shows the vector translation field on the Yosemite sequence. We used the Farneback algorithm in this thesis, due to its performance and wide availability in the OpenCV framework [1].

[1] http://opencv.org

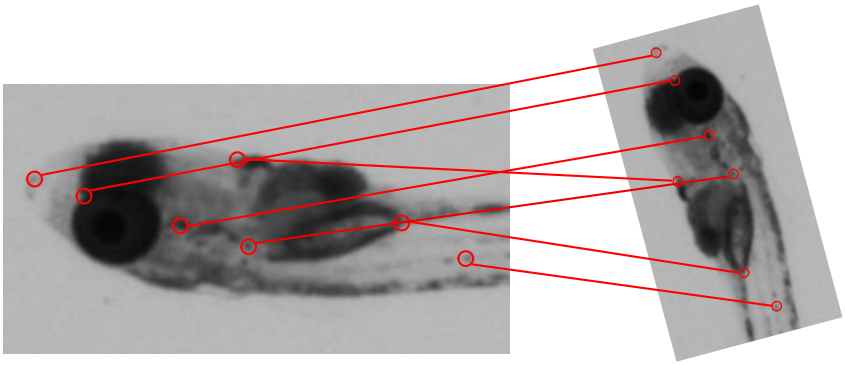


Figure 5.13: Example of key points extraction and matching by pair of corresponding points

5.2.6 Features

"Features" can be associated to descriptors. They are parameters that describe regions, that can be used for segmentation. They can correspond to a texture, a frequency, motion etc. We need to define the descriptors before doing any segmentation.

Keys points are representative areas with some special characteristics, defined by coordinates and features. Features can represent corners, edges or texture, or some other characteristics that are expected to be robust and relevant. These features represent the points and can be used to match them by pairs between frames. Fig. 5.13 illustrates pairs of matching points between two images, an initial image and its transform by rotation, axial symmetry and scale variation.

Several algorithm for key-points extraction exist.

- Scale-Invariant Feature Transform (SIFT) [81]

  SIFT builds a multi-resolution image pyramid, filters each layer with Gaussian kernels and computes gradient orientation at each level. This method consists of four major steps: scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor.

- Speeded Up Robust Features (SURF) [82]

  SURF uses a Hessian matrix approximation for point detection and integral images, which make the SURF calculations faster [83]. Specifically, authors that article create an image stack and filter it and filters it using a box filter approximation of second-order Gaussian partial derivatives. Integral images allow the computation of rectangular box filters in near constant time.

Luo Juan & Oubong Gwun presented in 2009 a comparison between the SIFT and SURF algorithm in [84].

5.2.7 Fourier Transform

The Fourier Transform [85] decomposes a signal into its constituants in the frequency domain. This means it can be used, among other things, to detect significant frequencies in a signal, i.e. the regular repetition of some patterns. Let f be a function of x, real or complex. We define the Fourier Transform of $f(x)$ the complex function of the real variable k:

$$F(k) = \mathcal{F}[f(x)] = \int_{-\infty}^{+\infty} f(x) \exp(-ikx)dx \tag{5.22}$$

If x is a length, k is an inverse length.
If f is a function of time, then x becomes t and we use ω instead of k, which represent the inverse of a duration, or period: a frequency.

$$\mathcal{F}[f(t)] = F(\omega) = \int_{-\infty}^{+\infty} f(t) \exp(-i\omega t)dt \tag{5.23}$$

The result of the Fourier Transform is a function describing the frequencies present in the signal. Important frequencies in a signal typically appear as peaks in the amplitude of the Fourier transform. The intensity of the peaks represents their relative importance in the original signal.

The peaks can be detected with simple peak-detection algorithms and a quadratic extrapolation can be performed to be more precise on the estimation [86] [2].

5.3 Applications

5.3.1 Definition of motion

Motion in a video sequence can be defined as grey-level intensity variation between frames, which differs from noise by its spatial and temporal characteristics. If we consider a fully denoised sequence with enough temporal resolution, we can define three kinds of components. A “non-moving" component is a pixel or a group of pixel whose grey-level intensity remains stable over time, a "slow-moving" component is a pixel or a group of pixel whose grey-level intensity slightly vary over time and a “moving" component will have significant variations over time.

5.3.2 Sensor pattern removal

Motivations

Our acquisitions show not only the subject under study but also a fixed-grid, texture-like pattern often present in high-speed sensors. This artifact is due to the sensor pattern of the camera itself. This pattern is a signature of the camera. It appears even when recording a blank field, leading to an image of this texture. A

[2] http://dspguru.com/dsp/howtos/how-to-interpolate-fft-peak

method for removing the grid pattern could then be to record a blank image, and subtract or divide this image from all subsequent recorded frames. However this method implies acquiring such a blank-field, which is not always desirable or even possible. For this reason we developed a method to obtain the pattern from the sequence itself, based on the idea of the unsharp masking method.

Methodology

The texture is, in fact, a high-spatial frequency, non moving component of the sequence. To obtain an image of the texture, we first computed the average image $\bar{\mathcal{S}}$ of the sequence $\mathcal{S}$. An average image of the sequence will represent the average variations of the sequence: it will preserve the slow and non moving components of the sequence including this pattern, and removes the components with larger motion. Indeed, an average will smooth the grey-level intensities. We hence recover all the "non-moving" and very "slow-moving" components of the sequence.
Because we require an image of the pattern, we filter the $\bar{\mathcal{S}}$ with a high-pass filter. We applied a Gaussian filter with $\sigma = 1$ on $\bar{\mathcal{S}}$ yielding to a "blurred" pattern-free image. Indeed, the pattern is composed of small independent pixels of various grey-level intensity, which are erased by the Gaussian filter. By subtracting the blurred $\bar{\mathcal{S}}$ from $\bar{\mathcal{S}}$, this high-pass filter conserves only the high-frequency "non-moving" elements. This is illustrated in Fig 5.14.

Once we have this image, we can obtain a pattern-free sequence $\mathcal{S}_{clear}$ of I_{clear} by subtracting the pattern from all frames I

$$\forall I \in \mathcal{S} : I_{clear} = I - \left(\bar{\mathcal{S}} - \mathcal{G}_\sigma(\bar{\mathcal{S}})\right) \tag{5.24}$$

with $\mathcal{G}_\sigma$ a gaussian filter with standard deviation $\sigma = 1$, $\bar{\mathcal{S}}$ is the average of the sequence $\mathcal{S}$. $\mathcal{S}_{clear}$ is the sequence of the I_{clear}.

5.3.3 Image stabilization

Motivations

Because we are most of the time analyzing physiological parameters, the organisms under study have to be kept alive. For this purpose and because of their characteristics, we kept them in a liquid medium while recording the video sequences. It implies a higher probability of observing undesirable motion induced by vibrations, or large-scale motion. We hence need to stabilize the object of interest on videos before going further in the analysis. The stabilization part takes place after the sensor pattern removal procedure if that one is needed. Indeed, after stabilization, if it were not removed beforehand, the pattern of Sec 5.3.2 would no longer be fixed but would appear to move due to the stabilisation.

Registration methodology for image stabilization

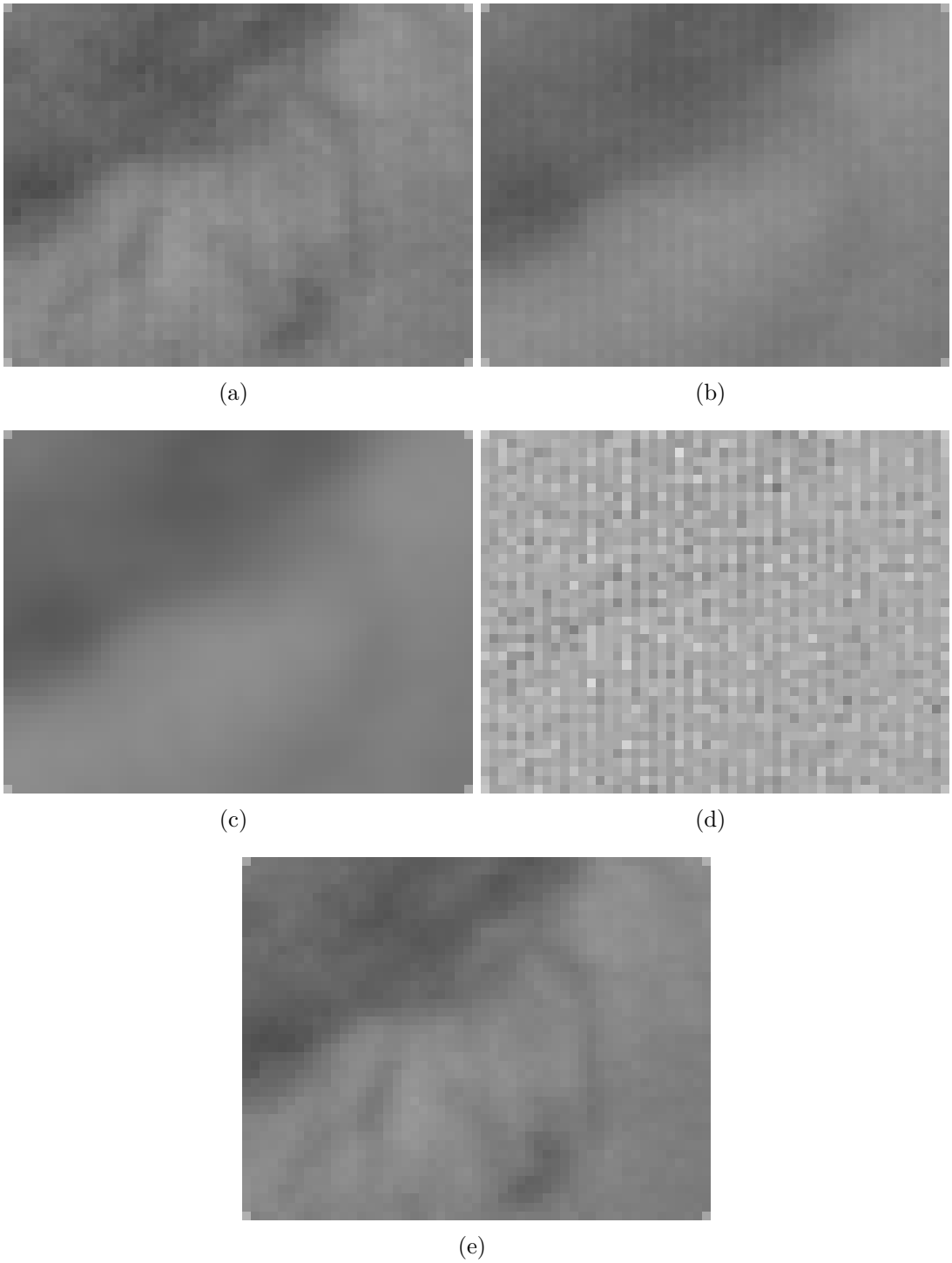
We developed an specific, adaptive registration. It is adaptive because of the parameters specified in our program: a "shift" and a binary mask. The "shift" of the

registration specifies the reference frame. If the shift is 1, each frame is registered according to the previous one. If it is 10, frames 1-10 are registered using the frame 0 as reference, 11-20 using 10 etc. The shift defines the frequency of selection of a new reference frame. The mask parameter is used to select a region of interest for the registration. Indeed, some frame parts are of no interest for the registration (featureless for instance). We calculate an adaptive binary mask which delimits the moving components of the sequence, or which considers the entire sequence if needed.
To find the transformation, we used a keypoint-based method. Key points were extracted in images using the SIFT method [81] or the SURF method [82]. A key point is a point associated with a rotation and scale-invariant set of descriptors, which are sufficiently specific to be distinguished from other points. Point descriptors were matched in pairs in a complete graph and sorted according to their distance coefficient: the smallest coefficient indicates the best matching pair. We kept at most the 10 best pairs of points belonging to the mask.
This yields the transform parameters after an iterative RANSAC-like method [87] to weed out the outliers. The translation (d_x, d_y) is estimated by the median of the difference of the points of the first image and their corresponding points in the second. The combination rotation/translation is determined using a linear least squares system resolution. This resolution is applied on each combination of pairs of points. Then, the median of all the transforms is considered as the rotation/translation. Since we use pairs of points to estimate the transform, it is always a similarity (translation/rotation + change of scale). We assumed that the angle of the similarity is always correctly estimated. We verify that the scale factor is small and we project it back to unity, yielding a rigid transform. This transform is applied and a new translation is then calculated. We then perform a model selection allowing three types of transformations between each pair of frame: identity, pure translation or translation plus rotation. The best transformation is the transformation yielding the smallest image covariance norm. As soon as the transformation matrix is known, it can be applied to all the pixels of a frame. This method is repeated for each frame, and is iterated twice to robustify it. It may happen that none of the transforms yield an improvement over the identity, in which case we decide that no transformation occurred. Matching pairs rather than individual points allow us to better constrain the result [88]. We neglect non-rigid deformations. We choose the best model taking into account both the error and the complexity of the model.
$\mathcal{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ is the rotation matrix. $\mathcal{T} = (d_x, d_y)$ is the translation;

$$\forall (x, y) \in I_{clear} : I_{reg}(x, y) = I_{clear}(x', y') \tag{5.25}$$

with the estimated $\mathcal{R}$ and $\mathcal{T}$, $[x'\ y']^\top = \mathcal{R}[x\ y]^\top + \mathcal{T}$. $\mathcal{S}_{reg}$ is the resulting stabilized sequence of I_{reg}

Now that the classical tools are presented, we will develop in the next chapter our simple motion analysis procedures.



(a) (b)

(c) (d)

(e)

Figure 5.14: Removing sensor pattern from the acquisition, illustration on a crop of a sequence.(a) is the initial frame of the sequence (b) The average of the sequence yields the non-moving components of the sequence. (c) Blurred average removes thin and textured elements. Computing (b)-(c)=(d) yields the sensor pattern. (e) Finally, we subtract this pattern from all the images of the sequence.

6

Simple motion analysis

In this chapter we present the (relatively) simple motion analysis tools that we have developed. We started from the notion of frame difference to enhance the areas where motion occurs, then we switched to more formalized notions based on temporal gradient and temporal variance. We also develop tools for eliminating spurious motion and for frequency estimation.

Contents

6.1 Motion enhancement

As a reminder (see section 5.3.1), we defined motion in a video sequence as grey-level intensity variation between frames, which differs from noise by its spatial and temporal characteristics. If we consider a sequence which has its pattern texture removed and is stabilized, we should have “non moving" components, which are not of interest, and moving component that are assumed to correspond to our objects of interest.

6.1.1 Motivations

We need to enhance the moving components of the sequence in order to obtain easily analyzable moving components. To this end, we need to get rid off all the background, including parts of the objects, which are of no interest to the problem at hand. This procedure takes place after the stabilization part (if needed), to ensure that the moving parts are the objects of interest.

6.1.2 Enhancement methodology

Assuming that the sequence is stabilized (directly from acquisition or after sequence processing), the moving components should mainly be the objects of interest. We again compute the average image of the sequence (Fig. 6.1 (b)), yielding an image of the non-moving parts of the sequence: the background and the parts we do not need for motion analysis. We subtract this result from each frame. This operation removes the static parts of the sequence and yields a sequence of only the moving elements.

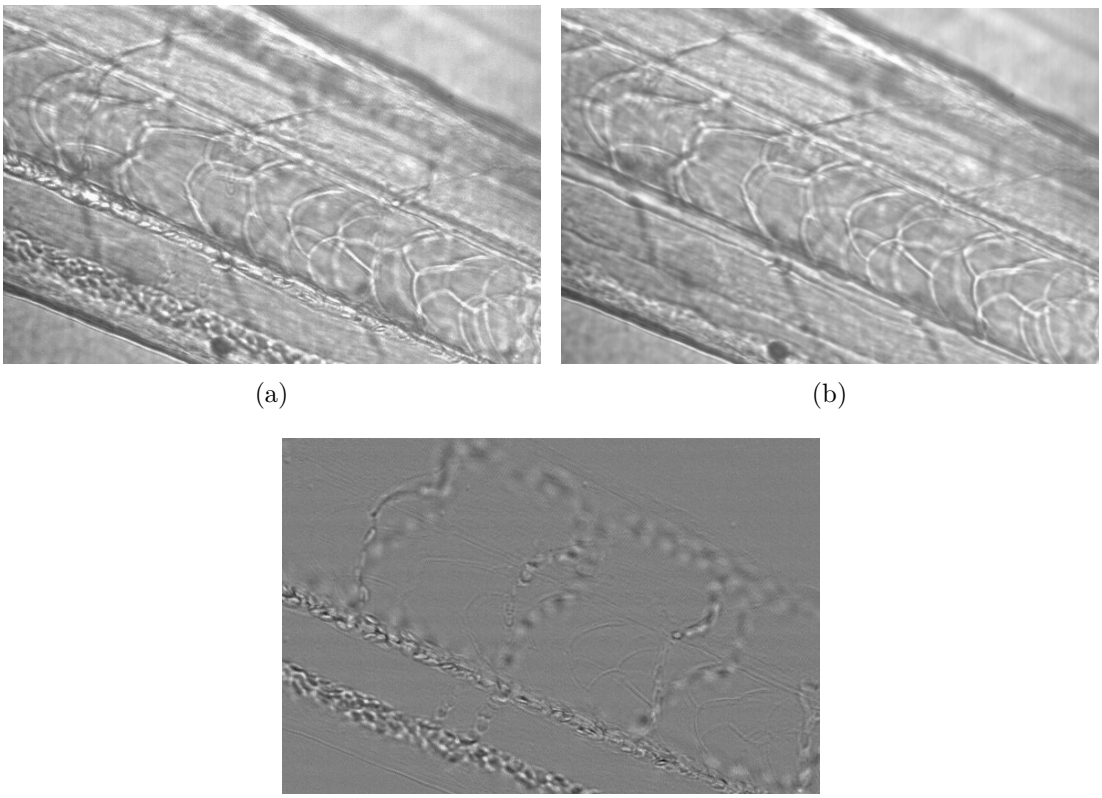
$$\forall I_{reg} \in \mathcal{S}_{reg} : I_{mov} = I_{reg} - \bar{\mathcal{S}}_{reg} \tag{6.1}$$

$\mathcal{S}_{mov}$ is the sequence of I_{mov}

6.2 Motion segmentation by temporal gradient

Our first segmentation protocol uses only motion criteria. Since motion components are sometime easily distinguishable from the background, and the motion is simple (e.g. a simple translation), sometime it may be efficient to segment motion areas regardless any other parameter. Let S_{mov} be the initial sequence. This sequence is precisely where the objects of interest are the moving components. The absolute value of the difference between two consecutive frames yields the motion between them.

$$\forall i \in [0:10], I^i_{diff} = |I^i_{mov} - I^{i+1}_{mov}| \tag{6.2}$$



(a) (b)

(c)

Figure 6.1: Detecting the moving motion components of an artifact-free sequence. (a) Initial frame. (b) Averaging of the sequence yielding the non-moving components, including all the parts that are not blood cells. (c) Areas with motion from the difference between frames and (b).

$\mathcal{S}_{diff}$ is the sequence of I^i_{diff}. We can simply integrate this sequence to obtain an image of the motion areas through the N first frames.

$$\mathcal{I}_{Sum} = \sum_{i=1}^{N} I^i_{diff} \tag{6.3}$$

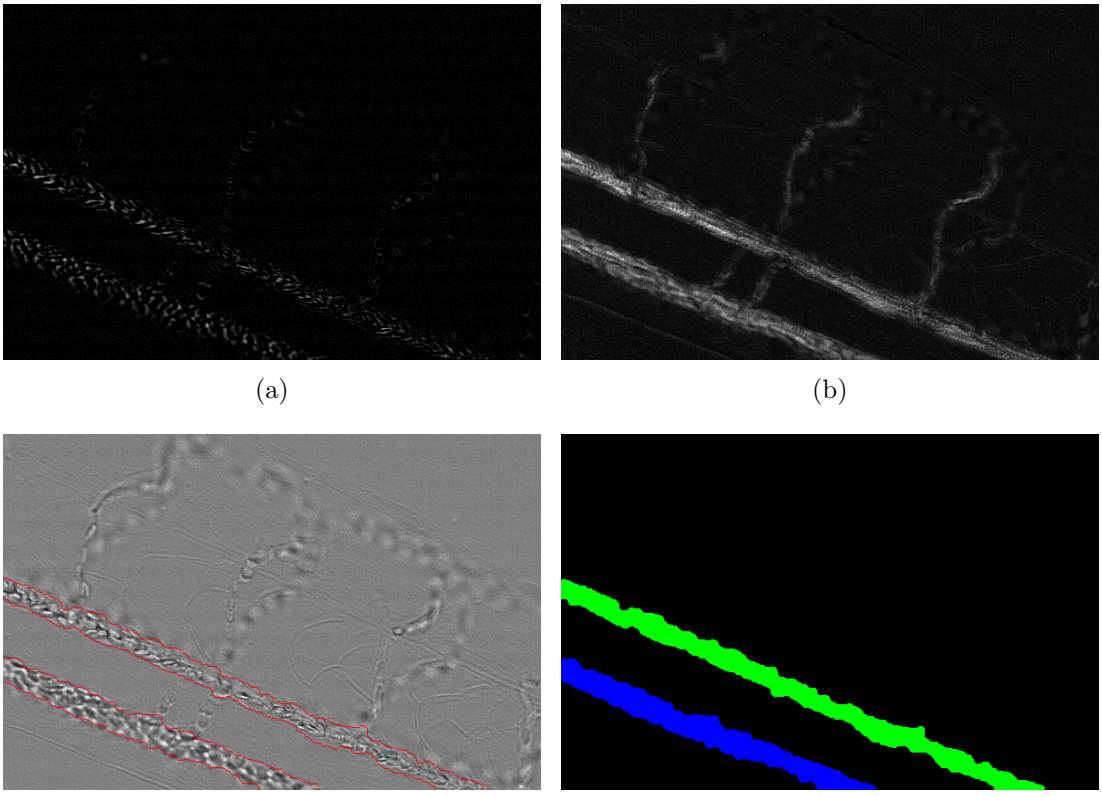
We binarize this image using the Otsu's criterion [89] θ_O: we threshold $\mathcal{I}_{Sum}$ at the θ_O value.

$$\mathcal{I}^T_{Sum} = (\mathcal{I}_{Sum})_{\geq \theta_O} \tag{6.4}$$

We thus obtain a binary image of the moving areas. Because the threshold criterion has no spatial regularity, the result can be noisy, hence we clean these areas with a morphological opening γ and a closing φ [53] with the same euclidian ball $\mathcal{B}$ of radius r to denoise $\mathcal{I}^T_{Sum}$, yielding $\mathcal{M}$.

$$\mathcal{M} = \gamma_{\mathcal{B}_r}(\varphi_{\mathcal{B}_r}(\mathcal{I}^T_{Sum})) \tag{6.5}$$

We then create $\mathcal{M}_L$ in which all connected component are distinguished with the attribution of a label (see fig. 6.2(d)). Obviously this method has many



(a) (b)

(c) (d)

Figure 6.2: Illustration of the sequence of operation for our motion-based segmentation. (a) Difference between two consecutive frames. (b) Integrating motion yields segmentation (c) and vessel labeling (d).

drawbacks. In particular, it requires the setting of many parameters that are problem-dependent and it is only suitable for simple motion. In our case we used it for vessel segmentation in the tail of alevins as in Fig. 6.2. In this example the motion is a simple translation, constant in direction and only changing in speed. As shown on Fig. 6.2(d) the result is quite satisfactory and stable in practice.

6.3 Motion segmentation by temporal variance

The previous method makes a number of assumptions with respect to the kind of motion that can be detected: motion is well confined in disconnected area, contours of motion are well defined and motion is easily detected. Here we propose a more general method based on the temporal variance that is less dependent on these assumptions. The temporal variance $\mathcal{V}$ of a sequence highlights the zones that exhibit motion. We first compute the asymmetric temporal gradient $\nabla_t I$, which is a sequence of the difference between consecutive frames. This gradient is very noisy but contains information about the motion components. We summarize the motion information by computing the variance on the sequence of temporal gradient. This yields a single image, which we blur using a Gaussian kernel in order to retain only

the most important motion. A formula for the whole process is as follows:

$$\mathcal{V} = \mathcal{G}_\sigma * \sum_{t=1}^{L-1} (\nabla_t I - \frac{1}{L-1} \sum_{t=1}^{L-1} \nabla_t I)^2, \tag{6.6}$$

where $\nabla_t I = I_{t-1} - I_t$ is the temporal gradient and $\sigma = 1$. For speed and simplicity, we used a marker-based watershed procedure [90] to segment the different areas of motion. This procedures require the computation of a gradient and of markers of the areas we want to segment. These marker areas can be small and their shape does not matter, however they need to be representative (or prototypic) of the final segmentation. Under this model it is possible to separate the *detection* of regions of interest and the *precise placement* of their contour. For marker detection, we use height maxima of the filtered variance image indicating the intensity of motion or a threshold. A high threshold indicates high motion, a low threshold, low motion, whereas the height maxima will keep even the slow motion area.

$$\mathcal{M} = (\mathcal{G}_\sigma(\delta_{\mathcal{B}_r}(\mathcal{V})))_{>p*\max(\mathcal{V})} \tag{6.7}$$

or

$$\mathcal{M} = H_h^{max}(\mathcal{G}_\sigma(\delta_{\mathcal{B}_r}(\mathcal{V}))) \tag{6.8}$$

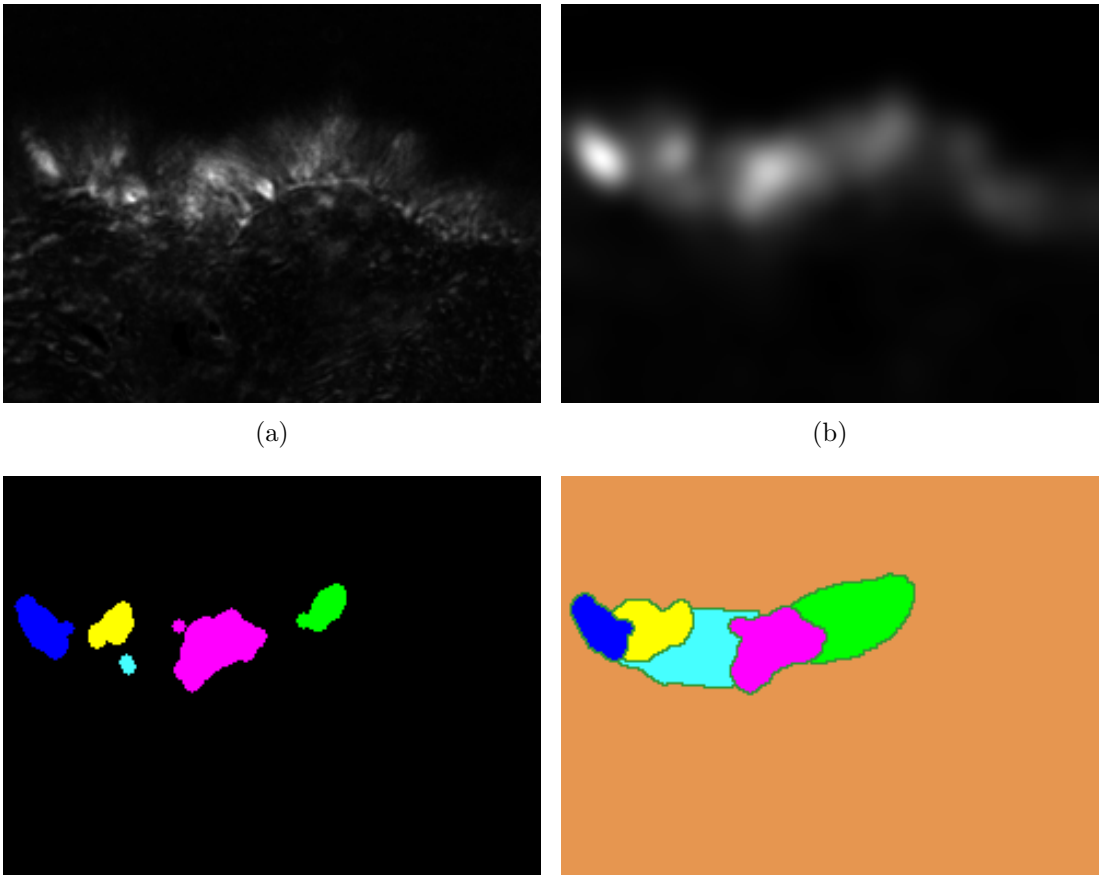
where $\delta_{\mathcal{B}_r}(I)$ is the dilation of I by an Euclidian ball $\mathcal{B}$ of radius $r = 3$. $I_{>\theta}$ denotes the thresholding of image I above θ, $p = 0.3$, $\sigma = 5$.
Internal markers are provided by $\varepsilon_{\mathcal{B}}(\mathcal{M})$ and external marker by $\varepsilon_{\mathcal{B}_3}(X \setminus \mathcal{M})$, where $\varepsilon_{\mathcal{B}}$ is the erosion by an Euclidian ball $\mathcal{B}$ and X is the image domain.

Computing the watershed of the gradient of the variance image leads to a mask denoted by $\mathcal{M}_L$. Figure 6.3 shows an example of the procedure and the segmentation, in which all the connected components are distinguished by a different (color) label. In these results we see that motion areas have been well segmented and that the contours of these regions are soft but sufficiently well defined for the watershed transform to provide a convincing result. However there remain a degree of arbitrariness to the thresholding criteria or the parameter of the height maxima, and solely the *quantity of motion* information is used. There is no characterization of the individual segmented areas, so their spatial differentiation could be an artifact. To counter this, we now seek a more sophisticated region-based method, where we develop and use descriptors of cyclic motion.

6.4 False motion elimination

6.4.1 Context

A cyclic motion is a motion that is repeated over time. If we take the motion definition from 5.3.1, cyclic motion can be described as a same grey-level variation that occurs periodically. Heart beat or motion of lungs during breathing are examples of cyclic motion. In this part, it is useful to differentiate cyclic motion from non-repeated motion. The former is typically something associated with biological processes (beating heart, blood flow, etc), whereas the former is more



(a) (b)

(c) (d)

Figure 6.3: Sequence of operations for segmentation. (a) Variance on the sequence of temporal gradient. (b) Blured variance (c) Markers determination (d) Segmentation result.

likely to be associated with motion artifacts.
Cyclic motion is very much at the core of the characterization of cilia motion for example, but also occurs when detecting a beating heart in a fish embryo.

However cyclic motion in living organisms is never perfectly periodic and can be subtle or irregular. Using a Fourier analysis would not necessarily be appropriate.

Here we propose a methodology for detecting approximately cyclic motion as opposed to noise or other types of motion that are not cyclic.

