

С.С. САДЫКОВ,
Ю.А. БУЛАНОВА,
Е.А. ЗАХАРОВА

**Диалоговая система анализа
маммографических снимков**

УДК 004.932

Муромский институт
(филиал) ФГБОУ ВПО
«Владимирский
государственный
университет имени
А.Г. и Н.Г. Столетовых»

В данной статье рассмотрен автоматизированный процесс обработки и анализа маммографических снимков. Приведена структура системы, описаны все входящие в нее подсистемы.

Приводится описание используемых алгоритмов предварительной обработки, таких как коррекция яркости/контраста, алгоритмов шумоподавления (медианная фильтрация, фильтр Ли, фильтр Кауна, сигма-фильтр), выделения контуров, текстурной сегментации методом водораздела, а также приведено описание «псевдоцветной» сегментации.

Приведены результаты работы системы на тестовом снимке.

This article describes the automated processing and analysis of mammographic images. Is described the structure of the system and all of its subsystems.

The description of the algorithms pre-processing of the image, such as brightness and contrast, noise reduction algorithms such as median filtering, the filter Lee, the filter Kaun, sigma-filter, edge detection. A method of texture segmentation using watershed algorithm, as well as a description of pseudocolor segmentation.

The results of the system tested for mammography images.

Одним из важных разделов в области охраны здоровья женщин является совершенствование системы скрининга, поскольку рак молочной железы занимает лидирующие позиции среди злокачественных опухолей у женщин. Ежегодно в мире регистрируется более 1 млн. новых случаев РМЖ. В структуре онкологической заболеваемости женщин во всем мире рак молочной железы стоит на первом месте и на втором месте по смертности от рака у женщин. Наиболее высокая заболеваемость наблюдается в США и Западной Европе, составляя 25-30% от всех новых случаев рака у женщин и 18-20% от всех смертей при опухолях у женщин. Промежуточное положение занимают Восточная Европа и Южная Америка, самая низкая заболеваемость в Азии. В США 1 из 8 женщин рискует заболеть РМЖ в течение жизни, в Европе — 1 из 12, в Японии 1 из 80. При этом заболеваемость постоянно растет во всем мире на 1-2% в год [1].

Маммография - специальный рентгенологический метод исследования молочных желез у женщин с использованием пониженной дозы рентгеновских лучей [2].



Рис. 1. Маммограф Mammodiagnost 1000.

На данный момент развитие технического оснащения рентгенодиагностической службы идет в направлении цифровой техники, что требует замены аналоговых рентгеновских аппаратов на цифровые или использование цифровых рентгеновских комплексов для обработки рентгеновских снимков. Не обошли эти изменения и маммо-

графию. Однако большинство медицинских учреждений еще не оснащены цифровой техникой.

Обычно врач-рентгенолог ставит диагноз на основе собственных знаний и опыта, а также индивидуальных характеристик организма пациента, таких как возраст, вес, количество беременностей, аборт, принятие гормональных препаратов, наличие гинекологических заболеваний и так далее. Поэтому на точность и своевременный диагноз существенное влияние оказывает субъективный фактор – опыт и знания врача.

Субъективная оценка для диагностики заболевания молочной железы, используемых рентгенологами, осуществляется на основе визуального анализа маммограммы, позволяющего выявить необходимый минимум диагностических характеристик пациента.

Целью настоящей статьи является описание предложенной диалоговой системы анализа маммографических снимков, используемой врачом-рентгенологом при постановке диагноза заболевания молочной железы на основе математических методов и информационных технологий, а также выявление основных диагностических характеристик для различных групп людей (мужчин и женщин, работающих в различных условиях, при наличии факторов риска).

Диалоговая система обработки и анализа маммографических снимков состоит из таких подсистем как:

- Подсистема ввода: изображение может быть открыто из файла, либо получено со сканера.
- Подсистема предварительной обработки – включает в себя коррекцию яркости и контраста, изменение резкости изображения, фильтры шумоподавления, функции выделения контуров.
- Подсистема сегментации – необходимый раздел системы, благодаря которому можно выявить новообразования.
- Подсистема печати – осуществляет печать снимков, медицинских заключений.

