

## Метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений

А.А. Захаров<sup>1</sup>, Д.В. Титов<sup>2</sup>, А.Л. Жизняков<sup>1</sup>, В.С. Титов<sup>2</sup>

<sup>1</sup> Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Муром, Россия,

<sup>2</sup> ФГБОУ ВО «Юго-Западный государственный университет», Курск, Россия

### Аннотация

В статье рассматривается разработка метода визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений. Целью исследований является создание метода, позволяющего с высокой точностью обнаруживать объекты на изображениях с низким цветовым контрастом выделяемых и фоновых областей. Для вычисления области значимости изображение предварительно сегментируется на регионы. На основе регионов строится граф. Каждый регион связан со смежными регионами, а также с областями, примыкающими к смежным регионам. Регионы являются вершинами графа. Вершины графа ранжируются по признакам соответствующих областей изображения. Область значимости выделяется на основе запросов фоновых областей. К фоновым областям относятся регионы, примыкающие к краям изображения. В существующем подходе визуального внимания на основе ранжирования вершин графа использовались только цветовые признаки изображения. В предлагаемом методе для повышения точности дополнительно используются текстурные признаки и признаки формы. Для вычисления текстурных признаков используется функция энергии Габора. При анализе формы рассчитывается расстояние между центрами регионов. Результаты экспериментов представлены на тестовых изображениях. Построены кривые точности-полноты, показывающие преимущество разработанного метода.

**Ключевые слова:** анализ изображений, визуальное внимание, граф, признаки изображений, ранжирование, компьютерное зрение.

**Цитирование:** Захаров, А.А. Метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений / А.А. Захаров, Д.В. Титов, А.Л. Жизняков, В.С. Титов // Компьютерная оптика. – 2020. – Т. 44, № 3. – С. 427-435. – DOI: 10.18287/2412-6179-CO-658.

**Citation:** Zakharov AA, Titov DV, Zhiznyakov AL, Titov VS. Visual attention method based on vertex ranking of graphs by heterogeneous image attributes. Computer Optics 2020; 44(3): 427-435. DOI: 10.18287/2412-6179-CO-658.

### Введение

Визуальное внимание («*visual attention*») представляет собой способность системы компьютерного зрения быстро находить нужные данные в наблюдаемой сцене. Считается, что поток визуальных данных размером  $10^8 - 10^9$  бит поступает в зрительную систему человека каждую секунду [1, 2]. Поэтому основная цель визуального внимания – быстрое обнаружение объектов в сцене с использованием наименее возможного количества информации для решения сложных высокоуровневых задач распознавания образов. Концепция систем визуального внимания стала использоваться в связи с появлением различных междисциплинарных подходов, что привело к взаимодействию учёных в областях психологии, физиологии, распознавания образов, искусственного интеллекта, компьютерного зрения, машинного обучения, робототехники.

В области компьютерного зрения NP-полные вычислительные задачи распознавания и группировки являются основным препятствием для реализации

приложений реального времени. Таким образом, для преодоления проблемы сложности в приложениях компьютерного зрения использование концепции визуального внимания очень актуально. Области применения моделей визуального внимания являются автономная навигация роботов, человеко-машинные интерфейсы, контроль технологических процессов, видеонаблюдение, дистанционное зондирование Земли, биометрические системы, медицинская диагностика и т.д. [2–5]. В последнее время разработано множество моделей визуального внимания на основе компьютерного зрения. Обычно эти модели используют признаки изображения для создания карты значимости («*saliency map*»). Выделяемая область представлена на карте значимости с помощью интенсивности пикселей.

Наиболее часто в моделях визуального внимания используются следующие признаки: цвет, яркость, ориентация [6, 7], направление движения [8], глубина сцены [9], структурные признаки [10], вейвлеты [11], центральное расположение [12], оптический поток [13], энтропия [14], симметрия [15], текстурный

контраст [16], текстурные неоднородности [17], морфлеты [18], спектр [19], контекстные особенности [20], ковариационные признаки [21], ошибки реконструкции [22], плотность ядра признаков [23], центр выпуклой оболочки [24], параметры цепей Маркова [25], признаки патчей [26], параметры условных случайных полей [27], ядра локальной регрессии [28], особенности низкоуровневой модели зрения [29], статистические признаки изображений [30], различие признаков областей в пространстве изображения [31] и т.д.

Модели визуального внимания разделяют на восходящие и нисходящие. Восходящие модели (*bottom-up*) основаны на низкоуровневых характеристиках визуальной сцены [6, 32]. Восходящие модели внимания характеризуются высокой скоростью анализа изображений. Однако восходящие модели соответствуют только небольшой части функций внимания человека. Нисходящие модели (*top-down*) определяются текущими целями и задачами [7]. Соответственно, реализация нисходящих моделей визуального внимания является более медленной по сравнению с восходящими подходами, но позволяет выделить смысловое содержание сцены.

Наиболее заметные достижения в области визуального внимания связаны с использованием моделей «глубокого обучения» [33–35]. Однако выделяют следующие недостатки подобных методов: высокая вычислительная сложность по сравнению с традиционными подходами, необходимость предварительного обучения на большом наборе данных, объём обучающей выборки сильно влияет на количество классов анализируемых изображений.

Несмотря на значительный прогресс, достигнутый в течение последних десятилетий, имеющиеся в настоящее время методы компьютерного зрения далеки от зрительной системы человека по надёжности и производительности. В настоящее время в области компьютерного зрения существуют проблемы, связанные с выделением объектов на изображениях с низким контрастом, присутствующим шумом и сложными текстурами. Часто существующие методы визуального внимания требуют априорных знаний, чтобы обеспечить их надлежащее функционирование. Это обстоятельство существенно ограничивает область применения подобных подходов. До сих пор для решения многих задач необходимы алгоритмическое и программное обеспечения, позволяющие точно выделять объекты на изображениях с низким контрастом и сложными текстурами без этапа предварительного обучения.