6.4.2 Methodology

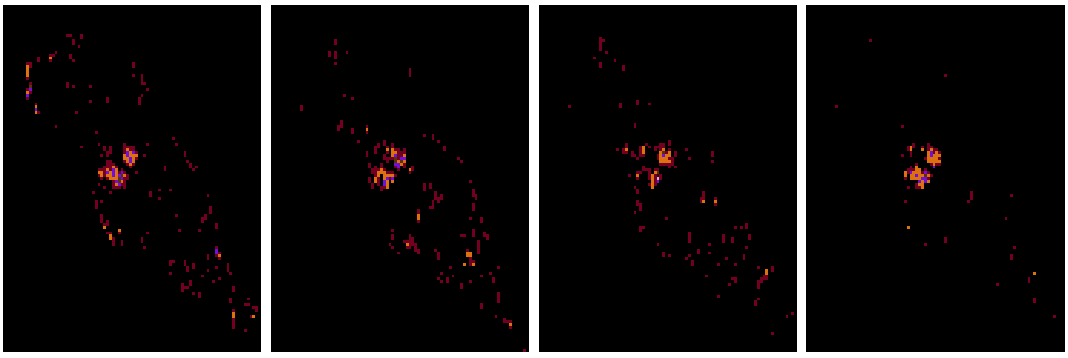
The particularity of cyclic motion is its recurrence. Let's consider a sequence $\mathcal{S}$. Computing the grey level variance at each pixel of a sequence along the time line show the grey-level variability for that pixel, however this variability may not be due to a cyclic pattern. If we split this sequence into sub-sequences long enough to contain one period, we will detect it in each sub-sequence. This is the idea behind the cyclic motion detection. We split $\mathcal{S}$ into n sub-sequences of l frames ($\mathcal{S}_1$, $\mathcal{S}_2$

... $\mathcal{S}_n$). The interval of l frames needs to be long enough to contain at least one period. We compute the pixels temporal variance on each sub-sequence, on the entire images or on a region of interest. This yields n variance images $\mathcal{V}_1$,$\mathcal{V}_2$ ... $\mathcal{V}_n$ (see Fig 6.4 (a)(b)(c)). We next compute the median of these n images $\mathcal{V}_i$. Residual variation occurring only once in the sequence will produce a large variance only once, and so the median of all the variances in this area will remain small (Fig 6.4(d)). On the other hand, non-spurious motion will exhibit significant variance in all n sub-sequences.

$$\mathcal{S} = \bigcup_{i=1}^{n} \mathcal{S}_i \tag{6.9}$$

$$\forall i \in [1, n],\ \mathcal{V}_i = \text{variance}(\mathcal{S}_i) \tag{6.10}$$

$$C = \text{median}(\mathcal{V}_i, i \in [1, n]) \tag{6.11}$$



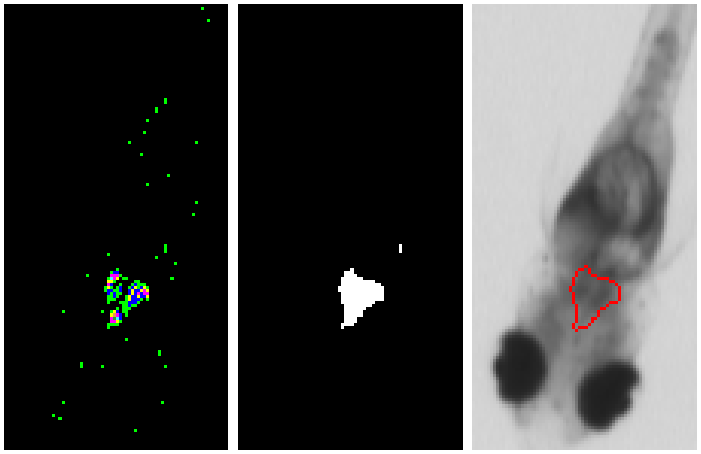
(a) (b) (c) (d)

Figure 6.4: False color rendering of the temporal variance. The brightest colors correspond to the highest values. (a), (b) and (c) are variances from sub-sequences, and (d) is the median of the three variances. We see that the areas that exhibit high variance in only one of the (a),(b) or (c) frames are eliminated in (d).

We can now segment the region where cyclic motion append with morphological operators. We want to connect the closest objects (for example the different parts of the heart) and to eliminate the small residual motion.

$$D = (\gamma^{\lambda_1}(\varphi_{\mathcal{B}_r}(\gamma^{\lambda_2}(C))))_{\geq 1}. \tag{6.12}$$

Here, $\mathcal{B}_r$ is a ball of radius $r = 4$, λ_1,and λ_2 are area parameters that depend on the objects of interest. This image can be thresholded at 1 to obtain a mask of the main cyclic variation areas and can be filtered again to remove areas that are not large enough. This filter reduces the false positive detection due to small remaining noise. Results are illustrated in fig. 6.5. In chapter 10, we apply this method for the detection of heartbeats. We have used this technique extensively and this has reduced the error rate by a factor of 2.



(a) (b) (c)

Figure 6.5: Segmentation of cyclic motion from C: (a) is the result of the median of the variance, (b) is the result of the thresholding before area filtering and (c) is the result of the segmentation

6.5 Frequency estimation

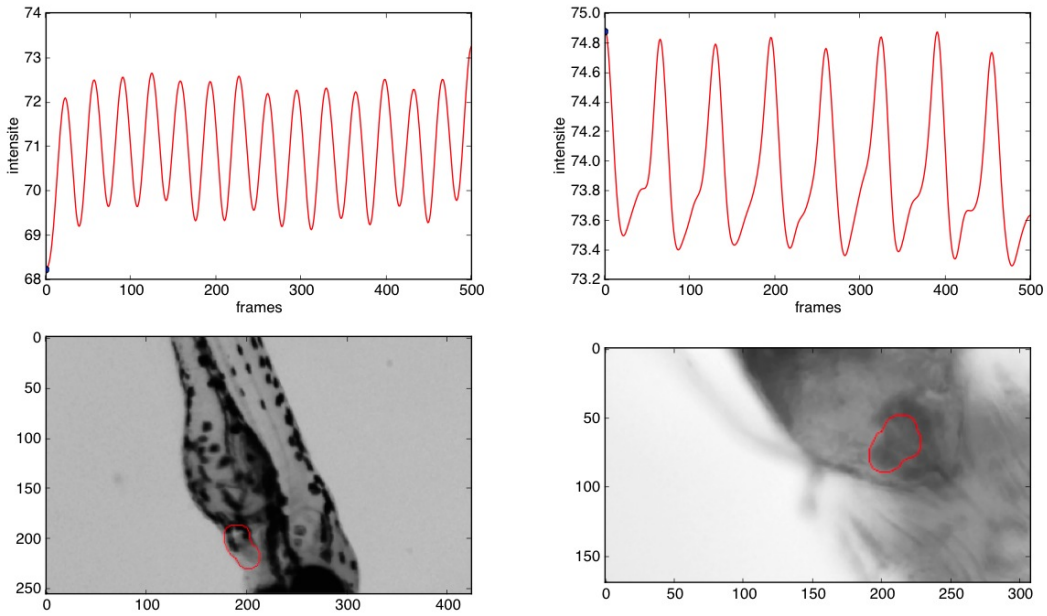
6.5.1 Semi-automatic grey-level intensity based frequency estimation

A simple way of estimating frequency from cyclic motion is to compute the average of grey level intensity in an area a and to follow this variation. Fig 6.6 illustrates the results on Zebrafish and Medaka hearts. Heart segmentation was first computed using the method of section 6.2.
However, this method is sensitive to noise and needs a spatially non-variant cyclic motion. We illustrated it in a feasibility study described in chapter 9.

6.5.2 Automatic optical flow based frequency estimation

As seen in the previous section, sometimes processes of interest manifest themselves via grey-level variations. For instance as the heart fills and empties, it becomes darker and lighter in appearance. However sometimes biological processes do not incur a variation in brightness, but it is still possible to detect variation in speed. To measure these, it is useful to consider the optical flow. Optical flow is the apparent motion of objects between two consecutive images. Three hypotheses are needed to use optical flow in the right way: object pixels intensities do not change between consecutive images; the acquisition rate is sufficiently high to ensure motion is smooth enough for differential calculus to be used, and neighbouring pixels show similar motion.

The dense optical flow methods provide motion information everywhere in the frame, even in areas with zero motion, whereas sparse techniques only show non-zero



(a) (b)

Figure 6.6: Grey level average intensity variation in the heart of (a) a Zebrafish and (b) a Medaka.

motion. Dense methods are known to be slower but may be more accurate than sparse methods. Because we are working on complex sequences of living models, various frequencies may be present all over the image. A dense optical flow method is preferable when coupled with a labelling of the zones with consistent motion and frequency. Indeed, before applying the optical flow, we first segment the areas of interest: we hence develop a combined method for frequencies estimation which allows the estimation of frequencies in each area of interest.

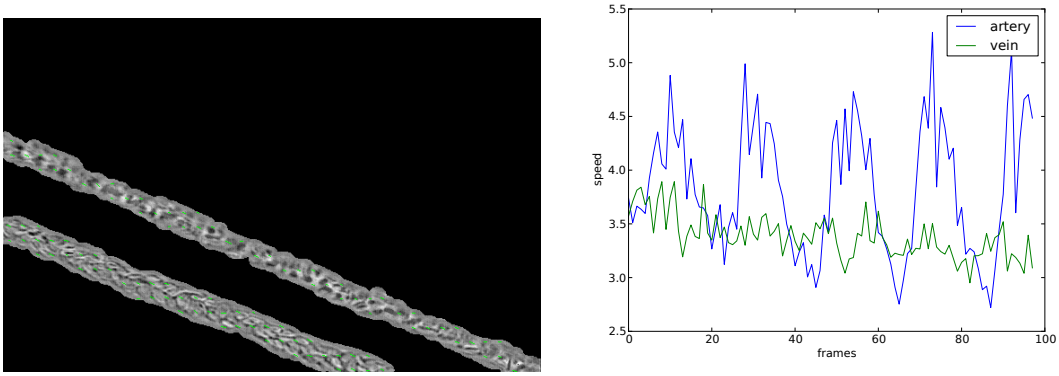
We use the Färneback's algorithm for computing the optical flow [79]. It works by dividing the frame with a grid, and calculating global motion in relation to its consecutive frame for each zone. The global motion is assigned to a point representing the center of that zone. For each point this method gives a translation vector that corresponds to the flow vector. We take the median value of the magnitude of all the speed vectors belonging to an area previously segmented to obtain the global translation in this area. We formulate the problem in this way:

$$V(i) = \bigcup_{\alpha \in M_L} V_\alpha(i) \tag{6.13}$$

$$\forall \alpha \in M_L, V_\alpha = \{v_i \in V / \mathcal{M}_L(x_{v_i}, y_{v_i}) = \alpha\} \tag{6.14}$$

$$\mathcal{V}_\alpha(i) = \text{median}(V_\alpha(i)) \tag{6.15}$$

$V(i)$ is the list of the translation vectors between two frames i and $i+1$. $V_\alpha(i)$ is the list of the translation vectors belonging to the component α of M_L, which is a



(a) (b)

Figure 6.7: (a) Representation of optical flow (in green) between two frames in vessels of a fish embryo tail and rough representation of the magnitude of translation vectors in vessel and artery (b)

labeled mask. For each region α, we have a median translation value $\mathcal{V}_\alpha(i)$ which is the median value of translation in the α region. We validate the choice of the median operator with speed vector clustering.

We identify average motion vectors with each segmented and labelled region. The plot of the magnitude of the speed vs. time in each labelled region allows us to identify useful physiological measurements, such as frequency, minimum and maximum speed, motion regularity and so on. Frequency analysis is performed using the Fast Fourier Transform.
Figure 6.7 illustrates the optical flow extraction in the tail of a fish embryo. If there is no significant differences between the blood flow representation between two frames on (a), by studying the evolution of that flow over time, we can distinguish the vein from the artery (b).
The main advantage of this method is that it can deal with several types of motion: the cyclic motion that is spatially non-variant such as heart that fills and empties, and spatially variant motion such as in blood vessels where blood cells move. We use this estimation on cilliated cells (presented in Chap. 8) and fish embryo heart and vessels where blood cells are visible (Chap. 11). For visual representation, the the frequency plots in these chapters are smoothed using a 1D Gaussian filter with a small σ.
On the three cases studies, we obtain precise frequency estimation using this method.

These tools were developed for the analysis of simple motion. They did not prove sufficient when we analyzed some more complex motion such as the description of cilia beating pattern. We hence developed two other procedures for this purpose. They are described in the next chapter.

7

Complex motion identification

The previously described motion analysis tools are useful and effective but do not provide a precise characterization of motion. We need to extract other complex parameters that will describe the motion such as period, amplitude, time pauses, intervals between two moving objects etc. These parameters are important in the cilia analysis. We present here tools that are designed for achieving such a complex motion estimation. We rely on a precise segmentation of moving areas that allows for a better analysis of object trajectories.

Contents

7.1 Feature-based region segmentation

Here we seek a motion segmentation method which is as much as possible parameter free and can distinguish zones of motion based on their characteristics. For periodic motion, which is of special interest to us in the case of cilia motion, we propose to use a multiscale Fourier domain analysis.
The challenge in this part is to obtain a segmentation of homogeneous moving areas. The watershed algorithm we used in the previous section yields plausible results, so it makes sense to continue using a marker-based algorithm in the same graph-based family. However simple watershed implementations does not lend themselves to a more sophisticated energy-based interpretation. In this section we formulate our motion-based segmentation as an optimisation problem. Remaining in the graph framework allows us to test various graph-based optimisation methods, namely the Graph Cuts, Random Walker, Shortest-path forest and Power Watershed. These were all shown to belong to the same framework in [68]. In this thesis, we only worked Random Walker.

7.1.1 Graph-based optimisation model

We take the definition of Section 5.2.2, using the Random Walker model (l=2 and k is arbitrary). We chose this model because it is easy to implement and can give good results while not requiring a high-quality gradient. The functional we wish to minimize is:

$$F(X) = \sum_{p \in V(I)} W^k(p)(x_p - y_p)^l + \sum_{(p,q) \in E(I)} {W^k}_D(p,q)(x_p - x_q)^l \tag{7.1}$$

7.1.2 Descriptors and weights

The processes we propose to study are assumed to be periodic. As we have seen, periodic motion in biological processes is only approximate. However, we assume here that the studied processes vary slowly, and so can be assimilated to true periodic processes for a short time. Hence we propose using windowed Fourier analysis, as is common in spectrograms in voice analysis, EEG, ECG and so on. The idea is to compare Fourier spectra in small neighboring regions. Let (p, q) be neighboring pixels. We consider a square window $\mathcal{R}_p$ (resp. $\mathcal{R}_q$) of size $(2r + 1)$-pixels centered on them (see Fig 7.1), and we calculate the FFT on each window. We named V_p and V_q the descriptors associated with pixels p and q, described below.
We write the weight W this way:

$$W_{pq} = \exp -\beta \|V_p - V_q\|^2 \tag{7.2}$$

where W_{pq} denotes the weight on the edge between p and q. W is the diagonal matrix of $\{W_{pq}\}$ used to calculate L

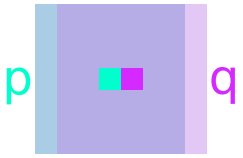


Figure 7.1: Illustration of neighbors centroids with their associated areas.

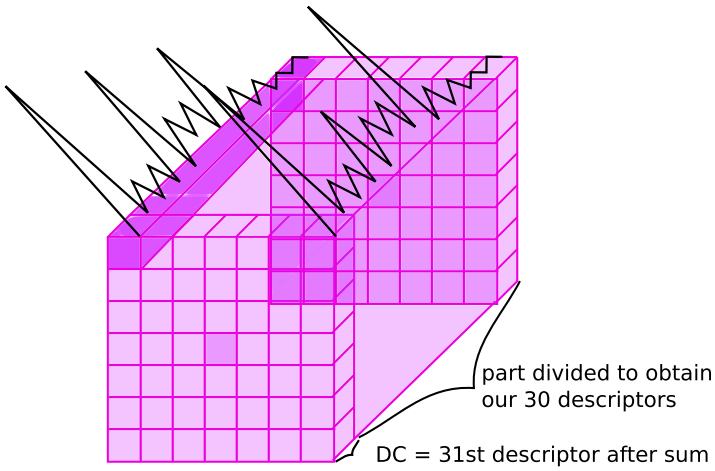


Figure 7.2: Decomposition of spectrum

Expression of V_p

V_p is a vector composed of N elements. From the local area $\mathcal{R}_p$ near a pixel, we propose to compute descriptors corresponding to subsets of the Fourier spectrum. We compute the FFT power spectrum on F frames for each pixel of $\mathcal{R}_p$. Each region R_p contains $(2r+1)^2$ spectra. S_i is the spectrum of the pixel number i. The final spectrum S_p of the centroid p is the sum of all the S_i limited to its first half (since the FFT spectrum is symmetrical) as illustrated on Fig. 7.2.

$$S_p = \sum_i S_i[1 : F/2] \tag{7.3}$$

The first descriptor is the DC component. We then compute E parameters from S_p. They are obtained by recursively dividing and averaging the vector S_p (see Fig. 7.3). In this way the resulting descriptor correspond to a wavelet-like multiscale representation of the average power spectrum of the region around the pixel of interest. If we subdivide S_p down to single values, we end up with N descriptors (DC + $N-1$). This vector is v_p. We then divided v_p by the total area of S_p to normalize the values.

$$v_p = [\int_a S_p da, \int_b S_p db, \int_c S_p dc, ..., S_p[0]] \tag{7.4}$$

$$V_p = \frac{v_p}{\int_{f_p} S_p df_p} \tag{7.5}$$

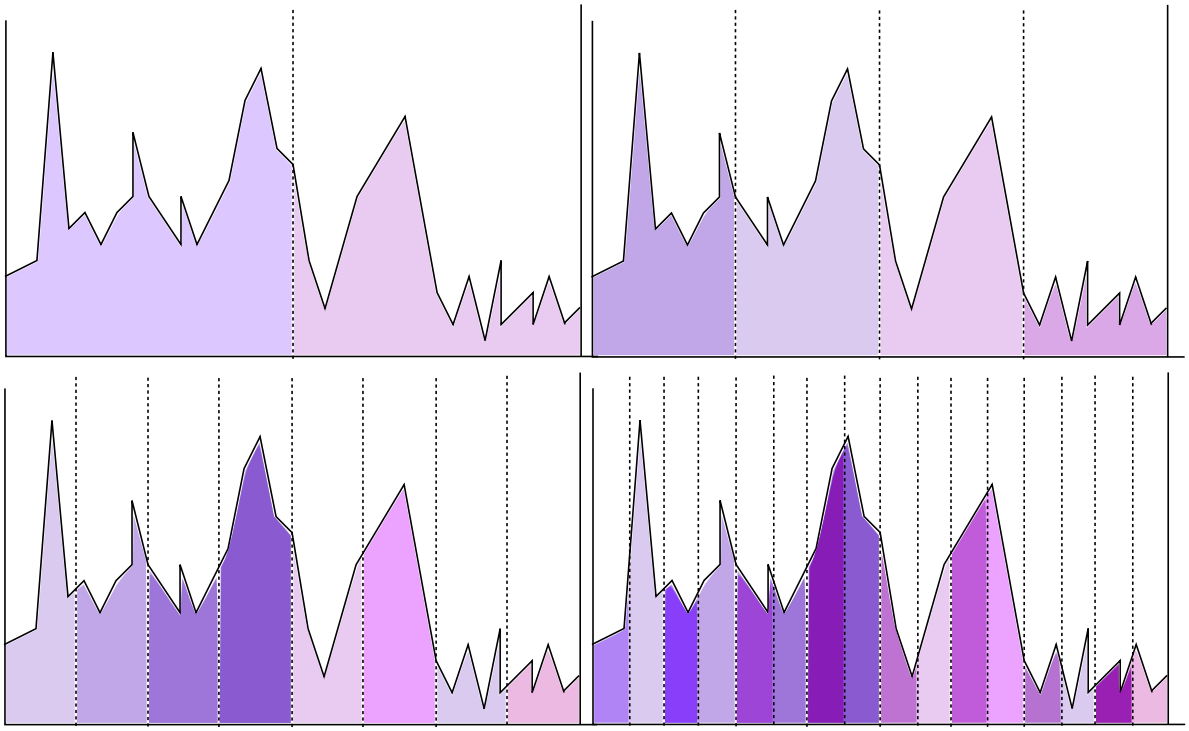


Figure 7.3: Decomposition of the FFT spectrum in 30 components for V_p

Comparison of V_p and V_q

In order to compare V_p and V_q, we subtract the two terms. We hence obtain:

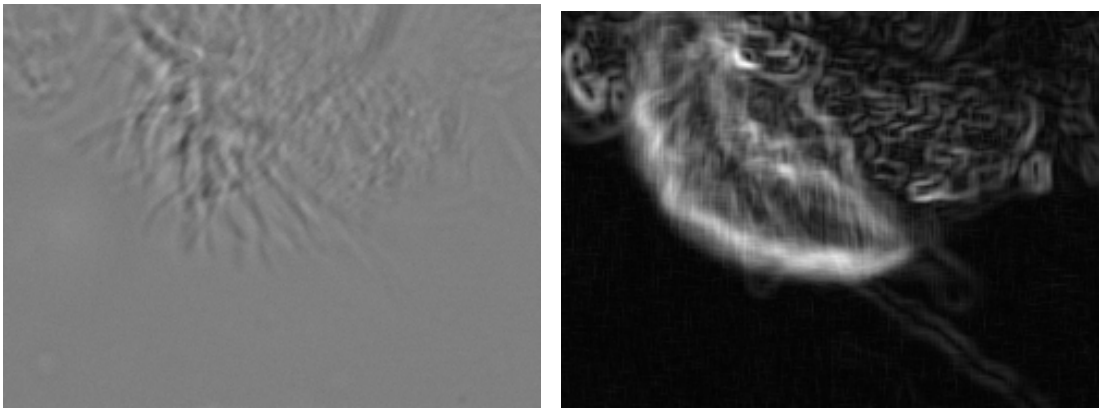
$$V_{pq} = \frac{v_p}{\int_{f_p} S_p df_p} - \frac{v_q}{\int_{f_q} S_q df_q} \tag{7.6}$$

Each term V_{pq} represents the similarity of the signal between p and q in the entire range of frequencies compatible with the time resolution of the FFT. We named V the image of $\{V_p\}$. V can be represented as a 3D image, of depth N. To deal with border effects, we only compute these descriptors r away from the border, resulting in a $(n - r \times m - r)$ image.
We can represent it using the finite difference vector gradient (right and bottom) G (see Fig 7.4). As we see, the regions with motion present are very well delimited by this gradient.

$$G[i,j] = \sqrt{(V[i,j] - V[i,j+1])^2 + (V[i,j] - V[i+1,j])^2} \tag{7.7}$$

Markers from unsupervised classification

The better the markers, the more accurate the segmentation. In particular, the markers have to be representative of their associated region.
We propose to use unsupervised classification on the descriptor image to determine the markers. Here we illustrate this idea with a k-means clustering method on the



(a) (b)

Figure 7.4: Representation of the non-centered gradient G (b) of the sequence of (a) using V_p

image V. The k-means clustering is useful in our case to obtain clusters. k-means clustering partitions data into k clusters according to their distance to a "reference" means. The k regions are associated to their centroïd or mean, which is built iteratively. Each point belonging to a particular region is closest to the centroid of that region than any other's. This assumes a clear-enough separation between groups when we plot two vectors belonging to two sets against each other. In our case, the k-means is a suitable choice such as shown in fig. 7.5 where we plotted each descriptor against the DC and attributed colors according to the k-means result for k=2.
The k-means method attributes a label to each pixel according to the number of classes we choose. The only criterion is the notion of descriptor similiarity, there is no notion of spatial consistency (see Fig 7.6(a)). We obtain clusters where each pixel is attributed to a class without spatial constraints. Due to the lack of spatial consistency, the result is not entirely satisfactory, with many isolated small areas attributed to the wrong cluster. However, this is sufficient for the purpose of obtaining markers. An erosion of the k-means image removes the small regions and provides good markers (see Fig 7.6(b)).

Segmentation

Since the non-supervised k-means already provides us with a plausible result, the choice of l in 5.7 is not critical. For simplicity, we chose $l = 2$ in equation 5.7, corresponding to the Random Walker algorithm. We now solve equation 5.15: $\mathrm{L}_X\mathrm{X} = -\mathrm{B}^\top\mathrm{Y}$ with Y and B obtained from the clusters computed before. The result X is a probability map where each pixel have a probability of belonging to one class or another.
We compare clusters in pairs. The matrix Y is composed only of 0 and 1. When there are only 2 classes, the result X can simply be thresholded at 0.5 to obtain the segmentation (see Fig 7.7). When we have z classes, we compare each class against all the others and the final result X contains z different values: each pixel is labeled

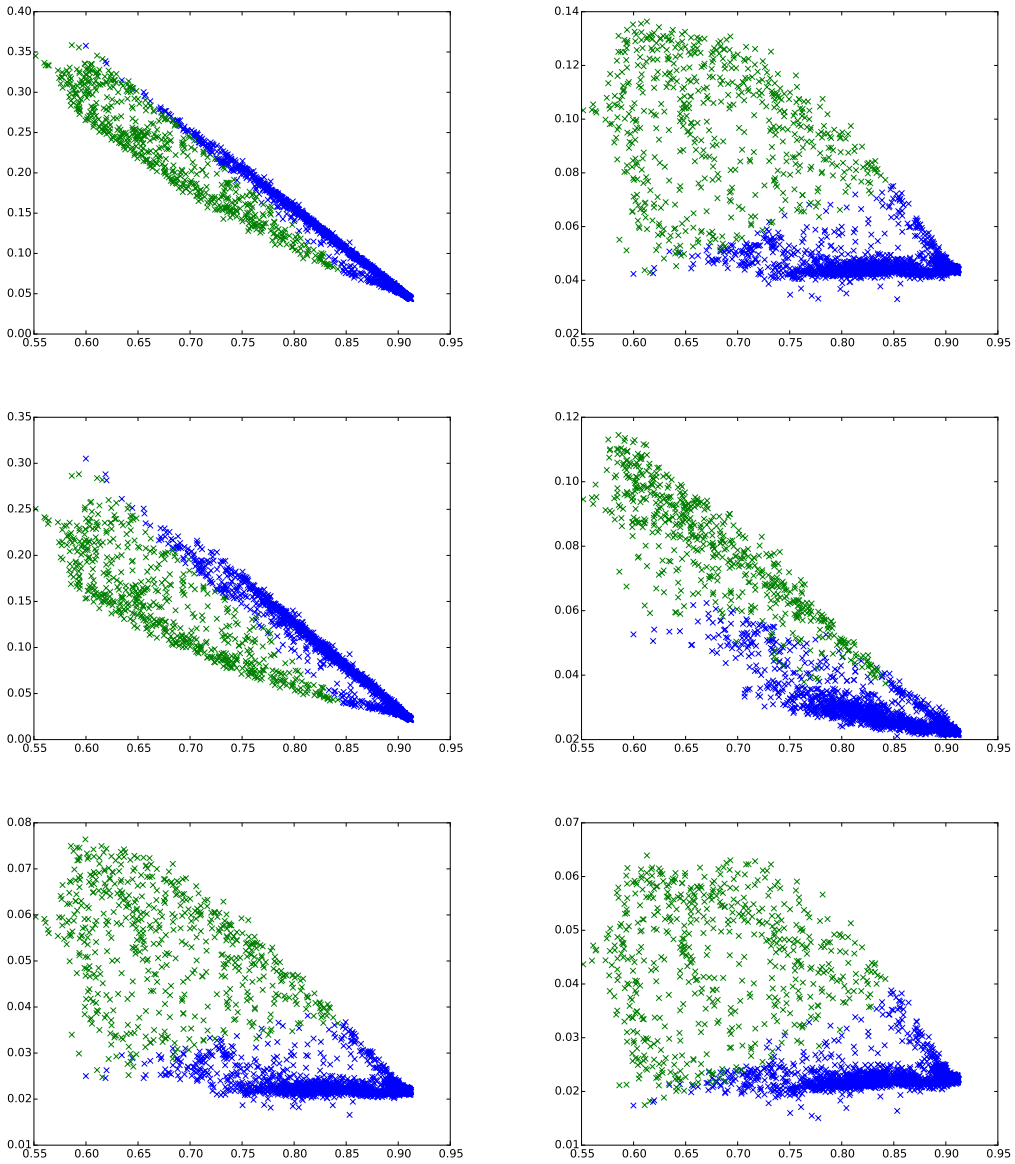
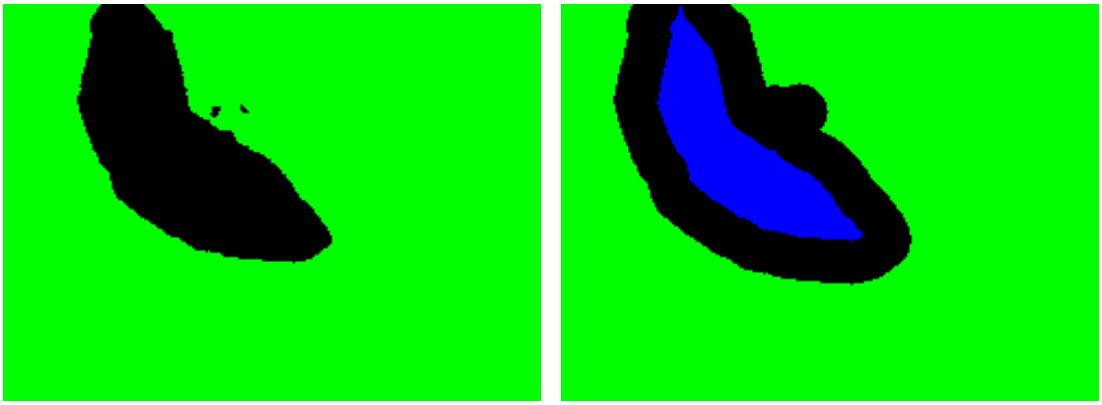


Figure 7.5: Validation of k-means procedure. These are the plot of the 6 first descriptors vs the DC component, and the colors are the two classes after the k-means algorithm.

according to the class with the highest affiliation probability.

7.2 Pattern extraction: Curvescan

After the classification step, we have segmented areas based on their motion content, but we have not identified this content according to parameters that clinicians find relevant. Particularly for cilia analysis, we need to identifiy various elements specific to the motion of cilia, that are not well captured by a pure Fourier analysis. Here we propose to extend the notion of the linescan, a semi-automated tool used in cilia



(a) (b)

Figure 7.6: Markers. (a) is the result of the k-means method with 2 classes asked, and (b) the result of the eroded (a).

(a) (b)

Figure 7.7: Result of segmentation. (a) is the probability map result of the random walker algorithm and (b) the thresholding of (a) at 0.5

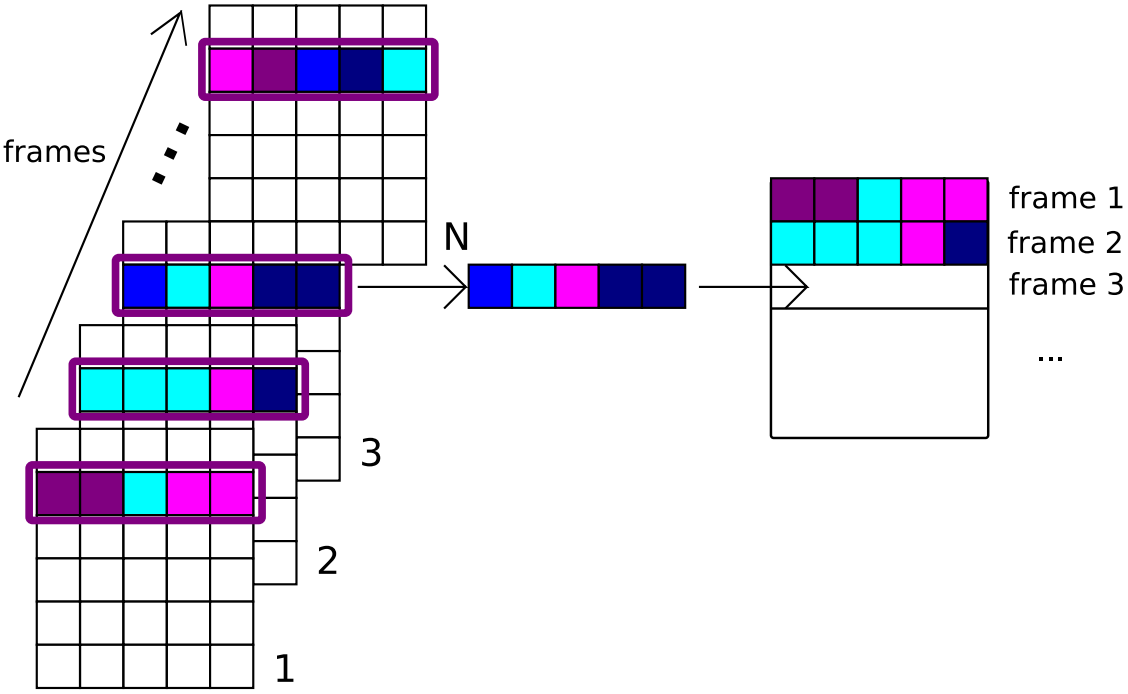
analysis, to something more flexible: the curvescan.

7.2.1 Principle

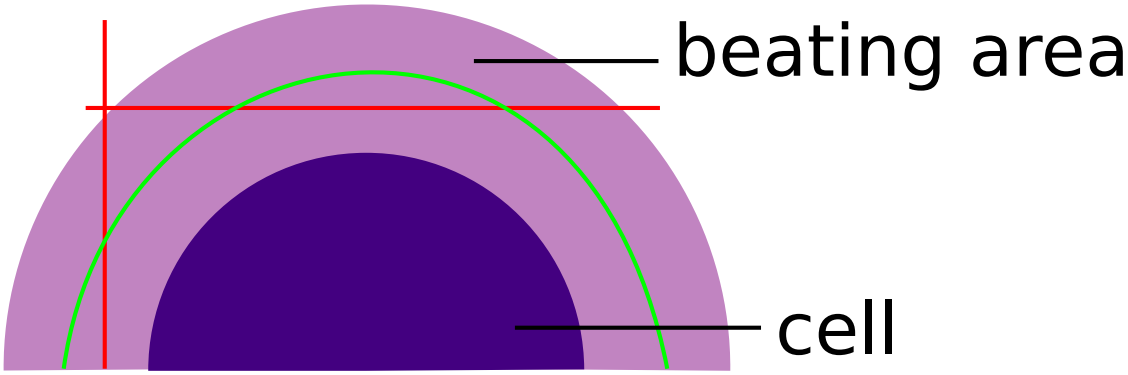
In this part, we are looking to extract the path of objects in a sequence after the area segmentation, using the shape of this segmentation. This method was developed for cilia analysis. The idea is to obtain a visualization of the objects over time along their path.