Для использования предлагаемой системы необходимо оцифровывать маммограммы для их последующей обработки. Для оцифровки снимков, используемых в данной статье, использовался сканер Epson Perfection V700 Photo [3], снимки взяты из НУЗ «Отделенческой больницы на станции Муром ОАО «РЖД».

В предлагаемой диалоговой системе врач-рентгенолог, используя подсистему предварительной обработки [4] снимка, может: корректировать яркость и контраст несколькими методами, такими как бинаризация [5], гауссиан [6], косинус [6]; увеличить резкость изображения [7]. Эти изменения маммографического снимка необходимы для увеличения качества снимка, выделения скрытых областей, например на фоне мастопатии, так как часто получаются нечеткие, малоконтрастные маммограммы (например, неплотный прижим крышек кассеты или экрана к пленке приводит к появлению нерезкости изображения, использование экранов с повышенной зернистостью приводит к повышенной зернистости фона, что затрудняет поиск микрокальцинатов, механическое загрязнение экранов приводит к появлению мелких белых точек, симулирующих наличие мелких кальцинатов).

Для визуального объема, получения рельефности маммографического снимка используется градиент [8]. Если быть точнее, то градиент показывает, насколько «резко» или «плавно» меняется яркость изображения в каждой точке, а значит, вероятность нахождения точки на грани, а также ориентацию границы. Исходя из уникальности маммографических снимков, диалоговая система включает целый набор фильтров шумоподавления:

- медианная фильтрация - этот метод наиболее эффективен, если шум на изображении имеет аддитивный и импульсный характер и представляет собой ограниченный набор пиковых значений на фоне нулей.

Медианный фильтр реализуется как процедура локальной обработки снимка скользящим окном различной формы (рисунок 1), которое включает нечетное число отсчетов. Суть фильтрации: точки, попавшие в окно фильтра, упорядочиваются по возрастанию, вычисляется медиана (средняя точка), эта медиана заменяет центральную точку обрабатываемого маммографического снимка.

В результате применения медианного фильтра наклонные участки и резкие перепады значений яркости на изображениях не изменяются. Это очень полезное свойство именно для изображений, на которых, как известно, контуры несут основную информацию. В то же время импульсные помехи, протяженность которых составляет менее половины окна, будут подавлены [9, 10];

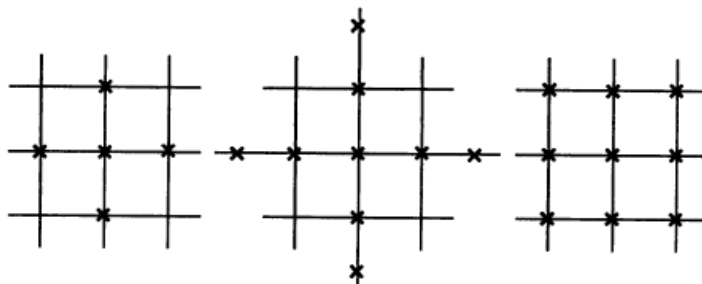


Рис. 2. Примеры скользящих окон медианного фильтра.