В работе предлагается метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений. Изображение сегментируется на регионы (суперпиксели). На основе регионов строится граф. Регионы являются вершинами графа. Каждой вершине графа присваивается ранг

на основе признаков регионов изображения. В соответствии со значением ранга определяется вхождение региона в область значимости. В представленной работе новыми являются следующие положения: 1) для ранжирования вершин графа в качестве признаков региона используется не только цветовые, но и текстурные характеристики; 2) имеется возможность изменять форму выделяемой области значимости, оценивая расстояние между регионами.

### 1. Разработка метода визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений

#### Построение графа

В работе [36] предложен метод ранжирования, который использует внутреннюю структуру данных для маркировки графов. Для построения карты значимости предлагается использовать метод ранжирования вершин графа, представленных регионами изображения. На первом этапе изображение сегментируется на регионы с использованием простой линейной итеративной кластеризации (*simple linear iterative clustering – SLIC*) [37].

На основе регионов строится граф  $G=(V, E)$ , где  $V$  – набор вершин, а  $E$  – набор неориентированных рёбер (рис. 1). Каждый регион связан как со смежными регионами, так и с регионами, примыкающими к смежному. Смежными регионами называются регионы, имеющие общую границу. Вершины, представленные регионами с подобными связями, соединяются рёбрами.

На рис. 1 показано построение рёбер между некоторой вершиной (обозначена точкой большого размера) и другими вершинами графа (обозначены точками малого размера). Подобным образом осуществляется построение всех рёбер графа.

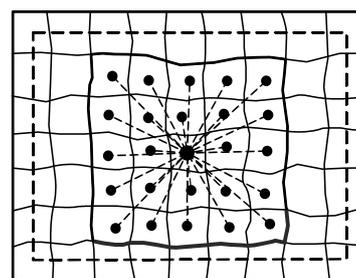


Рис. 1. Построение рёбер графа между отдельной вершиной и другими вершинами графа (все регионы, связанные с рассматриваемым регионом, ограничены сплошным контуром; через регионы, примыкающие к границам изображения, проходит пунктирная линия)

#### Цветовые признаки

Для описания цветовых свойств регионов используется цветовая модель CIE Lab. В модели CIE Lab используются три составляющих цветового зрения человека: светлота (*Lightness*) и две хроматические компоненты (канал *a* кодирует цвета от тёмно-зелё-

ного до пурпурного цвета, канал  $b$  представляет цвета от синего до жёлтого). Значения каналов  $a$  и  $b$  меняются от  $-128$  до  $127$ , а параметр  $L$  – от  $0$  до  $100$ . При вычислении цветовых характеристик каждого региона происходит усреднение цветовых значений всех входящих в него пикселей [38].

### Текстурные признаки

Текстурные признаки часто игнорируются в вычислительных моделях визуального внимания. Для извлечения текстурных признаков предлагается использовать функцию энергии Габора [39]. Фильтр Габора может обнаруживать края определённой ориентации и масштаба.

Фильтр Габора записывается в следующем виде:

$$g(x, y, \gamma, \theta, \lambda, \sigma, \psi) = \exp\left(-\frac{\bar{x}^2 + \gamma^2 \bar{y}^2}{2\sigma^2}\right) \cos(2\pi i \bar{x} + \psi), \quad (1)$$

где  $\bar{x} = x \cos \theta + y \sin \theta$ ;  $\bar{y} = -x \sin \theta + y \cos \theta$ ;  $x, y$  – координаты пикселя на изображении;  $i$  – радиальная частота;  $\theta$  – угол, который определяет ориентацию нормали параллельных полос функции Габора;  $\psi$  – сдвиг фаз в градусах;  $\sigma$  – среднеквадратичное отклонение ядра Гаусса;  $\gamma$  – коэффициент сжатия, характеризующий эллиптичность функции Габора ( $\gamma = 0,5$ ).

Входное изображение обрабатывается группой фильтров Габора, которые имеют различные параметры ориентации и частоты. Полученная коллекция изображений называется банком. При создании банка различные значения ориентации и частоты выбираются таким образом, чтобы охватывалась вся пространственно-частотная область. В методе фильтры Габора вычисляются при следующих углах ориентации нормали:  $0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$ . Для изображения шириной  $N$  принято использовать следующие радиальные частоты для группы фильтров Габора [40]:

$$\frac{1\sqrt{2}}{N}, \frac{2\sqrt{2}}{N}, \frac{4\sqrt{2}}{N}, \dots, \frac{1}{4} N \sqrt{2}.$$

При вычислении фильтров Габора были использованы 4 значения радиальной частоты  $u$ :  $\sqrt{2}/8, \sqrt{2}/16, \sqrt{2}/32, \sqrt{2}/64$ . Таким образом, размер банка изображений, обработанных фильтром Габора, равен 24. Отношение  $\sigma/\lambda = 0,56$  является постоянным для всех фильтров в банке ( $\lambda = 1/u$  – длина волны). Этот выбор объясняется свойствами простых клеток зрительной коры, которые могут моделироваться фильтром Габора [39]. Энергия Габора связана с моделью сложных клеток в первичной зрительной коре человека [39]. Функция энергии Габора записывается следующим образом [39]:

$$Q = \sqrt{R(x, y, \gamma, \theta, \lambda, \sigma, 0)^2 + R(x, y, \gamma, \theta, \lambda, \sigma, \pi)^2}, \quad (2)$$

где  $R(x, y, \gamma, \theta, \lambda, \sigma, \psi) = I(x, y) * g(x, y, \gamma, \theta, \lambda, \sigma, \psi)$ ,  $I(x, y)$  – изображение,  $*$  – операция свёртки.

Для вычисления полной энергии Габора  $Z$  значения энергии  $Q_i$  всех изображений, обработанных фильтрами с различными параметрами ориентации и частоты, суммируются в каждой точке:

$$Z = \sum_{i=1}^K Q_i, \quad (3)$$

где  $K$  – размер банка фильтров.

При вычислении текстурных признаков каждого региона происходит усреднение значений полной энергии Габора всех входящих в него пикселей.