7.2.2 Linescan definition

The curvescan method is directly inspired from the linescan method, itself inspired from kymography. In these methods we record the variation of the pixels of a line over time, and we create a synthetic image of these pixels: each line l of the



(a)



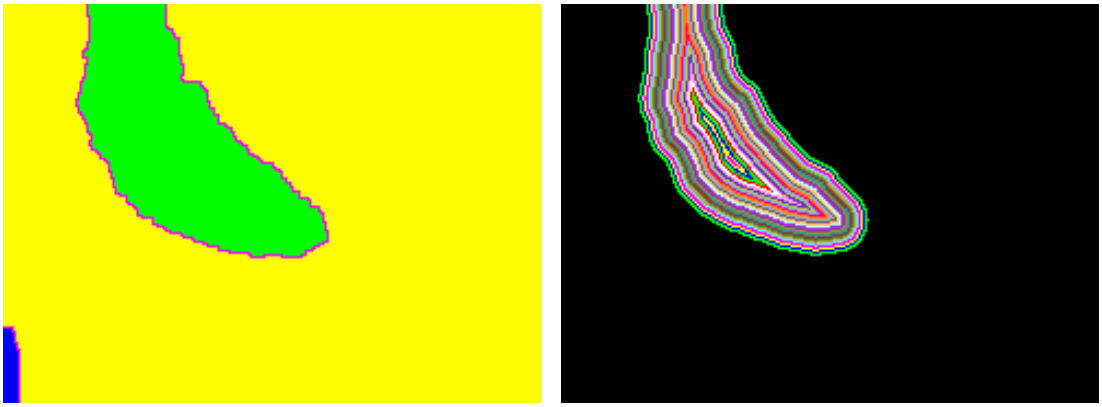
(b)

Figure 7.8: Illustration of linescan.(a) explanation of linescan and (b) linescan in red vs curvescan in green.

synthetic image corresponds to the line in the frame number l. This method is very useful to observe the pattern of trajectories of objects, but has a major limitation: it is applied only on straight lines whereas we study circular objects.
In our studies, we analyze objects with a smooth, curved boundaries, and so we needed to consider the path of cilia along “curves" instead of "lines" (see 7.8). Hence we developed the “curvescan", described hereafter.

7.2.3 Methodology

We first use the segmentation M_L obtained either in the previous section or in chapter 6 and we smoothed each motion zone a by a closing followed by an opening



(a) (b)

Figure 7.9: Extraction of level lines. (a) is the result of the previous segmentation and (b) the corresponding distance map result

of radius r, resulting in $a_f = \gamma_{\mathcal{B}_r}\varphi_{\mathcal{B}_r}(a)$, where γ is the opening and φ the closing. The value of r must be smaller than the diameter of the smallest object that we need to follow, so they are all preserved by the filtering. For each region a_f, an interior Euclidean distance map [91] D_f is computed. Each pixel of D_f belonging to a_f has for value the minimal euclidian distance to the complement of a_f (see Fig 7.9).

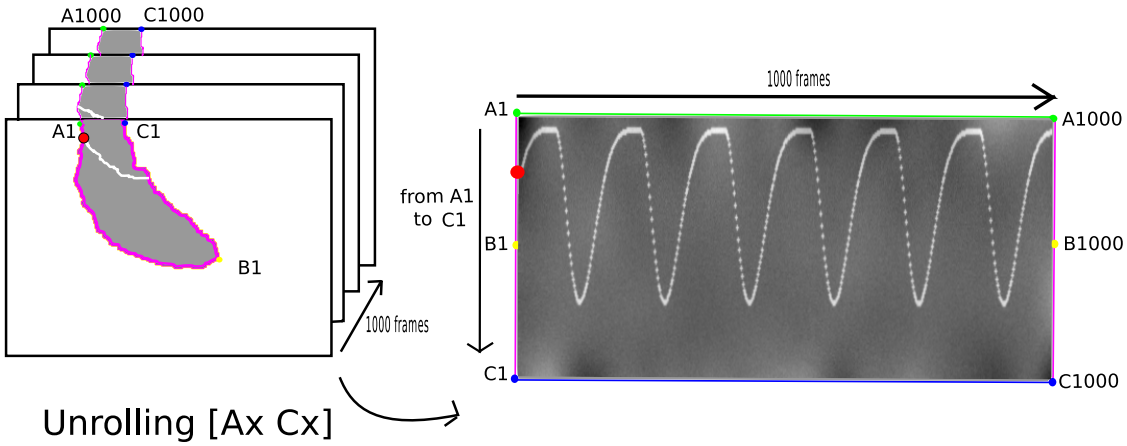
$$\forall x \in a_f, D_f[x] = \arg\min_p(\text{distance}(x, p)) \tag{7.8}$$

where p are the external contour of a_f. This global filtering and distance map simplifies the region contours, removes regions that are too small to contain your object of interest, separates regions and guarantees that the first r level lines of D_f are connected.
Each level-line of D_f allows us to "unroll" the region a_f at a different distance from the complement of a_f (see Fig. 7.10.a). Starting from an extremity point of a level line L_v at distance v in D_f, we follow this curve around a_f in an 8-connected fashion, recording the pixel coordinates as we go. We stop when we encounter the edge of the image. We create a new image I_{L_v} of size (length of the curve) $\times$ (number of frames in the sequence), where each column m corresponds to the grey level values of L_v in the frame m.

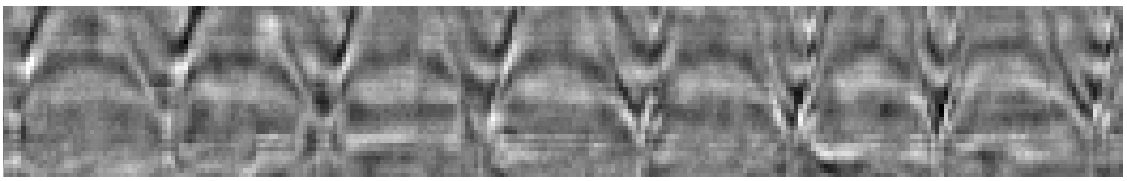
$$I_{L_v}[n, m] = I^m_{mov}[\mathcal{C}_v[n]] \tag{7.9}$$

Where I^m_{mov} corresponds to the frame number m of I_{mov}, $\mathcal{C}_v[n] = [n', m']$ is the coordinates of the point number n of the line L_v in I^m_{mov}. We subtract the output of a spatial median filter with a window size of 11×11 to eliminate illumination variation and only retain thin objects. We denote I_v this resulting image (see Fig. 7.10.b).

In the case where there are no extremities, we begin at an arbitrary point, and we stop when we encounter this point after two loops. The second loop ensures



(a)



(b)

Figure 7.10: Example of grey level extraction.(a) Schematic example of unrolling of the first line in distance map (b) I_{Lv}^{med} $(= I_v)$ is an example of curvescan result after median filtering subtraction.

that we record complete trajectories.

We presented here our proposal for a complex motion analysis procedure. Results associated with this methods are shown in chapter 8. In the next chapters, we described our proposed analysis pipelines corresponding to our applications.

Part III

Application: cilia motility evaluation

8

Cilia Beating Analysis

In this chapter we propose a fully automated method to characterize the motion of cilia from high-speed video microscopy. Our methodology is mostly based on Fourier analysis and so readily estimates frequencies.
The idea is to analyze not only beating frequencies but also their pattern. Indeed, simple frequency analysis is not sufficient to provide practitioners a practical tool. The advantage of our approach is its capacity to automatically compute robust, adaptive and regionalized measurements, *i.e.* associated with different regions in the image. We describe two different methods, which nevertheless have identical initial steps.

These two works were respectively published at ISBI 2015 [1] and selected for oral presentation at ICIP 2016[2].

Contents

8.1 Pipelines

We present here the two pipelines were developed. The purple items are in common.

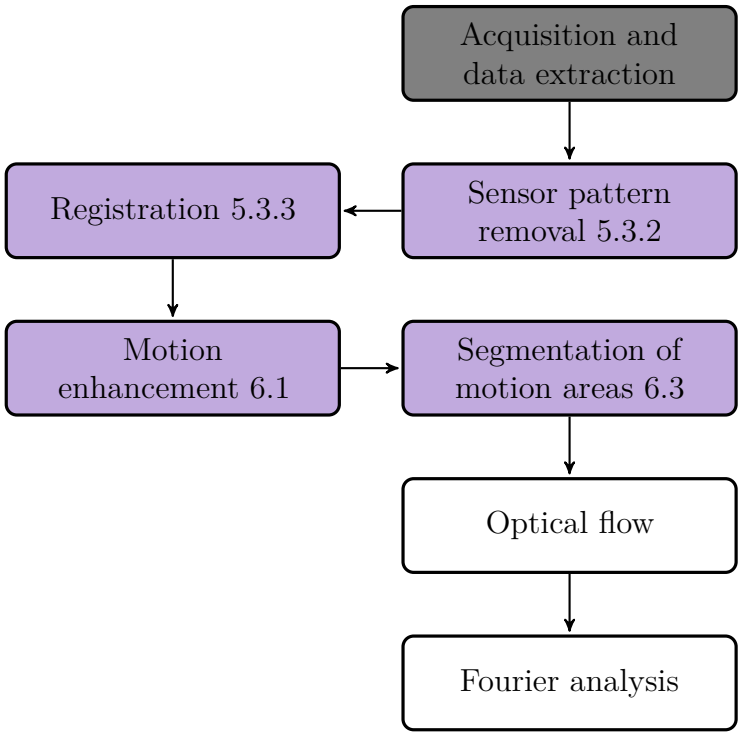


Figure 8.1: Flowchart of our cilia beating frequency estimation.

8.2 Details of the methodology: common parts

Let $\mathcal{S}$ be the sequence acquired under the microscope. We use a pre-processing procedure that consists of three main steps: sensor pattern removal, registration and motion enhancement. We have to first remove the sensor pattern due to the camera before stabilizing the sequence. Indeed, if we stabilize first, the fixed sensor pattern becomes noise.

Acquisition and data extraction.

We acquire videos from nose biopsies recorded under a microscope and we extract the frames with purpose-written software. We obtain the sequence $\mathcal{S}_{1800}$. In the remainder, we use the first 300 frames for analysis. We name this sub-sequence $\mathcal{S}$. The samples were obtained by nasal brushing of 10 patients with suspected cilia disease for diagnosis purpose, and from healthy volunteers. The brushing was performed by the ENT department under local anesthesia and sequences were acquired by the team 13 of Inserm U955.

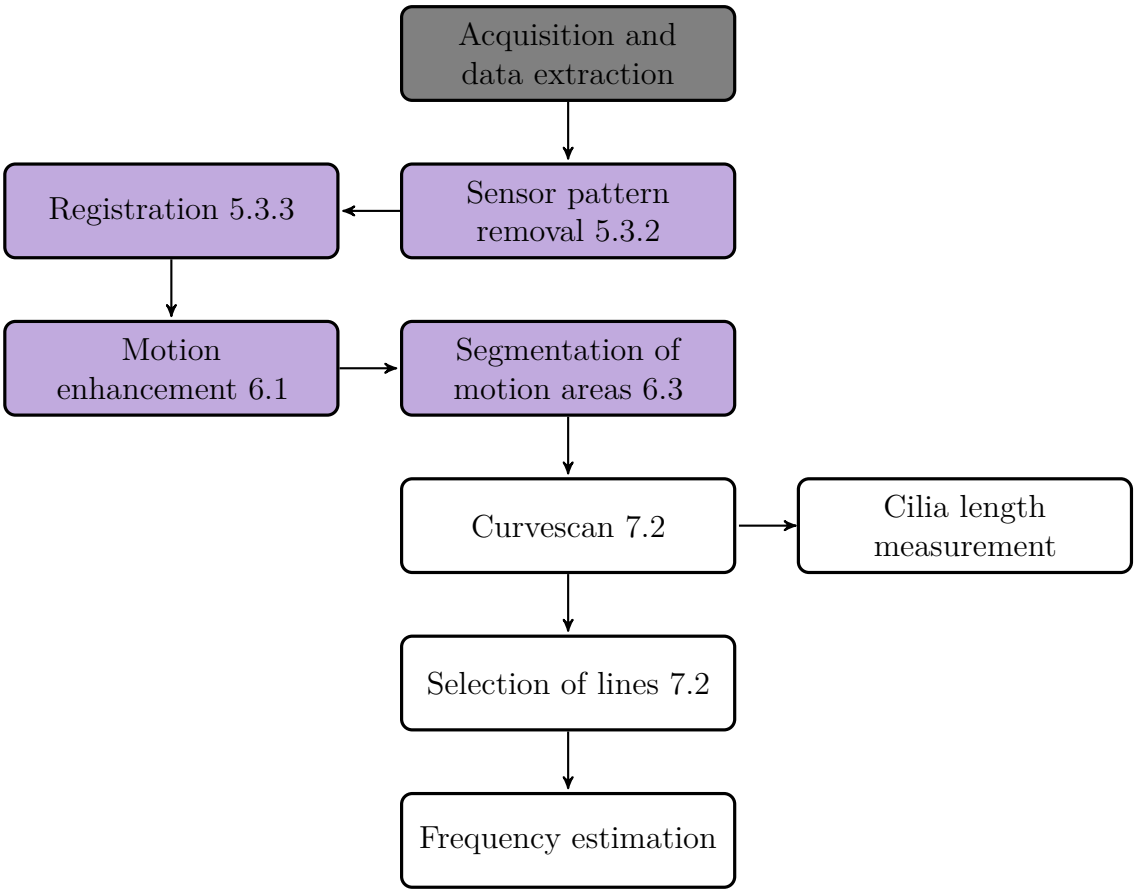


Figure 8.2: Flowchart of our cilia beating characterization steps.

Sensor-pattern removal (Chap. 5.3.2).

We apply our sensor pattern removal procedure described in 5.3.2 on the sequence $\mathcal{S}$ to obtain $\mathcal{S}_{clear}$

$$\forall I \in \mathcal{S} : I_{clear} = I - \left(\bar{\mathcal{S}} - \mathcal{G}_\sigma(\bar{\mathcal{S}})\right) \tag{8.1}$$

where $\mathcal{G}_\sigma$ is a gaussian filter with standard deviation $\sigma = 1$ and $\bar{\mathcal{S}}$ is the average of the sequence $\mathcal{S}$. This sequence $\mathcal{S}_{clear}$ is free from this artifact and contains the cells with beating cilia. (see Fig. 8.3)

Registration (Chap. 5.3.3).

Then we stabilize the sequence with our proposed registration procedure 5.3.3 to obtain $\mathcal{S}_{reg}$, the pattern-free stabilized sequence.

$$\forall (x, y) \in I_{clear} : I_{reg}(x, y) = I_{clear}(x', y') \tag{8.2}$$

and $[x'\ y']^\top = \mathcal{R}[x\ y]^\top + \mathcal{T}$.

We estimate $\mathcal{R}$ and $\mathcal{T}$ using our robust iterated regression:

$$P_1 = P_2 \times \mathcal{R} + \mathcal{T} \tag{8.3}$$

where P_1 and P_2 are matching sets of points, extracted using SIFT [81] or SURF [92]. We call $\mathcal{S}_{reg}$ the resulting stabilized sequence of I_{reg}. This sequence contains stabilized denoised cells.

Motion enhancement.

The final step of the pre-processing steps is the motion enhancement. Indeed, once the two previous steps are processed, the only remaining moving parts of the sequence $\mathcal{S}_{reg}$ should be the beating cilia. Applying our procedure yields the sequence of moving parts $\mathcal{S}_{mov}$ and so the sequence of cilia. The result is shown in Fig. 8.3.

$$\forall I_{reg} \in \mathcal{S}_{reg} : I_{mov} = I_{reg} - \bar{\mathcal{S}}_{reg} \tag{8.4}$$

$\mathcal{S}_{mov}$ is the sequence of I_{mov}. These steps are illustrated on Fig. 8.3.

At this stage, we have a stabilized sequence containing the highlighted beating cilia.

Segmentation of motion areas (Chap. 6.3).

After pre-processing, beating cilia are the only moving part in the sequence. Our objective is to analyse the beating pattern of cilia. We seek to apply our curvescan method, which should be the best procedure for our purpose. Hence, we need first to segment the beating areas. The temporal gradient highlights the pixels that move between two consecutive frames.

$$\forall I_{mov} \in \mathcal{S}_{mov} : I_{sub} = \mathcal{S}_{mov}[i] - \mathcal{S}_{mov}[i+1] = \nabla_t I \tag{8.5}$$

The temporal variance $\mathcal{V}$ of this sequence (Fig. 8.9.b) highlights the zones with motion. The temporal variance provides qualitative information about the beating of cilia: the variance is higher where cilia are beating faster. In this way we can obtain a preview of the beating areas of ciliated cells. We then apply a Gaussian filter on the variance leading to an image in which each zone with a similar beating intensity forms a white blob.

$$\mathcal{V} = \mathcal{G}_{\sigma=1} * \sum_{t=1}^{L-1} (\nabla_t I - \frac{1}{L-1} \sum_{t=1}^{L-1} \nabla_t I)^2, \tag{8.6}$$

where $\nabla_t I = I_{t-1} - I_t$ is the temporal gradient. We use a classical watershed-based morphological procedure [90] to segment the different areas of motion. The markers are selected from $\mathcal{M}$, a thresholding of a smoothed dilated image of the variance. We threshold at 30% of the maximum intensity, and the markers of the blobs are obtained by morphological erosion and dilation on $\mathcal{M}$.

$$\mathcal{M} = (\mathcal{G}_{\sigma=5}(\delta_{\mathcal{B}_3}(\mathcal{V})))_{>0.3*\max(\mathcal{V})} \tag{8.7}$$

where $\delta_{\mathcal{B}_3}(I)$ is the dilation of I by an Euclidian ball $\mathcal{B}$ of radius 3. $I_{>\theta}$ denotes the thresholding of image I above θ.

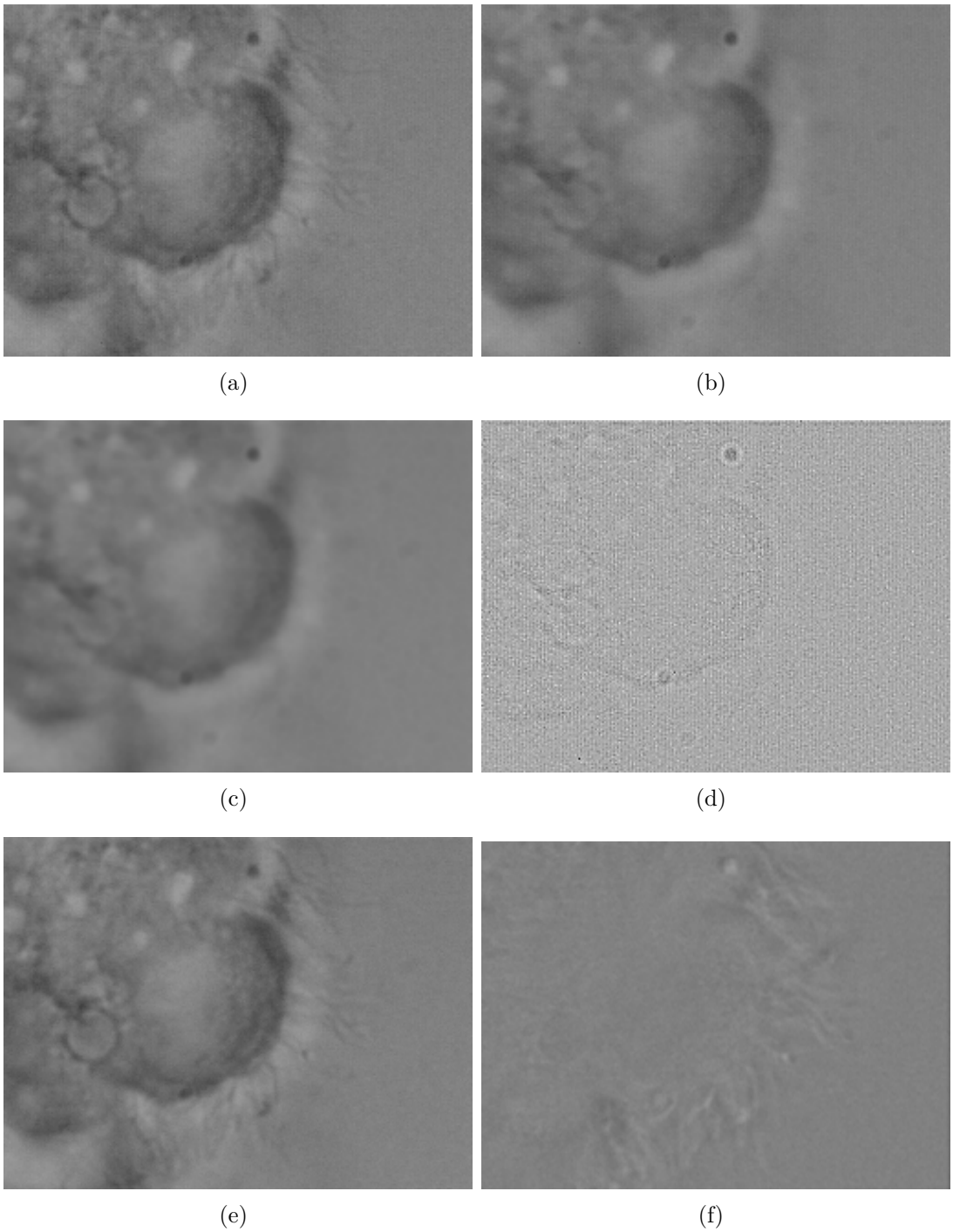


Figure 8.3: Removal of sensor pattern. (a) is the initial frame I of the sequence $\mathcal{S}$, (b) is the average frame of the sequence $\bar{\mathcal{S}}$ (c) is the blurred average $\mathcal{G}_{\sigma=1}(\bar{\mathcal{S}})$ (d) is the sensor pattern obtained by subtracting $\mathcal{G}_{\sigma=1}(\bar{\mathcal{S}})$ to the average $\bar{\mathcal{S}}$, (e) is the result I_{clear} of the subtraction between the sequence and the sensor pattern $I - (\bar{\mathcal{S}} - \mathcal{G}_{\sigma=1}(\bar{\mathcal{S}}))$ (f) is the extraction of the moving parts of the sequence $I_{mov} = I_{reg} - \bar{\mathcal{S}_{reg}}$, using registered images I_{reg}.

Internal markers are provided by $\varepsilon_{\mathcal{B}}(\mathcal{M})$ and external marker by $\varepsilon_{\mathcal{B}_3}(X \setminus \mathcal{M})$, where $\varepsilon_{\mathcal{B}}$ is the erosion by an Euclidian ball $\mathcal{B}$ and X is the image domain.

The watershed of the variance yields a mask denoted by $\mathcal{M}_L$, in which all the connected components are distinguished by a different (color) label. This segmentation shows the areas in which cilia have a similar intensity of displacement.

We can also segment the regions and obtain $\mathcal{M}_L$ using the feature-based segmentation presented in 7.1. We solve:

$$F(X) = \sum_{p \in V(I)} W^k(p)(x_p - y_p)^l + \sum_{(p,q) \in E(I)} W^k{}_D(p,q)(x_p - x_q)^l \tag{8.8}$$

using $k = 1$, $l = 2$. $\mathcal{M}_L$ is the result of the segmentation. Results are illustrated in Fig. 8.4. Here, we have two distinct beating areas.

8.3 Methodology for frequency estimation

In this section, we estimate cilia beating frequency using optical flow and Fourier analysis.

8.3.1 Methodology after the segmentation

Optical flow.

We compute the dense Farneback optical flow the entire sequence. In each of the previously segmented regions, the median of the vectors associate with a region provides an estimate of the displacement of that region.

Fourier analysis.

Frequency is then estimated via a Fourier analysis of the speed variation (*i.e.*, the norm of the displacement vector) over time (Fig. 8.5).

8.3.2 Results and Validation.

We analyzed 10 annotated nasal brushing samples from patients of the ENT department of *Centre Hospitalier Intercommunal et CHU H Mondor* (Créteil, France). The correlation between our estimations and the ground truth is illustrated in Fig. 8.6. The “Ground truth" is the plot of the Kymography versus the Cinematic analysis that are our references. We can see that one outlier is present in our estimations, but even with the two reference standards we can observe outliers from the 5% confidence interval, meaning that our measurement are as reliable as the two other methods.

This is confirmed with the Bland-Altman diagram when comparing the methods in pairs. Bland-Altman diagrams show a repartition of the distance between our method and the two methods of reference (cinematic analysis and kymography, see Fig. 8.7). We can observe that our frequency estimations are all contained in the 95% confidence interval when compared with the kymography method. By

comparison with the cinematic anaysis, only one measure is out of the interval, which remains acceptable.

Limitations.

In all of our sequences, we can only validate one frequency and so one estimation. For now, we only have one reference frequency per sequence, corresponding to the one cilium that was manually studied in that sequence. Further annotations will help with the validation of every area.
With the presented method, we only estimate invidual frequencies, which is not sufficient for a precise beating analysis.

8.4 Methodology for cilia beating characterization

Our objective is to characterize the beating of cilia. We wish to estimate or estimate all the parameters currently measured by practitioners: cilia beating frequency, cilia length, amplitude of beating, distance cross by cilia, pauses during the beating. We applied our curvescan procedure to measure these parameters (exemplified in Fig. 8.8). Because cilia beat in 3D, they come in and out of focus during their motion. They also beat in groups. This makes tracking an individual cilium very challenging. The idea of the "linescan" technique is to only observe the variations of intensity in a given narrow, elongated region (typically a line) encompassing the perceived cilia motion. Since the cilia motion is pseudo-periodic, so should be the intensity variation. Here we seek to specify this scanning region automatically, and we do not limit our acquisitions to a line segment. Cilia located on sampled cells are rooted along the surface of the cell. Contrary to the experimental conditions of [24], where cultured ciliated cells are located on a flat surface, the beating extremity of cilia form a curve. Since cilia may be of varying lengths, it is useful to consider not just a single curve but several from the root to the tip of the beating region. By analogy to the linescan technique, we call this new one the adaptive *curvescan*. In this section, we present the frequency estimation and cilia measurement.

8.4.1 Methodology after the segmentation

Curvescan and selection of lines (Chap. 7.2)

Each segmented motion zone a of M_L is smoothed by a closing followed by an opening of radius r, resulting in $a_f = \gamma_{\mathcal{B}_r}\varphi_{\mathcal{B}_r}(a)$, where γ is the opening and φ the closing. An interior Euclidean distance map [91] D_f is then computed from a_f. The value of r has to be chosen to correspond to a diameter smaller than the smallest cilia, so they are all preserved by the filtering. We chose $r = 15$ which corresponds to a 4μm cilium. This filtering simplifies the region contours, removes regions that are too small to contain cilia, separates regions containing different groups of cilia and guarantees that the first 15 level lines of D_f are

connected. Each level-line of D_f allows us to "unroll" the region a_f at a different distance from the ciliated cell. Starting from an arbitrary point of a level line L_v at distance v in D_f, we follow this curve around a_f in an 8-connected fashion, recording the pixel coordinates as we go. We stop when we encounter the edge of the image or the initial point after one loop. We create a new image I_{L_v} of size (length of the curve) $\times$ (number of frames in the sequence), where each column m corresponds to the grey level values of L_v in the frame m.

$$I_{L_v}[n, m] = I^m_{mov}[\mathcal{C}_v[n]] \tag{8.9}$$

Where I^m_{mov} corresponds to the frame number m of I_{mov}, $\mathcal{C}_v[n] = [n', m']$ is the coordinates of the point number n of the line L_v in I^m_{mov}. We subtracted the output of a spatial median filter with a window size of 11×11 to eliminate illumination variation and only retain thin objects. We denote I_v this resulting image (see Fig. 8.10).

Frequency estimation

We estimate the power spectrum $\mathcal{S} = |\mathcal{F}|^2$ by a power spectral density method, exemplified in [24] as the curvescan is an adaptation of their method. It consists of computing the Fourier transform $\mathcal{F}_n$ on each line n of I_v and averaging the square modulus of all the $\mathcal{F}_n$. The cilia beating frequency is then estimated using a parabolic approximation of the main peak and its potential neighbors, in the range of plausible frequencies for cilia (i.e. 0-30Hz).

$$\forall k \in [0, M-1] : \mathcal{F}_n(k) = \sum_{g=0}^{M-1} I_v[n, g] e^{-j2\pi \frac{gk}{M}} \tag{8.10}$$

$$\mathcal{S}(k) = |\mathcal{F}(k)|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |\mathcal{F}_n(k)|^2 \tag{8.11}$$

where M is the number of columns of I_v (number of frames of the sequence), and N is the number of pixels of L_v. For robustness evaluation, we analyze 3 lines of each region, to compare them (each line should return around the same beating frequency), and to take their average as result for frequency estimation and validation. For each sequence, we estimate the cilia beating frequencies over the level lines L_v 2, 7 and 14 of D_f, which are all connected by construction thanks to the precautions taken during the curvescan building, and we take the average as our estimation. These three lines are representative of the difference between inner, outer and middle areas of the beating region. Note that the fact that different level lines have differing length has no bearing on the frequency estimation since the time axis has the same length. Also note that we do not need to discriminate on the image I_v along the vertical axis for areas that correspond to zones close to the root or near the tip of the cilia, since due to the continuity of the cilia, the pseudo-periodic motion occurs consistently.

Cilia length measurement

The regions of cilia motion segmented in $\mathcal{M}_L$ represent the zones where cilia are located. The width of the mask is representative of the length l of the beating cilia. In order to obtain precise measurements, we compute a distance map D_f^Q on labels of M_l using a quadratic Euclidian distance [91]. The maximum value of D_f^Q inside the region of interest is the square of the radius of the largest inscribed disk, and thus using this value we can obtain a good approximation of the width of the region, and therefore of the length of the cilia beating in it, by using the following formula:

$$l = \sqrt{\max(D_f^Q)} \times 2 \times 0.13 \tag{8.12}$$

where 0.13 is the spatial resolution (1 pixel = 0.13μm).

8.4.2 Results

Frequency

We computed our curvescan results on 11 annotated sequences of beating cilia. We tested our method on cilia exhibiting a variety of beating frequencies, between 6Hz and 14Hz. As stated above, for each sequence, we measured frequencies using 3 different level curves of the distance map, one at the border, one in the center and one in between them. We hence ended up with 33 measurements for 11 sequences. The frequency analysis of the power spectrum may reveal several frequencies present in a region. It may be a single real frequency and its harmonics, or several similar frequencies corresponding to cilia group beating at different frequencies (see Fig. 8.11). In all cases we only considered the frequency with the highest power in the spectrum.

We validated our results in two steps: we first checked the robustness of the method by comparing the values obtained for each line (see Fig. 8.12(b)). Then we validated the accuracy of our method by comparing our frequencies with those carefully estimated by experts (see Fig. 8.12(c)). We measured the average error rate in the classical way with the formula AverageError $= \frac{1}{J}\sum_{j=1}^{J}\frac{|m_j - t_j|}{t_j}$ with m_j our measurement and t_j the corresponding ground truth .
We measured the error rate of each measurement separately first, and then the error rate of the averaged estimations. Results obtained are shown in Fig 8.12(a).The mean errors are similar, about 2.2%.
Among our 33 measurements, 30 (or 91%) are within the confidence interval of $\pm$5%. However, if we consider the averaged frequency estimations and compare them to the expert estimations, the main frequency is estimated correctly in 100% of sequences. These results are promising for both the robustness and accuracy of our proposed method.

Cilia Length We computed cilia length measurement on our sequences using the method described before. A unique value is obtained for each sequence. An expert manually measured the length of cilia in a region. Their manual measurement was repeated 3 times. We took the average as ground truth for our validations.

We compared our measurements in two different ways (see Fig 8.13). The first is the equivalent of the frequency validation: we considered the expert measurement as perfect and we measured the *AverageError* rate. In this way, we obtained an error rate of 2.13%, with one measurement out of 11 outside the 95% confidence interval and a standard deviation on the error of 2.91%. The second way is the total least square regression method, which seeks a proportionality factor between the manual and automated measurements, taking into account the fact that there could be variability in the expert measurements too. With this method, assuming equal variance between the automated and expert measurements, we achieved an error rate of only 0.8%. The real error rate may be in between these two values. In any case these error rates are reasonably low and so our results are again promising.