- адаптивная медианная фильтрация [10] – устраняет недостатки медианной фильтрации, например при фильтрации флуктуационного шума, а также не размывает контуры изображения и сохраняет его четкость;

- сигма-фильтр [9] – выполняет взвешенное усреднение только тех точек в окне, которые не слишком сильно отличаются от значения яркости центрального пикселя обрабатываемого окна,

$$\hat{f}(n_1, n_2) = \sum_{(k_1, k_2) \in D} h \cdot (k_1, k_2) \cdot g(n_1 + k_1, n_2 + k_2) \quad (1)$$

где обрабатываемая окрестность формируется следующим образом:

$$D = \{(k_1, k_2) : |g(n_1 + k_1, n_2 + k_2) - g(n_1, n_2)| \leq \sigma\} \quad (2)$$

- фильтр Ли – используется для фильтрации спекл-шума, не изменяет контуры объектов и контрастность отдельных частей. При его реализации выполняется оценка локального среднего $\mu_g(n_1, n_2)$ и дисперсии $\sigma_g(n_1, n_2)$ значений яркости снимка, расположенных в текущем окне D . Выходное значение фильтра формируется следующим образом:

$$\begin{aligned} \hat{f}(n_1, n_2) &= g(n_1, n_2) + [1 - \alpha(n_1, n_2)] \\ \alpha(n_1, n_2) &= \max\left\{0, \frac{\sigma_g^2(n_1, n_2) - \sigma_v^2}{\sigma_g^2(n_1, n_2)}\right\}, \end{aligned} \quad (3)$$

где σ_v^2 - оценка дисперсии шума на обрабатываемом снимке [8, 9, 10];

- фильтр Кауна - также не изменяет контуры объектов и контрастность отдельных частей. Идея фильтра в том, что мультипликативная модель спекл-шума преобразуется к аддитивной модели шума [10, 11].

$$R = I_c \cdot W + I_m \cdot (1 - W),$$

$$W = \frac{1 - \frac{C_u^2}{C_i^2}}{1 + \frac{C_u^2}{C_i^2}}, \quad (4)$$

$$C_i = \frac{S}{I_m},$$

где I_c – центр фильтрующего окна, I_m – среднее значение яркости фильтрующей маски, S – стандартное отклонение яркости в окне.

Врач-рентгенолог, осуществляя подбор фильтра к каждому изображению, удаляет с изображения некоторые артефакты, шум, отпечатки пальцев и так далее, гауссово размытие [11] также использует перед текстурной сегментацией методом водораздела [12, 13].

$$G(u, v) = \frac{1}{2 \cdot \pi \cdot \sigma^2} \cdot e^{-(u^2 + v^2) / (2 \cdot \sigma^2)}, \quad (5)$$

$$r^2 = u^2 + v^2,$$

где r^2 – радиус размытия, σ – отклонение распределение Гаусса.

Врач-рентгенолог может выделить область интереса на снимке, а затем корректировать ROI (изменяя яркость и контраст, резкость, выполняя псевдоокрашивание и так далее). В диалоговой системе имеются функции измерения новообразований, например диаметр новообразования.

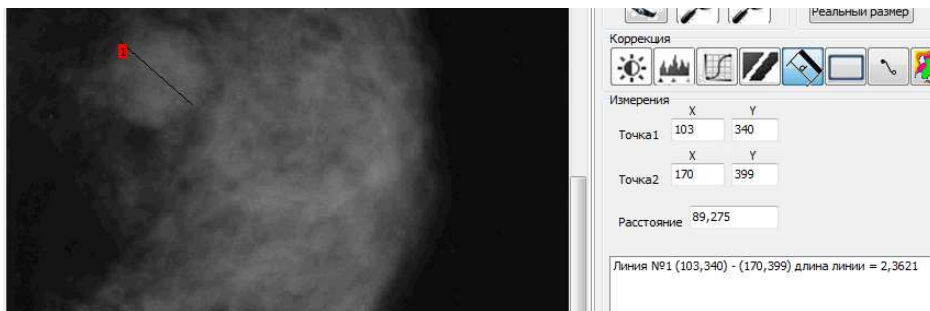


Рис. 3. Измерение диаметра новообразования.

Длина линии определяется в сантиметрах.

