

Сравнительный анализ алгоритмов сегментации для выделения микрокальцинатов на маммограммах

Ю.А. Подгорнова¹, С.С. Садыков¹

¹Муромский институт (филиал) Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых, Орловская 23, Муром, Россия, 602264

Аннотация. Самым частым заболеванием текущего века у женского населения всего мира является рак молочной железы. Основной задачей исследований большинства ученых является обнаружение этой патологии на ранней стадии (размер опухоли менее 7 мм), когда женщине еще можно помочь. Индикатором данного заболевания является наличие мелкоточечных обызвествлений, расположенных группами внутри или в непосредственной окружности опухоли. При раке обызвествления носят мелкоточечный характер, напоминая песчинки неправильной формы, размеры которых составляют от 100 до 600 мкм. С увеличением числа микрокальцинатов на единицу площади вероятность рака возрастает. Так, при более 15 микрокальцинатах на 1 кв. см вероятность рака достигает 80%. Обнаружение микрокальцинатов даже при отсутствии опухолевого узла весьма подозрительно на рак, поскольку кальцинаты нередко являются единственным признаком, позволяющим заподозрить рак молочной железы. Проведенные исследования позволили выбрать оптимальные алгоритмы сегментации для выделения областей микрокальцинатов для дальнейшего анализа их скоплений, размеров и так далее.

1. Введение

Молочная железа – сложный, чувствительный орган, требующий постоянного контроля в связи с ежегодным увеличением заболеваемости раком молочной железы и его «омоложением» [1]. Часто симптомом серьезных заболеваний являются микрокальцинаты (отложения солей кальция) [1]. Самостоятельно выявить патологию невозможно, так как обычно единичные микрокальцинаты или их скопления имеют маленькие размеры. Для обнаружения таких новообразований требуется проводить аппаратную диагностику: УЗИ и маммографию.

Маммография – это неинвазивный метод обнаружения патологий молочных желез [2].

Микрокальцинаты различаются по локализации, размеру, форме, скопленности, количеству. Примеры микрокальцинатов представлены на рисунке 1. Для оценки всех указанных параметров необходимо их найти и выделить на маммографическом снимке. Таким образом, сегментация [3] играет важную роль в обработке медицинских снимков. Основная идея процесса сегментации состоит в следующем: каждый пиксел изображения может быть связан с некоторыми визуальными свойствами, такими как яркость, цвет и текстура. В пределах одного объекта или одной части объекта эти атрибуты изменяются относительно мало, тогда как при переходе через границу от одного объекта к другому обычно происходит существенное изменение одного или другого из этих атрибутов.

На сегодняшний день разработано достаточное количество алгоритмов сегментации [4], поэтому основной задачей данной публикации является анализ существующих методов сегментации и выбор оптимального алгоритма для выделения микрокальцинатов на маммографических снимках.

На рисунке 1 представлены примеры маммограмм с различными микрокальцинатами, на рисунке 1(в) кроме микрокальцинатов присутствует еще злокачественное новообразование, имеющее нечеткие спиклообразные контуры.

