

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО
Кафедра «Прикладная математика»

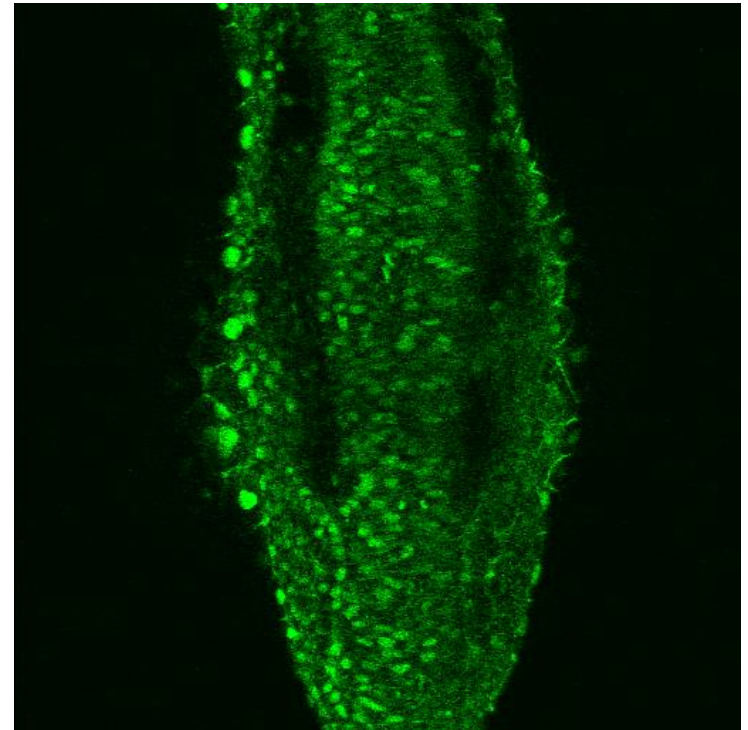
Алгоритм распознавания клеточных структур на снимках электронного микроскопа

Диссертация на соискание степени магистра

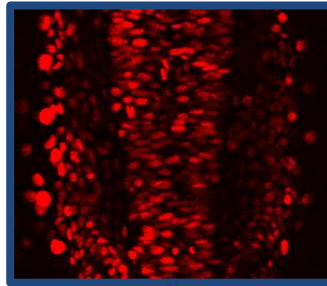
Выполнила студентка гр. 63601/2 М.К. Вердина
Научный руководитель, к.т.н., проф. И.В. Штурц

Флуоресцентная микроскопия

- *Флуоресцентная микроскопия* — метод получения увеличенного изображения с использованием люминесценции возбужденных атомов и молекул образца
- Использование флуоресцентных меток на базе зеленого флуоресцентного белка и его аналогов позволяет получать снимки контрастно окрашенных внутриклеточных белковых структур
- Использование флуоресцентных зондов дает возможность детально исследовать химический состав организмов, а также его изменение во времени и пространстве.



Обработка биомедицинских изображений



Выделение
клеток/ядер

Построение
трехмерной
модели

Определение
моментов
деления

Отслеживание
траекторий

Результат работы критически зависит от исходных данных

Исходные данные

- Снимки развития организма рыбы данио-рерио, полученные с использованием флуоресцентной микроскопии
- Снимки получены с помощью микроскопа с малой глубиной резкости
- Полутоновые изображения разрешения 512x512 пикселей с плотностью 0.284 пикселя на мкм
- Наличие нечетких ядер с выраженной текстурой (рис. 2)
- Аддитивный шум (рис. 3)