8.5 Conclusion

8.5.1 Discussion

Preprocessing steps are meaningful in our context. Removing camera artefacts is necessary for the subsequent image analysis to succeed. The importance of stabilization can be highlighted by some of the image sequences. Indeed, one example shows a dead cell sequence. If the stabilization is not performed, both our algorithm and the classical techniques estimate a beating frequency of 12Hz. A closer look at the video seems to indicate that this frequency corresponds to the motion of the cell due to vibrations. After stabilization, no motion was longer visible and the recorded frequency was 0Hz, as expected. After thie pre-processing, any remaning motion should be that of beating cilia, that can be segmented and analyzed according to their motion pattern.
We have proposed a regionalized automated measurement of the ciliary beating frequency, capable of coping with several cell groups, each with their own beating pattern. Comparison with semi-automated kymography and cinematic analysis for frequency measurement in the same segmented areas shows that our results are significant.
We still need to confirm that our approach is indeed able to process more than one cell group in each field of view.
We believed that we can measure other characteristics beyond frequency from the segmented regions. In particular, an analysis of the qualitative components of beating patterns seemed achievable, including a full description of range, rhythm and structure. This is the reason why we have proposed another method.

This new method, the curvescan, currently records only the main frequency in a beating region but can be easily extended to record all relevant frequencies. At present we only estimate fairly simple characteristics, however our curvescan results opens the way for more interesting analyses. We are currently working on estimating the beating amplitude, regularity, symmetry, and so on from a better segmentation. Indeed, we need a more precise segmentation to analyze the beating pattern. We are currently combining our segmentation method presented in Chap.7.1 using the

Table 8.1: Comparison of our two methods.

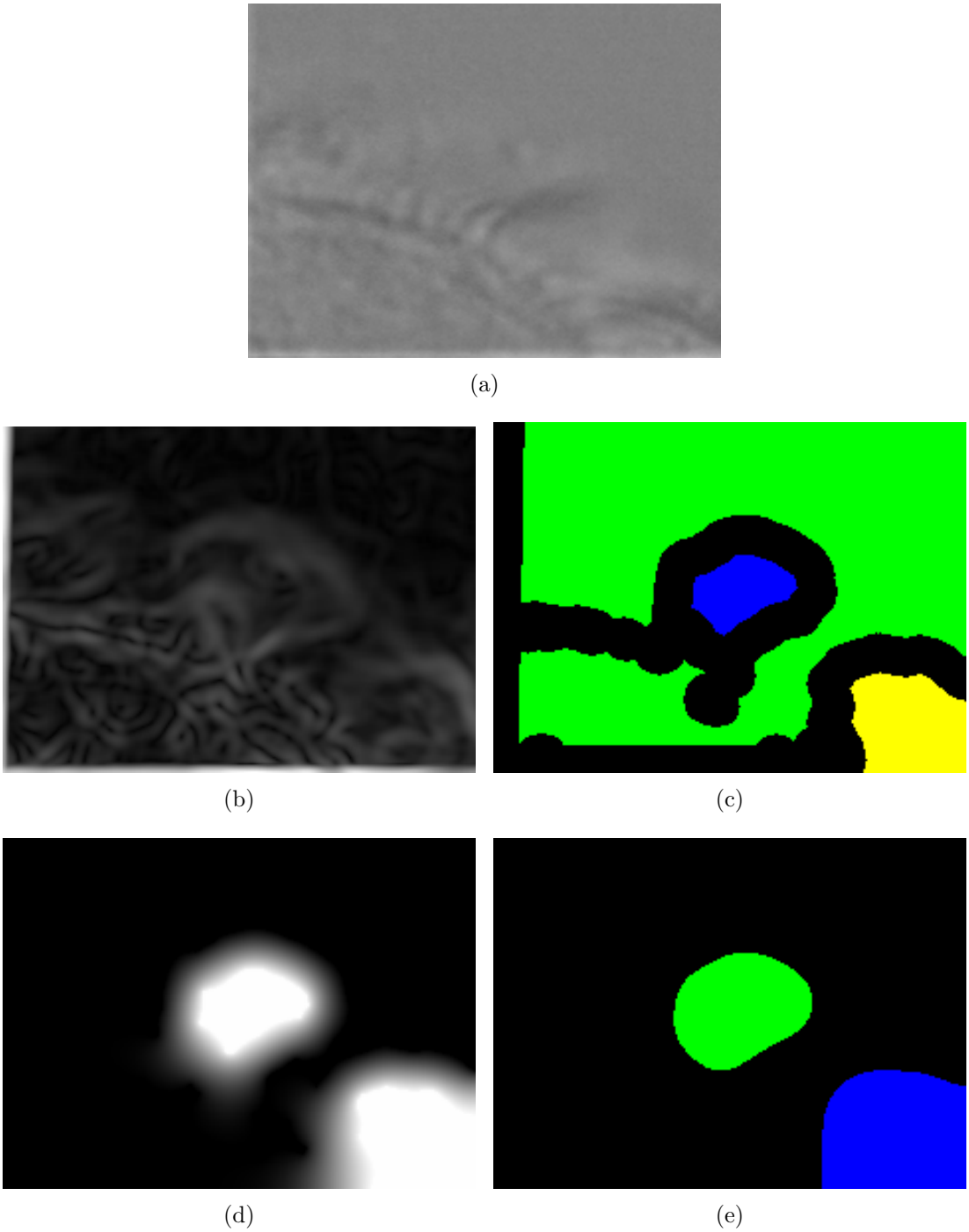
	Optical flow	Curvescan
Frequency estimation	Yes	Yes
Other parameters	No	Possible
Easy-to use	Yes	Needs quality segmentation
Detection of dead cilia	Yes	No

frequency into account to obtain more homogeneous and more precisely defined beating regions. For the analysis, the idea is to analyze the tracks of cilia (see Fig. 8.8) in the curvescan. With respect to this diagram, we are already measuring T_x and a subset of D_x, the distance crossed by cilia. For a more complete set of measurements, we would need to estimate A_x corresponding to the amplitude of beating cilia, and P_x the pause duration in the cilia motion. The combination of these parameters with T_x and the cilia length that we already measure, will provide most of the parameters necessary for disease characterization currently used in clinical practice [15].
A procedure of clinical validation will be necessary to complete this study.

8.5.2 Comparison of the two methods

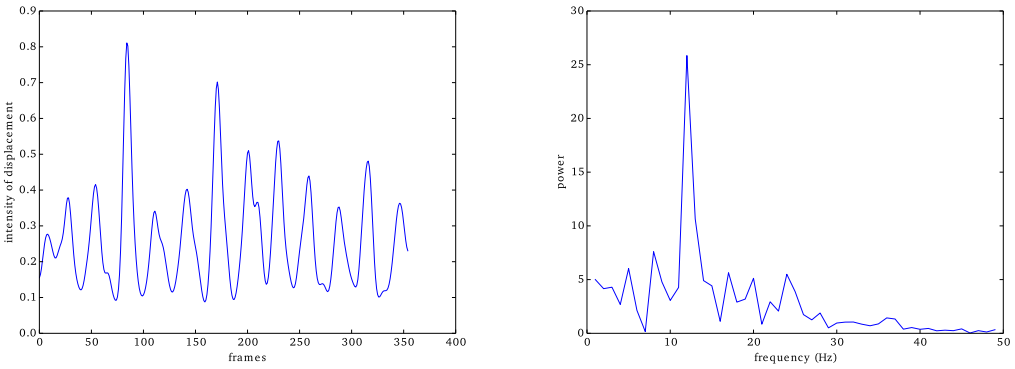
The two methods have different pros and cons. Table 8.1 helps to highlight them.

The curvescan is more complete than the optical flow method. But it requires a precise segmentation. The optical flow is less precise and estimates only the beating frequency, but can be used with a segmentation that is less precise and is more versatile, in the sense that the curvescan is only suitable for spatially cycly motion. However, the curvescan has a great deal of potential use in clinical practice, if it can replace the painstaking manual measurements constituting the current gold standard. For dead cilia, the optical flow can easily detect them, regardless of the segmentation (there will always be a segmented region), whereas the linescan may fail in detecting such specific cases. For now, we suggest to combine both methods: the optical flow first as pre-analysis (mostly for the dead cilia), and then the curvescan for more precise analysis.



(a) (b) (c) (d) (e)

Figure 8.4: Sequence of operations for segmentation. (a) Is the first frame of the sequence. (b) Gradient of the descriptors Vp (c) markers of regions (d) Probability map (e) Segmentation result $\mathcal{M}_L$.



(a) (b)

Figure 8.5: Fourier analysis of speed variation for one of the sample yields to a frequency of 12.10 Hz.

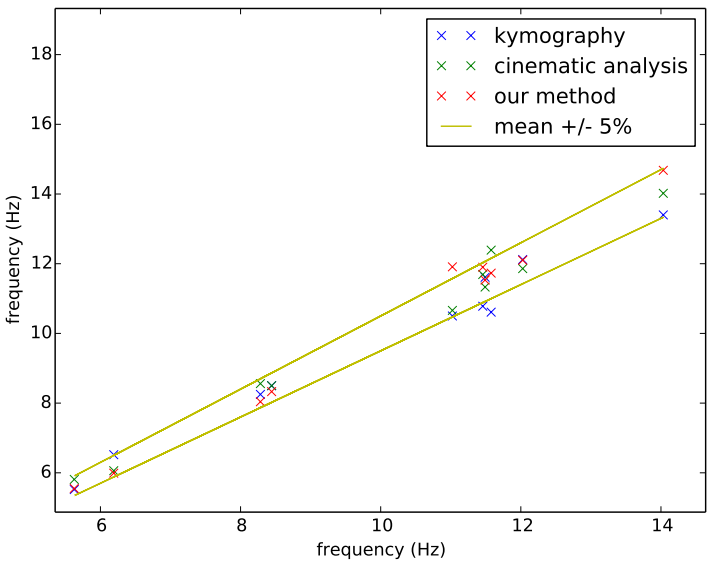


Figure 8.6: Correlation between our measurements and the ground truth

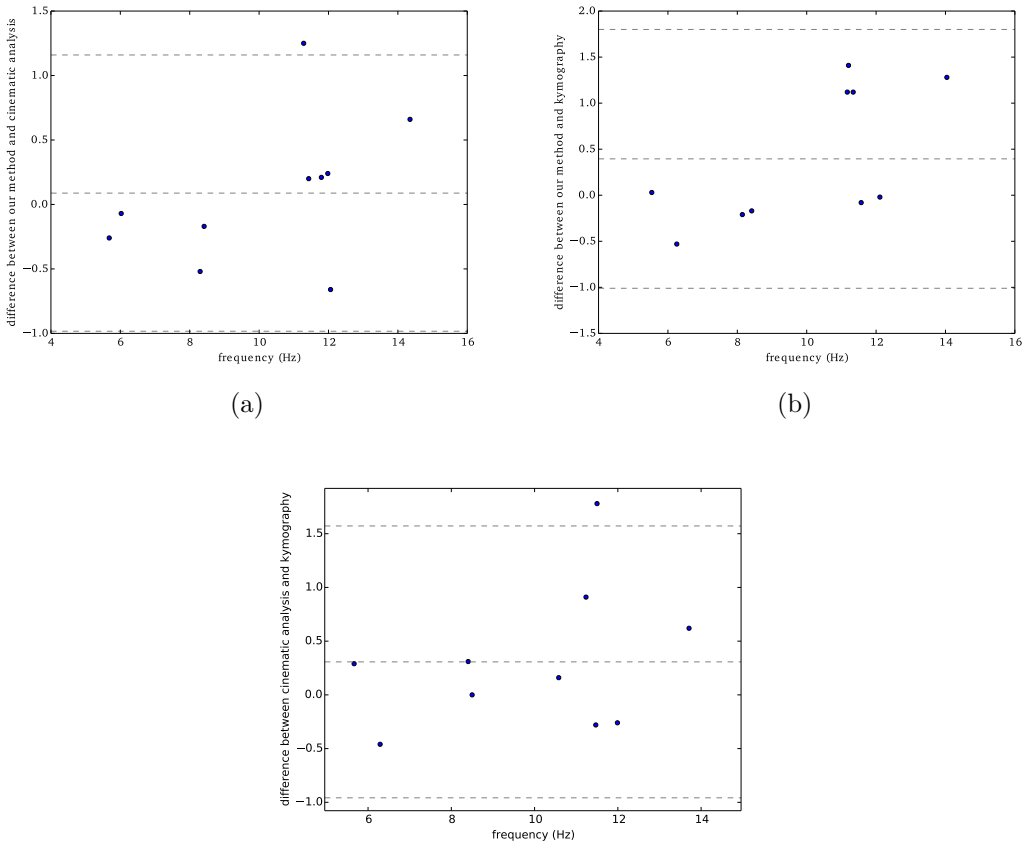


Figure 8.7: Bland-Altman plots show the consistency between our proposed approach vs. cinematic analyis and kymography.

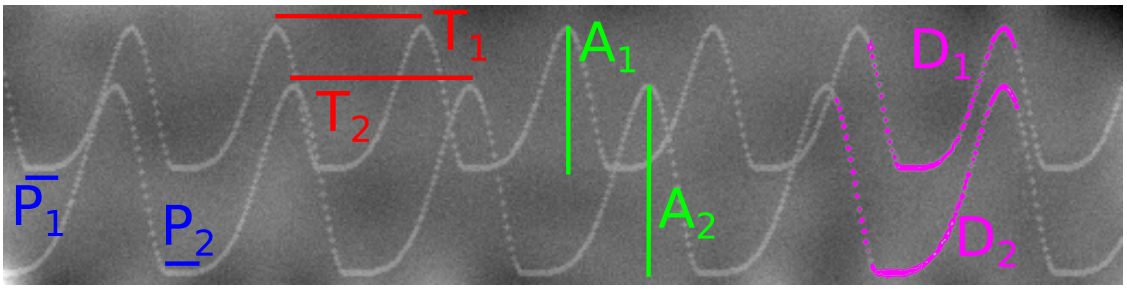
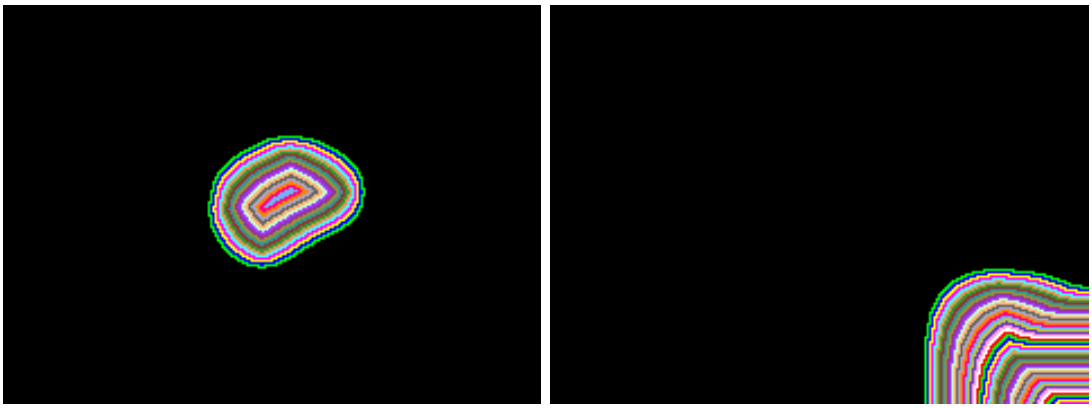
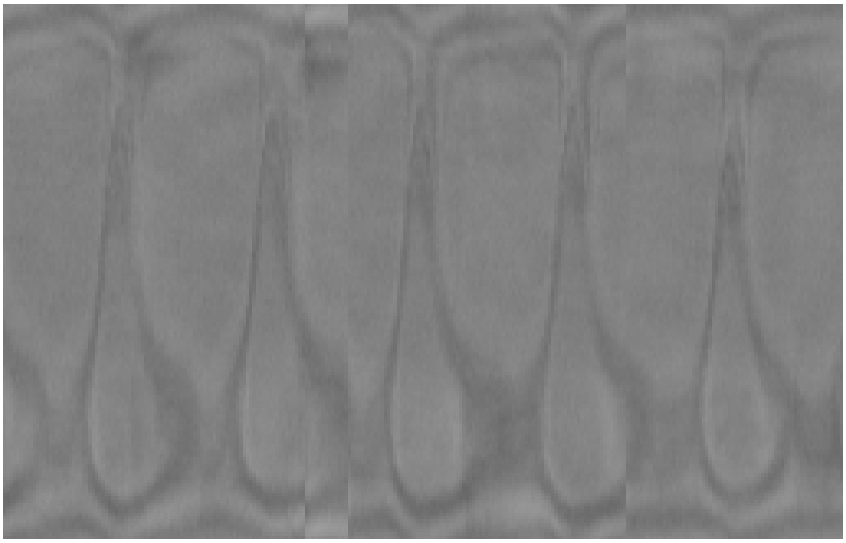


Figure 8.8: Ilustration of curvescan and parameters, on synthetic image where we can see two patterns.

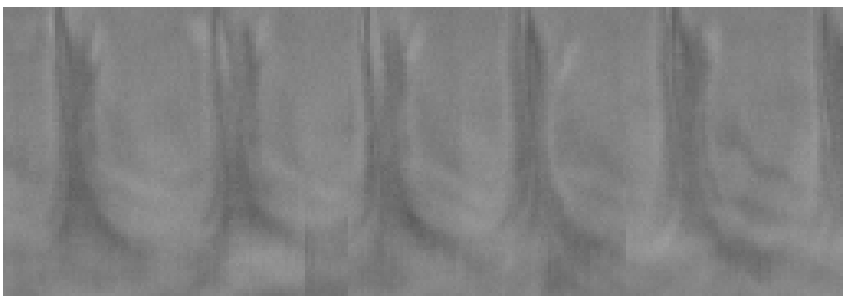


(a) (b)

Figure 8.9: Adaptative curvescan. (a) Distance map D_f inside the first region of 8.4(e) . (b) Distance map D_f inside the second region of 8.4(e), yielding concentric lines.

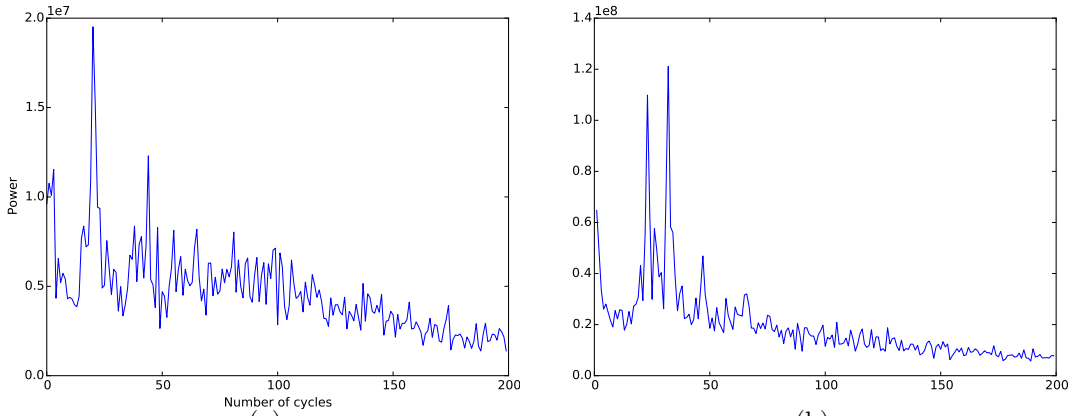


(a)



(b)

Figure 8.10: Example of grey level extraction (a) from the lines of 8.9(a) and (b) from the lines of 8.9(b)

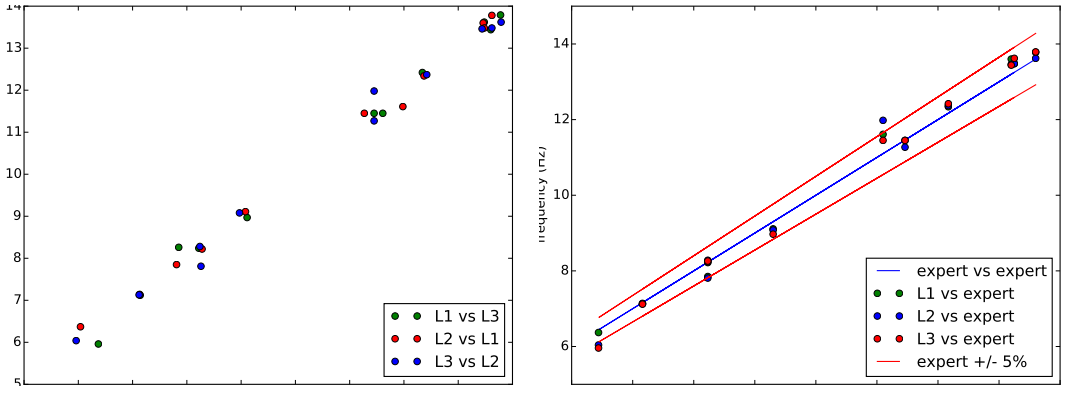


(a) (b)

Figure 8.11: Power spectra examples (a) a single frequency present, with harmonics, (b) Two distinct frequencies present.

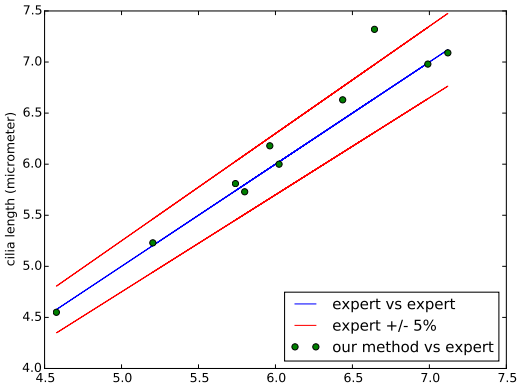
	Mean error	Standard deviation on error
Single line	0.0223	0.0207
Averaged frequency	0.0219	0.0162

(a)

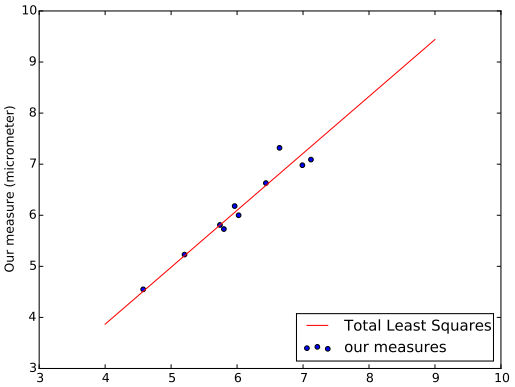


(b) (c)

Figure 8.12: Validations of our method for frequency estimation. (a) Mean and standard deviation of our error measurement (b) Validation of the robustness : frequencies estimated on several lines of D_f^Q are similar. (c) Validation of the accuracy : frequencies obtained with our method correspond to ground truth.



(a)



(b)

	Mean error	Standard deviation on error
Classical	0.0213	0.0291
TLS	0.0080	-

(c)

Figure 8.13: Validations of our method for cilia length measurement. (a) Validation of the robustness : frequencies estimated on several lines of D_f are similar. (b) Validation of the using expert variation with total least squares regression method yields a line which fits our measurements better. (c) Error estimation with the two methods of validation : our mean error may lay somewhere between 0.08% and 2.13%.

9

In vivo assessment of cilia motility evaluation

The main part of the medical application concerns ex-vivo analysis. Sampling is invasive and may damage cilia, and ex-vivo measurements may not always reflect the in-vivo cilia function. In this work we investigate the possibility of assessing cilia motility in vivo in humans. This work has been published in [3].

Contents

9.1 Existing tools and solutions proposed

Classical endoscopes are not suitable to evaluate cilia motion due to the lack of spatial resolution. We hence looked for the state of the art in existing micro-endoscopy. Micro-endoscopy is already in use in various applications, particularly for in-vivo histology.

Potential Co-development

One possibility was to develop a tool in partnership with a research group. We found a team in Stanford, the team of Joseph Khan, who has developed a protoype for a high-resolution endoscope. Currently, it is not designed for human experiment [93]. The current prototype does not have the required resolution at present. Their team is looking to improve this point. In their latest article, they present the design of their apparatus [94]. Unfortunately, these developments would imply long authorization procedure and pre-clinical research before becoming available for medical practice.
We hence surveyed the existing approved material for clinical use in various fields.

Olympus instrument

Olympus is currently developing a micro-endoscope. We do not have a lot of information about its level of development, but some tantalizing articles have been published. We contacted them, and the answer was that their equipment was not ready for testing outside of Olympus.

Pentax and Optiscan partnership

Optiscan is an Australian company that focuses on RD of optical products for medical use. They have a partnership with Pentax. They provide a confocal micro-endoscopy system that seems to have the required specification. However, their activities seem limited to Australia.

Mauna Kea CellVizio instrument

Mauna Kea is a pionneer in the area of micro-endoscopy. They have a commercial product ready and approved for pre-clinical and clinical use, called the Cellvizio, which is a confocal micro-endoscope. We contacted them and organized a feasibility study using their instrument.

9.2 The Cellvizio properties

The Cellvizio is a confocal endomicroscopy system used for in-vivo « histology and pathology » to evaluate tissues and cells morphology and abnormalities directly in situ during endoscopy of various organs such as esophagus. This technology is based on optical fibers and fluorescence response. According to the objects studied,

different probes with different resolutions are provided. The objects of interest are labeled with a fluorescent dye.
It is of interest as it can be used to observe respiratory epithelium at cellular scale in-vivo, with minimal invasive procedure. For the respiratory tract, this device has only been used for bronchoscopy. It had never been used for the analysis of mobile organites such as cilia. We sought to develop a partnership with Mauna Kea to develop the device and protocols adapted to our problematic.
Cellvizio exists in two models: the clinical and the pre-clinical device.
As ours was a feasability study, we used only the pre-clinical device. The difference between the two models are the name and durability of the probes, and the software which can reach the temporal resolution from 8-12Hz to 100Hz available in the pre-clinical model. This device can have a 1μm spatial resolution , which is at the lower limit of what is necessary for our needs.

9.3 Experimental runs

We conducted two experimental runs:
The first one was on nasal human brushing samples to observe cilia beating ex-vivo with the Cellvizio probes with a fluorescent dye, Octadecyl Rhodamine B Chloride (R18). It showed it was possible to observe cilia with that endoscope. The second experiment lasted a week and allowed us to explore the limitations of the device, to set an experimental protocol and to evaluate the analysis of sequences acquired this way.
During this week of experiment, ex-vivo pig trachea samples, harvested human cells and human nasal biopsy samples were labeled either with fluorescent marker R18 and with fluoresceine. We stabilized the samples and the probe with a shelf and an elevator jack (see Fig. 9.1), and recorded sequences using several probes and temporal resolution. We acquired the sequences with the default settings and with the software that boosts the frame rate acquisition.
The pig trachea samples were provided by the ENVA (National Veterinary School of Alfort).

9.4 Results ans analysis

Beating cilia were easily identified allowing the acquisition of 200 sequences. Those sequences showed cilia, cells and impurities labeled by R18. For sequences with fluoresceine, the cilia and cells are not labeled, but the medium is. Hence we can see the movement of cilia with the fading of the fluorescence while they evacuate the labeled medium. We developed specific software to evaluate frequencies of beating in sequences. Out of the 200 sequences, we could analyze only 19 because the stability of the cells were not sufficient, the temporal resolution was too low, or the duration of acquisition was not long enough.

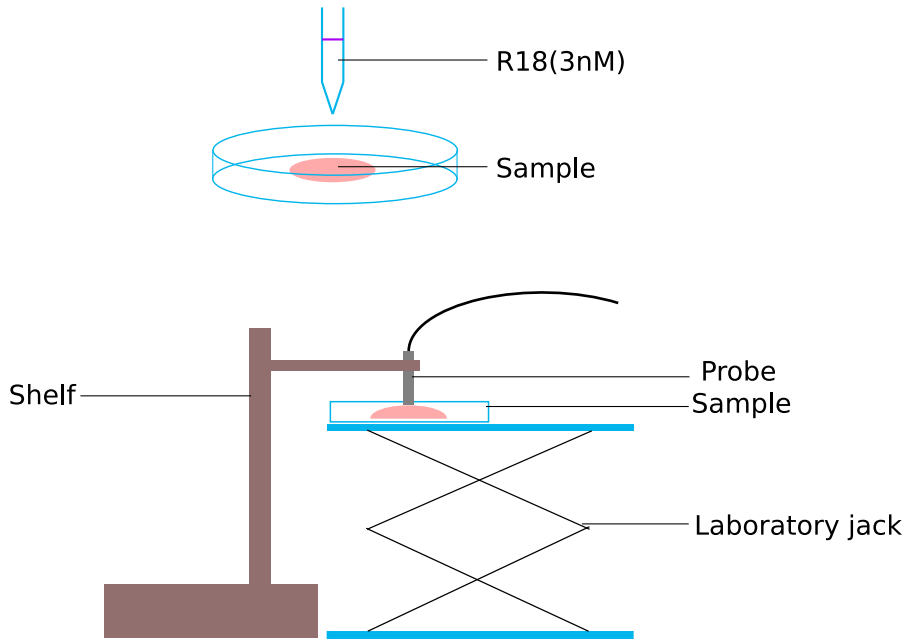


Figure 9.1: Experimental apparatus.

Markers

The R18 marker gave really good results for visualization of cilia. Indeed, as a membrane marker, it labeled cilia but also the cells. For the fluoresceine, it is complicated to analyze the sequences and to quantify motion. We eliminated this dye to focus on R18 for the experiment. Figure 9.2 shows example of ciliated cells with the two markers. With the R18, cilia are visible, and relatively well defined whereas with the fluoresceine they remain fuzzy.

Frequency estimation: acquisition settings.

We validated our frequency estimations on videos with a sufficiently long and stable sequence of beating cilia to allow measurements over several cycles. With an acquisition frequency under 30Hz, it was impossible to estimate the frequency of beating, so we eliminated all these sequences.

- Estimations were validated on sequences acquired between 30 and 90Hz;
- We estimated frequencies both in human and pig samples;
- 5 areas were selected in each sequence, yielding 95 measurements.

On each sequence we could analyze, we estimated frequency in 5 areas. We selected a pixel in each area and we computed the pixel intensity average in a 4-pixels square containing that pixel. Variations of pixels intensities were then analyzed using a FFT analysis to estimate frequency (see Fig 9.3). We validated these estimations by comparing them with expert human assessment.
Among the 95 measurements, in 5 cases, we incorrectly confused harmonic frequencies with the fundamental (see Fig 9.4(a)).

