

Интернет-журнал «Наукоедение» ISSN 2223-5167 <http://naukovedenie.ru/>

Том 8, №4 (2016) <http://naukovedenie.ru/index.php?p=vol8-4>

URL статьи: <http://naukovedenie.ru/PDF/30TVN416.pdf>

Статья опубликована 29.07.2016.

**Ссылка для цитирования этой статьи:**

Ревякин А.М., Скурнович А.В. Подходы к разработке системы распознавания для решения задачи определения контента цифровых изображений // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 8, №4 (2016) <http://naukovedenie.ru/PDF/30TVN416.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ.

**УДК 004.932.2**

**Ревякин Андрей Михайлович**

ГКОУ ВПО «Академия Федеральной службы охраны Российской Федерации», Россия, Орел<sup>1</sup>  
Сотрудник  
E-mail: [dronrev@mail.ru](mailto:dronrev@mail.ru)

**Скурнович Алексей Валентинович**

ГКОУ ВПО «Академия Федеральной службы охраны Российской Федерации», Россия, Орел  
Сотрудник  
Кандидат технических наук  
E-mail: [alexeymail2010@mail.ru](mailto:alexeymail2010@mail.ru)

## **Подходы к разработке системы распознавания для решения задачи определения контента цифровых изображений**

**Аннотация.** В настоящей статье приводится описание множества методов распознавания образов на цифровых изображениях (контента изображения). Обосновывается актуальность разработки систем распознавания контента изображений, так как это находит применение в большинстве интеллектуальных систем, а также суть задачи распознавания. Выделены цели и задачи разработки систем распознавания контента изображений. Представлена подробная и обобщенная классификация методов обработки цифровых изображений с выделением наиболее эффективных.

В частности, в статье приведен разбор основных методов сегментации изображений. Раскрыто содержание наиболее эффективных и современных методов. Проведен анализ подходов к построению систем распознавания, формированию признакового пространства при проектировании сложных автоматизированных систем распознавания с обучением.

На основе проведенных исследований авторы делают вывод, что обобщенная в статье информация может служить теоретической основой для разработки системы распознавания при решении задачи определения контента цифровых изображений. Наиболее эффективным современным методом сегментации текста на изображении можно выделить метод, основанный на применении диаграмм Вороного, на который можно опираться в процессе разработки и создания системы понимания изображений, но и он, в свою очередь, требует адаптации к современным условиям ввиду ограниченного ресурса времени на решения данного типа информационных задач.

---

<sup>1</sup> 302034, Российская Федерация, город Орел, улица Приборостроительная, дом 35

**Ключевые слова:** компьютерное зрение; контент изображения; системы распознавания образов; сегментация изображения; морфологические методы анализа изображений; методы обработки цифровых изображений

## 1. Вариант классификации методов обработки изображений

Новейшие технические разработки в области компьютерных и информационных технологий, в том числе позволяют достаточно хорошо видеть реальный мир, распознавая объекты. Технология построения автоматизированных систем, которые могут производить обнаружение, отслеживание и классификацию объектов носит название «компьютерное зрение». Такие системы также можно назвать «автоматизированными системами обработки изображений и распознавания объектов» (АСОИРО)<sup>2</sup>. Эти искусственные системы получают информацию из изображений. Они могут делать заключения о типах объектов, об их взаимоотношениях, классифицировать по принадлежности к определенному классу, например, определить наличие текста и распознать его и т.д. Суть задачи распознавания – определить, обладают ли исследуемые объекты фиксированным конечным набором признаков, который позволяет отнести эти объекты к определенному классу.

АСОИРО технически постоянно совершенствуются за счет внедрения инновационных технологий в производство комплектующих для них, тем самым повышая их производительность на аппаратном уровне. Но на программном уровне также требуется оптимизация процессов обработки информации, в том числе цифровых изображений, в том числе для оперативного анализа контента изображения<sup>3</sup>.