Рис. 1

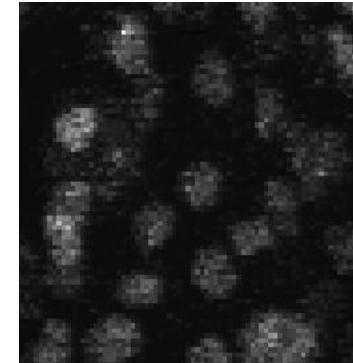


Рис. 2

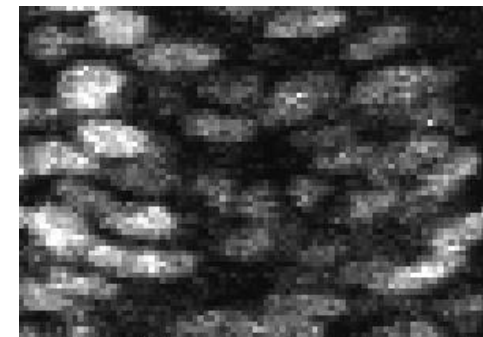


Рис. 3

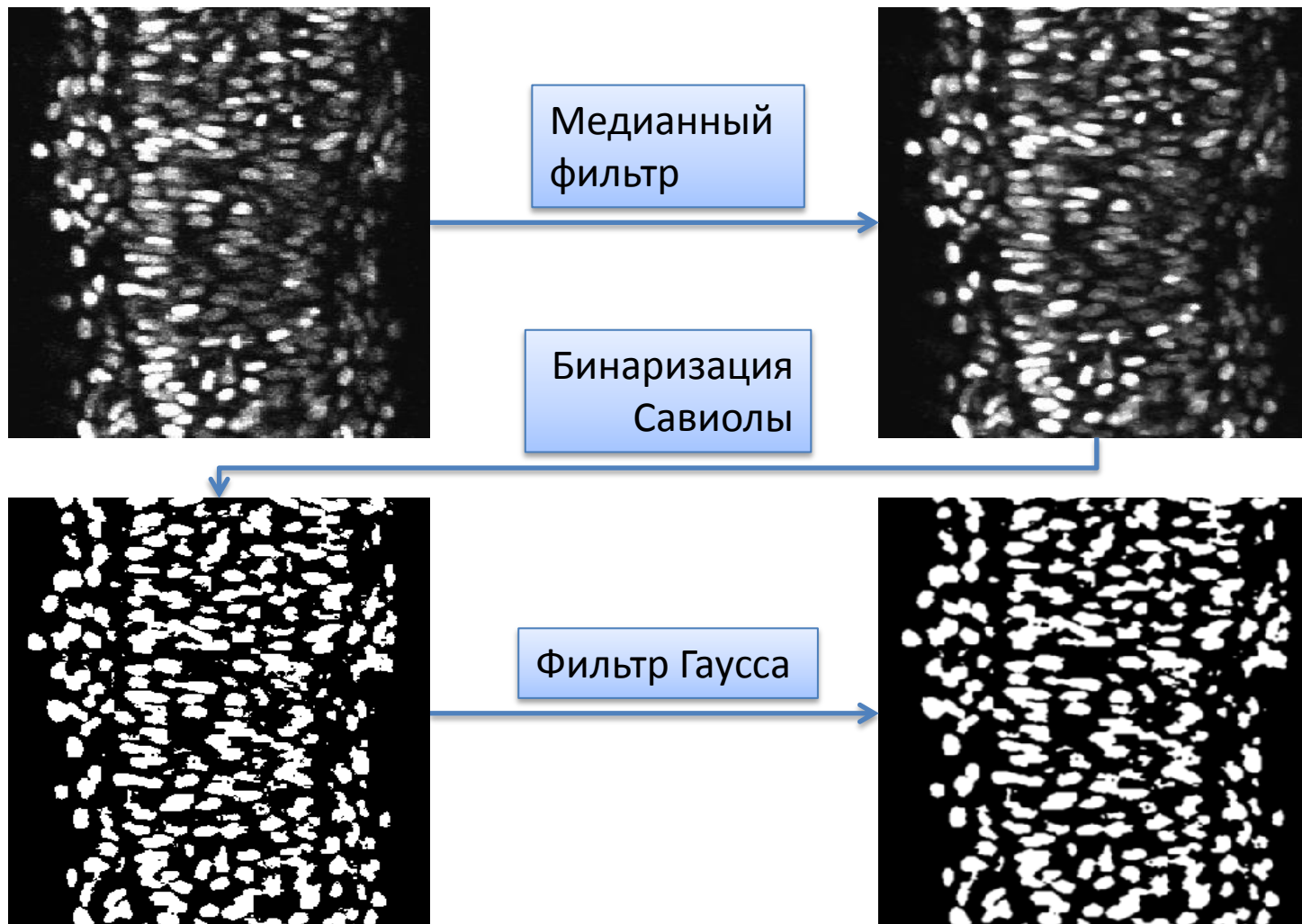
Классификация методов выделения клеточных ядер

Методы водораздела	Методы, основанные на модели клеточного ядра
Изображение – карта местности. Метод моделирования погружения. Существует большое количество модификаций.	Используют априорные знания о свойствах ядер: выпуклость границ, радиальная симметрия, геометрическая форма клеток и др.
Хорошо изучены теоретически, имеют большое количество реализаций.	Как правило, устойчивы к качеству изображения.
Склонны к чрезмерной сегментации, неустойчивы к изображениям низкого качества.	Показывают результат только на определенных классах изображений.