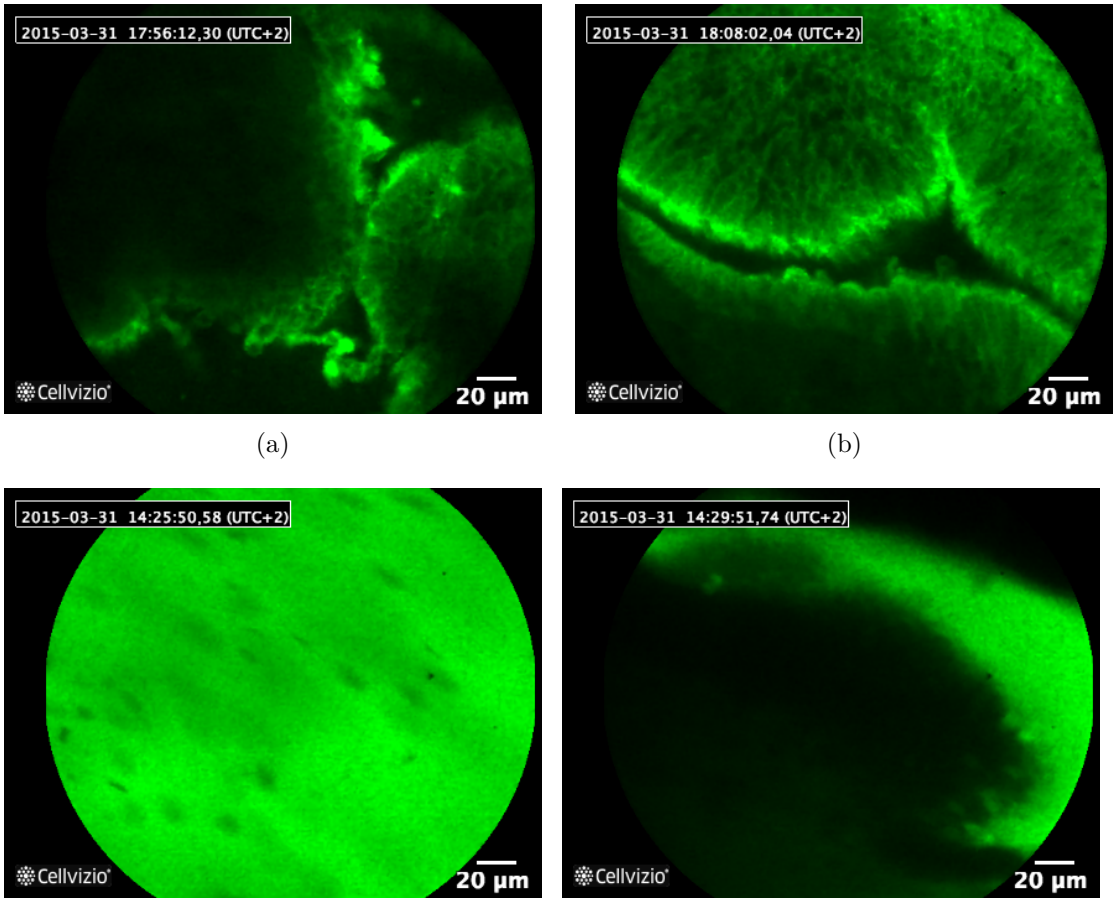
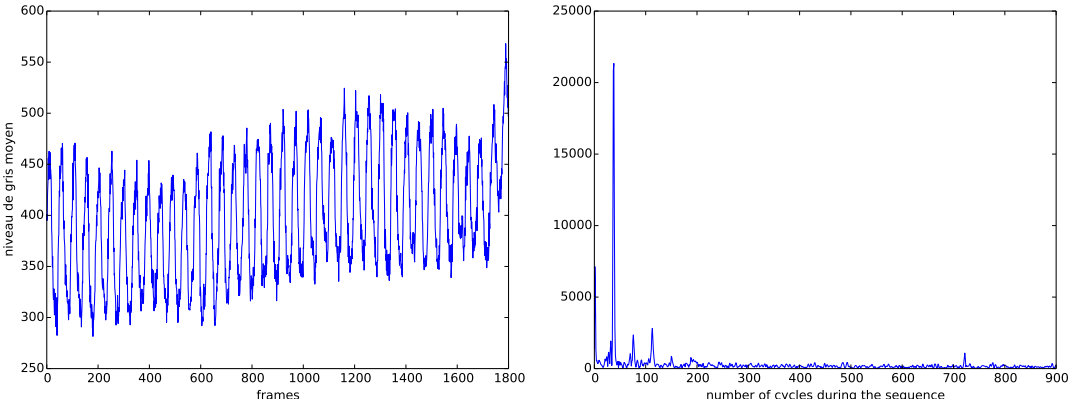


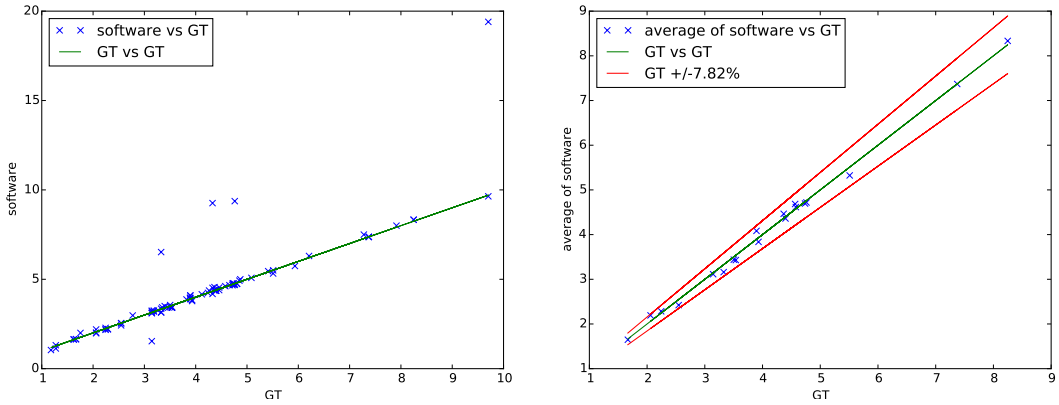
Figure 9.2: Images of samples labeled with R18 ((a) and (b)) and fluoresceine ((c) and (d)) aquired with Cellvizio.

a rough error rate is so about 5%. However this says nothing of the precision of our measurements. For this, we we propose a global error rate, that takes into account multiple acquisition sequences. We take the median of the estimation for each sequence and compare it with the median of the ground truth. We also compute a confidence interval: by allowing of +/- 1 in the sum of cycles. Indeed, the ground truth is made by manually counting the number of cycles in the sequence. It appears that, even if the same person counts the cycle, they can make an error of 1 or 2 cycles for the same sequence analyzed at different moments. We hence chose to include this variability to define our confidence interval. While comparing the different medians, the total average error rate is 2.72% while the confidence interval is +/- 7.62%. We represented this comparison in Fig 9.4(b).
These results shows the necessity of estimating 5 frequencies per sequence and to take the median value to have a reliable frequency determination.
We cannot analyze the beating pattern because of limits in spatial resolution.



(a) (b)

Figure 9.3: Variation of pixel intensity and Fourier Transform.



(a) (b)

Figure 9.4: Correlations between software measurement and ground truth. (a) is the individual measurements (95 measures) and (b) is the average of each sequence.

Probes of the endoscope.

We eliminated 2 probes that are too slow to permit frequency estimation (in red in Fig 9.5). We selected the 2 probes which gave the best sequences (in green in Fig 9.5).

Probes		
Ref	Speed	Spatial resolution
S-1500	8 fps	3.3μm
Z serie	8 fps	3.5μm
MiniO-30	8-50 fps	1.4μm
UltraMiniO	8-90 fps	1.4μm
MiniO-100	8-100 fps	1.4μm

Figure 9.5: Probe specifications.

Limitations

Even if these experiment are promising, R18, the dye we used, is cytotoxic and cannot be used in-vivo. Finding a bio-compatible marker with comparable marking qualities is a challenge that must be addressed for further studies.
The spatial resolution is sufficient to evaluate beating frequency, but not to analyze the beating pattern. It has to be improved for a complete analysis. Moreover, the temporal resolution of Cellvizio without the software that boosts the frame rate acquisition is not sufficient. Unfortunately, this software comes only with the pre-clinical instrument and not the clinical one. The usage of the pre-clinical software in a clinical context is a costly certification issue, and is not in the short-term plans of Mauna-Kea. Thus, at this moment, the usage of Mauna-Kea instruments for cilia beating evaluation requires further work.

9.5 Perspectives

This study constitutes a first feasability step for the in-vivo cilia motility evaluation. Using the Mauna-Kea confocal endoscope, we are able to visualize beating cilia and to accurately measure beat frequency. We pointed out some limitations due to the available probes: for reliable results, we need the highest frame rate and spatial resolution. For now, we have to compromise between these parameters. Consequently, current sequences are suitable for frequency estimation but not for further analysis (study of beating pattern for example). Crutially, a less cytotoxic fluorescent dye approved for human use than R18 is necessary for future in-vivo experiments. We found two potential dyes for that purpose that seem to have very little toxicity, the Sir-tubilin and Sir-actin that are respectively markers of tubulin and actin. They are already used for cilia labelling, and could be a good starting point for further work.
Olympus is currently developing a device that could be used in our study, even if the specifications remain confidential.
We are pretty confident that adequate technologies exist and can be adapted to in-vivo analyses with comparatively little work, and will hopefully be available in a a reasonably near future.

Part IV

Application: fish embryo based assays

10

Fish embryo mortality evaluation

This work represents the development of an eco-toxicity assay based on a fish embryo model. The aim of this assay is to determine the life or death status of fish embryo using image processing in an industrial environment using heartbeat detection. In this chapter, we propose the design of an image processing pipeline based on mathematical morphology [95, 96, 54] to assess the presence or absence of a beating heart pattern in fish embryo. We describe, investigate and propose solutions for all the problems and challenges we have encountered in this study. Another objective of this chapter is to showcase the design of an automated pipeline for processing sequences of fish embryo and more generally to provide some guidelines for the design of robust image analysis pipelines, through the particular example of a real-world problem. This work was published in ISMM 2015 [4] and in the journal Computers in Biology and Medicine [5].

Contents

10.1 Aim

Because of throughput and storage constraints, a major aim is to achieve complete sequence analysis online within the same time frame as the acquisition, i.e. roughly under 10s. We therefore propose a robust pipeline suitable for production usage. It consists of simple operators, which are for the most part readily available and fast. Figure 10.1 presents the flowchart of our assay. It is split into two phases: a pre-processing step which consists of sequence stabilization and denoising, while the actual processing step consists of detecting significant periodic changes in the embryo (variation of grey level values) assuming they are caused by its beating heart.

10.2 Pipeline

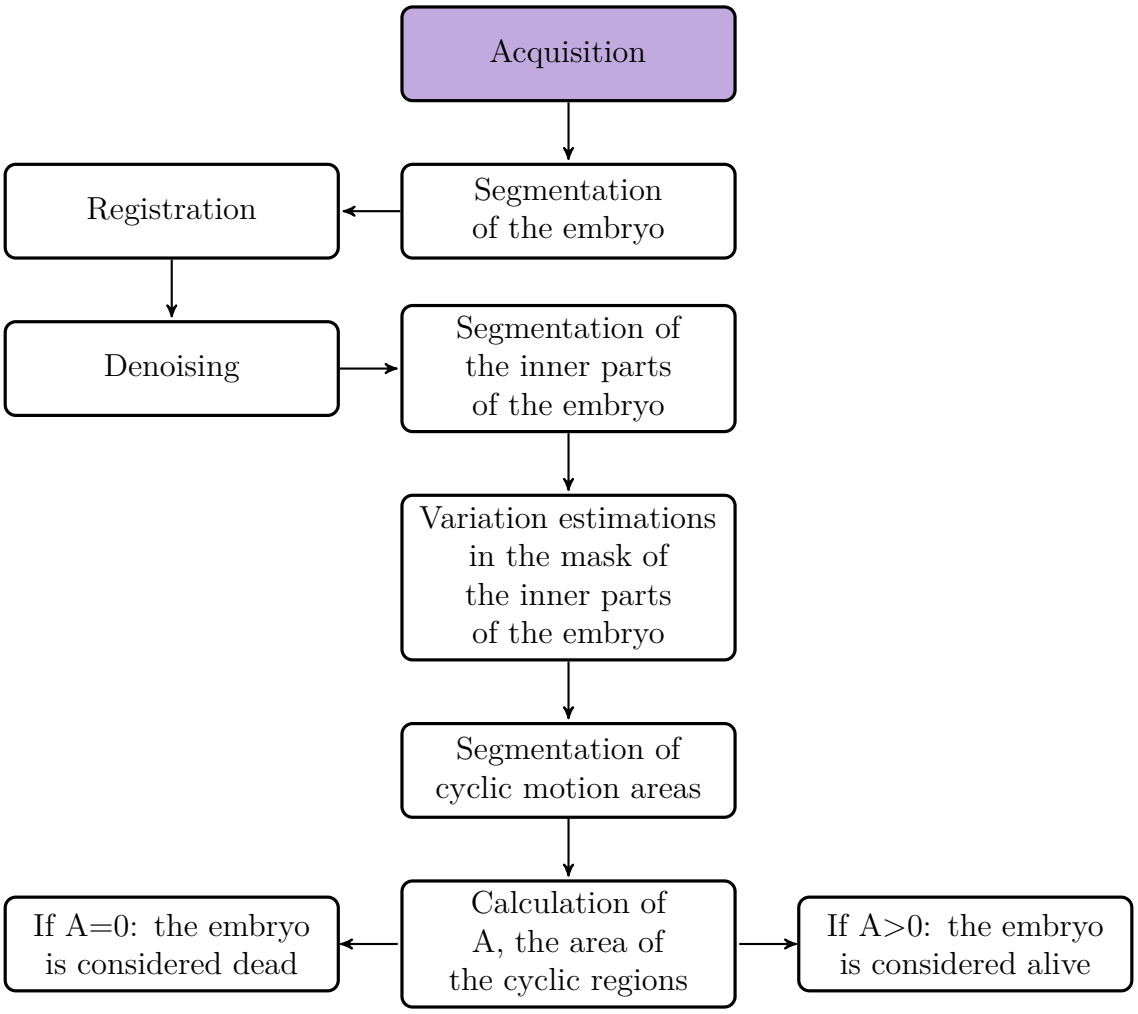


Figure 10.1: Flowchart of our embryo mortality image processing assay.

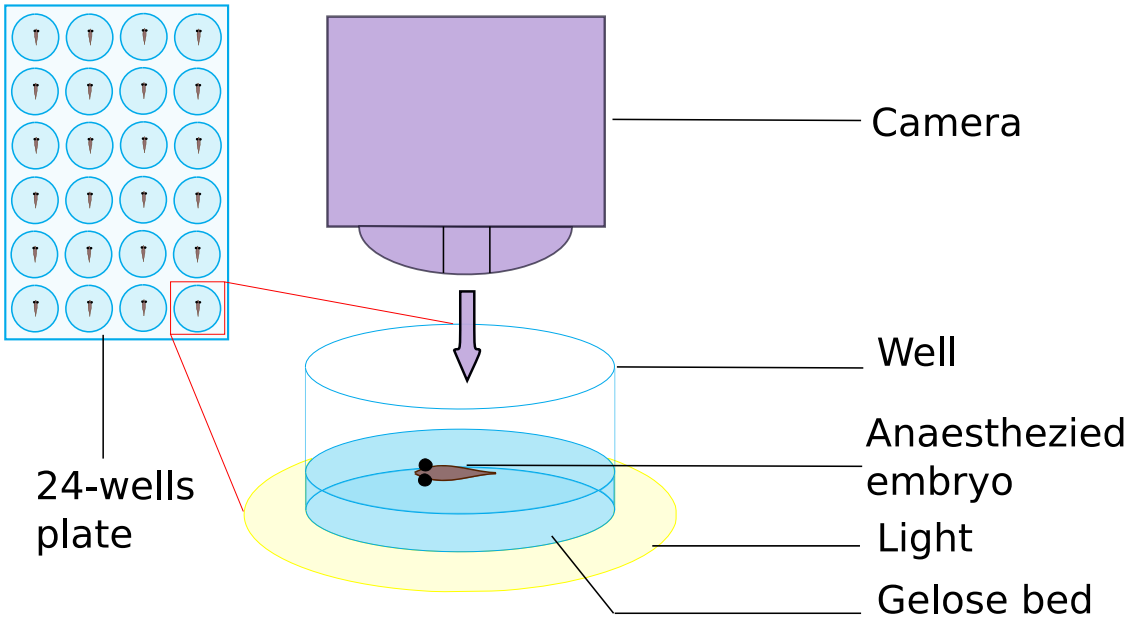


Figure 10.2: Schematization of our acquisition procedure

10.3 Details of the methodology

Acquisition

Fish embryos were anesthetized with tricaine (0.1g/L) and placed on a support gel in a 24-well plate, one embryo per well. The plate was then placed under the connector board and the acquisition was automatically performed by the software package [1]. For each well, we record 2s of uncompressed video sequence at 15 frame per second and a resolution of 800×600 pixels. This protocol is illustrated in Fig. 10.2.

Segmentation of the embryo

The first step is an initial segmentation of the embryo. Indeed, this segmentation is crucial for several reasons. In order to reduce the memory usage, we need to crop the area of interest to a small window centered on the embryo. This step allows us to weed out the sequences where an embryo is not visible or intersects the acquisition window. Moreover, we generally need to stabilize the sequence, and this stabilization must be performed on the embryo itself, and not on other elements on the field of view like the contours of the well. The background appears white whereas edges of the well and embryo are darker (see Fig. 10.3(a)), particularly the eyes are very dark. Therefore the embryo is easy to segment as the largest connected component in the min-tree associated with the darkest minimum not connected to the edges of the field of view, simultaneously maximizing the inter-class variance between foreground and background (i.e, following the Otsu criterion [89]). We implemented it as follows. If θ_O is the Otsu threshold for the first input frame S^0,

[1] https://www.fei.com/software/visilog/

then

$$A^0 = \mathbb{1} - S^0_{\geq(\theta_O+\theta_b)}, \tag{10.1}$$

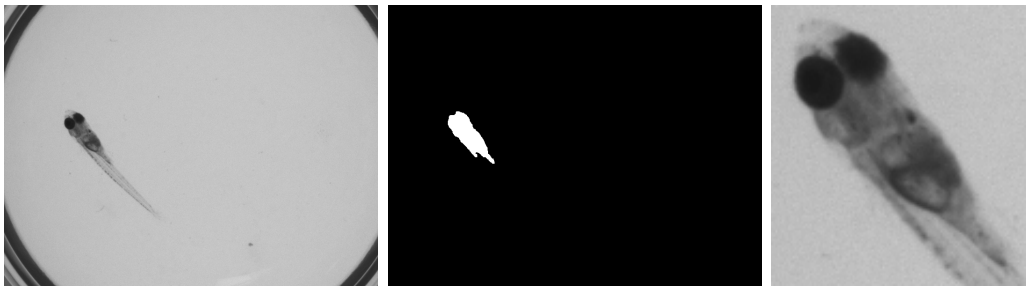
where $\mathbb{1}$ is a boolean image consisting only of ones, and θ_b is a baseline value, optimized to 30 experimentally.

Because we are not interested in the thin elements of the embryo (i.e. the tail), a morphological opening is applied, and because a single connected component is desired, an area criterion is also used:

$$A^1 = \varphi_{\mathcal{B}_1}(\varphi_{\lambda_1}(A^0)). \tag{10.2}$$

Experimentally, we determined the average area of an healthy embryo to 3000 pixel, so we set λ_1 conservatively to 600 pixel. The embryo's body is between 20-60 pixel wide, while artifacts in the well are usually much smaller, so $\mathcal{B}_1$ is set to a discrete Euclidean ball of radius 3. Since the embryo is expected to be near the center of the field of view, we remove all objects connected to the frame of the image, calling the result A^2. For speed, all these operators are implemented on the min-tree structure of the image, except the opening with $\mathcal{B}_1$, which is approximated by a fast, separable dodecagon [97].

This result A^2 is expected to represent the mask of the embryo (see Fig 10.3(b)). However, if the result is empty, this means that the embryo is not centered in the well and then the sequence cannot be reliably analyzed. If we do find an embryo, we crop the sequence by defining a bounding box around our segmentation, dilated by 10 pixels. The result is a new sequence S^1 centered on the thorax of the embryo (see Fig 10.3(c)).



(a) Initial frame. (b) Embryo segmentation. (c) Cropped frame.

Figure 10.3: Segmentation of the initial frame to locate and the embryo in the well.

Registration

Because of vibrations associated with other equipment around the acquisition platform, the embryo can appear to move around slightly in its well. The sequence needs to be stabilized to avoid false positives. For registration, we used the method described in chapter 5.

We then apply the estimated transform between the first cropped frame of the sequence S^1 and all subsequent frames, taking the first as reference. We compute $d_{xmax} = \max|d_x|$ and $d_{ymax} = \max|d_y|$ for all frames, and we crop the bounding

box of the whole sequence further by these amounts in x and y respectively. We call $\mathcal{S}^2$ the stabilized, cropped sequence of the embryo.

Denoising

To eliminate global illumination variation during a sequence, which occurs sometimes, we equalize the average intensity of all frames. We obtain a sequence $\mathcal{S}^3$ where the average value of all frames is constant throughout the sequence.

Depending on the average illumination, the sequence can be quite noisy, so we used the 2D+t bilateral filter described in chapter 5.

We have optimized the parameters that removed the noise without removing the heartbeat. These are: window size=$5 \times 5 \times 5$, σ_r=0.75, σ_d=0.9. The result is a denoised sequence $\mathcal{S}^4$. Due to border effects, we remove the first three and the last three frames from this sequence, which is then 24-frames long.

Segmentation of the inner parts of the embryo

If the embryo is alive, the heart should be detected in its thorax region. To ascertain this, cyclic motion should be detected in this region of the sequence, but not anywhere else where remaining noise or motion might be present. Because eyes are very dark, noise causes detectable variations in them, and remaining motion may be detectable near the contours of the embryo. We developed a mask $\mathcal{M}^1$ of the inner part of the embryo to avoid false detection in these areas.

We define B^1 as the minimum image of the sequence: $B^1 = min(\mathcal{S}^4)$. In this image, where the heart and vessels are darker because of the presence of blood, we apply the same procedure as before:

$$B^2 = \varphi_{\mathcal{B}_1}(\varphi_{\lambda_1}(\mathbb{1} - B^1_{\geq(\theta_O+\theta_c)})) \tag{10.3}$$

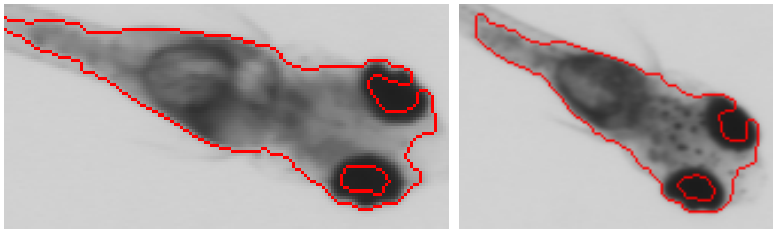
In this instance, because of the change of contrast, θ_c is set to 20; θ_O is again obtained using the Otsu criterion; λ_1 and $\mathcal{B}$ are unchanged. The resulting B^2 is a binary mask (with values in $\{0, 1\}$) of the registered body of the embryo. Considering B^2 as a geodesic mask, we now segment the eyes as the one or two more prominent minima from the min-tree. We cannot rely on the eyes being separated. Depending on the pose of the embryo, they might get merged. We write:

$$\mathcal{M}^1 = \epsilon_{\mathcal{B}_3}(\gamma_{\mathcal{B}_2}((B^2.B^1)_{\geq(\theta_O-\theta_c)})) \tag{10.4}$$

Here the dot in $B^2.B^1$ denotes the pixelwise multiplication. Note that θ_O is re-estimated from the grey-level distribution of the embryo within the geodesic mask. In this formula, $\mathcal{B}_2$ is a ball of radius 1 and $\mathcal{B}_3$ a ball of radius 3. The outline of the resulting mask $\mathcal{M}^1$ are exemplified on Fig. 10.4.

Variation estimations in the mask of the inner parts of the embryo

Our challenge is now to detect cyclic motion-induced variations in the sequence assuming it corresponds to a heartbeat. The heartbeat is noticeable by changes of contrast due to blood cell concentration in the heart region. Computing the grey level variance at each pixel of a sequence along the time line show the grey-level



(a) (b)

Figure 10.4: Inner parts segmentation on two embryos. (a) is alive whereas (b) is dead.

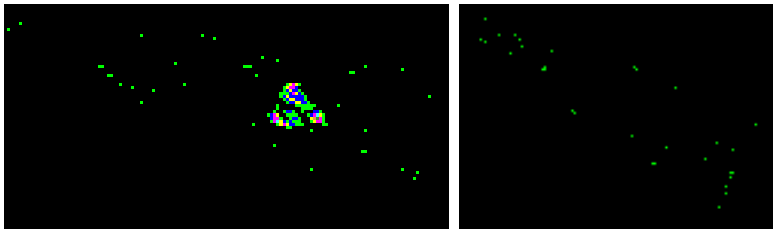
variability for that pixel, however this variability may not be due to a cyclic pattern. To account for cyclic variations, we split $\mathcal{S}^4$ into 3 sub-sequences of 8 frames ($\mathcal{S}^4_1$, $\mathcal{S}^4_2$ and $\mathcal{S}^4_3$). An interval of 8 frames is similar to the expected period of a single heart beat. We compute the pixels variances inside ($\mathcal{M}^1$) on each sub-sequence, yielding three variance images $\mathcal{V}_1$,$\mathcal{V}_2$,$\mathcal{V}_3$. We next compute the median of these three images $\mathcal{V}_i$. In this way, some spurious, potentially even large, residual variation occurring only once in the sequence will produce a large variance only once, and so the median of all the variances in this area will remain small. On the other hand, cyclic variation will exhibit significant variance in all three sub-sequences.

$$\mathcal{S}^4 = \{\mathcal{S}^4_1 \cup \mathcal{S}^4_2 \cup \mathcal{S}^4_3\} \tag{10.5}$$

$$\forall i \in [1,3],\ \mathcal{V}_i = \text{variance}(\mathcal{S}^4_i) \tag{10.6}$$

$$C^1 = \mathcal{M}^1.\,\text{median}(\mathcal{V}_i, i \in [1,3]) \tag{10.7}$$

The result of this process is shown in Fig. 10.5.



(a) (b)

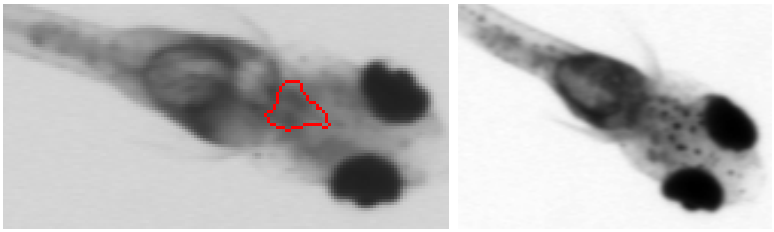
Figure 10.5: False color rendering of the temporal variance. (a) is for a living embryo, (b) a dead one.

Segmentation of cyclic motion areas

Binarizing the image of the cyclic variance is simple with area openings and a small closing in classical alternating sequence.

$$D^1 = (\gamma^{\lambda_3}(\varphi_{\mathcal{B}_4}(\gamma^{\lambda_2}(C^1))))_{\geq 1}. \tag{10.8}$$

Here, $\lambda_2 = 3$, $\mathcal{B}_4$ is a radius-2 ball, and $\lambda_3 = 10$. This image is then thresholded at 1.



(a) (b)

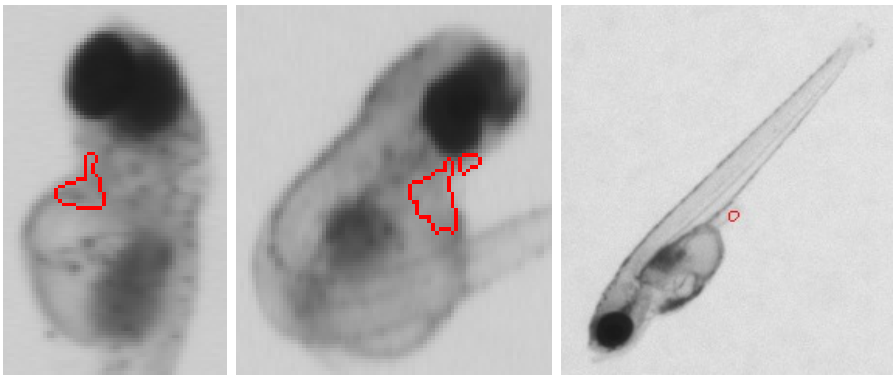
Figure 10.6: Segmentation of cyclic motion detection on embryos. (a) corresponds to the segmentation of 10.5 (a) (living embryo), and (b) corresponds to the segmentation of 10.5 (b) (dead embryo)

Calculation of A, the area of the cyclic regions

A is the number of non-zero pixels in D^1. It corresponds to the area of the cyclic regions. If this area is null, it means the software didn't find any cyclic region in the sequence, and so no heartbeat is detected. The alevin is considered dead. If A is positive, a cyclic motion considered as the heart is found and the embryo is considered alive.

10.4 Results and validation

We tested our program on 651 videos taken over several experimental runs. The protocol was identical each time but conditions such as illumination or marker concentration could change. This number of videos is significant and reflects production usage. We have also tested our protocol on unusual embryos: some with oedemas and with other malformations to ensure the robustness of our protocol (see Fig. 10.7(a) and (b)). The program processes a sequence in less than 10 seconds, which is in accordance with our initial constraints.



(a) (b) (c)

Figure 10.7: Heart segmentation in the presence of oedema (a) or malformation (b). Incorrect segmentation in the fin due to fluttering (c).

In some cases, even actually dead embryos can appear to move. This can be due to fluctuations in the gel or the presence of shadows (figure 10.7(c)). In dark areas, electronic noise is proportionally more present, and may appear as motion and in some cases, the embryo may appear to slide on the gel. This happens for example if we incorrectly do not compensate for rotation in the sequence stabilization phase. The main remaining cause of problem is ambiguity: in some cases, the heart beats so slowly or weakly that we cannot detect it. In most cases, a human operator would also have difficulties in detecting it. We could also argue that embryo with weak heartbeats are not very healthy and likely to die, however this is unconfirmed.

We have run experiments on many sequences: some darker, some lighter, some with varying amounts of tricaine, some with embryo abnormalities, and so on. In total, we processed 651 such sequences, which might be considered sufficient for a first validation step.

From an initial set of 660 videos, 9 were eliminated as incomplete. From the remaining set of 651 sequences, 100 were used for training, i.e. optimizing the parameters of the pipeline. The remaining 551 were used for testing. There was 1 remaining error in the training sample (1% error rate). There were 3 errors in total in the test set, for an error rate of 0.54%. Such an error rate is effectively quite low.

10.5 Pipeline improvements for enhanced automation.

The procedure described in the previous section is essentially a feasibility study. In a second step, changes in the acquisition protocol were introduced to allow for a larger degree of automation in the handling of fish embryos. Under this protocol, fish embryos live all their life in the same well: fish eggs are put in their well unhatched once. They develop, hatch and grow for 11 days in there, with occasional changes of liquid medium. Since they are embryo, they do not need to feed.

This change is significant because the medium is liquid and so embryo are much more mobile. Also illumination changes were introduced. Indeed with the liquid medium, it is no longer possible to move the platform, so the camera needs to move instead. The camera also features a higher resolution and faster speed.

Here we describe the new protocol and the changes in image analysis that were introduced to improve the pipeline and keep it robust. Fish eggs are placed manually in 24-well plate, one embryo per well, immediately after spawning on day 0. Each well contains 2mL of incubation medium, together with a predetermined concentration of a water-soluble compound under study. Embryos are incubated in this medium in their respective well until day 9, when image acquisition is performed. Medium replacements are automatically performed on days 2, 5 and 7. On day 9, we removed 1.5 mL of incubation medium from each well, to limit embryo movements, and fish embryos are anesthetized with 70 μL of tricaine (0.1 g/L) in a total volume of 0.57 mL. The plate is then placed under the connector board and the acquisition is automatically performed under the control of a Visilog

Visual Basic script. For each well, 30 uncompressed video frames at a resolution of 1500×1500 pixels are recorded for a duration of 1 s.
These changes induced some problems with our procedure: the morphological parameter had been optimized for the previous camera. Now the definition is higher, and so the size of structuring elements have to be changed. Moreover, the embryo in its well yields several problems: the well can be dirty as the embryo is left in it while it was still an egg, the embryo can be erroneously vacuumed off by the system, causing some wells to become empty, and without the gelose bed, the embryo can move more easily and appears deformed by the refraction of light in water. Finally, we can encounter both eggs or hatched alevins, making mandatory a differentiation step in our image processing assay.

10.6 Modifications : details

Quality test of the video sequence

We begin by detecting the videos that present important and undesirable changes during the sequence. These may be due to the presence of black frames, shadows or large uncontrolled motion. For this, on each difference d_i between two successive frames of the video sequence $\mathcal{S}^0$, we compute the statistical variance (V_i):

$$\forall i \in [1, 29], d_i = S_i^0 - S_{i+1}^0 \tag{10.9}$$

$$V_i = \text{variance}(d_i) \tag{10.10}$$

In the case of a correctly recorded video, two successive frames should be very similar, and their pixelwise difference should yield a near-zero output, so its variance remains small. On the other hand, if a large motion appears on a frame, the difference will show a high contrast. If at least one of all computed variances is higher than an experimentally determined threshold (set to 30), the video sequence is deemed non-exploitable.