### Признаки формы

Вес рёбер графа измеряет сходство между вершинами. В существующих методах обнаружения объектов на основе графов веса рёбер обычно вычисляются на основе цветового различия между регионами [38]. Недостаток подобных подходов состоит в том, что при цветовом сходстве областей удалённого фона и обнаруживаемого объекта получаются некорректные результаты. Для решения этой проблемы предлагается учитывать расстояние между регионами. Использование расстояния между регионами обусловлено тем, что значимые области объекта обычно являются относительно компактными (с точки зрения пространственного распределения) и однородными по внешнему виду (с точки зрения распределения признаков). Расстояние между регионами рассчитывается следующим образом:

$$H = \sqrt{(x_{Ci} - x_{Cj})^2 + (y_{Ci} - y_{Cj})^2}, \quad (4)$$

где  $(x_{Ci}, y_{Ci})$  и  $(x_{Cj}, y_{Cj})$  – координаты центров  $i$ -го и  $j$ -го регионов, вычисляемые как среднее арифметическое координат всех точек региона.

### Ранжирование вершин графа по разнородным признакам изображений

Разрабатываемый метод визуального внимания использует графы для представления данных. Каждая вершина графа ранжируется на основе признаков регионов изображения. В соответствии с рангом вершины регион относится либо к фону, либо к выделяемому объекту [38].

Пусть изображение содержит  $n$  регионов. Тогда вектор признаков для  $i$ -го региона можно обозначить  $x_i \in R^m$ . С учётом векторов признаков регионов формируется набор данных,  $X = \{x_1, \dots, x_q, x_{q+1}, \dots, x_n\} \in R^{m \times n}$ , в котором первые  $q$  элементов являются запросами, а остальные элементы должны быть ранжированы в соответствии с их релеванностью запросам.

Пусть  $f: X \rightarrow R^n$  является функцией ранжирования, которая связывает значение  $x_i$  со значением ранга  $f_i$ . Таким образом,  $f$  можно рассматривать как вектор  $f = [f_1, \dots, f_n]^T$ . Пусть  $y = [y_1, y_2, \dots, y_n]^T$  является индикатором

торным вектором, в котором  $y_i = 1$ , если  $x_i$  является запросом, и  $y_i = 0$  в противном случае.

На основе регионов строится граф  $G = (V, E)$ , в котором вершины  $V$  являются набором данных  $X$ . Рёбра  $E$  имеют веса  $w_{ij}$ , вычисленные с помощью матрицы  $W = [w_{ij}]_{n \times n}$ :

$$w_{ij} = \begin{cases} \exp\left(-\frac{\|Lab_i - Lab_j\|}{\sigma_{Lab}^2} - \frac{\|Z_i - Z_j\|}{\sigma_Z^2} - \frac{\|H_{ij}\|}{\sigma_H^2}\right), \\ \text{если } (v_i, v_j) \in E; \\ 0, \text{ в другом случае,} \end{cases} \quad (5)$$

где  $Lab_i$  и  $Lab_j$  – средние значения регионов в цветовом пространстве CIE LAB;  $Z_i$  и  $Z_j$  – значения полной энергии Габора  $i$ -го и  $j$ -го регионов;  $H_{ij}$  – расстояние между центрами  $i$ -го и  $j$ -го регионов;  $\sigma_{Lab}$ ,  $\sigma_Z$ ,  $\sigma_H$  – константы, управляющие значением веса;  $E$  – множество рёбер графа.

Для матрицы смежности графа  $G$  вычисляется матрица степеней  $D = \text{diag}\{d_{11}, \dots, d_{nn}\}$ , где  $d_{ii} = \sum_j w_{ij}$ .

Таким образом, ранжирование вершин сводится к решению задачи оптимизации [36]:

$$f^* = \arg \min_f \frac{1}{2} \left( \sum_{i,j=1}^n w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^n \|f_i - y_i\|^2 \right), \quad (6)$$

где  $\mu$  – параметр, управляющий гладкостью и совместностью.

Функция ранжирования не должна сильно изменяться между близлежащими вершинами (ограничение гладкости) и не должна сильно отличаться от начального значения запроса (ограничение совместности). Функция ранжирования (6) может быть записана следующим образом:

$$f^* = (I - \alpha L_N)^{-1} y, \quad (7)$$

где  $I$  – единичная матрица;  $\alpha = 1/(1 + \mu)$  – коэффициент;  $L_N = D^{-2} W D^{-2}$  – нормализованная матрица Лапласа.

Выражение (7) можно записать другим образом:

$$f^* = (D - \alpha W)^{-1} y. \quad (8)$$

#### Вычисление карты значимости

Значимость каждого узла определяется на основе функции ранжирования (8) с использованием графа изображения и вершин запроса. Для упрощения вычислений выражение (8) можно записать  $f^* = Ay$ . Матрицу  $A$  можно рассматривать как матрицу смежности, которая равна  $(D - \alpha W)^{-1}$ . Величина значимости для каждого региона изображения определяется на основе признаков других регионов. При вычислении значимости конкретного региона его собственные при-

знаки не учитываются. С этой целью при вычислении ранга диагональные элементы матрицы  $A$  устанавливаются в 0. При вычислении карты значимости  $i$ -го региона на основе фоновых запросов используется выражение  $1 - f^*(i)$ . Регионы на границах изображения используются в качестве фоновых областей (на рис. 1 обозначены пунктиром). Если нормализовать вектор  $f^*(i)$  в диапазоне от 0 до 1, то значение карты значимости  $S(i)$  можно записать в виде:

$$S(i) = 1 - f_N^*(i); \quad i = 1, \dots, n, \quad (9)$$

где  $f_N^*(i)$  – нормализованный вектор;  $i$  – номер региона;  $n$  – количество регионов на изображении.

#### Алгоритм вычисления карты значимости

Шаг 1. Сегментация входного изображения на регионы. Построение графа  $G$ . Вычисление матриц  $D$  и  $W$ .

Шаг 2. Вычисление матрицы  $A = (D - \alpha W)^{-1}$ . Присвоение диагональным элементам матрицы  $A$  значения 0.