Задачи

- Разработать и реализовать алгоритм выделения клеточных ядер на снимках микроскопа, удовлетворяющий следующим требованиям:
 - быть устойчивым к изображениям низкого качества;
 - распознавать как отдельно стоящие ядра, так и ядра в областях их скоплений.
- На основании достоверных данных оценить качество распознавания

Предобработка



Выделение ядер с использованием графа отношений их границ

- Метод выделения клеточных ядер с использованием графа отношений их границ основан на *априорных знаниях* о выпуклости границ ядер
- **Идея:** смоделировать процесс распознавания ядер человеком. Клеточное ядро – область изображения в форме эллипса, ограниченная слева, сверху, справа и снизу.
- Этапы:
 1. Выделение границ ядер
 2. Построение графа отношений границ ядер
 3. Поиск ядер на графе

Выделение границ

- Свертка операторами Собеля:

$$s_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad s_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad G = \sqrt{G_x^2 + G_y^2}$$

- Бинаризация Оцу: пороговое значения бинаризации максимизирует дисперсию между двумя классами пикселей: «фоновыми» и «полезными»
- Алгоритм Оцу:

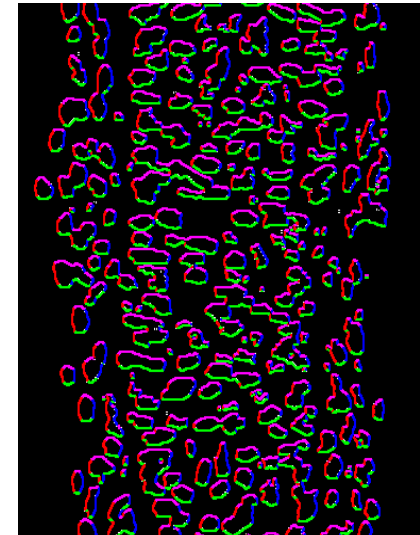
1. Вычислить гистограмму h уровней интенсивности пикселей
2. Для каждого уровня интенсивности t вычислим дисперсию $\sigma^2(t)$ между классами: $\sigma^2(t) = \omega_1(t) * \omega_2(t) * (\mu_1(t) - \mu_2(t))^2$, где

$$\omega_1(t) = \sum_{i=0}^t \frac{h(i)}{M^2}, \quad \omega_2(t) = \sum_{i=t+1}^N \frac{h(i)}{M^2}, \quad \mu_1(t) = \frac{\sum_{i=0}^t \frac{h(i)}{M^2} * i}{\omega_1(t)}, \quad \mu_2(t) = \frac{\sum_{i=t+1}^N \frac{h(i)}{M^2} * i}{\omega_2(t)}$$

3. Пороговое значение бинаризации $T = \operatorname{argmax}_t \sigma^2(t)$

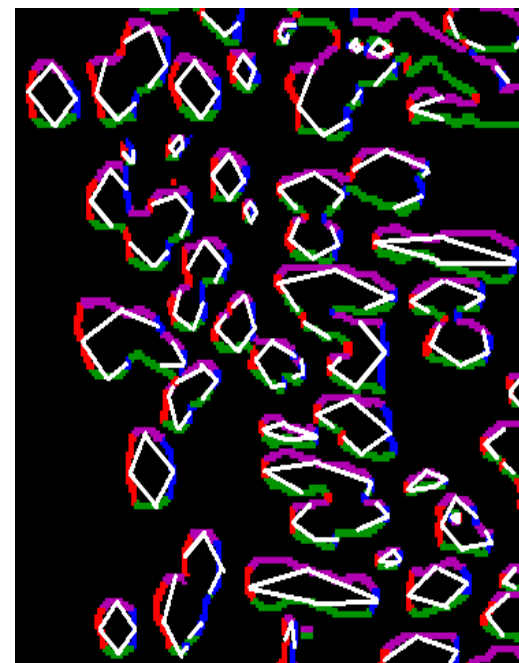
- Определение типа границ:

$$g_{ij} \in \begin{cases} I_{left}, & \text{если } |g_{ij}^x| > |g_{ij}^y| \text{ \& } g_{ij}^x > 0 \\ I_{top}, & \text{если } |g_{ij}^x| \leq |g_{ij}^y| \text{ \& } g_{ij}^y > 0 \\ I_{right}, & \text{если } |g_{ij}^x| > |g_{ij}^y| \text{ \& } g_{ij}^x < 0 \\ I_{bottom}, & \text{если } |g_{ij}^x| \leq |g_{ij}^y| \text{ \& } g_{ij}^y < 0 \end{cases}$$