Распознавание изображений представляет собой задачу преобразования входной информации в графическом виде в определенных форматах, в качестве которой рассматриваются некоторые параметры и признаки распознаваемых образов в выходную, представляющую заключение о том, к какому классу относится то или иное изображение. За последнее десятилетие достигнуты определённые результаты в решении задач распознавания контента изображения. Такие задачи заключаются в классификации некоторой группы объектов на основе конкретных требований. Объекты, относимые к одному классу, обладают общими свойствами. В частности, в рамках таких задач реализовано множество методов распознавания контента изображений, например, лиц, номерных знаков, различных предметов и т.д. Одна из основных трудностей в создании хороших программных продуктов для распознавания объектов на изображениях состоит в определении того, какие пиксели распознавать, а какие игнорировать. Для решения этой проблемы необходимо выделить необходимые данные при помощи их компактного представления. Получение такого представления называется *сегментацией* [1], то есть разбиение изображения на непохожие по некоторому признаку, но однородные по некоторому критерию области. Результатом сегментации является множество сегментов, которые покрывают всё изображение. Другими словами, каждый пиксель отмечен определенной меткой определенного класса. Сегментация изображений находит широкое применение во многих отраслях науки и техники, в том числе для анализа текста на изображении, например, для получения предварительной разметки рукописных и машинописных документов. Методы сегментации можно разделить на два класса: *автоматические* – не требующие взаимодействия с пользователем и *интерактивные* – использующие пользовательский ввод непосредственно в процессе работы.

---

<sup>2</sup> Сформулировано авторами.

<sup>3</sup> Информационно-значимое наполнение графического представления данных (сформулировано авторами).

При детальном анализе методов обработки цифровых изображений сталкиваемся с достаточно сложной их структурой, которая обобщена в виде блок-схемы (рисунок 1).



**Рисунок 1. Методы обработки цифровых изображений**  
 (разработано авторами на основе данных из источников [2-9])

Стандартные методы анализа изображений состоят в применении к ним преобразований, известных в математике или специально разработанных для тех или иных задач. Результатом этих операций является некоторая система признаков, характерных для заданного класса изображений; далее методами теории распознавания образов решаются многие практические задачи. Однако, как выбор признаков, так и способы оценки их значений – достаточно трудоемкие задачи. Методы обработки изображений могут существенно различаться в зависимости от того, каким путем изображение было получено – синтезировано системой машинной графики, либо, например, путем оцифровки черно-белых или цветных изображений. В последнем случае, например, возникает необходимость шумоподавления.

Для того чтобы решить проблему распознавания контента изображений используются, в том числе, методы морфологического анализа, которые способствуют созданию адекватного математического описания изображений, передающего их содержание и смысл. Впервые этот подход был предложен профессором МГУ Пытьевым Ю.П. в 70-х годах прошлого века. Под морфологией понимают внешнее строение объекта, а также форму, размеры и взаимное расположение образующих его структурных элементов на поверхности и в объеме. А под морфологическим анализом изображений принято понимать методы решения задач узнавания, классификации объектов, выделения отличий в сценах по их изображениям, оценивания параметров объекта по его изображению, основанные на исследовании математических моделей изображений. Другими словами, морфологические методы анализа изображений основаны на математических моделях, связывающих изображения с объектами изображаемой сцены и условиями их регистрации. Центральным понятием морфологических методов анализа изображений служит понятие формы изображения, понимаемой как часть информации, сохраняющаяся при вариациях условий формирования изображения [3]. Задача построения формы изображения (кусочно-гладкого; размытого; контурного; ламбертова объекта, произвольного объекта, освещаемого конечным числом источников света произвольной интенсивности) является существенной частью морфологического анализа. От

того, насколько качественно построена форма, существенно зависит и результат решения задачи морфологического анализа.

Задачи морфологического анализа изображений решаются путем узнавания объекта по форме его изображения, классификации объектов по форме их изображений, выделения отличий по форме и оценивания объекта по форме его изображения. Морфологические методы анализа изображений нашли широкое применение на практике:

- 1) морфологическое подавление случайного шума;
- 2) выделение неизвестного объекта на фоне, форма изображения которого известна;
- 3) аппроксимация формы текстурнозначного изображения;
- 4) морфологический метод сжатия изображения текста;
- 5) фильтрация гладкого фона;
- 6) поиск области интерлейсинга;
- 7) изменение микрорельефа поверхности по набору его изображений;
- 8) классификация изображений с оценкой параметров системы регистрации;
- 9) цветовая сегментация на основе морфологического фильтра;
- 10) поиск отличий по форме в цветных изображениях.

Один из способов построения формы состоит в задании областей постоянной яркости по физическим свойствам объекта, т.е. по расположению однородно светящихся или отражающих граней, или границ относительно наблюдателя. Приписывая этим областям всевозможные яркости, получим форму изображения как множество изображений. Если же мы не имеем столь подробных сведений об объекте исследования, то мы можем построить форму по какому-либо одному изображению, зная, к каким преобразованиям яркости данного изображения могут привести изменившиеся условия наблюдения.