Segmentation of the well

The segmentation of the embryo is crucial for several reasons. For speed and reduced memory usage, we crop the area of interest to a small window centered on the embryo. During this step, we also detect sequences where the embryo is not fully visible, for instance too close to the edges of the well. Moreover, motion stabilization must be performed on the embryo itself, and not on other elements of the field of view. We first delimit the region of interest by finding the disk area corresponding to the inner part of the well. This step also removes all objects connected to the edges of the well. The procedure for finding this disk area is as follows: edges of the disk appear dark, so we first compute a so-called *bottom-hat* filter: see Fig. 10.9(a) and Eq. (10.11) for a definition. In this equation, B_{20} is chosen to remove small artifacts in the well. This yields image A^0 (Fig. 10.9(b)), which we binarize via an automated thresholding operation to obtain image A^1 in Fig. 10.9(c). The automated thresholding procedure is described here: we determine

the threshold which maximizes the inter-class variance between foreground and background (i.e. following the Otsu criterion [98]), with the added constraint that the area of the foreground must be in the interval [20 000, 40 000]. This interval was determined experimentally to guarantee the presence of the well edges in the foreground. This constraint is convex and so easily implemented: we consider all thresholds from the highest to the lowest in order. The area of the foreground necessarily increases during this process. In the acceptable interval for the area of the foreground, we select the threshold with the highest Otsu criterion. We call this threshold θ_{co} (for constrained-Otsu). $\mathcal{F}^0$ is the first frame of the sequence.

$$A^0 = \varphi_{B_{20}}(\mathcal{F}^0) - \mathcal{F}^0 \tag{10.11}$$

$$A^1 = (A^0)_{\geq \theta_{co}} \tag{10.12}$$

We weed out small components in the well with an area opening γ^{α}_{100} of parameter $\lambda = 100$, followed by a closing with a ball B_{40} to reconstruct fragmented edges of the well. Then a radial opening γ^{ρ}_{100} with linear element ρ of length $\tau = 100$ is applied to remove short artifacts in the well, while retaining the thin well borders [99]. This yields image A^2:

$$A^2 = \gamma^{\rho}_{100}(\varphi_{B_{40}}(\gamma^{\alpha}_{100}(A^1))) \tag{10.13}$$

From this result shown on Fig.10.10(a), we wish to keep only the internal ring that represents the separation between the interior of the well and its edges. For this, we use a well-established morphological approach to segmentation, based on the Watershed transform. This operator is intuitively defined as in hydrology, in a grayscale image whose intensity can be assimilated to a 3D terrain, as delimiting the borders between adjacent catchment basins [100]. We compute the magnitude of the Derivative of Gaussian filter (DoG): $\mathrm{DoG}_\sigma = \nabla \star G_\sigma$ using the Deriche recursive implementation of the gradient operator for speed [101] with parameter $\sigma = 10$ and we use a markers-based Watershed algorithm [102] on the magnitude of this gradient. A disk at the center of the frame is taken as internal marker m^1_{int} and the frame corners are the external marker m^1_{ext}. We write:

$$A^3 = \text{watershed}(\| \mathrm{DoG}_{10}(\mathcal{F}^0)\|, m^1_{\text{int}}, m^1_{\text{ext}}) \tag{10.14}$$

The resulting contour is shown on Fig.10.10(b). The result may be incorrect if the embryo is too close to the edge of the well. To help with this, we expand the contour using the smallest convex set that contains A^3 [103]. We call A^4 the resulting image (see Fig. 10.10(c)) and G^4 the set of points contained in the central component of A^4. We compute the barycenter C of coordinates (a, b) and the diameter $2R$ of G^4 as the largest width or height of its bounding box. The final well segmentation is the disk $\mathcal{D}$ centered in C and of radius R. Its contour is shown on Fig.10.10(d).

$$\mathcal{D} = \left\{(x, y), (x-a)^2 + (y-b)^2 \leq R^2\right\} \tag{10.15}$$

$$\text{with } (a, b) = \text{barycenter}(G^4) \text{ and } R = \frac{\max(\text{width}(G^4), \text{height}(G^4))}{2}$$

Segmentation of the embryo: location in the well

Our image analysis procedure is intended to work for both alevins and eggs, but some eggs do not develop at all and differ markedly from healthy eggs and alevins (see Fig. 10.11(a) and (c)). They feature low contrast, which make them look like empty chorions or impurities we can have in the well too. An early challenge in the pipeline is to detect and identify the embryo in each well reliably. To achieve this, we begin by performing an initial segmentation adapted to all components in the well, whatever the intensity level or variance they may have. In the previously calculated bottom-hat image A^0 (see Fig. 10.9(b)), all the components of interest are easy to detect as connected components located strictly inside the segmented well. We call h the *contrast significance*, understood as the intensity variation that connected components should at least have to be significant [104]. The value h is set experimentally to ignore the irrelevant intensity variations for instance due to noise, while still detecting the dimmest components we cannot ignore (e.g. the undeveloped eggs). We call p^i_{peak} the local maximum of intensity in the neighborhood of the pixel p^i. Image B^1 contains the so-called *h-maxima* of $A^0 \cap \mathcal{D}$, defined as follows:

$$B^1 = \left\{ p^i \in A^0 \cap \mathcal{D} \text{ with } p^i = \left[\begin{array}{ll} 0 & \text{if } (p^i_{peak} - p^i) > h \\ p^i & \text{if } (p^i_{peak} - p^i) \leq h \end{array} \right. \right\} \tag{10.16}$$

The image B^1 of the h-maxima can be efficiently computed using a morphological reconstruction operator, as explained in [56].

Several components can be detected in the resulting frame B^1, potentially matching with the embryo, an empty chorion, or some impurity. To identify the embryo, we combine several criteria: presence of eyes, minimal and average intensities, variance and circularity. A component is considered as including an eye if it presents a very dark part representing less than the quarter of the total component area. Since impurities are generally homogeneous, this procedure filters out dark impurities because these are uniformly dark, and light impurities and chorions because they are evenly bright. However it can also filter out under-developed eggs. To avoid that, we enhance the classification with other criteria: a high average intensity or a low variance, we also check the circularity to differentiate between undeveloped eggs and chorions or bright impurities. Finally this process allows us to associate each component with one of the following classes: “under-developed eggs” , “other embryos” or “impurities and chorions”. We delete every components identified as an impurity or a chorion. As we know that it must be only one embryo per well, if after this step several components of the class “under-developped eggs" remain, we only keep the largest. If several components of both remaining classes remain, we only keep the largest component among those from the class “other embryo”. Indeed, we have found experimentally that it is more difficult to distinguish under-developed eggs from chorions than other embryos from impurities. Thus the probability of making a mistake from the class “undeveloped egg” is higher. The result $\mathcal{M}^0$ is a binary mask (with values in $\{0,1\}$) containing only one component expected to locate the embryo in the well (see its red contour in Fig. 10.11(a,b)). If the result is empty, this means that the embryo intersects the edges of the well and then the sequence cannot be reliably analyzed. If instead we

find an embryo, we crop the sequence and the mask $\mathcal{M}^0$ by defining a bounding box around our segmentation, dilated by B_{20}, a ball of diameter 20 pixels. This results in a new sequence denoted $\mathcal{S}^1$ centered on the embryo (see Fig. 10.11(c)). However, because of contrast variations and the large variability of grey levels between embryos, the mask $\mathcal{M}^0$ is only approximate. In particular, for alevins, it delimits a rough area with the tail included (see Fig. 10.11(b)) and potentially some shadows and impurities if they are too close to the embryo. For the purpose of heartbeat detection, we need to restrict the search field to the thorax region.

Differentiation (egg or alevin)

Because eggs and alevins have different visual properties, we need to identify the type of embryo for further processing. The differentiation step is based on the morphological analysis of the embryo contours previously detected. The previous segmentation provides a reliable localization of the embryo, but only a rough approximation of its contours (see Fig.10.12(a) and (b)), so they need to be refined. For this, we consider the first frame $\mathcal{F}^1$ of the cropped sequence $\mathcal{S}^1$. In particular, it is crucial to weed out potential shadows and impurities, which may have been segmented with the embryo, while retaining the tail segmentation for the alevins. We apply the same bottom-hat procedure as in Eq. (10.11) to eliminate the background. Then we define experimentally an adaptive threshold slightly above the average pixel intensity near the border of the cropped frame. For our images, with a 8-bit depth, an increment of 5 was experimentally determined as appropriate: $\theta = \text{average}(\mathcal{F}^1) + 5$. We obtain a binary image, whose small components are filtered out with an area opening with parameter $\lambda = 5$. We apply the morphological gradient of an image I defined as $\text{grad}_M(I) \equiv \delta_{B_1}(I) - \varepsilon_{B_1}(I)$ to the resulting image, to obtain image C^1.

$$C^1 = \text{grad}_M(\gamma_5^\alpha((\varphi_{B_{40}}(\mathcal{F}^1) - \mathcal{F}^1)_{\geq\theta})) \tag{10.17}$$

In order to properly extract the contours of the embryo without confusing them with residual artifacts that can still be present in the background, we again use a markers-based watershed methodology on the image C^1, as follows:

$$m^2_{\text{int}} = \text{skeleton}(\gamma_{B_8}(\varepsilon_{B_{15}}(\mathcal{M}^0))) \tag{10.18}$$
$$\mathcal{M}^1 = \text{watershed}(C^1, m^2_{\text{int}}, m^2_{\text{ext}}) \tag{10.19}$$

The outline of the image is set as the external marker m^2_{ext} and the ultimate binary skeleton [105] of the eroded and opened mask $\mathcal{M}^0$ is set as the internal marker m^2_{int}. The erosion and the opening are respectively performed with a radius-15 and a radius-8 disk, in order to remove potential thin impurities linked to the previous embryo segmentation.

The outline of the resulting binary mask $\mathcal{M}^1$ is shown on Fig. 10.12(c) and (d).

We now use the shape of $\mathcal{M}^1$, to differentiate between eggs and alevins. Eggs are highly circular, so we could use the classical circularity attribute:

$$\text{circ} = \frac{4\pi A}{P^2} \tag{10.20}$$

Where A is the area and P is the perimeter of the binary shape under study. This ratio is at most equal to 1 for a disk, and decreases as the elongation becomes more pronounced. However it is still possible for some alevins to be so tightly wound that their associated binary shape presents a high circularity. Eggs also possess hairs on their chorion that may reduce the circular aspect of their associated mask. To correctly differentiate both cases, we also consider two other criteria. We have determined experimentally that a healthy well-segmented egg has a radius around 60 pixels. Therefore, allowing for some margin of error, we apply a morphological opening $\gamma_{B_{40}}$ that deletes the mask of every alevins, which are much thinner than eggs. If the component under study is filtered out during this step, it is considered an alevin. If it is not, we determine the minimum enclosing disk of the mask and we calculate the area difference d between this disk and the mask $\mathcal{M}^1$. Indeed, as eggs hairs are globally distributed uniformly on the chorion, the difference between the mask area and the area of its minimum enclosing disk is higher for hatched alevins than for eggs. Below a threshold of 3000 pixels, determined experimentally, we considered the component under study as an egg. Otherwise, we conclude it is an alevin.

Segmentation adjustment and frames cropping

For the purpose of heartbeat detection, it is essential to restrict the region of interest to the thorax region, in order to minimize the probability of false detection due to electronic noise or blood motion in the bright tail regions. Therefore, after the differentiation step, we refine the embryo segmentation in the alevin case. Alevins are darker than the background and in particular, their eyes are very dark. They are fairly easy to segment as a large connected component associated with the darkest minima in the thorax region. We apply the same threshold process as in Eq. (10.12). Because we determined experimentally the minimal area of the thoracic region of an alevin to be around 600 square pixels, we apply a morphological are opening using a criterion $\lambda = 100$ to eliminate small components. Moreover, we limit the thorax region to the $\mathcal{M}^0$, by computing the intersection.

$$\mathcal{M}^2 = \mathcal{M}^0 \cap (\gamma^{\alpha}_{100}((\mathcal{F}^0)_{\leq \theta_{oc}})) \tag{10.21}$$

The result $\mathcal{M}^2$ is a new binary mask representing thorax of the alevin (see Fig. 10.13(b)). We crop the sequence and obtain a new sequence $\mathcal{S}^2$ centered on this area (Fig. 10.13(c)). For egg sequences, these remain unchanged and $\mathcal{M}^2$ and $\mathcal{S}^2$ are respectively equivalent to $\mathcal{M}^1$ and $\mathcal{S}^1$.

Registration, denoising

We used the same procedure as before the changes in acquisition protocol, described in section 10.3.

Segmentation of the inner parts of the embryo

To ascertain the presence of a heartbeat in the thorax region, we look for cyclic motion in this region only. We develop a mask $\mathcal{M}^3$ of the inner part of the alevin

to avoid false detection in area most subject to noise, like the eyes. We define D^1 as the sequential average image of the sequence $\mathcal{S}^4$:

$$D^1 = \text{seq\_average}(\mathcal{S}^4). \tag{10.22}$$

In this image, the heart and vessels are darker because of the presence of blood, so they are easy to segment as a large connected component associated with the darkest minima, simultaneously maximizing the inter-class variance.

$$D^2 = (D^1)_{\leq(\theta_o+\theta_c)} \tag{10.23}$$

Here θ_o is obtained using the Otsu criterion, and we experimentally optimized θ_c to 20. The resulting D^2 is a binary mask of the registered body of the alevin. Considering D^2 as a geodesic mask, we now extract the eyes from D^2 as the one or two most prominent minima from its min-tree [57]. We cannot rely on the eyes being separated. Depending on the pose of the embryo, they might be merged. We write:

$$\mathcal{M}^3 = \varepsilon_{B_1}(\gamma_{B_3}((D^2.D^1)_{\geq(\theta_o-\theta_d)})) \tag{10.24}$$

Note that θ_o is re-estimated from the grey-level distribution of the embryo within the geodesic mask. Here we subtract an experimentally optimized constant θ_d, which turns out to be equal to 20 as well. The outline of the resulting mask $\mathcal{M}^3$ in alevins is exemplified on Fig. 10.14(a,b). this procedure is used only on alevins in order to restrict the region of interest for detecting heartbeats. It is not suitable for eggs, due to the folded aspect of the embryo. For these, we compute $\mathcal{M}^3$ using the same procedure for segmenting the eyes but considering $\mathcal{M}^2$ as the geodesic mask. (Fig. 10.14(c,d)).

Variation estimations in the mask of the inner parts of the embryo

Our challenge is now to detect cyclic motion-induced variations in the sequence assuming it corresponds to heartbeats. The heartbeat is noticeable by contrast variations due to blood concentration in the heart region. Computing the grey level variance at each pixel of a sequence along the time line shows the grey-level variability for that pixel, however this variability may not be due to a cyclic pattern. To account for cyclic variations, we split $\mathcal{S}^4$ into 4 sub-sequences of 7 frames ($\mathcal{S}^4_1$, $\mathcal{S}^4_2$, $\mathcal{S}^4_3$ and $\mathcal{S}^4_4$). An interval of 7 frames is similar to the expected period of a single heartbeat. We compute the sequential variance images $V_i = \text{seq\_variance}(\mathcal{S}^4_i)$ on each sub-sequence, yielding four variance images $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{V}_3$, $\mathcal{V}_4$. We next compute the sequential median of these four images seq_median($\{\mathcal{V}_i\}$). In this way, some spurious, potentially even large, residual variation occurring only once in the sequence will produce a large variance only once, and so the median of all the variances in this area will remain small. On the other hand, cyclic variation will exhibit significant variance in all four sub-sequences. The result is shown on 10.15.

$$\mathcal{S}^4 = \mathcal{S}^4_1 \cup \mathcal{S}^4_2 \cup \mathcal{S}^4_3 \cup \mathcal{S}^4_4 \tag{10.25}$$

$$\forall i \in [1,4], V_i = \text{seq\_variance}(\mathcal{S}^4_i) \tag{10.26}$$

$$E^1 = \mathcal{M}^3.\,\text{seq\_median}(\{V_i, i \in [1,4]\}) \tag{10.27}$$

Segmentation of cyclic motion areas and calculation of A, the area of the cyclic regions

We used the same procedure as before the acquisition changes, described in section 10.3.

10.7 Results

Our heartbeat detection method described in the previous sections returns three possible results : “unexploitable”, “alive” or “dead” and processes a sequence in around 10 seconds, which is in accordance with our initial constraints. We tested this procedure on 3192 videos acquired over several experimental runs. The protocol was always the same but conditions such as Methylene blue marker concentration changed. This number of videos is significant and concerns healthy as well as diseased embryos: some with edemas and with other malformations. Thus, this set reflects production usage and allows us to validate the robustness of our protocol (see Fig. 10.16).

The strongest way of establishing ground truth is to observe embryos under a microscope. Indeed, it is the most reliable way to assess the quality of the complete method of embryos analysis. We originally performed this study on the 2532 videos: embryos were observed under a microscope before the acquisition. However, we noticed that the discrepancy between video and microscope analysis was much higher (282 videos, for a rate of 11%). Several reasons linked to the experimental protocol or the acquisition method can explain this fact. (i) Video quality is such that some weak heart beats may be undetectable on a video even when it is visible under a microscope. (ii) Observations with a microscope also depend on operator fatigue and subjectivity. (iii) Because there is some gap between the microscope observations and videos acquisition, an embryo can also die during the interval. To avoid this, in future work we plan to assess the embryo twice: before and after the video analysis to detect heart failure for example. (iv) Observations under a microscope make it possible to observe embryos in a favorable position to see the heart, whereas in videos the embryo posture may obscure the heart. Finally, because of these differences, observations under microscope did not appear as the most relevant way to assess the quality of the program itself, as we wanted to do in this chapter. That is why it was decided to provide ground truth basing on the videos observations of several operators.

Firstly, all the 3192 video sequences were observed by an expert (named here "expert 1") who manually assessed the exploitability of the video by checking that it was complete, well-recorded, that the well was not empty and that the embryo was not too close from the well boundary. Then, for the usable sequences, the health status of the embryo was assessed by checking for the presence of a heartbeat. Whereas the determination of the videos usability is easy and thus reliable, the detection of a beating heart is sometimes difficult and therefore subject to errors. For this reason, the same work was performed by two other, independent observers ("experts 2" and "expert 3") on a subset of 200 exploitable videos. A consensus was then reached between the three observers on videos where the assessment differed.

Table 10.1: Results and error rates. (a) shows the distribution of dead and living embryos in the program results compared to ground truth and (b) the error rates calculated for each expert and for the program versus ground truth.

	Program: Dead	Program: Alive
Ground truth: Dead	38	0
Ground truth: Alive	3	159

(a)

	Expert 1	Expert 2	Expert 3	Program
Ground truth	1.5%	2.5%	1.5%	1.5%

(b)

The 200 data resulting from this consensus represent the ground truth of our assay.

Among the 200 ground truth videos, we identified 3 errors by the program for a corresponding error rate of 1.5% (see Table 10.1). Moreover, this error rate only corresponds to false negatives i.e living embryos detected as dead by the program. In toxicity tests, this type of error is more acceptable since they do not provide a false sense of security over the tested molecule. We present in Table 10.1(b) the error rates calculated for each expert compared to this final set of 200 ground truth videos. Expert 1, who processed the 3192 video set, has a similarity rate of 98.5% with the ground truth. We can consider their observations as sufficiently reliable to analyze the entire program results. Results of this analysis (program vs expert 1) are described below.

From an initial set of 3192 videos, 660 were flagged by the program as unusable (20.7%). 655 of them correctly so (99.25% of the subset "unusable"), and 5 (or 0.75% of the subset) incorrectly so due to some error in the program itself. If we consider the entire set, 3187 videos were well-flagged, leading to a success rate of 99.85%. The remaining 2532 were used for mortality test validation. There were 45 errors in this set, for an error rate of 1.77%. Such an error rate is low and can be considered satisfactory. We noticed that 11 of these 45 errors are due to embryos that had died a long time before the acquisition, and had consequently absorbed the blue marker. These embryos appear very dark on the video and so are more affected by noise, which was incorrectly labeled as periodic motion. This is something we can improve in a future version of our software pipeline.

In some cases, even actually dead embryos can appear to move. This can be due to motion in the water, fluttering, shadows or embryo rotation inside the well (see Fig. 10.17 for an illustration). In dark areas, acquisition noise is proportionnally more troublesome [106], and may be confused for cyclic motion. Sometimes, the embryo may appear to slide on the water. This can happen if we do not correctly compensate for rotation in the sequence stabilization phase. The main remaining cause is ambiguity: in some cases, the heart beats so slowly or weakly that we cannot detect it. In most cases, a human operator would also have difficulties in detecting

it. In our pipeline, all parameters, were hand-optimized on a training sample of 100 sequences. However, several parameters such as the size of structuring elements or thresholds can be optimized further than they already have, in particular those that are dependent on the size of the embryo.

The purpose of this chapter is to present and assess an image analysis pipeline for detecting a beating heart on 1 second long videos of fish embryos. Our results on 2532 usable sequences show an acceptably low error rate, near 1.5% overall. This proves the efficacy and reliability of our image analysis method. However, when considering its integration in the whole system of embryos preparation, image acquisition and processing, several points remain to be discussed, in particular concerning the validation phase and the ground truth.

Since we are discussing living organisms, establishing a ground truth is not always easy. We rely on multiple visual observations of a subset of the video sequences, which turned out not to be always consistent: expert observers did not always come to same conclusion. Indeed, we noticed 6 differences between them for 200 assessed videos, a rate of 3%. A second viewing of these videos was consequently performed with all observers present to achieve a consensus. With respect to this consensus, each observer had made between 3 and 5 errors, a rate between 1.5% and 2.5%. We note that our program also had made 3 errors. We conclude that the rate of assessment subjectivity is near 1.5%. Such a rate is acceptable, because experiments can easily be repeated three times. Assuming errors are random and independent, the final error rate on repeated experiments is negligible (2.25×10^{-4}).

However, this low error rate is based on video-based ground truth and is only representative of the quality of the program itself. To assess the reliability of the whole procedure, including preparation, acquisition and treatment, we need to establish ground truth by observing embryos under a microscope. Nevertheless, as we explained in 10.4, we still face to videos quality and accuracy issues with the current acquisition procedure. To improve this, we are currently investigating the use of the VAST system [2], which is however currently not available for Medaka, only for Zebrafish.

Many sequences (20.5%) are correctly detected by the program as unusable. Some of them are due to an empty well, and so are not an issue, but the majority are due to embryos being too close to the well boundary. This represents an actual problem for the efficiency of the global procedure. To solve it, we are currently experimenting with wells with a rounded, rather than flat bottom. These are not widely available commercially, and are not compatible with our Hamilton MICROLAB automated filling system. However, they are compatible with our acquisition device and they do solve the problem of the proximity with the edges. We hope that with some experimental protocol adjustments, we will be able to use them in production.

In spite of these issues, our current pipeline is used in production at L'Oreal. We are confident to eventually solve all of the stated experimental problems and reach an error rate comparable with a human observer under the best conditions,

[2] http://www.unionbio.com/vast/

using a microscope.

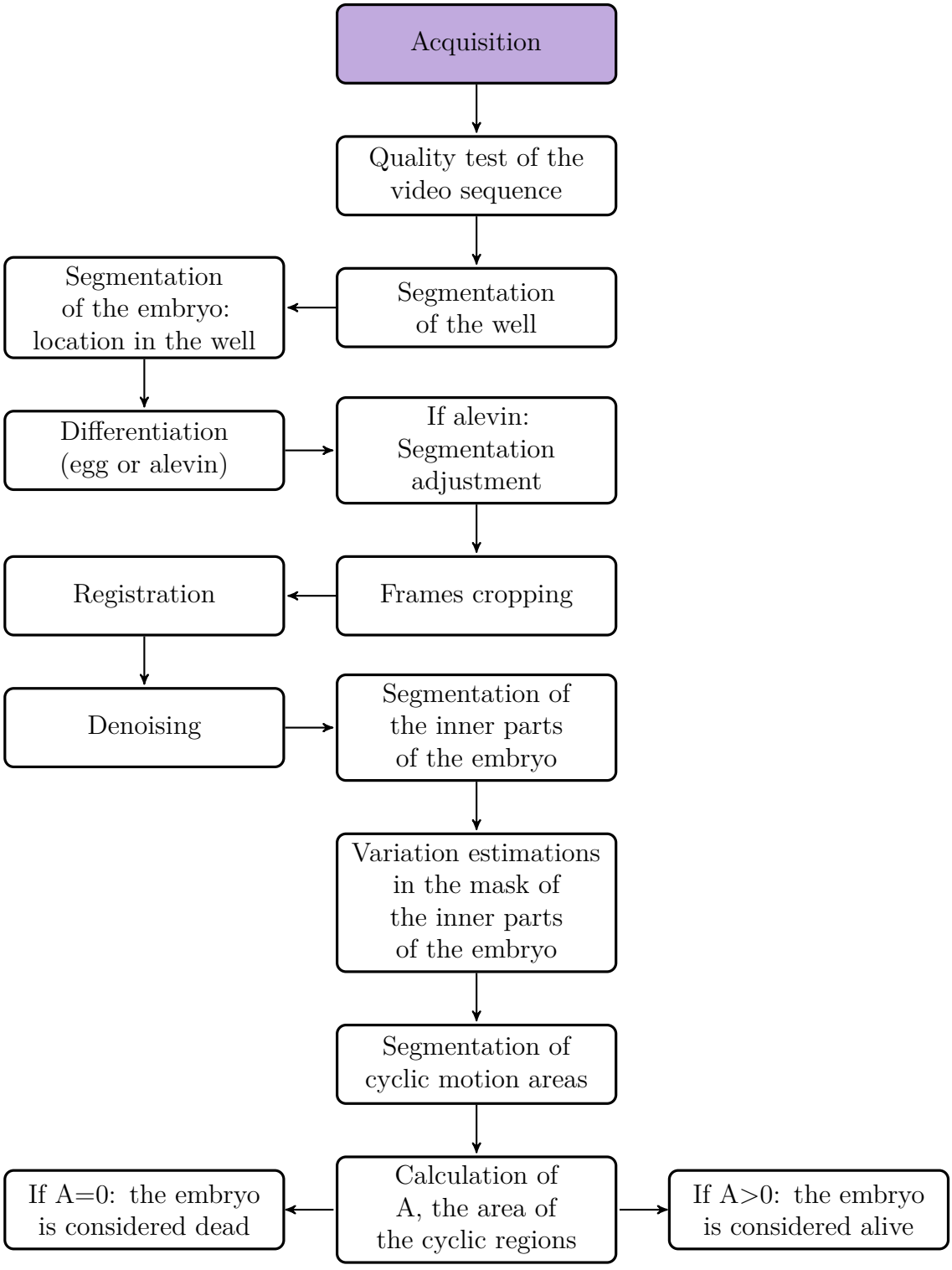
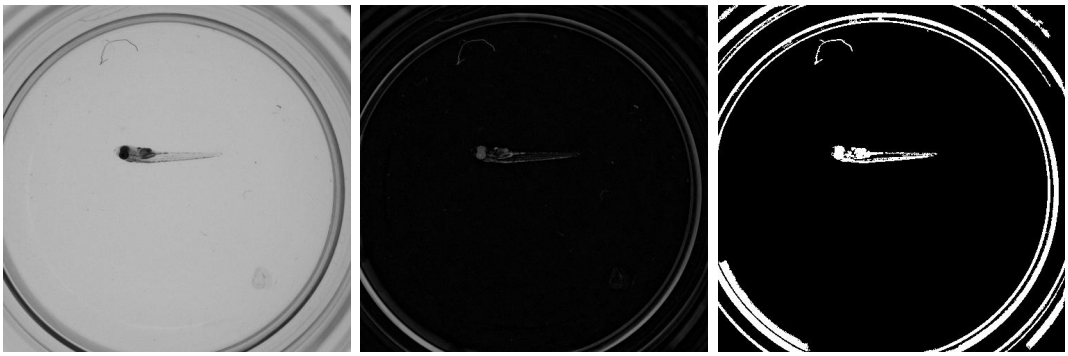
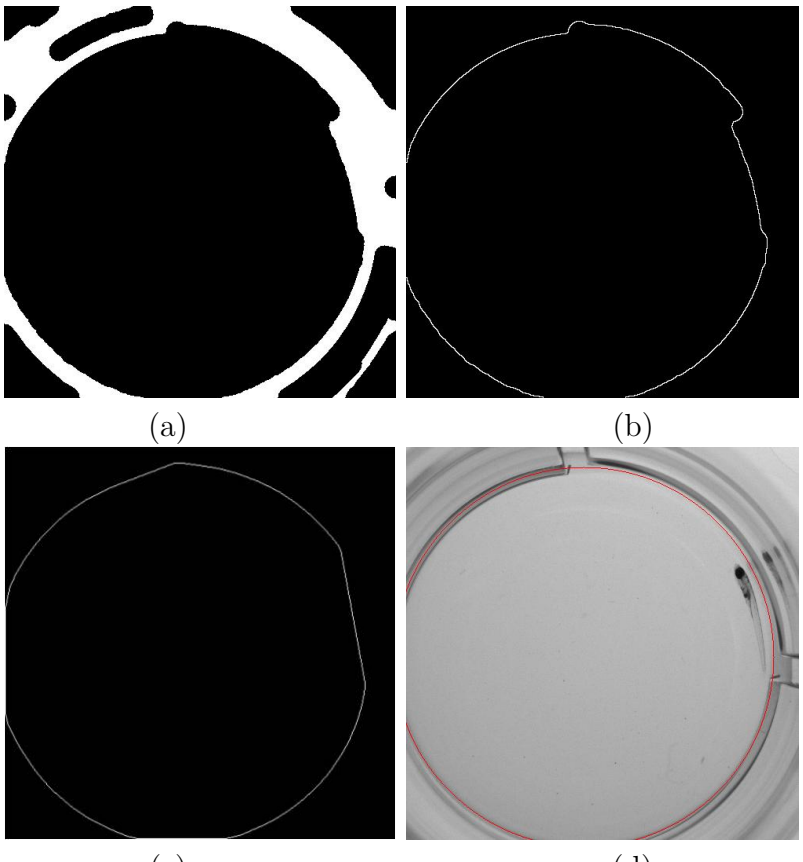


Figure 10.8: Flowchart of our embryo mortality image processing assay 2.



(a) (b) (c)

Figure 10.9: Bottom-hat application. (a) is the frame $\mathcal{F}^0$ before bottom-hat, (b) is the result A^0 and (c) is the subsequent thresholded image A^1.



(a) (b)

(c) (d)

Figure 10.10: Segmentation of the inner part of the well. (a) is before applying the watershed algorithm (image A^2); (b) is the result (image A^3); (c) is the result of convex hull (image A^4) and (d) is the outline of the final result $\mathcal{D}$ superimposed on $\mathcal{F}^0$.

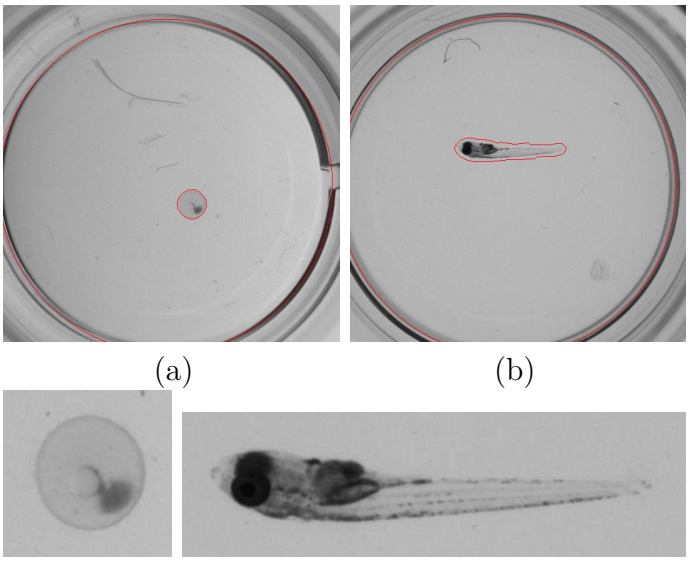


Figure 10.11: Segmentation of the well and location of the embryo. (a) and (b) show the outlines of $\mathcal{D}$ and $\mathcal{M}^0$ on the initial frame $\mathcal{F}^0$. (c) and (d) show the first frames of cropped sequences $\mathcal{S}^1$.

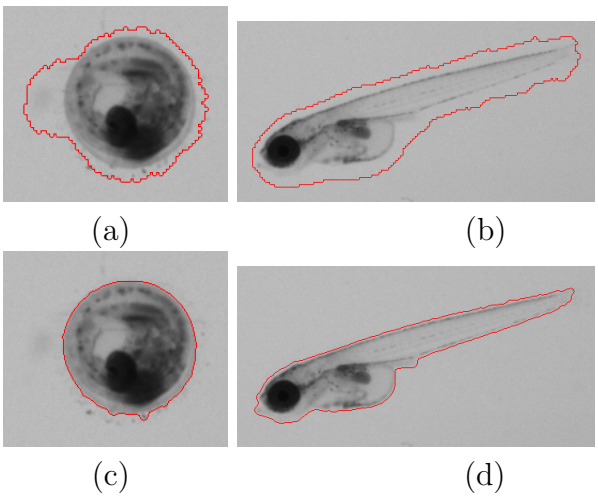
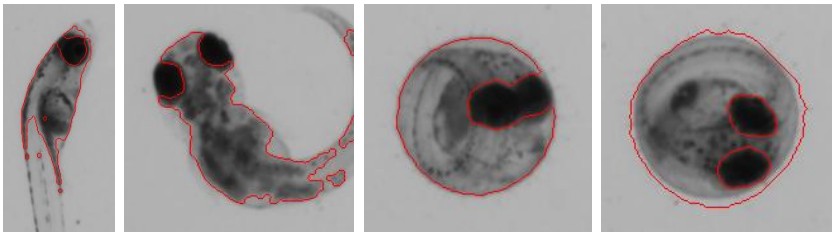


Figure 10.12: Segmentation of the embryo. (a,b) are the outline of the mask $\mathcal{M}^0$ (before adjusting) ; (c,d) are those of $\mathcal{M}^1$ (after adjusting).