Шаг 3. Формирование четырёх индикаторных векторов  $u$ . Запросам в каждом из четырёх случаев соответствуют регионы, примыкающие к одной из сторон изображения (левой, правой, верхней, нижней). На основе выражения (8) вычисляются четыре карты значимости. Вычисляется общая карта значимости:

$$S_{bq}(i) = S_l(i) \times S_b(i) \times S_t(i) \times S_r(i), \quad (10)$$

где  $S_l(i)$ ,  $S_b(i)$ ,  $S_t(i)$ ,  $S_r(i)$  – карты значимости, вычисленные на основе фоновых запросов для верхней, нижней, левой и правой границ изображения.

Шаг 4. Для уточнения карты значимости используются запросы на основе регионов переднего плана. Карта значимости  $S_{bq}(i)$ , вычисленная на шаге 3, используется для формирования запросов переднего плана и вектора  $u$ . Вычисляется область значимости на основе фоновых запросов переднего плана изображения:

$$S_{fq}(i) = f_n^*(i); \quad i = 1, \dots, n. \quad (11)$$

### 2. Исследование метода визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений

Исследование метода проводилось на изображениях, входящих в набор данных MSRA-1000. Анализ текстурных признаков необходим при выделении объектов на изображениях с низким цветовым контрастом и сложными текстурами. На рис. 2 показано использование текстурных признаков в методе визуального внимания на основе ранжирования вершин графа [38].

На рис. 3 показано использование признаков формы для создания карты значимости. При использова-

нии информации о расстоянии между центрами регионов выделяемая область становится более компактной, удаляются мелкие включения на карте значимости (рис. 3г).

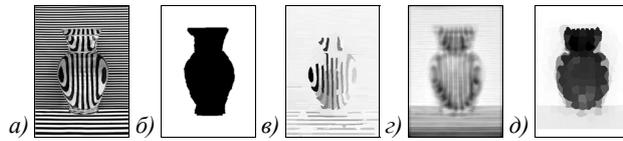


Рис. 2. Использование текстурных признаков для получения карт значимости: а) тестовое изображение из набора данных MSRA-1000, б) эталонная карта значимости тестового изображения, в) карта значимости, полученная на основе метода ранжирования вершин графа по цветовым признакам изображения [38], г) изображение, обработанное с использованием функции энергии Габора, д) карта значимости, полученная на основе метода ранжирования вершин графа по цветовым и текстурным признакам изображения

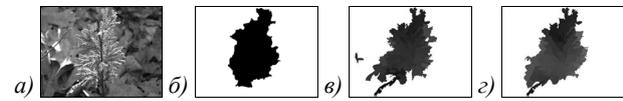


Рис. 3. Использование признаков формы для получения карт значимости: а) тестовое изображение из набора данных MSRA-1000, б) эталонная карта значимости тестового изображения, в) карта значимости, полученная на основе метода ранжирования вершин графа по цветовым признакам изображения [38], г) карта значимости, полученная на основе метода ранжирования вершин графа по признакам формы и цветовым признакам изображения

Сравнение разработанного метода проводилось со следующими моделями визуального внимания CA [19], COV [20], DSR [21], FES [22], GR [23], MC [24], PCA [25], SEG [26], SeR [27], SIM [28], SR [29], SUN [30], SWD [31], MR [38] (для сравнения были использованы изображения из набора данных MSRA-1000), приведены карты значимости, полученные с использованием различных методов (табл. 1, см. ниже).

Для демонстрации преимуществ разработанного метода были выбраны изображения с низким цветовым контрастом и сложными текстурами.

Для количественной оценки результатов работы предложенного метода используются кривые точности-полноты (*precision-recall* – PR) [41]. Карту значимости  $S$  можно преобразовать в двоичную маску  $M$  и вычислить точность и полноту.

Полученная двоичная маска  $M$  сравнивается с истинной значимой областью  $P$  конкретного изображения:

$$\text{Точность} = \frac{|M \cap P|}{|M|}, \quad \text{Полнота} = \frac{|M \cap P|}{|P|}. \quad (12)$$

Для построения кривой точности-полноты карту значимости  $S$  рассматривают как изображение в градациях серого. Осуществляется бинарная сегментация изображения в диапазоне  $[0, 255]$  с шагом 1. На каждом шаге вычисляются значения точности и пол-

ноты. На основе вычисленных значений формируется кривая точности-полноты.

На основе полученных карт значимости были построены кривые точности-полноты (рис. 4).

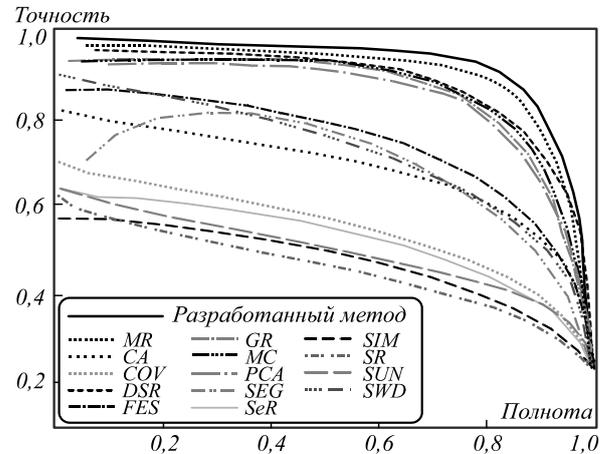


Рис. 4. Кривые точности-полноты существующих и разработанного методов

На основе полученных карт значимости и кривых точности-полноты можно сделать вывод, что разработанный метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений превосходит многие существующие подходы. Особенно хорошо это видно при выделении области значимости на изображениях с низким цветовым контрастом и сложными текстурами.

### Заключение

В работе предложен метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений. Метод имеет существенные преимущества по сравнению с другими подходами при выделении карт значимости на изображениях с низким цветовым контрастом и сложными текстурами. К достоинствам метода относится высокая скорость вычислений, обусловленная использованием низкоуровневых признаков. Кроме того, метод не требует этапа предварительного обучения, что позволяет использовать его в различных прикладных областях, связанных с компьютерным зрением.