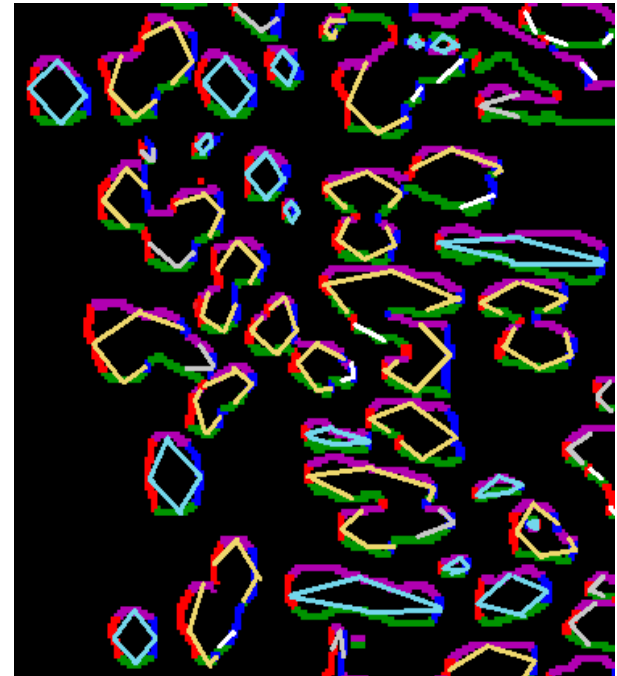
Граф отношений границ

- Граф $G = (V, E)$ отношений границ клеточных ядер:
 - вершины графа – границы ядер.
 - две вершины u и v соединены ребром $e = (u, v)$, если выполняются следующие условия:
 - 1) у границ, соответствующих вершинам u и v , есть соседние пиксели;
 - 2) одна из границ является вертикальной, а другая – горизонтальной;
 - 3) каждая граница должна иметь достаточно длинный участок, расположенный с определенной стороны относительно другой границы.