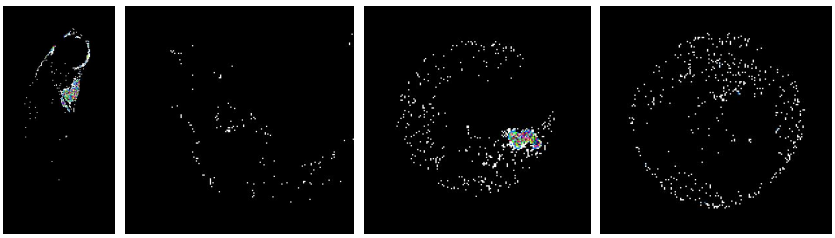
(a) (b) (c)

Figure 10.13: Segmentation of the initial frame to locate the thorax of the alevin. (a) is the initial frame $\mathcal{F}^0$, (b) the thorax mask $\mathcal{M}^2$ and (c) the first frame of $\mathcal{S}^2$



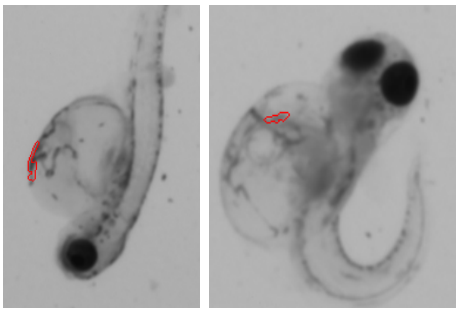
(a) (b) (c) (d)

Figure 10.14: Inner parts segmentation on two alevins and two eggs. (a) and (c) are alive whereas (b) and (d) are dead.



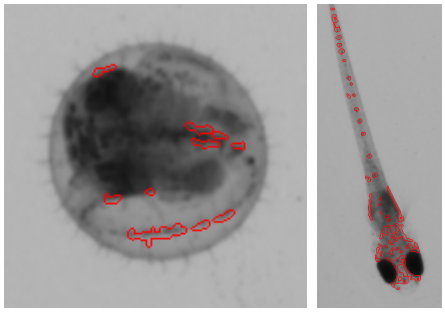
(a) (b) (c) (d)

Figure 10.15: False color rending of the temporal variance. (a) and (c) are for the living embryos in Fig.10.14(a,c); (b) and (d) for the dead ones in Fig.10.14(b,d).



(a) (b)

Figure 10.16: Heart segmentation in the presence of malformations. (a) is in the presence of edema and (b) in the presence of axial malformation.



(a) (b)

Figure 10.17: Incorrect segmentations due to fluttering.

11

Heart frequency estimation

In this chapter we study elements that exhibit cyclic behavior, such as a beating heart, or the speed of blood cells in vessels. Various methods already exist for estimating frequencies based on motion: for instance tracking methods [107, 108], integrated box methods [17], measuring the amount of changing pixels between frames [51], and optical flow [78]. In our case, we developed a fully automatic method based on segmentation and optical flow to reliably estimate heart frequency directly in heart or in vessels of Medaka embryos.
This work was given an oral presentation at IPTA 2016 [6]

Contents

11.1 Pipeline

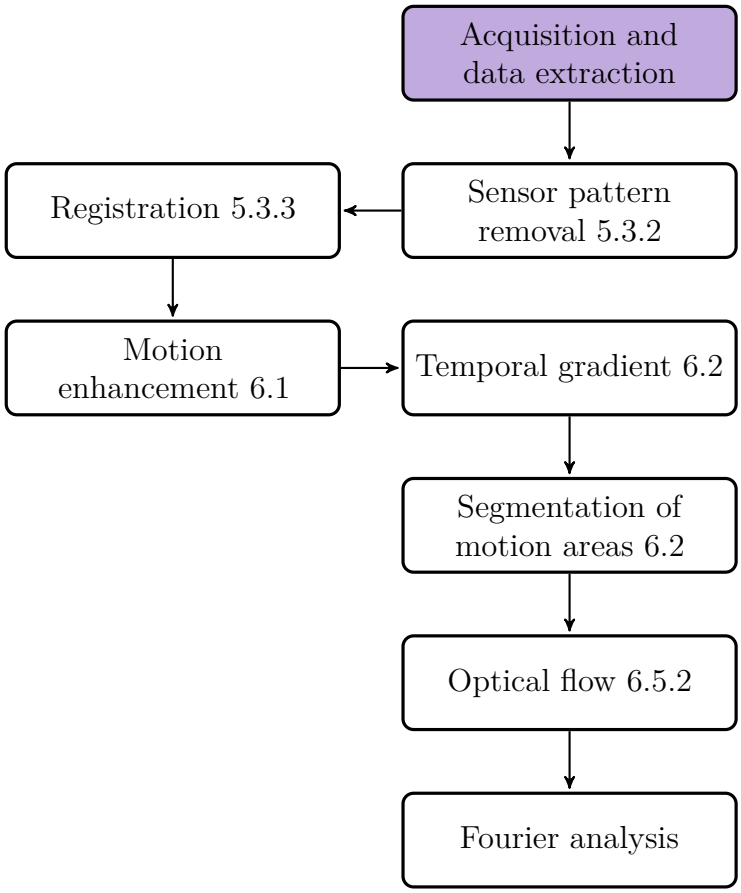


Figure 11.1: Flowchart of our Heart frequency estimation.

11.2 Details of the methodology

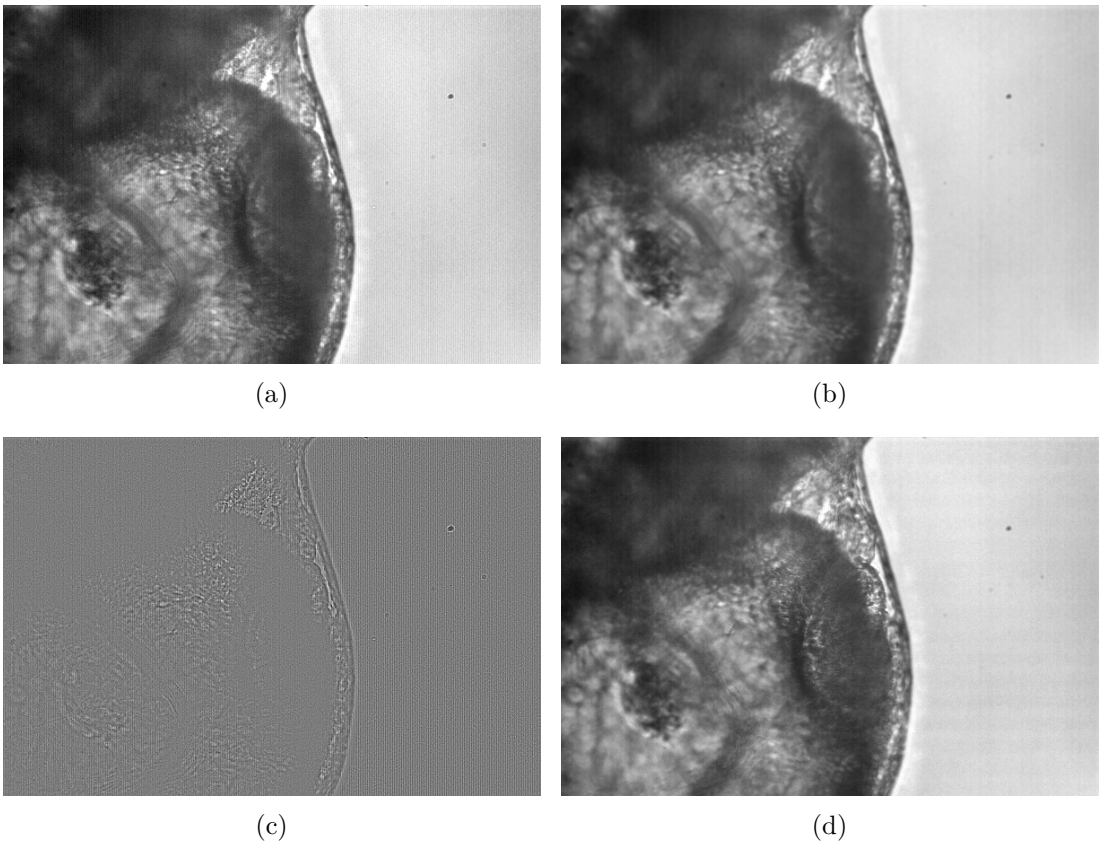
Medaka embryos were anaesthetized with tricaine for immobilization, placed under a Leica DMi8 inverted brightfield microscope and 600×480 2D video were acquired with a ×20 objective and a ×10 focusing lens, at a frame rate of 100 frames per second via a high-speed camera. The calibration was 2μm per pixel, sufficient to identify the motion of blood cells in the bloodstream. Embryos were captured at various stages of their development between 1-7 days after hatching, and exposed to various substances, some of which induced either bradycardia or tachycardia. The acquired video stream exhibit various artifacts, including noise, global motion due to vibrations, remaining embryo motion and spurious stage motion, as well as sensor pattern. After sequence stabilization, this grid would no longer be fixed and so would impair further processing. We therefore need to remove this texture beforehand.
We used the three pre-processing steps described in Chap. 5 and 6.

Sensor pattern removal (Chap. 5.3.2)

As presented in Chap. 5.3.2, we obtain an image of this texture by computing the average image of the unstabilized sequences, and we subtracted it from the same image on which we applied a Gaussian filter with $\sigma = 1$. This texture is then subtracted from each image of the whole sequence. This is illustrated in Fig 11.2.

$$\forall I \in \mathcal{S} : I_{clear} = I - \left(\bar{\mathcal{S}} - \mathcal{G}_{\sigma=1}(\bar{\mathcal{S}})\right) \tag{11.1}$$

with $\mathcal{G}_{\sigma=1}$ a gaussian filter with standard deviation $\sigma = 1$, $\bar{\mathcal{S}}$ is the average of the sequence $\mathcal{S}$. $\mathcal{S}_{clear}$ is the sequence of the I_{clear}.



(a) (b)

(c) (d)

Figure 11.2: Removing sensor pattern from the acquisition. (a) The average of the sequence yields the non-moving parts of the sequence. (b) Blurred average removes thin and textured elements. Computing (a)-(b)=(c) yields the sensor pattern. (d) Finally, we subtract this pattern from all the images of the sequence.

This removes the grid without blurring the sequence, yielding to the sequence $\mathcal{S}_{clear}$.

Registration (Chap. 5.3.3)

Next, we use our robust rigid registration method (see Chap. 5.3.3) to stabilize our structures of interest, based on detection of key-points and RANSAC-like method [87] to weed out the outliers.

$$P_1 = P_2 \times \mathcal{R} + \mathcal{T} \tag{11.2}$$

Where P_1 and P_2 are the sets of points of the two frames we seek to match, $\mathcal{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ is the rotation matrix. $\mathcal{T} = (d_x, d_y)$ is the translation;

$$\forall (x, y) \in I_{clear} : I_{reg}(x, y) = I_{clear}(x', y') \tag{11.3}$$

with the estimated $\mathcal{R}$ and $\mathcal{T}$, $[x'\ y']^\top = \mathcal{R}[x\ y]^\top + \mathcal{T}$. $\mathcal{S}_{reg}$ is the resulting stabilized sequence of I_{reg}

Motion area highlighting (Chap. 6.1)

Then we compute the average image of the sequence (Fig. 11.3 (b)), yielding an image of the non-moving parts of the sequence as presented in Chap. 6.1. We subtract this result from each frame. This operation removes the static parts of the sequence and yields only a sequence of moving elements.

$$\forall I_{reg} \in \mathcal{S}_{reg} : I_{mov} = I_{reg} - \bar{\mathcal{S}}_{reg} \tag{11.4}$$

with $[x'\ y']^\top = \mathcal{R}[x\ y]^\top + \mathcal{T}$. $\mathcal{S}_{mov}$ is the sequence of I_{mov}

We add a new step after this pre-processing procedure: an illumination correction step. Indeed, illumination variations may occur during the sequences. To correct for these, we compute the average value of the sequence, and multiply each pixel by this value divided by the average of the frame. This operation ensures that the average value of each frame is constant, and so will normalize the sequence in term on grey-level intensity value. For each pixel :

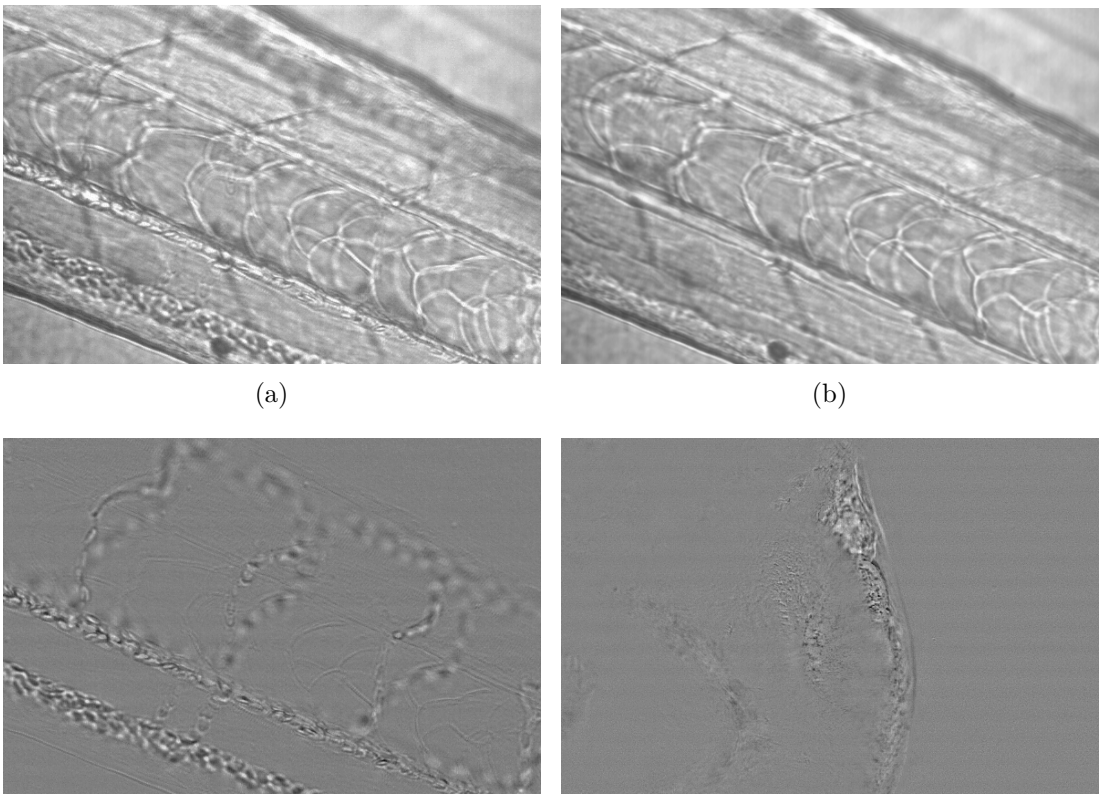
$$\forall (x, y) \in I_{mov} : I_{corr}[x, y] = I_{mov}[x, y] \frac{\mu_{\mathcal{S}_{mov}}}{\mu_{I_{mov}}} \tag{11.5}$$

where $\mu_{\mathcal{S}_{mov}}$ is the average value of the sequence and $\mu_{I_{mov}}$ is the average grey-level of the frame. We obtain a sequence S_{corr} where the average value of all frames is constant throughout the sequence.

Temporal gradient and segmentation of motion areas (Chap. 6.2)

At the end of our pipeline, the areas where motion is present are our regions of interest: the blood cells. When we observed the tail of the embryos, we notice severan "moving" areas: the main artery, the main vein and capillaries. We are only interested in the artery in which blood cells will exhibit a cyclic motion due to the natural elasticity of this particular vessel. We hence segmented the vein and the artery using the methodology developed in Chap. 6.2. Indeed, vein and artery are the parts of the sequence where the motion is the most important. On heart sequences, we segment only the heart region as it is the part of the image with the most motion present.

We segment the areas according to connectivity criteria and rough motion. To achieve this, we integrated the absolute value of the first few frames to obtain an



(a) (b)

(c) (d)

Figure 11.3: Detecting the moving areas of the sequence. (a) Initial frame. (b) Averaging of the sequence yielding the non-moving parts, including all the parts that are not blood cells. (c) Areas with motion from the difference between frames and (b). (d) Equivalent result on a beating heart.

image of the motion areas. Then we computed a labelled mask using a morphological closing and connectivity criteria [54].

$$\text{sumdiff} = \sum_{i=0}^{10} I_{corr}^{i} - I_{corr}^{i+1} \tag{11.6}$$

$$\mathcal{M} = \gamma_{\mathcal{B}_5}(\varphi_{\mathcal{B}_5}((\text{sumdiff})_{\geq\theta_O})) \tag{11.7}$$

Where θ_O is the Otsu threshold [89]; $\gamma_{\mathcal{B}_5}$ and $\varphi_{\mathcal{B}_5}$ are respectively the morphological opening and closing [53] by structuring element $\mathcal{B}_5$, a discrete Euclidean ball of radius 5. $\mathcal{M}$ is the mask of the vessels. We created $\mathcal{M}_L$ in which all connected component are distinguished with the attribution of a label. We should expect a maximum number of 2 labels (corresponding to the artery and the vein) for vessels sequences, whereas in heart we should expect a unique label. Figure 11.4 shows the steps of this segmentation.

Optical flow (Chap. 6.5.2)

Tracking methods follow one or more elements of a sequence. Hence they can

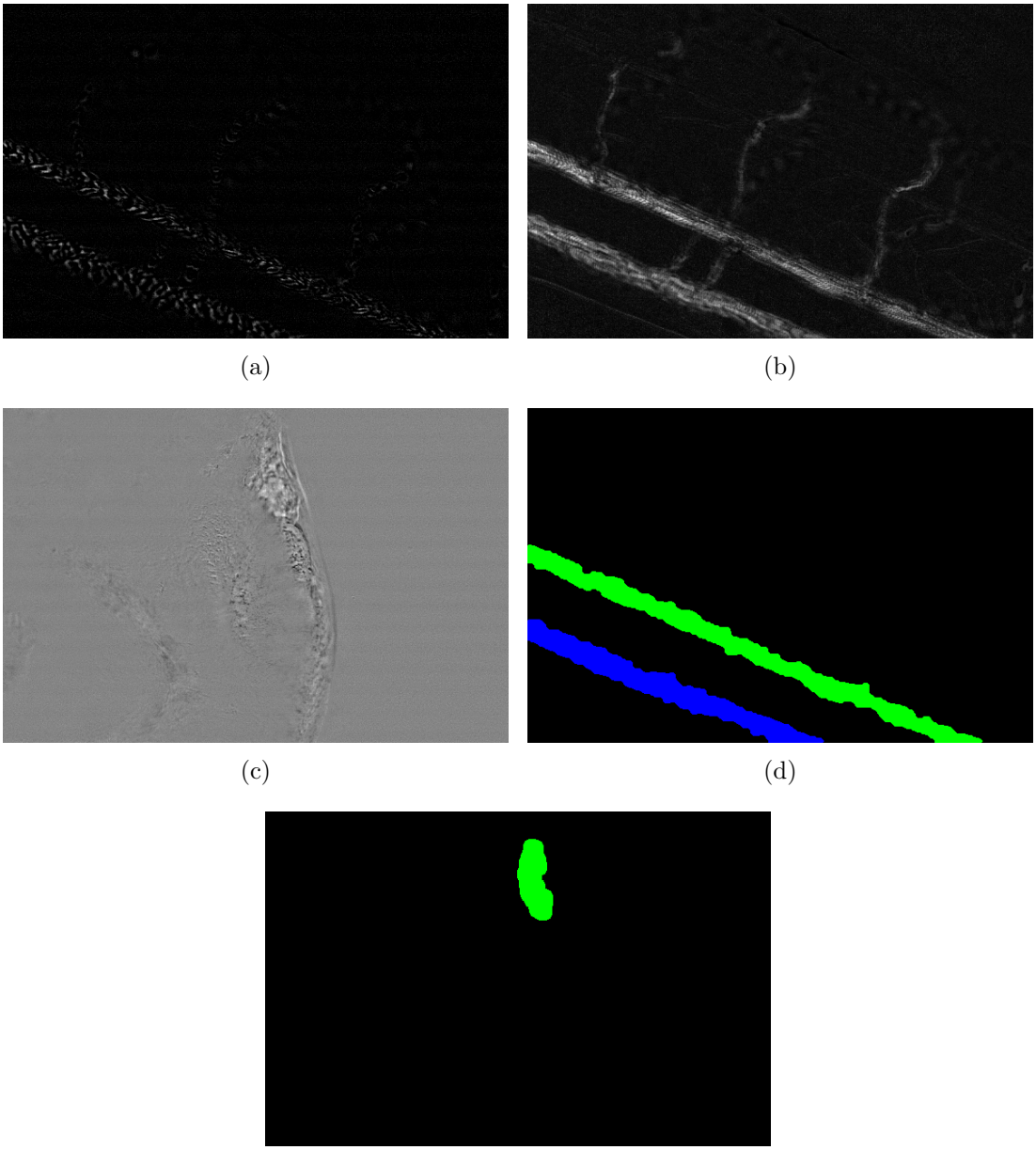


Figure 11.4: Sequence of operations for segmentation. (a) Difference between two consecutive frames. (b) Integrating motion yields a segmentation (c) and vessel labeling (d). Equivalent labeling in a heart (e): a single area is segmented.

estimate a proper motion together with a position, and so can be used to estimate frequencies associated with variations in motion. Integrated boxes, such as linescans, measure the change of luminosity in a fixed image subset. Change of luminosity in the box over time yields a time series associated with the motion. Variations in grey level in the box through the sequence can be used to estimate the frequency of spatial cycles. Similarly, measuring the number of changing pixels between frames is correlated to the speed of the change, under ideal conditions [51].

In our case, tracking methods are in practice unusable, particularly since moving elements of interest can be difficult to identify and segment. Indeed, the appearance of blood cells may change significantly during motion, since it occurs in 3D and we only measure 2D information. We cannot use integrated box methods either because motion may not be spatially cyclic, is 3D and may not involve large illumination variations. Similarly, measuring changing pixels between frames is simple, but does not correspond to a physical measurement, is sensitive to noise and spurious motion. This measurement is sufficient for the vessel segmentation presented above but not for motion analysis. In contrast, optical flow can be rendered insensitive to global motion via image stabilization, is generally robust to noise and does indeed correspond to a physical measurement. While the study of speed variation can clearly be connected to heart motion and so heart rate, this is not so clear in the blood vessel case. In reality we assume that artery wall elasticity converts changes in pressure to changes in speed, as in the human vessel case. For this reason, we estimate blood cells motion using the optical flow method as described in Chap. 6.5.2.

We used the Färneback's algorithm for computing the optical flow [79]. It works by dividing the frame with a grid, and calculating global motion in relation to its consecutive frame for each zone. The global motion is assigned to a point representing the center of the zone. For each point we have a displacement vector (see Fig. 11.5(a,b)). This method is suitable even though is not possible to track individual blood cells in our videos.

$$V(i) = \bigcup_{\alpha \in M_L} V_\alpha(i) \tag{11.8}$$

$$\forall \alpha \in M_L, V_\alpha = \{v_i \in V / \mathcal{M}_L(x_{v_i}, y_{v_i}) = \alpha\} \tag{11.9}$$

$$\mathcal{V}_\alpha(i) = \text{median}(V_\alpha(i)) \tag{11.10}$$

$V(i)$ is the list of the displacement vectors between two frames $I_{corr}(i), I_{corr}(i+1)$. $V_\alpha(i)$ is the list of the displacement vectors belonging to the component α of M_L. For each region α, we have a median displacement value $\mathcal{V}_\alpha(i)$ which is the global value of displacement in the α region. We validate the choice of the median with speed vector clustering (see Fig. 11.6).

We can observe two different behaviours in blood vessels: the speed profiles of veins and arteries are different. We observed wide variations of speed in the arteries while in the veins the speed is almost constant (see Fig. 11.5(d)). This is to be expected since arteries are closer to the ventricular expulsion, whereas veins convey blood that has gone through the whole body and so speed variations are damped. For this reason, we focus only on measurements in arteries to extract physiological parameters such as heart rate.

We identify average motion vectors with each segmented and labelled region. The plot of the magnitude of the speed vs. time in each labelled region allows us to identify useful physiological measurements, such as frequency, minimum and maximum speed, motion regularity and so on.

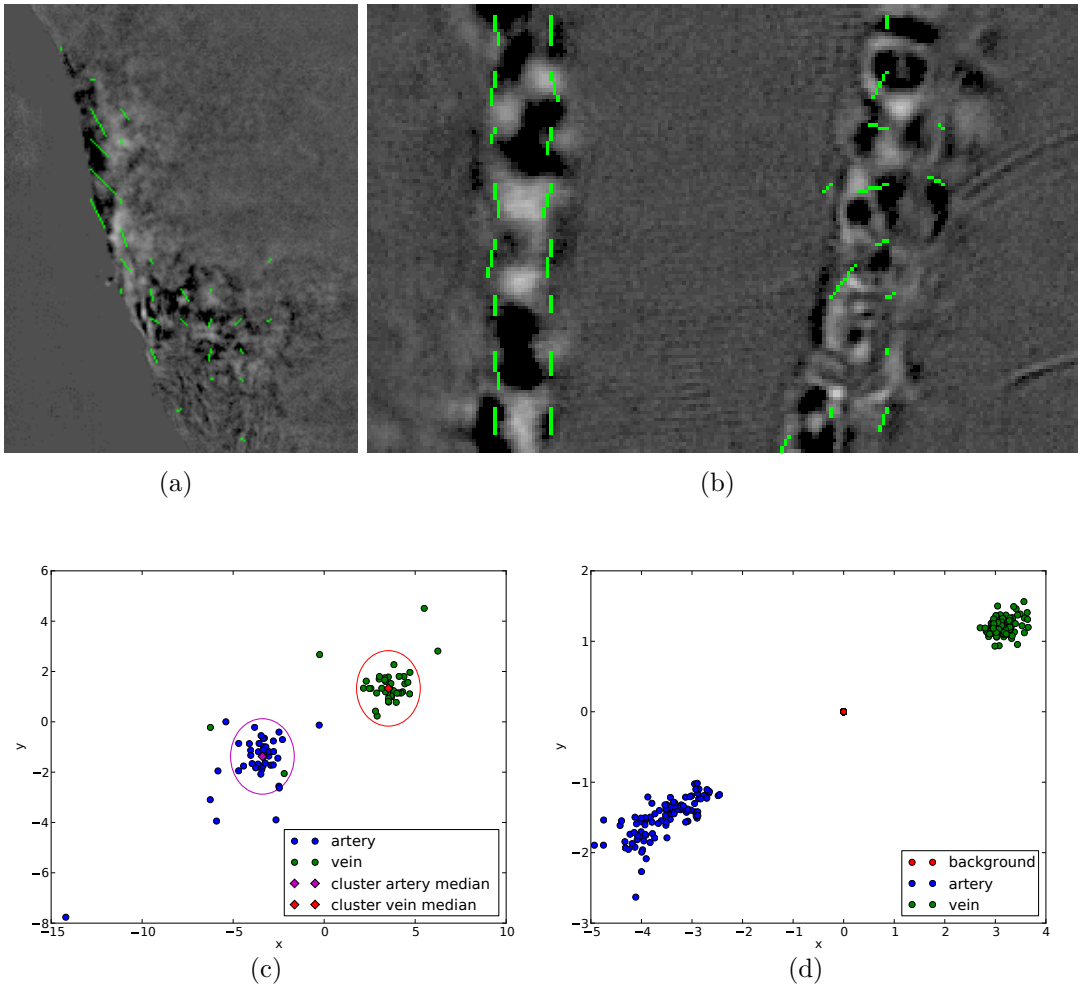


Figure 11.5: Speed analysis on a sequence. Optical flow estimation in a heart (a) and vessel (b). Left vessel is the artery, right is the vein. (c) Speed vector clustering and median values, notice speed consistency in each labeled region. (d) Median speeds during a whole sequence, showing direction and speed variation.

Fourier analysis

Frequency analysis is performed using the Fast Fourier Transform on the magnitude vector obtained before.

11.3 Results and validations

Fig. 11.6 shows examples of speed variation and associated Fast Fourier Transform in heart and vessel.

Our optical flow-based method can be used to measure speed variations in arteries and around the heart, and so should be correlated with heartbeats. To validate our approach, we first consider frequency estimations in heart sequences. Then we compared frequency estimation in vessels with frequency estimation in the heart. Finally, we compare our speed estimations with another method.

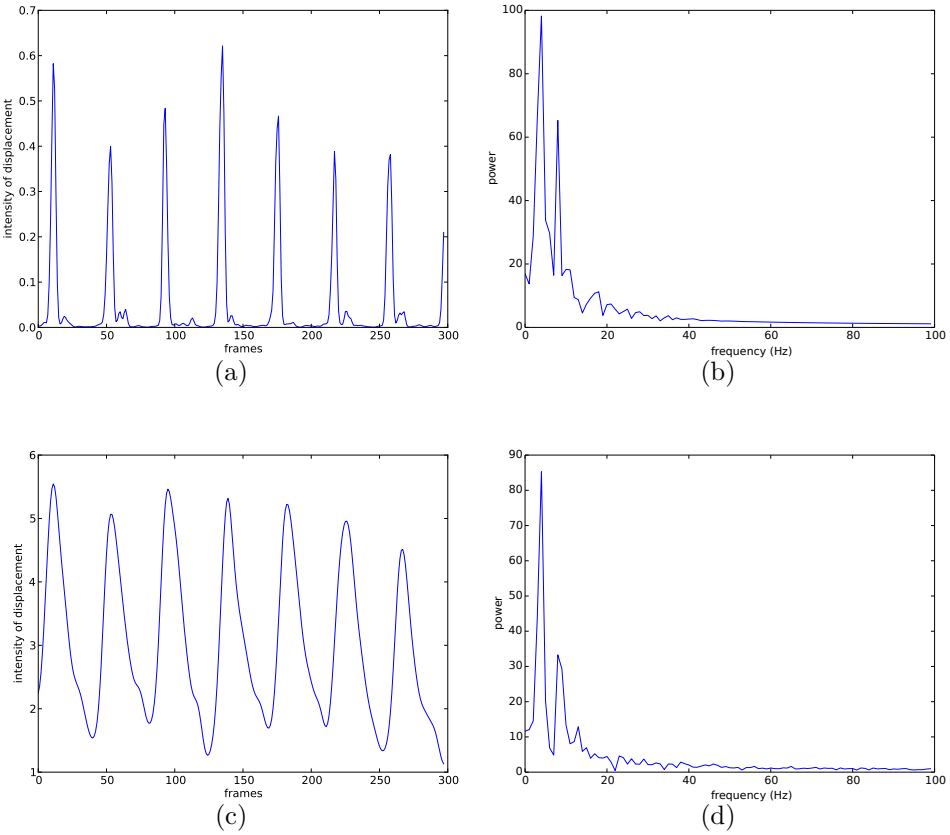


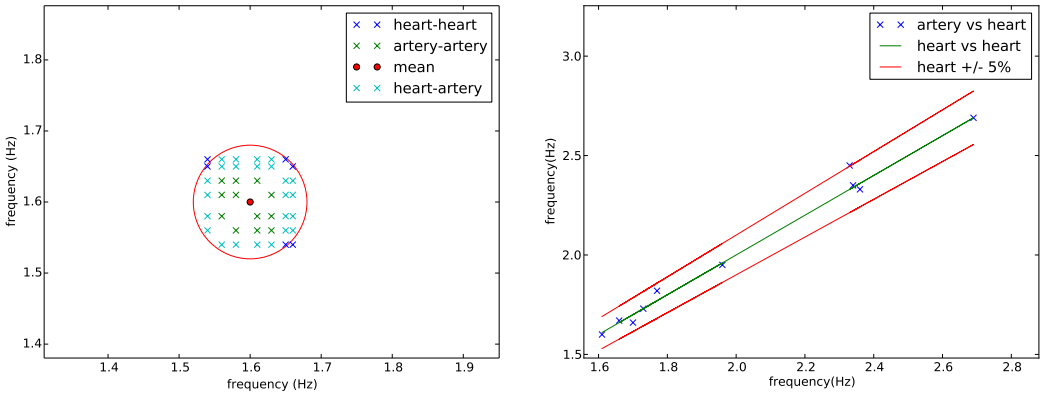
Figure 11.6: Speed frequency analysis on a sequence. Speed variations vs. time over a complete heart sequence (a) and vessel sequence (c). Corresponding Fast Fourier Transform (b) and (d) (cropped to 100 Hz).