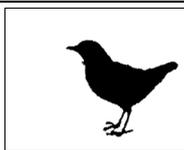
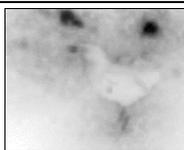
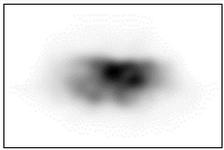
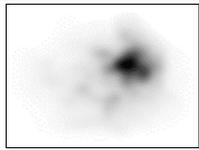
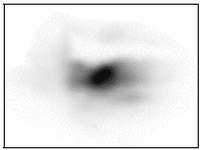
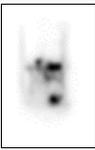
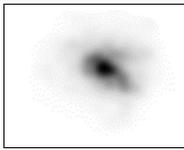
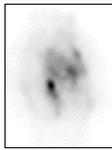
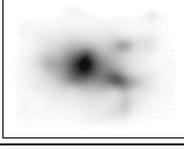
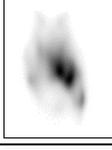
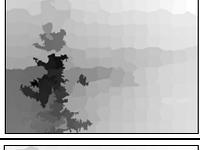
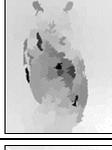
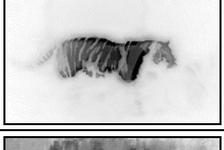
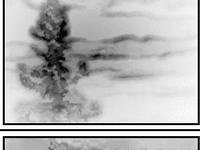
### Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (Госзадание ВлГУ ГБ-1187/20).

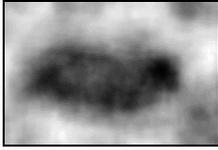
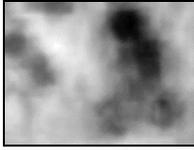
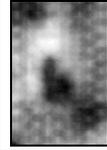
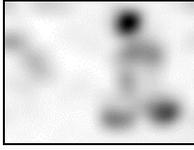
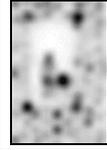
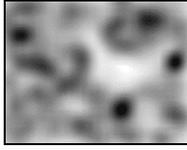
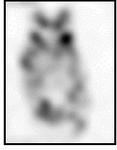
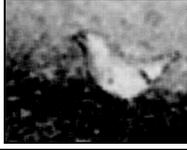
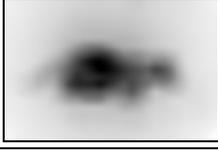
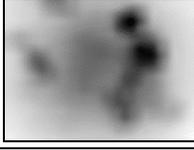
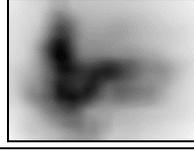
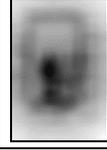
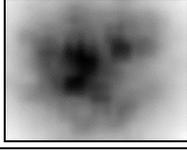
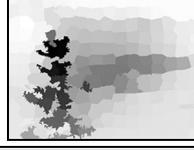
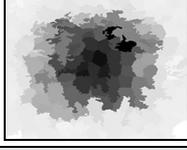
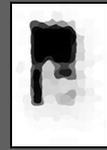
### Литература

1. Koch, K. How much the eye tells the brain / K. Koch, J. McLean, R. Segev, M.A. Freed, M.J. Berry, V. Balasubramanian, P. Sterling // *Current Biology*. – 2006. – Vol. 16, Issue 14. – P. 1428-1434.
2. Borji, A. State-of-the-art in visual attention modeling / A. Borji, L. Itti // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2013. – Vol. 35, Issue 1. – P. 185-207.

Табл. 1. Карты значимости, полученные с использованием представленного и существующих методов визуального внимания на основе тестовых изображений из набора данных MSRA-1000

Методы	Тестовые изображения из набора данных MSRA-1000					
						
	Эталонные карты значимости тестовых изображений					
						
Методы	Карты значимости, полученные с использованием различных методов					
CA [19]						
COV [20]						
DSR [21]						
FES [22]						
GR [23]						
MC [24]						
PCA [25]						
SEG [26]						

Продолжение табл. 1

SeR [27]						
SIM [28]						
SR [29]						
SUN [30]						
SWD [31]						
MR [38]						
<b>Карты значимости, полученные с использованием разработанного метода</b>						
						

- Begum, M.** Visual attention for robotic cognition: A survey / M. Begum, F. Karray // *IEEE Transactions on Autonomous Mental Development*. – 2011. – Vol. 3, Issue 1. – P. 92-105.
- Mahdi, A.** A comparison study of saliency models for fixation prediction on infants and adults / A. Mahdi, M. Su, M. Schlesinger, J. Qin // *IEEE Transactions on Cognitive and Developmental Systems*. – 2018. – Vol. 10, Issue 3. – P. 485-498.
- Garg, A.** A survey on visual saliency detection and computational methods / A. Garg, A. Negi // *International Journal of Engineering and Technology*. – 2017. – Vol. 9, Issue 4. – P. 2742-2753.
- Itti, L.** A model of saliency-based visual attention for rapid scene analysis / L. Itti, C. Koch, E. Niebur // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1998. – Vol. 20, Issue 11. – P. 1254-259.
- Frintrop, S.** VOCUS: a visual attention system for object detection and goal-directed search / S. Frinrop. – Heidelberg, Germany: Springer-Verlag, 2006. – 216 p.
- Itti, L.** Realistic avatar eye and head animation using a neurobiological model of visual attention / L. Itti, N. Dhavale, F. Pighin // *Proceedings of SPIE*. – 2003. – Vol. 5200. – P. 64-78.
- Wang, J.** Computational model of stereoscopic 3D visual saliency / J. Wang, M.P. Da Silva, P. Le Callet, V. Ricordel // *IEEE Transactions on Image Processing*. – 2013. – Vol. 22, Issue 6. – P. 2151-2165.
- Harel, J.** Graph-based visual saliency / J. Harel, C. Koch, P. Perona // *Neural Information Processing Systems*. – 2006. – Vol. 19. – P. 545-552.
- Salvucci, D.D.** An integrated model of eye movements and visual encoding / D.D. Salvucci // *Cognitive Systems Research*. – 2001. – Vol. 1. – P. 201-220.
- Tatler, B.W.** The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor bases and image feature distributions / B.W. Tatler // *Journal of Vision*. – 2007. – Vol. 14. – P. 1-17.
- Vijayakumar, S.** Overt visual attention for a humanoid robot / S. Vijayakumar, J. Conradt, T. Shibata, S. Schaal // *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. – 2001. – Vol. 4. – P. 2332-2337.

