



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

**SCHOOL OF SCIENCES
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS**

**POSTGRADUATE PROGRAM
"INFORMATION TECHNOLOGIES IN MEDICINE AND BIOLOGY"**

MASTER THESIS

Bacterial image analysis based on time-lapse microscopy

Athanasios D. Balomenos

SUPERVISOR: Elias Manolakos, Associate Professor
COSUPERVISOR: Panagiotis Tsakanikas, Postdoctoral Researcher

ATHENS

AUGUST 2013



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Επεξεργασία βακτηριακής εικόνας από time-lapse
μικροσκοπία**

Αθανάσιος Δ. Μπαλωμένος

ΕΠΙΒΛΕΠΩΝ: Ηλίας Μανωλάκος, Αναπληρωτής Καθηγητής
ΣΥΝΕΠΙΒΛΕΠΩΝ: Παναγιώτης Τσακανίκας, Μεταδιδακτορικός Ερευνητής

ΑΘΗΝΑ

ΑΥΓΟΥΣΤΟΣ 2013

MASTER THESIS

Bacterial image analysis based on time-lapse microscopy

Athanasios D. Balomenos

R.N.: PIV084

SUPERVISOR: **Elias Manolakos**, Associate Professor
COSUPERVISOR: **Panagiotis Tsakanikas**, Postdoctoral Researcher

EXAMINING COMMITTEE: **Katia Karalis**, Investigator - Associate Professor Level,
Biomedical Research Foundation of the Academy of Athens
Elias Manolakos, Associate Professor
Alexandros Eleftheriadis, Associate Professor

August 2013

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Επεξεργασία βακτηριακής εικόνας από time-lapse μικροσκοπία

Αθανάσιος Δ. Μπαλωμένος

A.M.:ΠΙΒ084

ΕΠΙΒΛΕΠΩΝ: **Ηλίας Μανωλάκος**, Αναπληρωτής Καθηγητής
ΣΥΝΕΠΙΒΛΕΠΩΝ: **Παναγιώτης Τσακανίκας**, Μεταδιδακτορικός Ερευνητής

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: **Κάτια Καραλή**, Ερευνήτρια Β', Ίδρυμα Ιατροβιολογικών
Ερευνών της Ακαδημίας Αθηνών (ΙΙΒΕΑΑ)
Ηλίας Μανωλάκος, Αναπληρωτής Καθηγητής,
Πανεπιστήμιο Αθηνών
Αλέξανδρος Ελευθεριάδης, Αναπληρωτής Καθηγητής,
Πανεπιστήμιο Αθηνών

Αύγουστος 2013

ABSTRACT

Time-lapse microscopy now enables detailed imaging data generation and monitoring of dynamic cellular processes at the single cell level. Recent studies have highlighted the usage and importance of this technology for investigating biological noise in gene regulation, cell growth and proliferation etc. Mathematical and statistical model development is of growing interest in capturing and testing hypothesis regarding the dynamical behavior of biological systems. Modeling bacterial communities forming biofilms relies on the efficient and accurate extraction of information from time-lapse microscopy data (image frame sequences) of growing bacterial colonies. However, the analysis of such "cell movies" data is currently very time consuming and error prone since it is essentially performed by human-experts. In this thesis we address this important limitation in a multi-resolution image analysis framework.

We have developed a methodology for identifying accurately the boundaries of individual bacterial cells and tracking them from frame to frame so as to construct the cells' genealogy (bacterial cell segmentation and lineage tree construction) even in large-size microbial communities where there is great difficulty in identifying the individual cell boundaries. The automated and novel pipeline of algorithms we have developed combines methods from image processing and machine learning to segment and track bacteria precisely.

The pipeline has been tested and evaluated with two different cell movies datasets and several images produced by different labs. The developed methodology has been shown to achieve high F-measure score (above 95%) in each evaluation case. It can be applied to different image modalities, such as phase contrast, bright field, and fluorescent, produced by optical and confocal microscopy. Using extensive experimentation we demonstrate the robustness and reliability of the proposed pipeline regardless of the image modality used. Our image processing pipeline is fully automated, computationally efficient and suitable for high throughput analysis of bacterial cell movies without any human intervention on its calibration.

SUBJECT AREA: Image Analysis and Machine Learning

KEYWORDS: bacterial segmentation, cell counting, cell lineage construction, cell feature extraction and visualization, expectation-maximization.