Manual heart rate validation

For a first validation of our frequency estimation procedures, we first estimated the beating frequency in the heart manually. This is acceptable since the fish embryo heartbeat is easy to follow in the sequence. We recorded several sequences on various embryos (the video acquisition lasted several minutes for each embryo) and we counted the number of heart cycle by watching the sequences at low speed. We compared it with the estimation of our approach. In all the cases, we obtained the same results for both methods up to rounding.

Heart vs. vessel validation

To validate our approach in vessels, we estimated heartbeat frequencies in a single fish embryo. Each estimation was performed on 18 different video acquisitions separated by several minutes, and we estimated the frequencies in both the heart and the arteries. Because heart estimation was validated previously, we considered it as our ground truth. In Fig. 11.7(a), we plotted all possible couples of measurements against each other. All measurements appear consistent and fit within a 5% uncertainty circle, so we can now consider them reliable.



(a) (b)

Figure 11.7: Correlation between heart and arteries sequences (a) in a single fish embryo with multiple observations and (b) for 10 subjects exposed to varying environmental conditions.

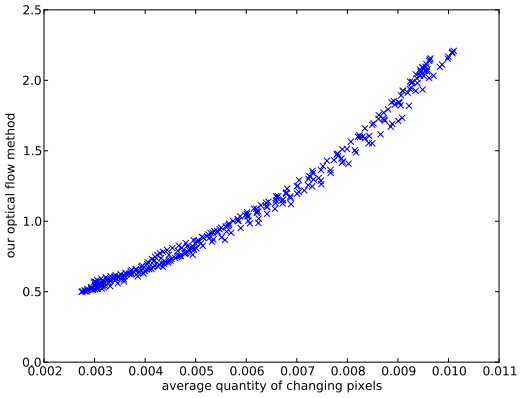


Figure 11.8: Comparison with the method of [51] in the ideal single-vessel case.

Heart rate variations validation

We also measured the heart rate in the heart and artery in 10 fish embryos subjected to varying environmental conditions, causing the heart rate to vary significantly. We see that our estimations of the heart rate and the frequency measured in the arteries correlate very well as illustrated in Fig. 11.7(b).

Speed measurement validation

Finally we compared our blood cell speed measurement vs. the method described in [51]. This latter, less sophisticated method relies on counting changing pixels between frames rather than estimating a flow. It is only usable in a sequence containing a single vessel, in a noise-free and vibration-free environment. The relationship between the number of changed pixels and speed is also not clear. However in this ideal case it is well-correlated, although non-linearly, with ours. The non-linearity in the correlation may be explained because our speed estimate uses a multi-scale approach, whereas [51] is single-scale. Consequently our speed

estimates have a wider range (a ratio of 5 vs. 3 between the highest and lowest values). However, reproducibility between cycles is good (see Fig. 11.8).

11.4 Further work

We have shown that we can successfully estimate frequencies from optical flow speed variations, both around the heart and in blood vessels after segmentation of these structures. The frequencies are consistently estimated and correlate with each other with a low margin of error. This ability to accurately estimate the heart rate from optical flow rather than intensity variations extends the range of possible physiological measurements. Indeed, it is often easier to locate an artery in the tail of a fish embryo than to obtain a clear view of the heart, especially in later stages of development. A possibility for for future work will be to study more sophisticated measurements such as estimation of ejection fractions.

Part V

Conclusion

12
Conclusion

In this manuscript, we have described some tools that we have developed for motion analysis studies in a bio-medical context. These tools have been illustrated through several applications on ciliated cells and fish embryos. Many perspectives are left open at the end of this work, and will be followed by new development and applications.

Contents

12.1 Contribution of this work

In this manuscript, we have proposed a number of motion analysis tools for biomedical applications, in particular for 2D+t sequences. While the processing of sequences is a common application in both the biomedical world and in computer vision, the entirely automated analysis of such sequences is much less so. These tools have been developed in the context of high-speed bright field and fluorescence microscopy, but can easily be adapted to other domains.

Our main area of application was the analysis of ciliated cells. We have sought to automate the characterization of cilia motion. As detailed in this document, ciliated cells in the body are critical for lung clearance. Many acquired and genetic diseases that affect clearance can be traced back to cilia motion disorders. Currently, existing tools for assessing cilia motion disorders are all semi-automated and require a significant amount of practitioner input. Our aim was to provide practitioners new protocols for cilia motion disorder assessment and thereby new tools for the diagnosis of diseases that affect clearance. For this we have proposed a number of increasingly sophisticated methods that are able to automatically pre-process images and provide validated, reliable measurements, such as the cilia beating frequency, in a regionalized manner (i.e per group of cilia independently). More generally, we have proposed a methodology for describing motion patterns, and segment image sequences into regions of similar motion, from which various descriptors can be derived. These can potentially be used further for as a direct input for a diagnosis system.

Importantly, we have also investigated the possibility of estimating the motion of cilia directly in vivo. This is particularly challenging given the size of cilia and their beating frequency. This task essentially requires a high power microscope combined with a high-speed camera. Since no instrument currently exist to perform this task, we formed a partnership with Mauna Kea Technology (MKT) and used their instrument on ex-vivo samples. We have conclusively shown that, even though the current MKT technology is not designed for this task, it is possible to at least automatically measure cilia beating frequencies with this instrument and that these measurements compare well with manual ones.

A secondary application was the analysis of fish embryo in the context of ecotoxicity. These living organisms are increasingly used to assess the toxicity status of test substances, for instance in cosmetic formulations. More generally, they can be used to assess the degree of pollution in waterways. For these organisms, we were able to provide validated sequence-based assays of the life/death status of fish embryo, with high accuracy, similar to that of trained human operators. We have also proposed various methods for estimating the heart frequency, from observing the heart of from blood circulation.

All of our applications have been validated, sometime extensively. Our automated life status assay was validated over more than 3000 videos, with an error rate estimated at 1.5%, comparable to human experts. Our error rates on the others assays and procedures are consistently in the 1-5% error range and have been published in one form or another, see the publication list in the introduction to this manuscript. We stress that all of our software was developed using open-source

tools, in order for our results to be reproducible.

12.2 Future work

Our current proposals allow us to be optimistic regarding the future of sequence analysis tools for bio-médical application. In particular, the following areas can be improved:

- Segmentation

  The current segmentation method that we have proposed in chapter 5 relies on a hierarchy of temporal Fourier descriptors. It would be interesting to add more spatial information and to allow other types of continuous wavelets (Gabor, Morlet) instead. Our segmentation model is a variation of the Random Walker algorithm. It would be interesting to explore other metrics (graph cuts, power watershed, etc), especially for speed and reliability in the presence of many regions.

- Pattern analysis

  From these segmentations and descriptors, it would be useful to be able the patterns of visible cilia motion and to ascribe these to various diseases, via a classification/machine learning approach. Some annotated data is available in this domain so at least an experimental study could be conducted.

- Clinical study

  While we have validated most of the measurements that we have proposed, we have done so only at a low level, i.e. we have shown that they do correspond to equivalent manual or semi-automated measurements. However their usefulness in a clinical setting has not yet been proven. This is a study that is well beyond the purpose of this thesis but is of obvious interest.

- More data from fish embryo sequences

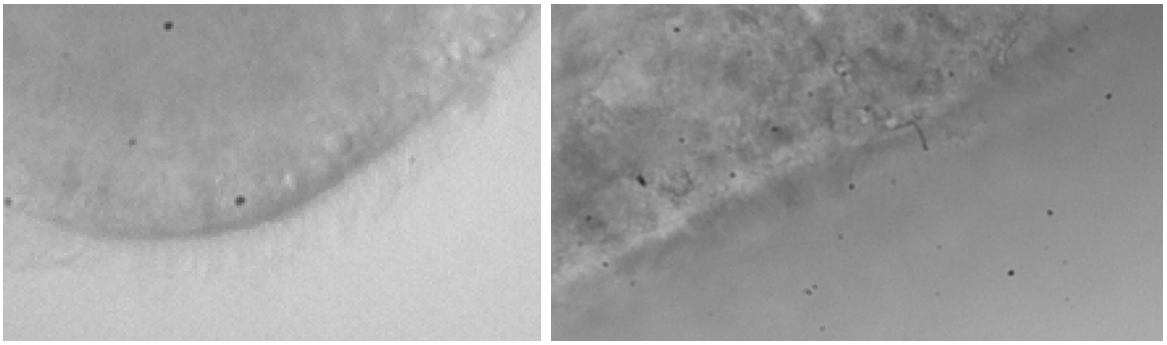
  More data can be extracted from sequences of beating hearts on fish embryo. We believe that physiologically relevant measurement can be extracted also, such as the ejection fraction. More sophisticated image acquisition devices also exist, e.g. the VAST system from Union-Biometrica, that allows practitioners to image embryo from many angles and not solely in 2D. Also the possibility exists of real 3D sequence analysis, for instance using fluorescence or bi-photon microscopy.

- Continuing the in-vivo study

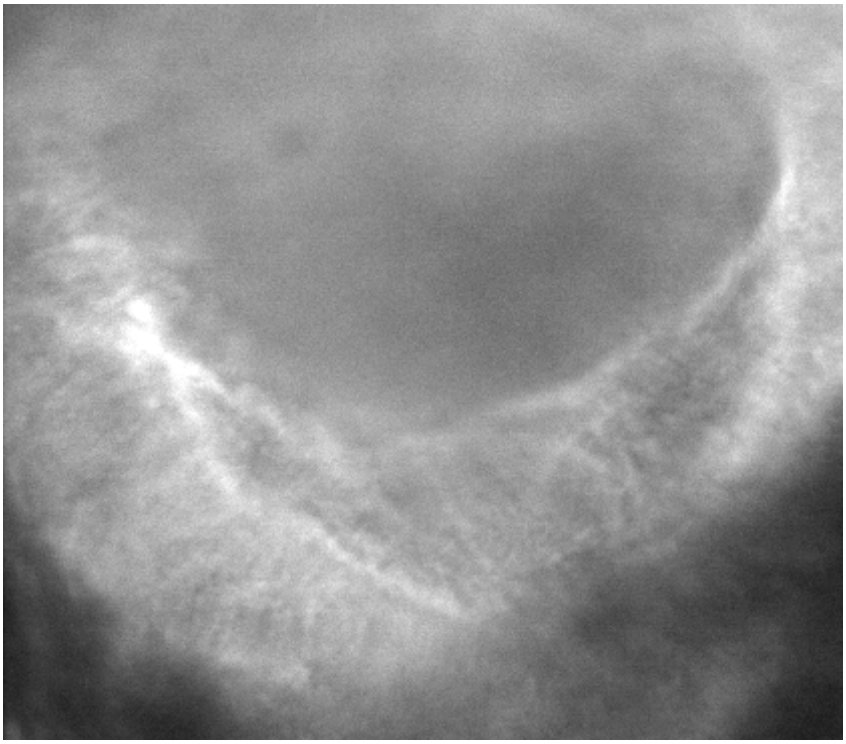
  At present the collaboration with MKT on in-vivo analyses has been halted, however other provider of similar instruments exist. For the moment we have not been able to raise sufficient interest, but this could change. The possibility of using rigid, bright-field, high-power endoscopes instead should also be investigated.

- Enlarging the field of applications of cilia analysis

  We obtained sequences of beating cilia from mouse brain and from zebrafish nasal cavities. One goal is to analyze these new data with our procedure to test our method and to adapt it if needed to propose a "universal" cilia analysis tool. The new data were provided by Nathalie Delgehyr (Ecole Normale Supérieure de Paris, France) for the brain, and Nathalie Jurisch-Yaksi (Kavli Institute for Systems Neuroscience, Sweden) for the zebrafish. They are different from the nasal ciliated cells we have been working with so far (see Fig. 12.1), but our procedures can certainly be adapted to them. A complementary study on flow analysis with particle tracking on living zebrafish embryo is also planned.



(a)



(b)

Figure 12.1: Ciliated cells images from brain sequences (a) and from zebrafish nasal cavity (b)

References

[1] E. Puybareau et al. “A regionalized automated measurement of ciliary beating frequency”. In: *Biomedical Imaging (ISBI), IEEE 12th International Symposium on*. IEEE. 2015, pp. 528–531.

[2] E Puybareau et al. “Automating the measurement of physiological parameters: a case study in the image analysis of cilia motion”. In: *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2016, pp. 1240–1244.

[3] E. Puybareau et al. “Towards the in-vivo automated assessment of cilia motility”. In: *Congress of the European Rhinologic Society*. Stockholm, July 2016.

[4] E. Puybareau, M. Léonard, and H. Talbot. “An Automated Assay for the Evaluation of Mortality in Fish Embryo”. In: *Mathematical Morphology and Its Applications to Signal and Image Processing*. Vol. 9082. Lecture Notes in Computer Science. Reykjavik: Springer, May 2015, pp. 110–121.

[5] E. Puybareau et al. “An automated assay for the evaluation of mortality in fish embryo”. In: *Computers in Biology and Medicine* (). In print.

[6] E. Puybareau, H. Talbot, and M. Leonard. “Automated heart rate estimation in fish embryo”. In: *Image Processing Theory, Tools and Applications (IPTA), International Conference on*. Orleans, Nov. 2015, pp. 379–384.

[7] PF Ceccaldi et al. “[Kartagener’s syndrome and infertility: observation, diagnosis and treatment]”. In: *Journal de gynecologie, obstetrique et biologie de la reproduction* 33.3 (2004), pp. 192–194.

[8] Chris McManus. *Right hand, left hand: The origins of asymmetry in brains, bodies, atoms and cultures*. Harvard University Press, 2004.

[9] Andrew J Thorley and Teresa D Tetley. “Pulmonary epithelium, cigarette smoke, and chronic obstructive pulmonary disease”. In: *International Journal of Chronic Obstructive Pulmonary Disease* 2.4 (Dec. 2007), pp. 409–428. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2699967/.

[10] Peter K Jeffery. “Structural and inflammatory changes in COPD: a comparison with asthma”. In: *Thorax* 53.2 (1998), p. 129.

[11] E Sapey and RA Stockley. “COPD exacerbations · 2: Aetiology”. In: *Thorax* 61.3 (2006), pp. 250–258.

[12] Duncan F Rogers. “Mucociliary dysfunction in COPD: effect of current pharmacotherapeutic options”. In: *Pulmonary pharmacology & therapeutics* 18.1 (2005), pp. 1–8.

[13] A. Barbato et al. “Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children”. In: *European Respiratory Journal* 34.6 (2009), pp. 1264–1276.

[14] W. A. Stannard et al. “Diagnostic testing of patients suspected of primary ciliary dyskinesia”. In: *American journal of respiratory and critical care medicine* 181.4 (2010), pp. 307–314.

[15] J.-F. Papon et al. “Quantitative analysis of ciliary beating in primary ciliary dyskinesia: a pilot study”. In: *Orphanet journal of rare diseases* 7.1 (2012), p. 78.

[16] M. A. Chilvers and C. O’Callaghan. “Analysis of ciliary beat pattern and beat frequency using digital high speed imaging: comparison with the photomultiplier and photodiode methods”. In: *Thorax* 55.4 (2000), pp. 314–317.

[17] R.T. Doyle et al. “Use of confocal linescan to document ciliary beat frequency”. In: *Journal of Microscopy* 223 (2006), pp. 159–164.

[18] T. Dalhamn and R. Rylander. “Frequency of ciliary beat measured with a photo-sensitive cell”. In: *Nature* 196 (1962), pp. 592–593.

[19] C. O’Callaghan, K. Sikand, and M. Chilvers. “Analysis of ependymal ciliary beat pattern and beat frequency using high speed imaging: comparison with the photomultiplier and photodiode methods”. In: *Cilia* 1.1 (2012).

[20] JH Sisson et al. “All-digital image capture and whole-field analysis of ciliary beat frequency”. In: *Journal of microscopy* 211.2 (2003), pp. 103–111.

[21] J Raidt et al. “Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia”. In: *European Respiratory Journal* (2014).

[22] CM Smith et al. “CiliaFA: a research tool for automated, high-throughput measurement of ciliary beat frequency using freely available software”. In: *Cilia* 1.1 (2012), pp. 1–7.

[23] W Kim et al. “An automated measurement of ciliary beating frequency using a combined optical flow and peak detection”. In: *Healthcare informatics research* 17.2 (2011), pp. 111–119.

[24] Olivier Meste, Frédéric Brau, and Alice Guyon. “Robust estimation of the motile cilia beating frequency”. In: vol. 53. 10. Springer, 2015, pp. 1025–1035.

[25] James W Cooley and John W Tukey. “An algorithm for the machine calculation of complex Fourier series”. In: *Mathematics of computation* 19.90 (1965), pp. 297–301.

[26] C. F. Gauss. “Nachlass, Theoria Interpolationis Methodo Nova Tractata”. In: *C.F. Gauss Werke*. Vol. 3. Göttingen: Königlichen Gesselshaft der Wissenschaften, 1866.

[27] M Bottier et al. “Characterization of upper airway ciliary beat by coupling isolated and collective cilia motion analysis”. In: *Cilia* 4.1 (2015), p. 1.

[28] S. P. Quinn et al. “Automated identification of abnormal respiratory ciliary motion in nasal biopsies”. In: *Science Translational Medicine* 7.299 (2015).

[29] J Buzek and B Ask. “Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products”. In: *Official Journal of the European Union L* 342 (2009).

[30] Tzutzuy Ramirez et al. “Toxicological testing of cosmetic ingredients using alternative methods - the perspective of a supplier”. In: *IFSCC Magazine* 16.4 (Dec. 2013), pp. 263–274.

[31] W. Lilienblum et al. “Alternative methods to safety studies in experimental animals: role in the risk assessment of chemicals under the new European Chemicals Legislation (REACH)”. In: *Archives of Toxicology* 82.4 (Apr. 2008), pp. 211–236. URL: http://link.springer.com/10.1007/s00204-008-0279-9.

[32] S. L. Oxendine et al. “Adapting the medaka embryo assay to a high-throughput approach for developmental toxicity testing”. en. In: *NeuroToxicology* 27.5 (Sept. 2006), pp. 840–845. URL: http://linkinghub.elsevier.com/retrieve/pii/S0161813X06000477 (visited on 01/19/2015).

[33] Norman A. Dial. “Methylmercury: Some effects on embryogenesis in the Japanese Medaka, *Oryzias latipes*”. en. In: *Teratology* 17.1 (Feb. 1978), pp. 83–91. URL: http://doi.wiley.com/10.1002/tera.1420170116 (visited on 01/19/2015).

[34] Johannes Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. In: *Nature Methods* 9.7 (June 2012), pp. 676–682. URL: http://www.nature.com/doifinder/10.1038/nmeth.2019 (visited on 01/19/2015).

[35] Sumitaka Hasegawa et al. “A medaka model of cancer allowing direct observation of transplanted tumor cells in vivo at a cellular-level resolution”. In: *Proceedings of the National Academy of Sciences* 106.33 (2009), pp. 13832–13837.

[36] Diana M Papoulias, Douglas B Noltie, and Donald E Tillitt. “An in vivo model fish system to test chemical effects on sexual differentiation and development: exposure to ethinyl estradiol”. In: *Aquatic toxicology* 48.1 (2000), pp. 37–50.

[37] H William Detrich, Monte Westerfield, and Leonard I Zon. *The zebrafish: Disease models and chemical screens*. Vol. 3. Academic Press, 2011.

[38] Lixin Yang et al. “Zebrafish embryos as models for embryotoxic and teratological effects of chemicals”. In: *Reproductive Toxicology* 28.2 (2009), pp. 245–253.

[39] Shosaku Kashiwada et al. “Silver nanocolloids disrupt medaka embryogenesis through vital gene expressions”. In: *Environmental science & technology* 46.11 (2012), pp. 6278–6287.

[40] Kerstin Howe et al. “The zebrafish reference genome sequence and its relationship to the human genome”. In: *Nature* 496.7446 (Apr. 2013), pp. 498–503. URL: http://dx.doi.org/10.1038/nature12111.

[41] Ralf Mikut et al. “Automated processing of zebrafish imaging data: a survey”. In: *Zebrafish* 10.3 (2013), pp. 401–421.

[42] BP Ruff, JA Marchant, and AR Frost. “Fish sizing and monitoring using a stereo image analysis system applied to fish farming”. In: *Aquacultural engineering* 14.2 (1995), pp. 155–173.

[43] Alexander K Morison, Simon G Robertson, and David C Smith. “An integrated system for production fish aging: image analysis and quality assurance”. In: *North American Journal of Fisheries Management* 18.3 (1998), pp. 587–598.

[44] NJC Strachan, Paul Nesvadba, and Alastair R Allen. “Fish species recognition by shape analysis of images”. In: *Pattern Recognition* 23.5 (1990), pp. 539–544.

[45] YH Toh, TM Ng, and BK Liew. “Automated fish counting using image processing”. In: *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*. IEEE. 2009, pp. 1–5.

[46] Gregory M Cahill, Mark W Hurd, and Matthew M Batchelor. “Circadian rhythmicity in the locomotor activity of larval zebrafish”. In: *Neuroreport* 9.15 (1998), pp. 3445–3449.

[47] Satoru Kato et al. “A computer image processing system for quantification of zebrafish behavior”. In: *Journal of neuroscience methods* 134.1 (2004), pp. 1–7.

[48] Wenhui Wang et al. “A fully automated robotic system for microinjection of zebrafish embryos”. In: *PloS one* 2.9 (2007), e862.

[49] Tianming Liu et al. “Computerized image analysis for quantitative neuronal phenotyping in zebrafish”. en. In: *Journal of Neuroscience Methods* 153.2 (June 2006), pp. 190–202. URL: http://linkinghub.elsevier.com/retrieve/pii/S0165027005004127 (visited on 01/19/2015).

[50] Shunren Xia et al. “Computational techniques in zebrafish image processing and analysis”. en. In: *Journal of Neuroscience Methods* 213.1 (Feb. 2013), pp. 6–13. URL: http://linkinghub.elsevier.com/retrieve/pii/S0165027012004657 (visited on 01/19/2015).

[51] S.H. Cheng and P.K. CHAN. *Method of in Vivo Screening for cardiac Toxic Agents Using Teleost*. US Patent App. 12/224,577. Sept. 2009. URL: http://www.google.com/patents/US20090232739.

[52] Masayuki Yoshida, Ruriko Hirano, and Takao Shima. "Photocardiography: a novel method for monitoring cardiac activity in fish". In: *Zoological science* 26.5 (2009), pp. 356–361.

[53] H. Heijmans. *Morphological image operators*. Advances in Electronics and Electron Physics Series. Boston: Academic Press, 1994.

[54] L. Najman and H. Talbot, eds. *Mathematical Morphology: from theory to applications*. ISBN 978-1848212152. London, UK: ISTE-Wiley, Sept. 2010.

[55] L. Vincent. "Grayscale area openings and closings, their efficient implementation and applications". In: *Proceedings of the conference on mathematical morphology and its applications to signal processing*. Barcelona, Spain, 1993, pp. 22–27.

[56] Luc Vincent. "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms". In: *IEEE transactions on image processing* 2.2 (1993), pp. 176–201.

[57] Philippe Salembier, Albert Oliveras, and Luis Garrido. "Antiextensive connected operators for image and sequence processing". In: *IEEE Transactions on Image Processing* 7.4 (1998), pp. 555–570.

[58] Ronald Jones. "Connected filtering and segmentation using component trees". In: *Computer Vision and Image Understanding* 75.3 (1999), pp. 215–228.

[59] Laurent Najman and Michel Couprie. "Building the component tree in quasi-linear time". In: *IEEE Transactions on Image Processing* 15.11 (2006), pp. 3531–3539.

[60] H Digabel and Christian Lantuéjoul. "Iterative algorithms". In: *Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*. Vol. 19. 7. Stuttgart, West Germany: Riederer Verlag. 1978, p. 8.

[61] S. Beucher and C. Lantuéjoul. "Use of watersheds in Contour Detection". In: *Int. Workshop on Image Processing*. CCETT/IRISA. Rennes, France, Sept. 1979.

[62] Fernand Meyer. "The watershed concept and its use in segmentation: a brief history". In: *arXiv preprint arXiv:1202.0216* (2012).

[63] Gilles Bertrand. "On topological watersheds". In: *Journal of Mathematical Imaging and Vision* 22.2-3 (2005), pp. 217–230.

[64] Michel Couprie, Laurent Najman, and Gilles Bertrand. "Quasi-linear algorithms for the topological watershed". In: *Journal of Mathematical Imaging and Vision* 22.2-3 (2005), pp. 231–249.

[65] Hieu Tat Nguyen, Marcel Worring, and Rein Van Den Boomgaard. "Watersnakes: energy-driven watershed segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.3 (2003), pp. 330–342.

[66] Cédric Allène et al. “Some links between min-cuts, optimal spanning forests and watersheds”. In: *Mathematical Morphology and its Applications to Image and Signal Processing* (2007), pp. 253–264.

[67] J. Cousty et al. “Watershed cuts: minimum spanning forests and the drop of water principle”. In: *IEEE transactions on Pattern Analysis and Machine Intelligence* (2008), pp. 1362–1374.

[68] Camille Couprie et al. “Power watershed: A unifying graph-based optimization framework”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7 (2011), pp. 1384–1399.

[69] Leo Grady. “Random Walks for Image Segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.11 (2006), pp. 1768–1783.

[70] Jean Cousty et al. “Watershed cuts: Thinnings, shortest path forests, and topological watersheds”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.5 (2010), pp. 925–939.

[71] Punam K Saha and Jayaram K Udupa. “Relative fuzzy connectedness among multiple objects: theory, algorithms, and applications in image segmentation”. In: *Computer Vision and Image Understanding* 82.1 (2001), pp. 42–56.

[72] Y. Boykov, O. Veksler, and R. Zabih. “Fast Approximate Energy Minimization via Graph Cuts”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.11 (2001), pp. 1222–1239.

[73] Robert Endre Tarjan. “Sensitivity analysis of minimum spanning trees and shortest path trees”. In: *Information Processing Letters* 14.1 (1982), pp. 30–33.

[74] Bernard Chazelle. “A minimum spanning tree algorithm with inverse-Ackermann type complexity”. In: *Journal of the ACM (JACM)* 47.6 (2000), pp. 1028–1047.

[75] Alexandre X Falcão, Jorge Stolfi, and Roberto de Alencar Lotufo. “The image foresting transform: Theory, algorithms, and applications”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.1 (2004), pp. 19–29.

[76] C. Tomasi and R. Manduchi. “Bilateral filtering for gray and color images”. In: *IEEE International Conference on Computer Vision*. 1998, pp. 839–846.

[77] Bruce D Lucas, Takeo Kanade, et al. “An iterative image registration technique with an application to stereo vision.” In: *IJCAI*. Vol. 81. 1. 1981, pp. 674–679.

[78] B. K. Horn and B. G. Schunck. “Determining optical flow”. In: *1981 Technical Symposium East*. International Society for Optics and Photonics. 1981, pp. 319–331.

[79] G. Farnebäck. “Polynomial Expansion for Orientation and Motion Estimation”. PhD thesis. Linköping, 2002.

[80] Gunnar Farnebäck. "Two-frame motion estimation based on polynomial expansion". In: *Scandinavian conference on Image analysis*. Springer. 2003, pp. 363–370.

[81] D. G. Lowe. "Object recognition from local scale-invariant features". In: *IEEE International Conference on Computer Vision*. Vol. 2. Kerkyra, Greece, Sept. 1999, pp. 1150–1157.

[82] Herbert Bay et al. "Speeded-up robust features (SURF)". In: *Computer vision and image understanding* 110.3 (2008), pp. 346–359.

[83] Paul Viola and Michael Jones. "Rapid object detection using a boosted cascade of simple features". In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE. 2001, pp. I–511.

[84] Luo Juan and Oubong Gwun. "A comparison of sift, pca-sift and surf". In: *International Journal of Image Processing (IJIP)* 3.4 (2009), pp. 143–152.

[85] Ron Bracewell. "The fourier transform and iis applications". In: *New York* 5 (1965).

[86] Alan V Oppenheim and Ronald W Schafer. *Discrete-time signal processing*. Pearson Higher Education, 2010.

[87] M. A Fischler and R. C Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography". In: *Communications of the ACM* 24.6 (1981), pp. 381–395.

[88] T. Collins, P. Mesejo, and A. Bartoli. "An Analysis of Errors in Graph-Based Keypoint Matching and Proposed Solutions". In: *European Conference on Computer Vision*. Vol. 8695. LNCS. Zurich, Switzerland: Springer, Sept. 2014, pp. 138–153.

[89] N. Otsu. "A threshold selection method from gray-level histograms". In: *Automatica* 11.285-296 (1975), pp. 23–27.

[90] F. Meyer and S. Beucher. "Morphological segmentation". In: *Journal of Visual Communication and Image Representation* 1.1 (Sept. 1990), pp. 21–46.

[91] Toyofumi Saito and Jun-Ichiro Toriwaki. "New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications". In: *Pattern recognition* 27.11 (1994), pp. 1551–1565.

[92] H. Bay, T. Tuytelaars, and L. Van Gool. "Surf: Speeded up robust features". In: *Computer vision–ECCV 2006*. Springer, 2006, pp. 404–417.

[93] Reza Nasiri Mahalati, Ruo Yu Gu, and Joseph M Kahn. "Resolution limits for imaging through multi-mode fiber". In: *Optics express* 21.2 (2013), pp. 1656–1668.

[94] Ruo Yu Gu, Reza Nasiri Mahalati, and Joseph M Kahn. "Design of flexible multi-mode fiber endoscope". In: *Optics express* 23.21 (2015), pp. 26905–26918.

[95] J. Serra. *Image analysis and mathematical morphology*. Academic Press, 1982.

[96] P. Soille. *Morphological Image Analysis, principles and applications*. Springer, 1999.

[97] R. Jones and P. Soille. "Periodic lines: Definition, cascades, and application to granulometries". In: *Pattern Recognition Letters* 17.10 (1996), pp. 1057–1063.

[98] N. Otsu. "A threshold selection method from gray-level histograms". In: *Automatica* 11.285-296 (1975), pp. 23–27.

[99] Pierre Soille and Hugues Talbot. "Directional morphological filtering". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.11 (2001), pp. 1313–1329.

[100] Fernand Meyer and Serge Beucher. "Morphological segmentation". In: *Journal of Visual Communication and Image Representation* 1.1 (Sept. 1990), pp. 21–46.

[101] Rachid Deriche. "Using Canny's criteria to derive a recursively implemented optimal edge detector". In: *International journal of computer vision* 1.2 (1987), pp. 167–187.

[102] J. Cousty et al. "Watershed cuts: minimum spanning forests and the drop of water principle". In: *IEEE transactions on Pattern Analysis and Machine Intelligence* (2008), pp. 1362–1374.

[103] Ray A Jarvis. "On the identification of the convex hull of a finite set of points in the plane". In: *Information Processing Letters* 2.1 (1973), pp. 18–21.

[104] P. Soille. *Morphological Image Analysis, principles and applications*. Springer, 1999.

[105] Renato Kresch and David Malah. "Skeleton-based morphological coding of binary images". In: *IEEE Transactions on Image Processing* 7.10 (1998), pp. 1387–1399.

[106] H. Talbot et al. "Efficient Poisson denoising for photography". In: *16th IEEE International Conference on Image Processing (ICIP)*. Cairo, Egypt, Nov. 2009, pp. 3881–3884. URL: http://dx.doi.org/10.1109/ICIP.2009.5414042.

[107] A. M. Plotnik and S. M. Rock. "Quantification of cyclic motion of marine animals from computer vision". In: *MTS/IEEE OCEANS'02*. Vol. 3. 2002, pp. 1575–1581.

[108] Dongmin Guo, Anne L. van de Ven, and Xiaobo Zhou. "Red Blood Cell Tracking Using Optical Flow Methods". In: *IEEE Journal of Biomedical and Health Informatics* 18.3 (May 2014), pp. 991–998. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6600731 (visited on 01/30/2015).