14. **Kadir, T.** Saliency, scale and image description / T. Kadir, M. Brady // International Journal of Computer Vision. – 2001. – Vol. 45, Issue 2. – P. 83-105.
15. **Kootstra, G.** Paying attention to symmetry / G. Kootstra, A. Nederveen, B. de Boer // British Machine Vision Conference. – 2008. – P. 1115-1125.
16. **Parkhurst, D.** Modeling the role of salience in the allocation of overt visual attention / D. Parkhurst, K. Law, E. Niebur // Vision Research. – 2002. – Vol. 42, Issue 1. – P. 107-123.
17. **Пластинин, А.И.** Обнаружение текстурных неоднородностей на микромасштабных изображениях материалов // А.И. Пластинин, А.Г. Храмов, В.А. Соيفер / Компьютерная оптика. – 2011. – Т. 35, № 2. – С. 158-165.
18. **Визильтер, Ю.В.** Поиск объектов на изображении с использованием морфлетных описаний / Ю.В. Визильтер, В.С. Горбачевич, Б.В. Вишняков, С.В. Сидякин // Компьютерная оптика. – 2017. – Т. 41, № 3. – С. 406-411. – DOI: 10.18287/2412-6179-2017-41-3-406-411.
19. **Goferman, S.** Context-aware saliency detection / S. Goferman, L. Zelnik-Manor, A. Tal // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2012. – Vol. 34, Issue 10. – P. 1915-1926.
20. **Erdem, E.** Visual saliency estimation by nonlinearly integrating features using region covariances / E. Erdem, A. Erdem // Journal of Vision. – 2013. – Vol. 13, Issue 4. – 11.
21. **Li, X.** Saliency detection via dense and sparse reconstruction / X. Li, H. Lu, L. Zhang, X. Ruan, M.-H. Yang // IEEE International Conference on Computer Vision. – 2013. – P. 2976-2983.
22. **Tavakoli, H.R.** Fast and efficient saliency detection using sparse sampling and kernel density estimation / H.R. Tavakoli, E. Rahtu, J. Heikkila // Scandinavian Conference on Image Analysis. – 2011. – P. 666-675.
23. **Yang, C.** Graph-regularized saliency detection with convex-hull-based center prior / C. Yang, L. Zhang, H. Lu // IEEE Signal Processing Letters. – 2013. – Vol. 20, Issue 7. – P. 637-640.
24. **Jiang, B.** Saliency detection via absorbing Markov chain / B. Jiang, L. Zhang, H. Lu, C. Yang, M.-H. Yang // IEEE International Conference on Computer Vision. – 2013. – P. 1665-1672.
25. **Margolin, R.** What makes a patch distinct? / R. Margolin, A. Tal, L. Zelnik-Manor // IEEE Conference on Computer Vision and Pattern Recognition. – 2013. – P. 1139-1146.
26. **Rahtu, E.** Segmenting salient objects from images and videos / E. Rahtu, J. Kannala, M. Salo, J. Heikkila // European Conference on Computer Vision. – 2010. – P. 366-379.
27. **Seo, H.J.** Static and space-time visual saliency detection by self-resemblance / H.J. Seo, P. Milanfar // Journal of Vision. – 2009. – Vol. 9, Issue 12. – 15.
28. **Murray, N.** Saliency estimation using a non-parametric low-level vision model / N. Murray, M. Vanrell, X. Otazu, C.A. Parraga // IEEE Conference on Computer Vision and Pattern Recognition. – 2011. – P. 433-440.
29. **Hou, X.** Saliency detection: A spectral residual approach / X. Hou, L. Zhang // IEEE Conference on Computer Vision and Pattern Recognition. – 2007. – P. 1-8.
30. **Zhang, L.** Sun: A Bayesian framework for saliency using natural statistics / L. Zhang, M.H. Tong, T.K. Marks, H. Shan, G.W. Cottrell // Journal of vision. – 2008. – Vol. 8, Issue 7. – 32.
31. **Duan, L.** Visual saliency detection by spatially weighted dissimilarity / L. Duan, C. Wu, J. Miao, L. Qing, Y. Fu // IEEE Conference on Computer Vision and Pattern Recognition. – 2011. – P. 473-480.
32. **Tsotsos, J.K.** Modeling visual attention via selective tuning / J.K. Tsotsos, S. Culhane, Y. Winky, L. Yuzhong, N. Davis, F. Nuflo // Artificial Intelligence. – 1995. – Vol. 78. – P. 507-545.
33. **Zhao, R.** Saliency detection by multi-context deep learning / R. Zhao, W. Ouyang, H. Li, X. Wang // IEEE Conference on Computer Vision and Pattern Recognition. – 2015. – P. 1265-1274.
34. **Almeida, A.F.** Deep networks for human visual attention: a hybrid model using foveal vision / A.F. Almeida, R. Figueiredo, A. Bernardino, J. Santos-Victor // ROBOT 2017: 3rd Iberian Robotics Conference. – 2017. – P. 117-128.
35. **Wang, W.** Deep visual attention prediction / W. Wang, J. Shen // IEEE Transactions on Image Processing. – 2018. – Vol. 27, Issue 5. – P. 2368-2378.
36. **Zhou, D.** Ranking on data manifolds / D. Zhou, J. Weston, A. Gretton, O. Bousquet, B. Scholkopf // NIPS'03: Proceedings of the 16<sup>th</sup> International Conference on Neural Information Processing Systems. – 2004. – P. 169-176.
37. **Achanta, R.** SLIC superpixels compared to state-of-the-art superpixel methods / R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2012. – Vol. 34, Issue 11. – P. 2274-2282.
38. **Yang, C.** Saliency detection via graph-based manifold ranking / C. Yang, L. Zhang, H. Lu, X. Ruan, M. Yang // 2013 IEEE Conference on Computer Vision and Pattern Recognition. – 2013. – P. 3166-3173.
39. **Andrysiak, T.** Image retrieval based on hierarchical Gabor filters / T. Andrysiak, M. Choras // International Journal of Applied Mathematics and Computer Science. – 2005. – Vol. 15, Issue 4. – P. 471-480.
40. **Randen, T.** Filtering for texture classification: A comparative study / T. Randen, J.H. Husoy // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1999. – Vol. 21, Issue 4. – P. 291-310.
41. **Borji, A.** Salient object detection: A benchmark / A. Borji, M.-M. Cheng, H. Jiang, J. Li // IEEE transactions on image processing. – 2015. – Vol. 24(12). – P. 5706-5723.