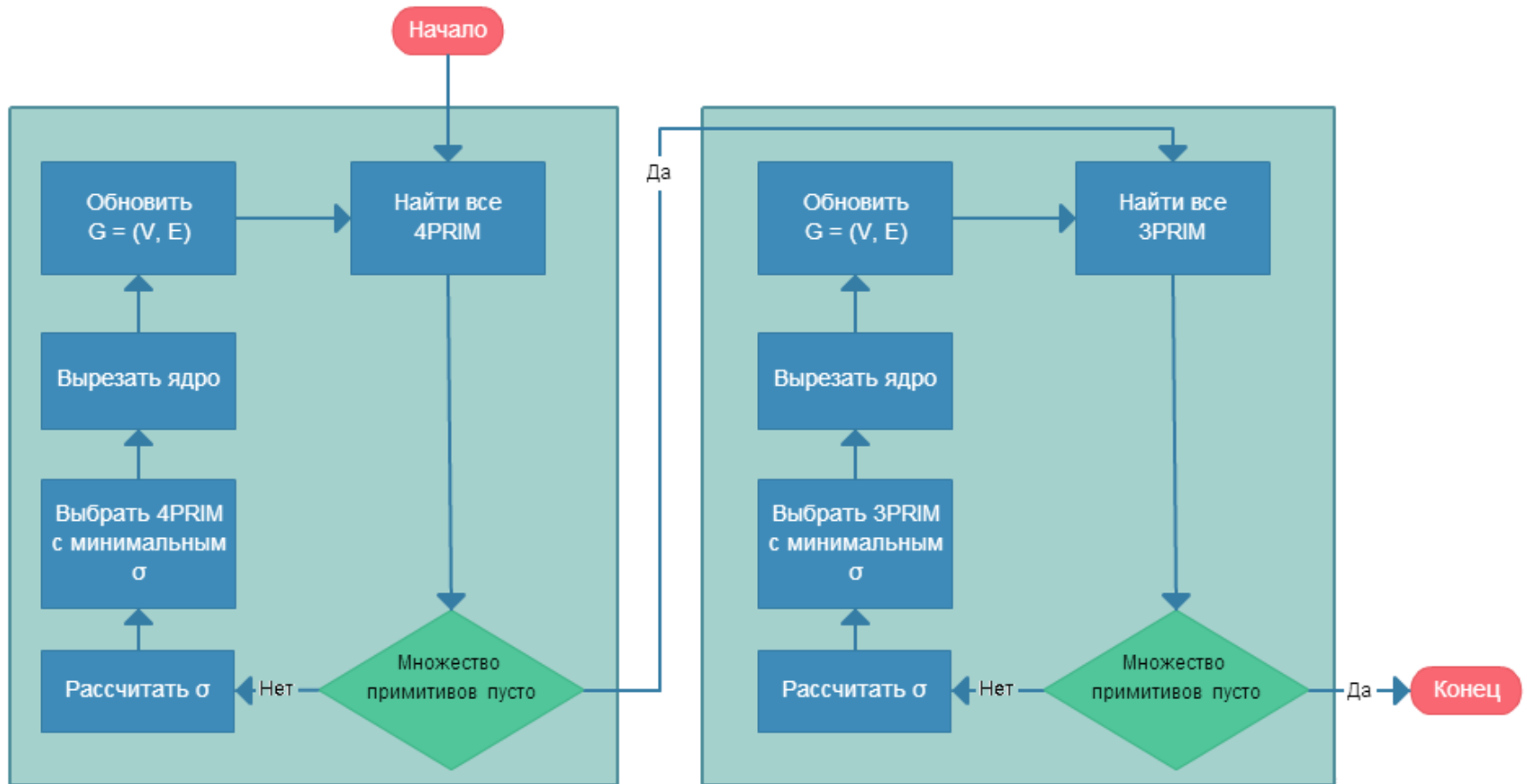


Примитивы графа отношений границ

- *4-примитив (4PRIM)* – структура на графе, состоящая из четырех вершин разного типа, соединенных ребрами так, чтобы эта структура представляла собой связный граф.
- 4-примитивы бывают двух типов:
 - замкнутые (цикл)
 - незамкнутые (цепь)
- *3-примитив (3PRIM)* – структура на графе, состоящая из трех вершин разного типа соединенных ребрами так, чтобы эта структура представляла собой связный граф.



Поиск клеточных ядер



Декомпозиция оболочки скопления ядер

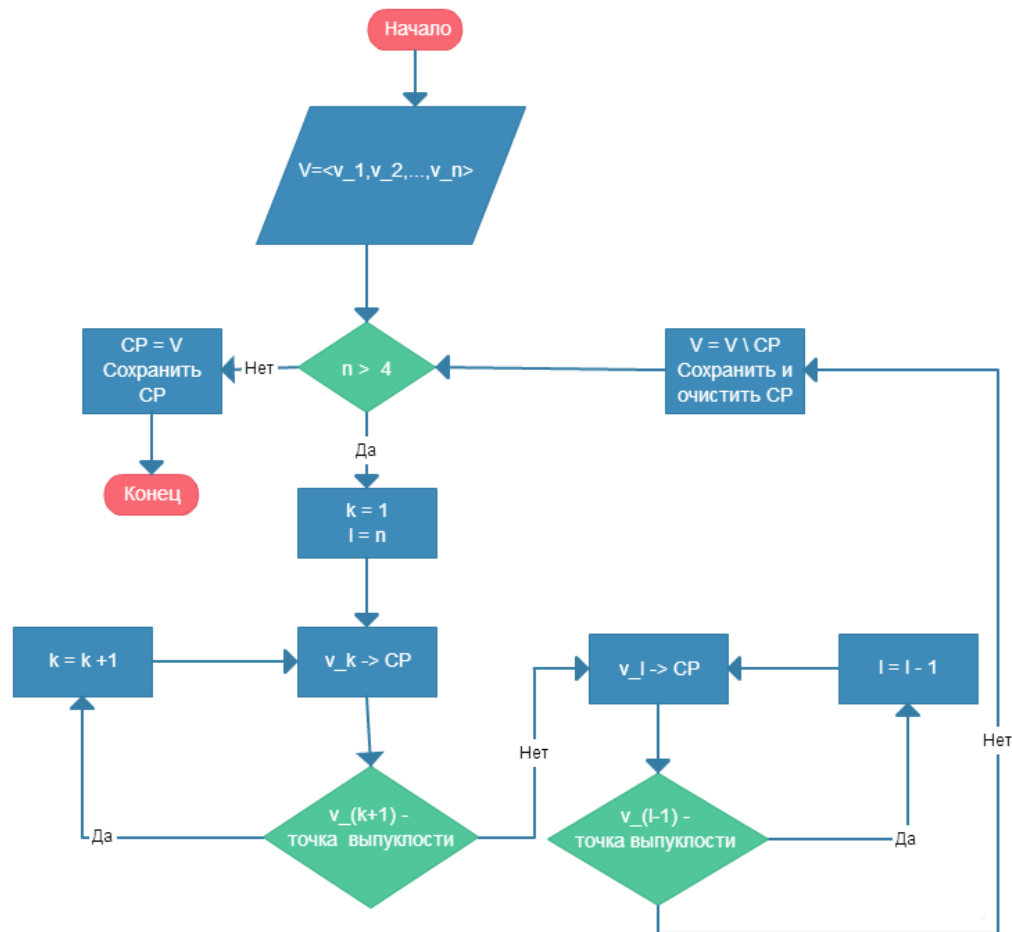
- Метод выделения клеточных ядер с использованием декомпозиции выпуклой оболочки их скоплений основан на *априорных знаниях* о выпуклости границ ядер.
- **Идея:** выделять отдельные ядра в областях их скоплений на основе выделения точек вогнутости на границе скоплений.
- Этапы:
 1. Выделение границ ядер (см. слайд 9 «Выделение границ»)
 2. Построение многоугольников, аппроксимирующих границы ядер
 3. Выделение выпуклых многоугольников
 4. Выделение клеточных ядер в областях их скопления