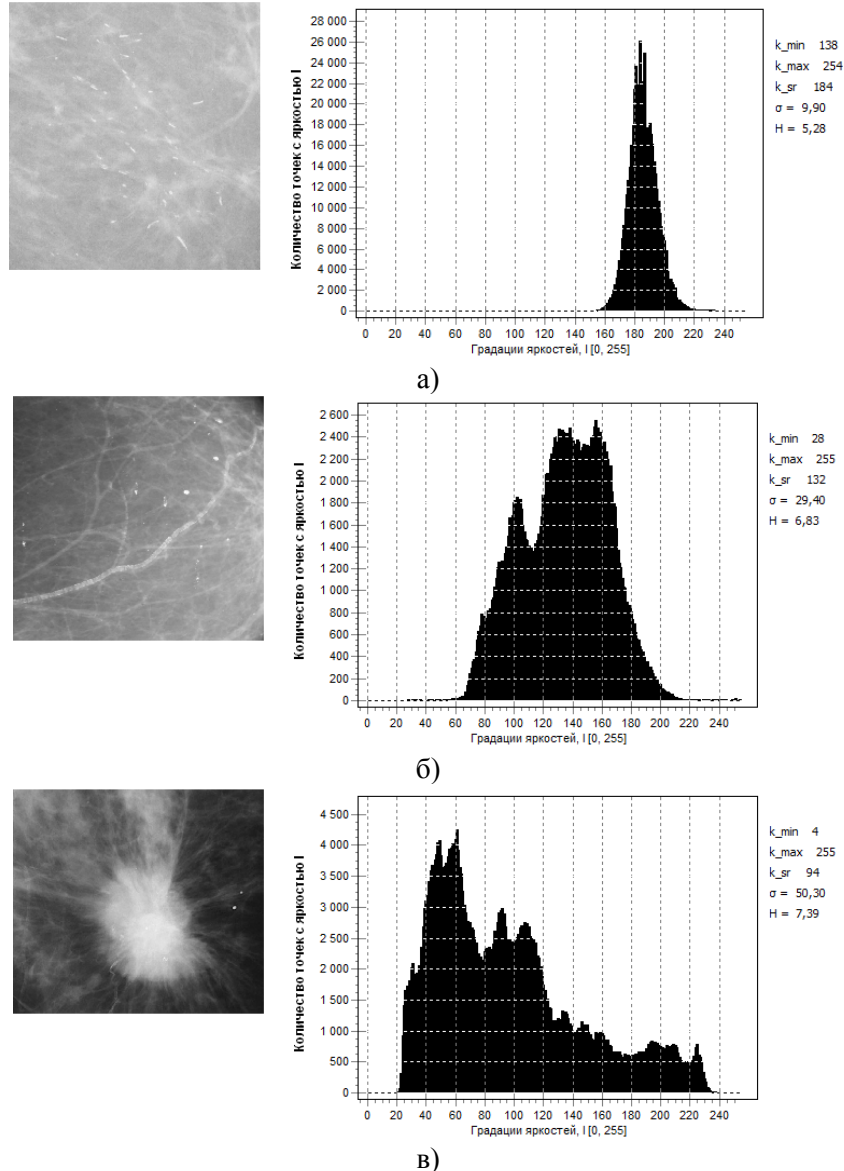


Рисунок 1. Примеры форм, локализации и количества микрокальцинатов на маммограммах.

2. Обзор алгоритмов сегментации

2.1. Алгоритм водоразделов

Понятие водораздела [5] основано на представлении изображения как трехмерной поверхности, заданной двумя пространственными координатами и уровнем яркости в качестве высоты поверхности (рельефа). В такой «топографической» интерпретации рассматриваются точки трех видов: (а) точки локального минимума; (б) точки, находящиеся на склоне, т.е. с которых вода скатывается в один и тот же локальный минимум; и (в) точки, находящиеся на гребне или пике, т.е. с которых вода с равной вероятностью скатывается более чем в один такой минимум.

Применительно к конкретному локальному минимуму, набор точек, удовлетворяющих условию (б), называется бассейном (или водосбором) этого минимума. Множества точек, удовлетворяющих условию (в), образуют линии гребней на поверхности рельефа и называются линиями водораздела.

Одним из важнейших применений сегментации по водоразделам является выделение на фоне изображения однородных по яркости объектов (в виде пятен). Области, характеризующиеся малыми вариациями яркости, имеют малые значения градиента. Поэтому на практике часто встречается ситуация, когда метод сегментации по водоразделам применяется не к самому изображению, а к градиенту этого изображения. В такой постановке локальные минимумы бассейнов хорошо согласуются с малыми значениями градиента, что обычно соответствует интересующим объектам.

2.2. Алгоритм сегментации MeanShift

Основная идея данного метода [6] заключается в том, что по входному изображению можно построить ядерную оценку для плотности вероятности распределения данных в пространстве признаков $RGBXY$. Далее делается естественное предположение о том, что локальные максимумы плотности вероятности соответствуют центрам кластеров. Из необходимого условия локального экстремума определяется выражение для вектора сдвига $m(p)$ точки пространства признаков $p \in RGBXY$, применяя который к точке p итеративно получаем последовательность точек, сходящуюся к локальному максимуму оценки плотности вероятности (т.е. к центру ближайшего кластера):

$$m(p) = \frac{\sum_{i=1}^n p_i \cdot g_i}{\sum_{i=1}^n g_i} - p,$$

где $g_i = g\left(\left\|\frac{(p - p_i)}{h}\right\|^2\right)$, $g(v) = -k'(v)$, h – параметр сглаживания, $K(v) = ck(v)$ – ядро оценки плотности вероятности.