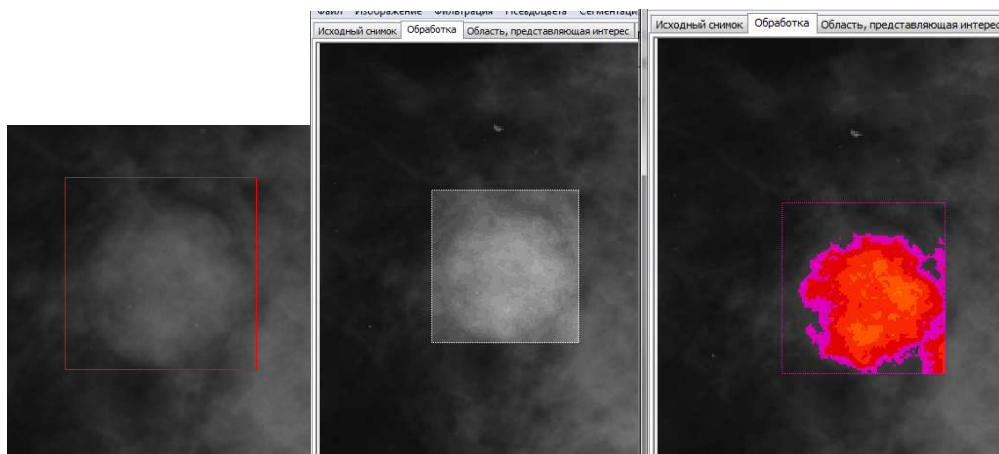


Рис.4. Обработка ROI.

Текстурная сегментация изображения маммограммы выполняется методом маркерного водораздела.

По исходному изображению строится гистограмма распределения. На основании гистограммы высчитываются основные структурные особенности, такие как:

- максимальная вероятность
- математическое ожидание (средняя яркость)
- дисперсия:

$$D = (\sum (i - \mu_i)^2 \cdot \sum P_{ij}) \cdot (\sum (j - \mu_j)^2 \cdot \sum P_{ij}). \quad (6)$$

- корреляция:

$$K = \sum_i \sum_j (i - \mu_x) \cdot (j - \mu_y) \cdot P_{ij} / \sigma_x \sigma_y, \quad (7)$$

где μ_x и μ_y - математические ожидания;
 σ_x и σ_y – СКО.

- энтропия:

$$H = \sum_i \sum_j P_{ij} \cdot \log(P_{ij}), \quad (8)$$

- энергия:

$$E = \sum_i \sum_j p_{i,j}^2, \quad (9)$$

На рис. 5 и 6 отображены схемы формирования линий водораздела, а также водосборных бассейнов.

$$L_i(x) = L(x, M_i) = \inf_{y \in M_i} L(x, y) \quad (10)$$

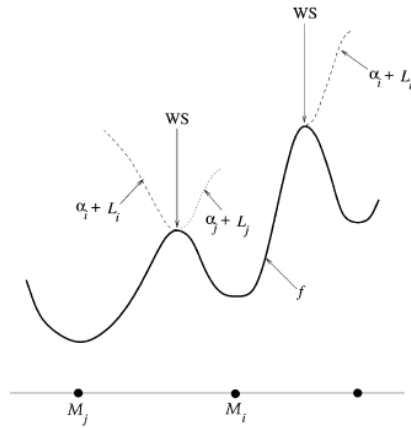


Рис. 5. Линии водораздела.

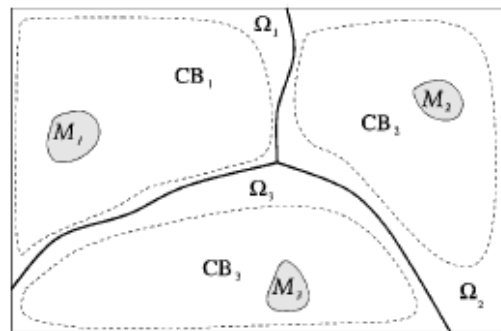


Рис.6. Линии водораздела (CB - Водосборный бассейн, M - Маркер).