#### Сведения об авторах

**Захаров Алексей Александрович**, 1978 года рождения, к.т.н., доцент, ведущий научный сотрудник кафедры программной инженерии МИ ВлГУ. Область научных интересов: компьютерное зрение, САПР. E-mail: [aa-zaharov@ya.ru](mailto:aa-zaharov@ya.ru).

**Титов Дмитрий Витальевич**, 1986 года рождения, д.т.н., директор Института международного образования Юго-Западного государственного университета. Область научных интересов: обработка изображений, компьютерное зрение. E-mail: [umsswsu@gmail.com](mailto:umsswsu@gmail.com).

**Жизняков Аркадий Львович**, 1975 года рождения, д.т.н., профессор, заведующий кафедрой программной инженерии МИ ВлГУ. Область научных интересов: компьютерное зрение, цифровая обработка изображений, САПР. E-mail: [lvovich1975@mail.ru](mailto:lvovich1975@mail.ru).

**Титов Виталий Семёнович**, 1943 года рождения, д.т.н., профессор, заведующий кафедрой вычислительной техники Юго-Западного государственного университета. Область научных интересов: теоретические и методологические основы построения адаптивных оптико-электронных систем, применяемых при автоматизации технологических процессов и производств различного назначения. E-mail: [tas\\_06@mail.ru](mailto:tas_06@mail.ru).

---

*ГРНТИ: 28.23.15*

*Поступила в редакцию 22 октября 2019 г. Окончательный вариант – 25 декабря 2019 г.*

---

---

# Visual attention method based on vertex ranking of graphs by heterogeneous image attributes

A.A. Zakharov<sup>1</sup>, D.V. Titov<sup>2</sup>, A.L. Zhiznyakov<sup>1</sup>, V.S. Titov<sup>2</sup>

<sup>1</sup> Murom Institute (branch), Vladimir State University named after Alexander and Nikolay Stoletovs, Murom, Russia,

<sup>2</sup> Southwest State University, Kursk, Russia

## Abstract

The paper discusses a method of visual attention based on vertex ranking of graphs on the basis of image features. The aim of the research is to develop a method that allows high-precision detection of objects in images with low color contrast between the selected and background areas. The image is pre-segmented into regions to calculate the saliency map. The graph is based on regions. Each region is associated with related regions, as well as with areas adjacent to adjacent regions. The regions are vertices of the graph. The vertices of the graph are ranked according to the characteristics of the corresponding image areas. The scope is highlighted based on requests from background areas. The saliency map is determined based on background area queries. Regions adjacent to the edges of the image belong to the background areas. Color features of the image were used in the existing approach of visual attention based on the manifold ranking. Texture features and shape features are additionally used in the proposed method to improve accuracy. Gabor's energy function is used to calculate texture features. The distance between centers of the regions is calculated by analyzing the form. The proposed method has shown good results for detecting objects in images in which the background color and object color are in similar ranges. The experimental results are presented on test images. Precision-recall curves showing the advantage of the developed method are constructed.

**Keywords:** image analysis, visual attention, graph, image attributes, ranking, computer vision.

**Citation:** Zakharov AA, Titov DV, Zhiznyakov AL, Titov VS. Visual attention method based on vertex ranking of graphs by heterogeneous image attributes. *Computer Optics* 2020; 44(3): 427-435. DOI: 10.18287/2412-6179-CO-658.

**Acknowledgements:** This work was financially supported by the Ministry of Science and Higher Education of the Russian Federation (State task of VISU GB-1187/20).

## References

- [1] Koch V, McLean J, Segev R, Freed MA, Berry MJ, Balasubramanian V, Sterling P. How much the eye tells the brain. *Current BIOLOGY* 2006; 16(14): 1428-1434.
  - [2] Borji A, Itti L. State-of-the-art in visual attention modeling. *IEEE Trans Pattern Anal Machine Intell* 2013; 35(1): 185-207.
  - [3] Begum M, Karray F. Visual attention for robotic cognition: A survey. *IEEE Trans Auton Ment Dev* 2011; 3(1): 92-105.
  - [4] Mahdi A, Su M, Schlesinger M, Qin J. A comparison study of saliency models for fixation prediction on infants and adults. *IEEE T Cogn Dev Syst* 2018; 10(3): 485-498.
  - [5] Garg A, Negi A. A survey on visual saliency detection and computational methods. *Int J Eng Technol* 2017; 9(4): 2742-2753.
  - [6] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Machine Intell* 1998; 20(11): 1254-1259.
  - [7] Frintrop S. VOCUS: a visual attention system for object detection and goal-directed search. Heidelberg, Germany: Springer-Verlag; 2006.
  - [8] Itti L, Dhavale N, Pighin F. Realistic avatar eye and head animation using a neurobiological model of visual attention. *Proc SPIE* 2003; 5200: 64-78.
  - [9] Wang J, Da Silva MP, Le Callet P, Ricordel V. Computational model of stereoscopic 3D visual saliency. *IEEE Trans Image Process* 2013; 22(6): 2151-2165.
  - [10] Harel J, Koch C, Perona P. Graph-based visual saliency. *Neural Inform Process Syst* 2006; 19: 545-552.
  - [11] Salvucci DD. An integrated model of eye movements and visual encoding. *Cogn Syst Res* 2001; 1: 201-220.
  - [12] Tatler BW. The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor bases and image feature distributions. *J Vis* 2007; 14: 1-17.
  - [13] Vijayakumar S. Overt visual attention for a humanoid robot. *Proceedings IEEE/RSJ international conference on intelligent robots and systems* 2001.
  - [14] Kadir T, Brady M. Saliency, scale and image description. *Int J Comp Vis* 2001; 45(2): 83-105.
  - [15] Kootstra G, Nederveen A, de Boer B. Paying attention to symmetry. *British Machine Vis Conf* 2008: 1115-1125.
  - [16] Parkhurst D, Law K, Niebur E. Modeling the role of saliency in the allocation of overt visual attention. *Vis Res* 2002; 42(1): 107-123.
  - [17] Plastinin AI, Khramov AG, Soifer VA. Texture defects detection on microscale images of materials. *Computer Optics* 2011; 35(2): 158-165.
  - [18] Vizilter YV, Gorbatshevich VS, Vishnyakov BV, Sidyakin SV. Object detection in images using morphlet descriptions. *Computer Optics* 2017; 41(3): 406-411. DOI: 10.18287/2412-6179-2017-41-3-406-411.
  - [19] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection. *IEEE Trans Pattern Anal Machine Intell* 2012; 34(10): 1915-1926.
  - [20] Erdem E, Erdem A. Visual saliency estimation by nonlinearly integrating features using region covariances. *J Vis* 2013; 13(4): 11.
-