ΠΕΡΙΛΗΨΗ

Η time-lapse μικροσκοπία επιτρέπει πλέον τη λεπτομερή δημιουργία δεδομένων από δυναμικές κυτταρικές διεργασίες σε επίπεδο μεμονωμένων κυττάρων (single cell level). Πρόσφατες μελέτες έχουν τονίσει τη χρήση και τη σημασία αυτής της τεχνολογίας για τη διερεύνηση του βιολογικού θορύβου στη ρύθμιση των γονιδίων, την ανάπτυξη και τον πολλαπλασιασμό κυττάρων κλπ. Τα μαθηματικά και τα στατιστικά μοντέλα παρουσιάζουν αυξανόμενο ενδιαφέρον για τη σύλληψη και τον έλεγχο υποθέσεων σχετικά με τη δυναμική συμπεριφορά των βιολογικών συστημάτων. Για τις βακτηριακές κοινότητες που σχηματίζουν βιοϋμένια αυτό εξαρτάται από την αποτελεσματική και ακριβή άντληση πληροφοριών από δεδομένα time-lapse μικροσκοπίας (ακολουθίες εικόνων) αναπτυσσόμενων βακτηριακών αποικιών. Ωστόσο, η ανάλυση τέτοιων δεδομένων «κυτταρικών ταινιών» σήμερα είναι πολύ χρονοβόρα και επιρρεπής σε λάθη, δεδομένου ότι ουσιαστικά πραγματοποιείται από έναν άνθρωπο-εμπειρογνώμονα. Στην παρούσα εργασία ασχολούμαστε με αυτό το σημαντικό περιορισμό σε ένα πολλαπλής ανάλυσης πλαίσιο επεξεργασίας εικόνας.

Έχουμε αναπτύξει μια μεθοδολογία που προσδιορίζει με ακρίβεια τα όρια μεμονωμένων βακτηριακών κυττάρων, τα ανιχνεύει από εικόνα σε εικόνα και κατασκευάζει τη γενεαλογία τους (κατάτμηση βακτηριακών κυττάρων και κατασκευή γενεαλογικών δέντρων) ακόμη και σε μεγάλες μικροβιακές κοινότητες όπου υπάρχει μεγάλη δυσκολία στον προσδιορισμό των επιμέρους ορίων των κυττάρων. Η αυτοματοποιημένη και καινοτόμος σωλήνωση (pipeline) των αλγορίθμων που αναπτύξαμε συνδυάζει μεθόδους από την επεξεργασία εικόνας και τη μηχανική μάθηση για την ακριβή βακτηριακή κατάτμηση και ανίχνευση.

Η προτεινόμενη μεθοδολογία έχει δοκιμαστεί και αξιολογηθεί με δύο διαφορετικά σύνολα δεδομένων κυτταρικών ταινιών και αρκετές ανεξάρτητες εικόνες από διαφορετικά εργαστήρια. Έχουμε αποδείξει ότι η μέθοδος που αναπτύχθηκε επιτυγχάνει υψηλή βαθμολογία F-measure (άνω του 95%) σε κάθε περίπτωση αξιολόγησης. Επίσης το λογισμικό μπορεί να εφαρμοστεί σε διαφορετικούς τύπους εικόνας, όπως phase-contrast, bright field, και φθορίζουσας (fluorescent), που προέρχονται τόσο από οπτική όσο και από συνεστιακή μικροσκοπία. Πραγματοποιώντας εκτεταμένα πειράματα που αποδεικνύουν την ευρωστία και την αξιοπιστία της προτεινόμενης σωλήνωσης, ανεξάρτητα από το τύπο της εικόνας που χρησιμοποιείται. Η μεθοδολογία επεξεργασίας εικόνας, που αναπτύξαμε, είναι πλήρως αυτοματοποιημένη, υπολογιστικά αποδοτική και κατάλληλη για υψηλής ρυθμαπόδοσης ανάλυση βακτηριακών κυτταρικών ταινιών χωρίς καμία ανθρώπινη παρέμβαση κατά τη βαθμονόμηση της.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Ανάλυση Εικόνας και Μηχανική Μάθηση

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Βακτηριακή κατάτμηση, απαρίθμηση κυττάρων, δημιουργία κυτταρικής γενεαλογίας, εξαγωγή και οπτικοποίηση βακτηριακών χαρακτηριστικών, expectation-maximization.