## 2. Обзор методов сегментации

Если говорить об эффективности различных методов и алгоритмов цифровой обработки изображений, выполняемой с целью распознавания объектов и оценки их параметров, то она в значительной степени зависит от результатов выполнения процедуры *сегментации* объектов в регистрируемых изображениях. Цель сегментации состоит в упрощении или изменении представления изображения, чтобы его было легче анализировать в дальнейшем. Результатом сегментации является множество сегментов, которые покрывают все изображение. Иначе говоря, каждый пиксель отмечен некоторой меткой некоторого класса. Поскольку сегментация обычно используется не самостоятельно, а как часть некоторой системы (например, системы машинного зрения), то с практической точки зрения, качество работы метода оценивается исходя из работы системы в целом. Поэтому один и тот же метод сегментации может оказаться хорошим для одной задачи и плохим для другой.

Разные методы сегментации ориентированы на разные свойства разбиения (однородность регионов (однородность цвета или текстуры), непохожесть соседних регионов, гладкость границы региона, маленькое количество мелких «дырок» внутри региона и т.д.). Поэтому при выборе метода сегментации для решения конкретной задачи, следует определиться, какие свойства разбиения действительно важны. Качество работы метода оценивается в зависимости от того, насколько полученная сегментация обладает этими свойствами. Методы сегментации делятся на два больших класса: автоматические и интерактивные. Автоматические методы сегментации изображений не очень хорошо

выделяют нужные объекты. Интерактивные методы сегментации позволяют достаточно точно выделить интересующие объекты. Однако они требуют участия человека при обработке каждого изображения.

Интерактивная сегментация изображений активно используется для редактирования изображений, а также является составной частью многих алгоритмов компьютерного зрения. Настоящий прорыв в разработке методов сегментации изображений произошел в 2000 г., когда Юрий Бойков и Мари-Пьер Жолли разработали алгоритм интерактивной сегментации ГРАФКАТ, который фактически стал эталонным. Большая часть новых алгоритмов интерактивной сегментации изображений является развитием ГРАФКАТ. Остальные алгоритмы, сравниваются в первую очередь с ним. Разрезы графов, на которые опирается ГРАФКАТ, стали активно использоваться и в других областях компьютерного зрения: сегментации видео, стереорекострукции и пр.

Данный алгоритм трактует все изображение, как граф. Но в данном случае к вершинам, соответствующим пикселям изображения, добавляются две терминальные вершины, называемые истоком и стоком. Вершины графа, соответствующие семенам объекта и фона, связываются соответственно с истоком и стоком ребрами с бесконечно большим весом. Пользователю необходимо указать несколько пикселей, принадлежащих объекту (семена объекта), и несколько пикселей фона (семена фона). В полученном графе находится минимальный разрез, который делит граф на две части. Пиксели, попавшие в один подграф с истоком, считаются объектом, остальные пиксели признаются фоном. Бесконечный вес ребер между семенами обеспечивает выполнение заданных пользователем ограничений: семена объекта будут отнесены к объекту, семена фона – к фону. Чем больше отличаются цвета соседних пикселей, тем вес ребра между ними меньше, а значит больше вероятность того, что разрез графа пройдет между ними. Это стимулирует прохождение разреза графа по наиболее контрастной границе.

Одно из наиболее активно развивающихся направлений для методов автоматической сегментации изображений – использование методов теории графов. Общая идея методов этой группы следующая. Изображение представляется в виде взвешенного графа, с вершинами в точках изображения. Вес ребра графа отражает сходство точек. Разбиение изображения моделируется разрезами графа. Как правило, в данных методах для отражения качества полученной сегментации вводят функционал «стоимости» разреза. Так задача разбиения изображения на однородные области сводится к оптимизационной задаче поиска разреза минимальной стоимости на графе. Этот подход позволяет помимо однородности цвета и текстуры сегментов управлять также формой сегментов, их размером, сложностью границ и т.п.

Процедура сегментации описывается, в том числе, методами выделения границ. Границей или контуром на изображении называется совокупность его пикселей, в окрестности которых наблюдается скачкообразное изменение яркости. Так как при цифровой обработке изображение представлено как функция целочисленных аргументов, то контуры представляются линиями шириной, как минимум, в один пиксель.