- 
- [21] Li X, Lu H, Zhang L, Ruan X, Yang M-H. Saliency detection via dense and sparse reconstruction. *IEEE Int Conf Comp Vis* 2013: 2976-2983.
- [22] Tavakoli HR, Rahtu E., Heikkila J. Fast and efficient saliency detection using sparse sampling and kernel density estimation. *Scandinavian Conf Image Anal* 2011: 666-675.
- [23] Yang C, Zhang L, Lu H. Graph-regularized saliency detection with convex-hull-based center prior. *IEEE Sign Process Lett* 2013; 20(7): 637-640.
- [24] Jiang B, Zhang L, Lu H, Yang C, Yang MH. Saliency detection via absorbing Markov chain. *IEEE Int Conf Comp Vis* 2013: 1665-1672.
- [25] Margolin R, Tal A, Zelnik-Manor L. What makes a patch distinct? *IEEE Conf Comp Vis Pattern Recog* 2013: 1139-1146.
- [26] Rahtu E, Kannala J, Salo M, Heikkila J. Segmenting salient objects from images and videos. *European Conf Comp Vis* 2010: 366-379.
- [27] Seo HJ, Milanfar P. Static and space-time visual saliency detection by self-resemblance. *J Vis* 2009; 9(12): 15.
- [28] Murray N, Vanrell M, Otazu X, Parraga CA. Saliency estimation using a non-parametric low-level vision model. *IEEE Conf Comp Vis Pattern Recog* 2011: 433-440.
- [29] Hou X, Zhang L. Saliency detection: a spectral residual approach. *IEEE Conf Comp Vis Pattern Recog* 2007: 1-8.
- [30] Zhang L, Tong MH, Marks TK, Shan H, Cottrell GW. Sun: A Bayesian framework for saliency using natural statistics. *J Vis* 2008; 8(7): 32.
- [31] Duan L, Wu C, Miao J, Qing L, Fu Y. Visual saliency detection by spatially weighted dissimilarity. *IEEE Conf Comp Vis Pattern Recog* 2011: 473-480.
- [32] Tsotsos JK, Culhane S, Winky Y, Yuzhong L, Davis N, Nuflo F. Modeling visual attention via selective tuning. *Artificial Intelligence* 1995; 78: 507-545.
- [33] Zhao R, Ouyang W, Li H, Wang X. Saliency detection by multi-context deep learning. *IEEE Conf Comp Vis Pattern Recog* 2015: 1265-1274.
- [34] Almeida AF, Figueiredo R, Bernardino A, Santos-Victor J. Deep networks for human visual attention: a hybrid model using foveal vision. *ROBOT 2017: 3rd Iberian Robotics Conference* 2017; 117-128.
- [35] Wang W, Shen J. Deep visual attention prediction. *IEEE Trans Image Process* 2018; 27(5): 2368-2378.
- [36] Zhou D, Weston J, Gretton A, Bousquet O, Scholkopf B. Ranking on data manifolds. *NIPS'03* 2004: 169-176.
- [37] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Machine Intell* 2012; 34(11): 2274-2282.
- [38] Yang C, Zhang L, Lu H, Ruan X, Yang M. Saliency detection via graph-based manifold ranking. *IEEE Conf Comp Vis Pattern Recog* 2013: 3166-3173.
- [39] Andrysiak T, Choras M. Image retrieval based on hierarchical Gabor filters. *Int J Appl Math Comp Sci* 2005; 15(4): 471-480.
- [40] Randen T, Husoy JH. Filtering for texture classification: A comparative study. *IEEE Trans Pattern Anal Machine Intell* 1999; 21(4): 291-310.
- [41] Borji A, Cheng M-M, Jiang H, Li J. Salient object detection: A benchmark. *IEEE Trans Image Process* 2015; 24(12): 5706-5723.
- 

#### *Authors' information*

**Alexei Alexandrovich Zakharov** (b. 1978), PhD, assistant professor, leading researcher of Software Engineering department of MI VISU. Research interests: computer vision, CAD. E-mail: [aa-zaharov@ya.ru](mailto:aa-zaharov@ya.ru).

**Dmitry Vitalievich Titov** (b. 1986), doctor of engineering, director of the Institute of International Education of SWSU. Research interests: image processing, computer vision. E-mail: [umsswsu@gmail.com](mailto:umsswsu@gmail.com).

**Arkady Lvovich Zhiznyakov** (b. 1975), doctor of engineering, professor, head of Software Engineering department of MI VISU. Research interests: computer vision, image processing, CAD. E-mail: [lvovich1975@mail.ru](mailto:lvovich1975@mail.ru).

**Vitaliy Semenovich Titov** (b. 1943), doctor of engineering, professor, head of the Computer Technology department of SWSU. Research interests: theoretical and methodological foundations of the construction of adaptive optoelectronic systems used in the automation of technological processes and industries for various purposes. E-mail: [tas\\_06@mail.ru](mailto:tas_06@mail.ru).

---

*Received October 22, 2019. The final version – December 25, 2019.*

---