Аппроксимация оболочки МНОГОУГОЛЬНИКОМ

- Вершины многоугольника компоненты связности границ:
 - группы пикселей пересечения границ
 - точки вогнутости границ
- Стороны многоугольников – отрезки, последовательно соединяющие вершины.
- Определение точек вогнутости:
 - $\theta(i) = \left| \tan^{-1} \frac{p_{i-1}.y - p_i.y}{p_{i-1}.x - p_i.x} - \tan^{-1} \frac{p_{i+1}.y - p_i.y}{p_{i+1}.x - p_i.x} \right|$
 - Пиксель p_i считается точкой вогнутости для границы, если выполнены следующие условия:
 - $\theta_{min} \leq \theta_i \leq \theta_{max}$
 - Отрезок $\overline{p_{i-1}p_{i+1}}$ не пересекает границу



Выделение выпуклых многоугольников

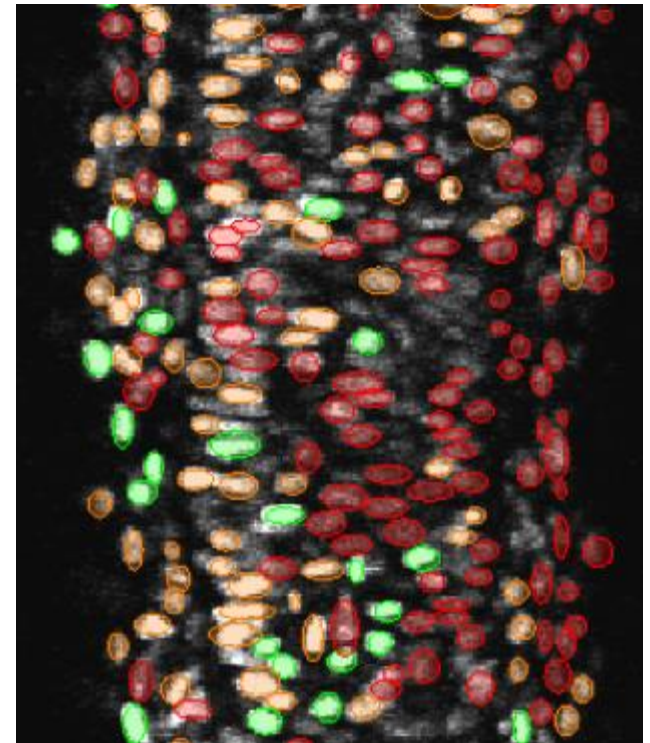


Выделение клеточных ядер в областях их скопления

- Центроид вершин выпуклого многоугольника $V^k = \langle v_1^k, v_2^k, \dots, v_{n_k}^k \rangle$:
$$C^k = \left(\frac{\sum_{i=1}^{n_k} v_i^k \cdot x}{n_k}, \frac{\sum_{i=1}^{n_k} v_i^k \cdot y}{n_k} \right)$$
- Центроиды – первые приближения ядер
- К каждой компоненте связности бинаризованного изображения применяется алгоритм наращивания областей
- **Идея:** «параллельный» запуск процедуры поиска в ширину для всех клеточных ядер