$$CB_i = \{x \in \mathcal{X} \mid \forall j \neq i, 1 \leq j \leq K : \alpha_i + L_i(x) < \alpha_j + L_j(x)\} \quad (11)$$

В качестве функции сегментации используется градиент изображения, вычисленный с помощью оператора Собеля [8]:

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * A, \quad (12)$$

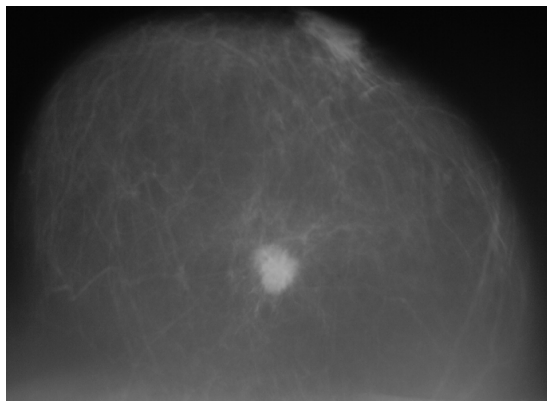
$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A,$$

где * - двумерная операция свертки.

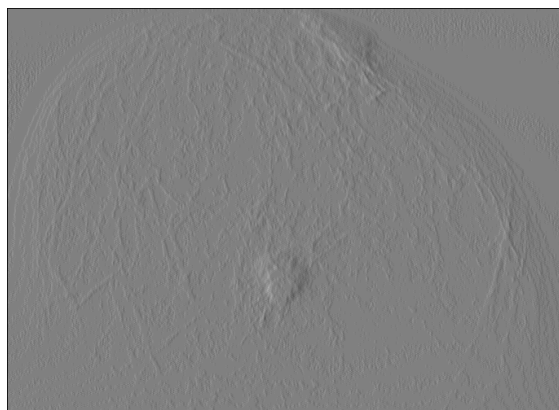
$$G = \sqrt{G_x^2 + G_y^2} \quad (13)$$

Формула 13 определяет значение градиента в каждой точке.

Для демонстрации работы алгоритма был взят снимок в прямой проекции и вычислено значение градиента:



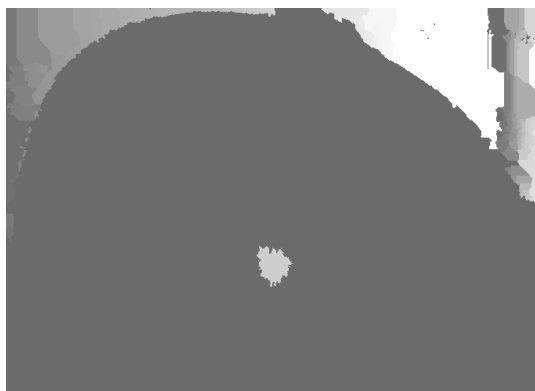
a)



б)

Рис. 7. Исходное изображение (а) и его градиент (б).

Далее, вычислив значение градиента, был применен метод водораздела, который дал следующий результат (Рис.8 (а)), а при использовании сглаживающего фильтра Гаусса получаем результат – Рис.8(б):



a)



б)

Рис. 8 – Результат текстурной сегментации

а) – текстурная сегментация изображения, не обработанного гауссовым размытием

б) текстурная сегментация изображения, обработанного гауссовым размытием

Тестирование алгоритма текстурной сегментации проводилось на 15 снимках, из них на 12 снимках эксперт подтвердил правильность выявления злокачественных новообразований.

Таким образом, результаты экспериментов доказали, что данный метод показывает хорошие результаты при предварительной обработке изображений. Серый участок на сегментированном изображении (Рис. 8) - выделенная искомая патология.

Для большей наглядности при отображении изображения точки из выбранных диапазонов могут быть закрашены каким-либо «псевдоцветом», что может также показать наличие тех или иных изменений в молочной железе.