2.3. Алгоритм сегментации FloodFill

С помощью метода FloodFill [7] можно выделить однородные по цвету области. Для этого нужно выбрать начальный пиксель и задать интервал изменения цвета соседних пикселей относительно исходного. Интервал может быть и несимметричным. Алгоритм будет объединять пиксели в один сегмент (заливая их одним цветом), если они попадают в указанный диапазон. На выходе будет сегмент, залитый определенным цветом, и его площадь в пикселях. Такой алгоритм может быть полезен для заливки области со слабыми перепадами цвета однородным фоном. Одним из вариантов использования FloodFill может быть выявление поврежденных краев объекта. Например, если, заливая однородные области определенным цветом, алгоритм заполнит и соседние регионы, то значит нарушена целостность границы между этими областями.

2.4. Алгоритм сегментации K-средних

K-средних (k-means) [4] — наиболее популярный метод кластеризации. Действие алгоритма направлено на минимизацию суммарного квадратичного отклонения точек кластеров от центров этих кластеров. Таким образом, это итеративный алгоритм, который делит данное множество пикселей на k кластеров точек, которых являются максимально приближенными к их центрам, а сама кластеризация происходит за счет смещения этих же центров. Следует учитывать тот факт, что метод k-средних очень чувствителен к шуму, который может существенно исказить результаты кластеризации.

3. Результаты экспериментов

В качестве критериев оценки работы алгоритмов сегментации можно использовать качество подавления фона и выделения объектов в виде связных областей.

Поскольку микрокальцинаты представляют собой сложный объект, то нельзя требовать точного выделения объекта, состоящего из нескольких частей разной яркости как одной связной области.

Для анализа методов были взяты реальные изображения микрокальцинатов на маммограммах. Изображения отличаются количеством и типом микрокальцинатов, яркостью фона и объектов, наличием повторяющихся текстур.

В таблице 1 представлены результаты работы алгоритмов на исходных изображениях. Как можно видеть только алгоритм FloodFill справился с выделением микрокальцинатов, все остальные алгоритмы выделили слишком много связных областей. Такие результаты эксперимента говорят о необходимости использования методов предварительной обработки перед применением сегментации.

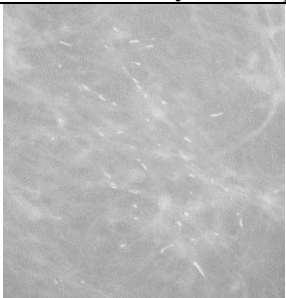

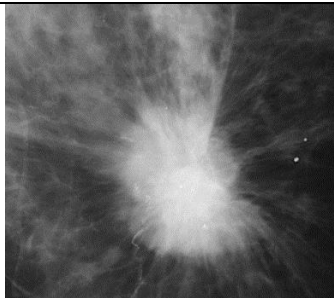


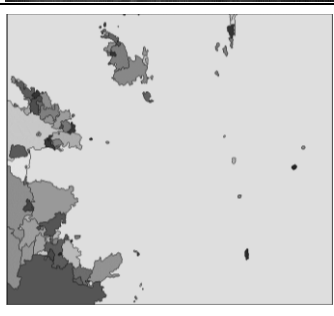

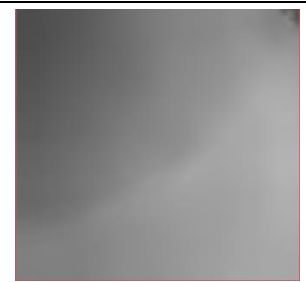
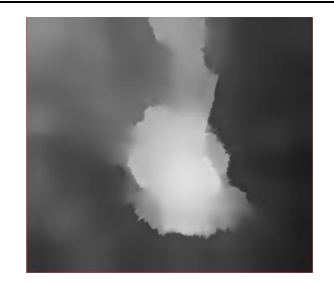
Далее к снимкам были применены методы контрастирования, подробно описанные в [2].