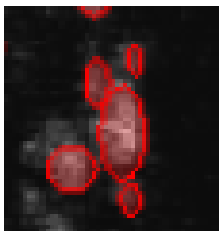
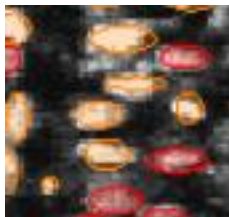
Достоверные данные

- *Достоверные данные* – полученные от экспертов изображения с отмеченными на них эллипсами клеточными ядрами
- Клеточные ядра на достоверных данных разделяются на три класса:
 - «отличные» – яркие ядра с четко выраженными границами;
 - «хорошие» – ядра средней яркости, как правило, находятся в областях скопления ядер;
 - «плохие» – плохо различимые ядра с сильно выраженной структурой.

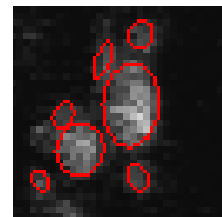
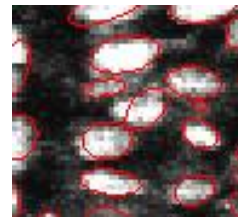


Визуальные результаты

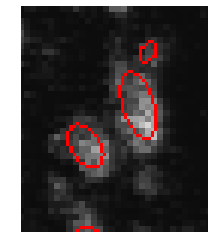
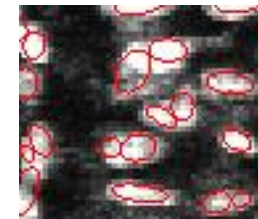
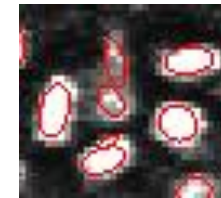
Достоверное
изображение



Граф отношений



Декомпозиция оболочки



Количественные результаты

- Выделенные клеточные ядра аппроксимируются эллипсами с помощью метода главных компонент
- Эллипс $E = \{C, r_x, r_y, \theta\}$, где C – координаты центра, r_x и r_y – длины большой и малой полуосей, θ – угол между большой осью и положительным направлением оси абсцисс
- Ядро считается распознанным, если среди достоверных данных существует эллипс, параметры которого совпадают с параметрами эллипса, аппроксимирующего ядро, с некоторой погрешностью

	«Отличные»	«Хорошие»	«Плохие»	«Лишние»
Граф отношений	95%	90%	83%	30%
Декомпозиция оболочки	80%	75%	75%	70%

Результаты

- Разработаны и реализованы алгоритмы выделения клеточных ядер на снимках микроскопа:
 - алгоритм выделения ядер с использованием графа отношений их границ
 - алгоритм выделения ядер с использованием декомпозиции оболочки их скоплений
- Проведены исследования качества распознавания ядер:
 - алгоритм выделения клеточных ядер с использованием графа отношений их границ оказался устойчивым к некачественным изображениям, а также способен распознавать как отдельно стоящие ядра, так и клеточные ядра в местах их скоплений.
 - алгоритм выделения клеточных ядер с использованием декомпозиции оболочки их скоплений оказался склонен к чрезмерной сегментации.
- Возможные улучшения алгоритма декомпозиции оболочки:
 - применение более продвинутых методов поиска точек вогнутости и выделения выпуклых многоугольников
 - применение постобработки, подобной используемой в методах водораздела

Приложение 1: выбор параметров

Параметр	Исследуемые значения	Выбранное значение
Размер окон медианного фильтра и фильтра Гаусса	3, 5, 7	3
Пороговое значение длины сегментов границ	3, 4, 5, 6	3
θ_{min}	$\frac{\pi}{12}, \frac{\pi}{10}, \frac{\pi}{7}$	$\frac{\pi}{10}$
θ_{max}	$\frac{7\pi}{12}, \frac{6\pi}{10}, \frac{4\pi}{7}$	$\frac{6\pi}{10}$