Для представления «псевдоцветного» изображения используется цветовая модель HSB (Оттенок, Насыщенность, Яркость), в которой координатами цвета являются:

Hue — цветовой тон, (например, красный, зелёный или синеголубой). Варьируется в пределах $0—360^\circ$, однако иногда приводится к диапазону $0—100$ или $0—1$.

Saturation — насыщенность. Варьируется в пределах $0—100$ или $0—1$. Чем больше этот параметр, тем «чище» цвет, поэтому этот параметр иногда называют чистотой цвета. А чем ближе этот параметр к нулю, тем ближе цвет к нейтральному серому.

Value (значение цвета) или Brightness — яркость. Также задаётся в пределах 0—100 и 0—1.

Считаем, что:

$$\begin{aligned} H &\in [0,360) \\ S, V, R, G, B &\in [0,1) \end{aligned} \quad (14)$$

Пусть MAX — максимальное значение из R , G и B , а MIN — минимальное из них.

$$H = \begin{cases} 0, & \text{если } MAX = MIN \\ 60 * \frac{G - B}{MAX - MIN} + 0, & \text{если } MAX = R \text{ \& } G \geq B \\ 60 * \frac{G - B}{MAX - MIN} + 360, & \text{если } MAX = R \text{ \& } G < B \\ 60 * \frac{B - R}{MAX - MIN} + 120, & \text{если } MAX = G \\ 60 * \frac{R - G}{MAX - MIN} + 240, & \text{если } MAX = B \end{cases} \quad (15)$$

$$S = \begin{cases} 0, & \text{если } MAX = 0 \\ 1 - \frac{MIN}{MAX} \end{cases} \quad (16)$$

$$B = MAX$$

Так, например, «псевдоцветное» изображение снимка на Рис.2 (а) выглядит следующим образом:

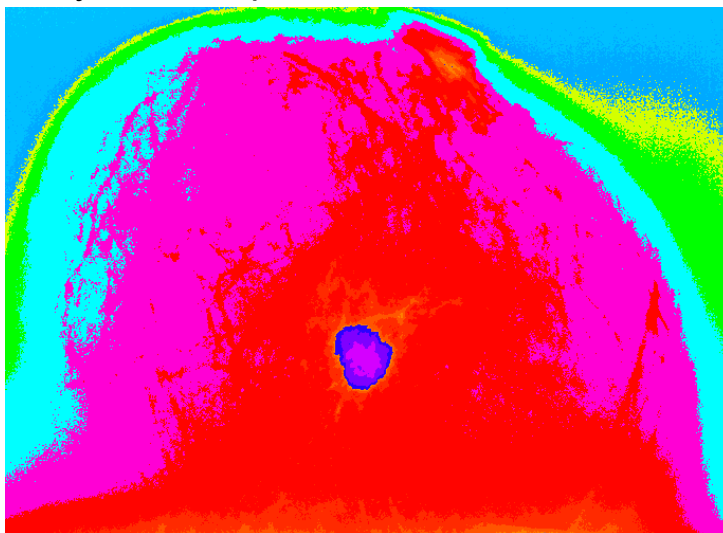


Рис. 9. «Псевдоцветное» изображение.

Для удобства врач может выключить отображение цветов фона, и оставить только те участки, которые ему наиболее интересны. Соответственно, на данном снимке, врач — эксперт выделил зону

оранжевого цвета, эта зона выделяет васкуляризацию вокруг раковой опухоли (Рис. 10).