В таблице 2 приведены примеры исследований работы алгоритмов сегментации на контрастированных снимках.

Результаты экспериментов таковы:

- 1 алгоритм водоразделов совершенно не подходит для решения поставленной задачи;
- 2 алгоритм MeanShift способен выделять требуемые объекты только на снимках без новообразований на фоне жировой инволюции;
- 3 алгоритм К-средних показал результаты аналогичные предыдущему методу.

Таблица 1. Результаты работы алгоритмов сегментации на исходных снимках.

Входные изображения			
Алгоритм водоразделов			
Алгоритм MeanShift			

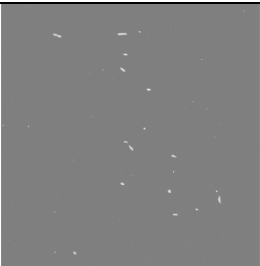

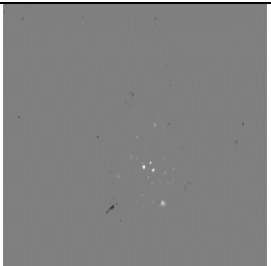
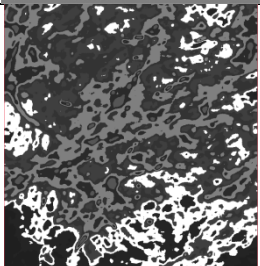
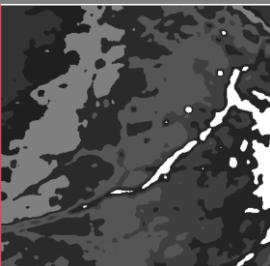
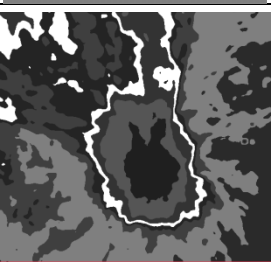
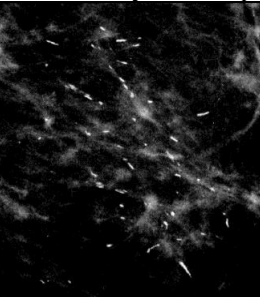


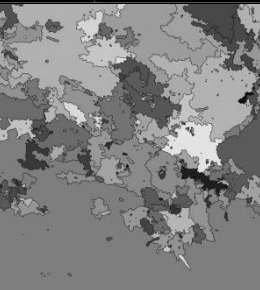

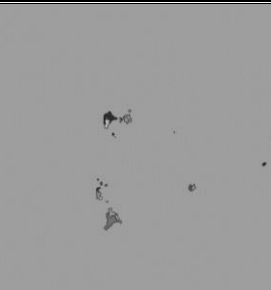




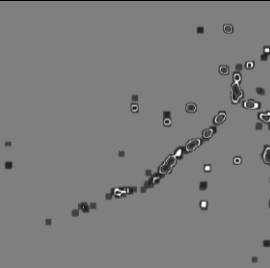
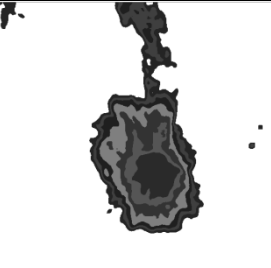
Алгоритм FloodFill			
Алгоритм K-средних			

Таблица 2. Результаты работы алгоритмов сегментации на обработанных снимках.

Входные изображения			
Алгоритм водоразделов			
Алгоритм MeanShift			
Алгоритм K-средних			

4. Заключение

Проведен сравнительный анализ различных методов сегментации изображений применительно к задаче выделения микрокальцинатов на маммографических снимках. Для сравнения методов сегментации были использованы критерии, основанные на оценке экспертом (на основе визуального анализа) качества подавления фона и выделения объектов в виде связных областей.

