



Universiteit
Leiden
The Netherlands

Pattern Recognition in High-Throughput Zebrafish Imaging

Nezhinsky, A.E.

Citation

Nezhinsky, A. E. (2013, November 21). *Pattern Recognition in High-Throughput Zebrafish Imaging*. Retrieved from <https://hdl.handle.net/1887/22286>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/22286>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/22286> holds various files of this Leiden University dissertation

Author: Nezhinsky, A.E.

Title: Pattern recognition in high-throughput zebrafish imaging

Issue Date: 2013-11-21

Russian Summary

В биологических исследованиях часто используют методику High Throughput, позволяющую с помощью большой серии экспериментов находить определенные характеристики биологических объектов. Конкретная реализация этой методики зависит от конкретных требований. В нашей работе мы применили методику High Throughput для измерения связанных с изображениями интересующих нас объектов.

Пример использования методики High Throughput приведён на рисунке 1.1.

Конкретной целью исследования в данной работе является создание ряда надежных методов Pattern Recognition в контексте методики High Throughput. Pattern Recognition можно рассматривать как распознавание объектов в изображении но также как Data Mining. Решения обеих задач детально описаны.

В нашем исследовании методом High Throughput обрабатывались изображения, полученные с помощью микроскопа. Изображения необходимо было и выделять (отделить от фона и случайных помех) и анализировать. Для человека разделение объектов переднего плана и фона (сегментация) несложно. Для компьютера же, - это нетривиальная задача. Методы сегментации могут быть разделены на две основные группы: снизу вверх (Bottom Up) методика основана на обнаружении эквивалентных свойств объектов и фона, что позволяет их разделить, но такое разделение не всегда достаточно, поэтому используется ещё и подход сверху вниз (Top Down), основанный на распознавании образов.

В главе 2 мы представляем два подхода к методике High Throughput: то есть, сегментацию и выделение объектов участвующих в эксперименте. В качестве реальных объектов были использованы рыбки данио (*Danio rerio*), т.е. осуществлялся поиск изображений эмбрионов рыбок. Мы выделяем два этапа: предварительная обработка, при которой фон отделяется от переднего плана (снизу вверх) и, второе, Template-Matching шаг, который заключается в нахождении деформируемого шаблона (сверху вниз). Для выбора Template-Matching шага были предложены два метода, опирающихся на выбор в качестве шаблона усреднённого силуэта объекта. При этом предполагалось, что изображение формируют один или несколько деформированных или смещенных объектов.

Поиск подходящих деформаций представляет из себя сложный вычислительный процесс, в котором используются различные виды оптимизаций. Также мы рассматриваем способы нормализации полученных результатов для возможности лучшего сравнения большего количества данных.

В главе 3 предложена дальнейшая разработка метода построения деформируемого шаблона. Поиск объект производится с помощью так называемых опорных точек вместо силуэта. Это способ более надежен и достоверен, т.к. исчезает необходимость

фиксировать размеры шаблона. Предложенный способ полностью автоматизирован, работает быстрее и точнее, чем методы предложенные ранее. Мы протестировали алгоритм, используя изображения эмбрионов рыбок данио и эмбрионов лягушки.

В главе 4 описан программный пакет ZFA, который может в контексте High Throughput анализировать изображения. Для разных пользователей представлен интерфейс соответствующих рабочих процессов.

Программный пакет включает: предварительную обработку, интерфейс, сегментацию и обработку данных. Предварительная обработка предусматривает выделение флуоресцентных и обычных изображений областей.

Интерфейс предоставляет возможность выбрать примеры, помещает изображения в выбранные папки. Форма эмбрионов данио локализована и доведена до алгоритма вычисления функции деформируемого шаблона. После определения формы предложен анализ характеристик различных инфекций. Требования к ZFA мы выбирали исходя из практических нужд. Известно, что туберкулез является одним из самых серьезных и распространённых заболеваний. Эффективное лечение его по-прежнему затруднено из-за его бактериальной устойчивости.

Поведение бактерии туберкулеза можно моделировать изучая поведение его близкого родственника - *Mycobacterium marinum*. При помощи использования данио как модель становится возможным отслеживать процесс инфекции. Для этого нужно локализовать рыбки данио и анализировать распределение инфекции в них. Наш программный пакет делает это возможным.

В главе 5 применяется ZFA в качестве программного обеспечения мониторинга исследования данио инфицированного бактериями *Mycobacterium marinum*. Такие инфекции характеризуется появлением гранул (скоплений иммунных клеток и бактерий). Гранулы легко визуализировать флуоресцентными агентами. Наш подход (использование алгоритма из 3 главы) позволяет автоматизировать анализ развития во времени инфекции в теле хозяина, данио. Для того, чтобы глубже разобраться в процессе развития инфекции необходимо выявить гены ответственные за развитие инфекции. Для этого были выбраны 1000 случайных мутантов этих бактерий, а из них выбраны те мутанты при которых рыбки не заболели. Таким образом 30 мутантов были выделены для углубленного анализа. Мы изучали, кроме того, отличаются ли гранулы по размерам и распределениям в определенных тканях инфицированные генами *Mycobacterium marinum* от инфицированных туберкулезом.

В главе 6 мы перечисляем и обсуждаем результаты исследований, проведённых в предыдущих главах.

Это и вопросы, связанные с сегментационным шагом, распознавания объектов в изображениях, особенности предложенных трех различных методов, мы сравниваем их производительность. Тест при сопоставлении разных методов мы проводим на одном и том же наборе данных. Исследования убедительно показали, что третий метод дает наилучшие результаты и поэтому используется в пакете ZFA.