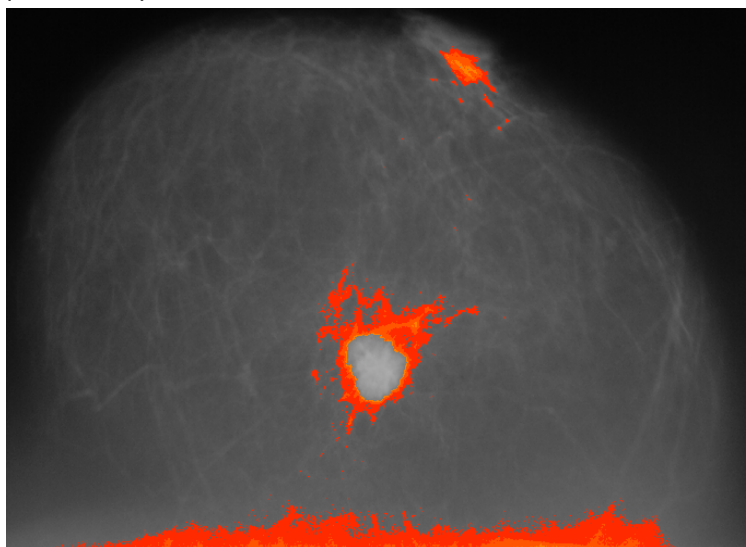


Рис. 10. Выделение с помощью «псевдоцвета» повышенной васкуляризации вокруг новообразования.

Основываясь на результатах экспериментов, можно считать, что предложенная методика наиболее подходит для выявления раковых опухолей в молочной железе. Это способствует раннему обнаружению рака молочной железы и скорейшему лечению этого заболевания.

Литература

1. *Н. И. Рожкова* Скрининг рака молочной железы, ФГУ «Российский Научный центр рентгенодиагностики Минздравсоцразвития России» «Медико-фармацевтический вестник Татарстана», 2009г.
2. *Корженкова Г.П.* Комплексная рентген-сонографическая диагностика заболеваний молочной железы, 1-е издание-М.: ООО «Фирма СТРОМ», 2004 -128 с, ил.
3. *С. Щербаков* [На стыке двух эпох](#): Планшетный сканер Epson Perfection V750 Pro // Foto&Video № 7, 2006 – с. 60-65
4. *С. С. Садыков, Д.Н. Стародубов* Инвариантные к повороту, переносу и изменению масштаба признаки площадных объектов // Системы и методы обработки и анализа информации: Сборник научных статей / Под ред. С.С. Садыкова, Д.Е. Андрианова – М.: Горячая линия – Телеком, 2005. – с. 32-38
5. *N. Otsu* (1979). «A threshold selection method from gray-level histograms». IEEE Trans. Sys., Man., Cyber. 9: 62-66.
6. *Р. Гонсалес, Р. Вудс, С. Эддинс* Цифровая обработка изображений в среде MATLAB. – М.: Техносфера, 2006. – 616 с.
7. *Р. Гонсалес, Р. Вудс* Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с. ISBN 5-94836-028-8

-
8. Методы компьютерной обработки изображений / Под ред. *В.А. Сойфера*. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 784 с. – ISBN 5-9221-0270-2
 9. Быстрые алгоритмы в цифровой обработке изображений / Под ред. *Т. Хуанга*. – М.: Радио и связь, 1984. – 320 с.
 10. *Харалик, Р.М.* Статистический и структурный подходы к описанию текстур / Р.М. Харалик. – ТИИЭР. – 1979. – Т. 67. – № 5. – С. 98-120.
 11. *Haralick R.M., Shanmugam K., Dinstein I.* Textural features for image classification. IEEE Transactions on systems, man and cybernetics, 1973, V. SMC-3. № 6, p. 610-621.
 12. *Кольцов П.П.* Оценка размытия изображения / Компьютерная оптика, том 35, №1. -2011. – с. 95-102
 13. *Прэнт У.К.* Цифровая обработка изображений. – М.: Мир, 1982. Т.1.-2. 792 с.

БУЛАНОВА Ю.А.

E-MAIL: YULIYABULANOVA@YANDEX.RU

НАУЧНЫЙ РУКОВОДИТЕЛЬ:

Д.Т.Н., ПРОФЕССОР САДЫКОВ С.С.

НАУЧНЫЙ КОНСУЛЬТАНТ:

ЗАХАРОВА Е.А.