Путем экспериментальных исследований обнаружено, что метод водоразделов некорректно сегментирует границы объектов и не является приемлемым при решении поставленной задачи. Наилучший результат сегментации был получен с помощью алгоритма FloodFill, заключающегося в выделении однородных по цвету областей. В ходе экспериментов установлено, что для повышения качества сегментации маммограмм целесообразно проводить предварительную обработку снимков, которая обеспечивает уменьшение количества анализируемых областей за счет объединения сегментов и удаления несущественных фрагментов с точки зрения рассматриваемой задачи. Использование тех же самых алгоритмов сегментации после обработки снимков показало, что алгоритмы MeanShift и K-средних способны выделять микрокальцинаты только на снимках без новообразований на фоне жировой инволюции.

Необходимо отметить, что требуются дальнейшие исследования, направленные на усовершенствование методов тематической сегментации, учитывающих пространственные свойства областей и обеспечивающих наилучший компромисс между недостаточной и чрезмерной сегментацией.

Полученные результаты позволяют наметить перспективы использования алгоритмов сегментации при построении автоматических систем обнаружения онкологических заболеваний на ранней стадии.

5. Литература

- [1] Корженкова, Г.П. Комплексная рентгено-сонографическая диагностика заболеваний молочной железы. – М.: ООО «Фирма СТРОМ», 2004. – 128 с.
- [2] Садыков, С.С. Компьютерная диагностика новообразований на маммографических снимках / С.С. Садыков, Ю.А. Буланова, Е.А. Захарова // Компьютерная оптика. – 2014. – Т. 38, № 1. – С. 131-138.
- [3] Eddaoudi, F. Microcalcifications detection in mammographic images using texture coding / F. Eddaoudi, F. Regragui // Applied Mathematical Sciences. – 2011. – Vol. 5(8). – P. 381-393.
- [4] Панченко, Д.С. Сравнительный анализ методов сегментации изображений / Д.С. Панченко, Е.П. Путьтин // Радиоэлектроника и информатика. – 1999. – Т. 4. – С. 109-114.
- [5] Садыков С.С., Буланова Ю.А., Захарова Е.А., Яшков В.С. Исследование маркерного водораздела для выделения области рака молочной железы // Алгоритмы, методы и системы обработки данных. – 2013. - №1. – С. 56-64.
- [6] Степанов, В.Н. Оптимизация алгоритмов сегментации с применением методов параллельных вычислений для систем анализа изображений // Цифровая обработка сигналов. – 2016. – № 3. – С. 75-78.
- [7] Comaniciu, D. Real-Time Tracking of Non-Rigid Objects Using Mean Shift / D. Comaniciu, V. Ramesh, P. Meer // Conference on CVPR. – 2000. – Vol. 2. – P. 1-8.

Благодарности

Подгорнова Ю.А. выражает искреннюю благодарность научному руководителю д-ру техн. наук, профессору Садыкову С.С. за оказанную помощь в проведении исследований, за ценные советы при их планировании и рекомендации по оформлению статьи, а также за моральную поддержку.

Detection of malignant breast tumors on the background of fibrocystic breast disease

Y.A. Podgornova¹, S.S. Sadykov¹

¹Murom Institute (branch) Vladimir State University named after Alexander Grigoryevich and Nickolay Grigoryevich Stoletovs, Orlovskaya street 23, Murom, Russia, 602264

Abstract. The most common disease of the current century in the female population of the world is breast cancer. The main objective of the research of most scientists is to detect this pathology at an early stage (tumor size less than 7 mm), when a woman can still be helped. An indicator of this disease is the presence of small-point calcifications located in groups within or in the immediate circumference of the tumor. In cancer, calcifications are of a small-point nature, resembling grains of irregular shape, the sizes of which are from 100 to 600 microns. With an increase in the number of microcalcifications per unit area, the likelihood of cancer increases. Thus, with more than 15 microcalcifications per 1 square cm, the probability of cancer reaches 80%. The detection of microcalcifications even in the absence of a tumor node is very suspicious of cancer, because calcinates are often the only sign that allows you to suspect breast cancer. Carried out researches have allowed to choose the optimal segmentation method to extract regions of microcalcifications for further analysis of their concentrations, sizes and so on.