Методы поиска границ хорошо разработаны для полутоновых изображений. Полутоновое изображение рассматривается как функция двух переменных, и предполагается, что границы регионов соответствуют максимумам градиента этой функции. Для их поиска применяется аппарат дифференциальной геометрии (в простейшем случае это фильтры Робертса, Собела, Кирша, Уоллеса) [26].

Одним из эффективных методов автоматической сегментации является метод СЮСАН. Основная идея СЮСАН в том, что соседи каждой точки в однородной области имеют близкую к ней яркость, а вблизи границы число соседей с одинаковой яркостью уменьшается.

Кроме того, этот метод обнаруживает и другие особенности на изображении, такие как углы, тонкие линии т.п. Вокруг каждого пикселя изображения строится маска, центральный пиксель которой называется ядром (используется круглая маска с радиусом 3.4 пикселя, которая включает 37 пикселей или традиционная квадратная маска 3x3). Пиксели в пределах маски, имеющие сравнимую с ядром яркость, образуют область ЮСАН (Univalue Segment Assimilating Nucleus – однородный сегмент, ассимилируемый ядром). Для обнаружения двумерных особенностей и границ используют размер, центр тяжести и вторые моменты ЮСАН. Такой подход обнаружения особенностей отличается отсутствием применения производных изображения, а значит, нет необходимости в предварительном подавлении шума. Площадь ЮСАН максимальна, когда ядро находится в однородной (или почти однородной) области изображения, она уменьшается до половины этого максимума вблизи прямой границы и уменьшается еще больше вблизи угла и достигает локальных минимумов в углах и точно на границе. Это свойство площади ЮСАН используется как главный критерий присутствия границ и двумерных особенностей.

Еще один подкласс методов автоматической сегментации – это *методы семантической сегментации*. Анализ семантики является вершиной иерархической процедуры обработки изображений. В основании пирамиды лежат методы формирования первичной системы информативных признаков. На следующем уровне иерархии формируется система вторичных признаков с примерно одинаковым уровнем значимости. На верхнем уровне пирамиды параметрическими или непараметрическими методами решается задача семантической классификации образов. Существует большое разнообразие способов и стратегий решения задач семантической классификации изображения при реализации этапов известными методами, сохраняя при этом типичную схему. Первые работы, посвященные проблеме распознавания семантики изображения (в зарубежной литературе используется аббревиатура CBIR – Content based image retrieval), берут свое начало в 80-х годах прошлого века. Однако наиболее существенное развитие данное направление получило в последнее десятилетие. Преобладающее число исследований посвящены развитию стратегий последовательного уточнения запроса и оптимизаций поисковых процедур для изображений в конкретных базах данных большого объема. В контексте семантической классификации весьма значимыми являются алгоритмы адаптивной сегментации изображений, алгоритмы нечеткой метрической классификации сегментов изображения, алгоритмы информативной оценки системы первичных признаков и формирования сложных вторичных признаков, алгоритмы нечеткой иерархической классификации изображений по результатам сегментной классификации [27].

Из представленных на рисунке 1 методов сегментации, авторы статьи [6] считают наиболее эффективным современным методом сегментации текста на изображении – метод, основанный на применении диаграмм Вороного.

Данный метод сегментации также заключается в разбиении изображения на области объектов, области фонов, границы между объектами и фоном или другие информативные фрагменты. Для данных составных частей изображения выполняется некоторый критерий однородности. Результатом сегментации является множество сегментов, которые покрывают все изображение. Иначе говоря, каждый пиксель отмечается некоторой меткой некоторого класса.

Математически задачу сегментации описывают следующим образом. Пусть  $f(x, y)$  – функция распределения освещенности в изображении, определенная на замкнутом множестве  $D: D = \{D_1, D_2, \dots, D_k\}$  – разбиение  $D$  на  $k$  непустых связных подмножеств  $D_i = (i = 1, 2, \dots, k)$ .  $LP$  – предикат, определенный на множестве  $D$ , принимающий значение *True* – истина, тогда и только тогда, когда любая пара точек из каждого подмножества  $D_i$  удовлетворяет некоторому критерию однородности.

Сегментацией изображения  $f(x, y)$  по предикату  $LP$  называется разбиение, удовлетворяющее условиям [2, 8, 9, 11]:

- 1)  $\bigcup_{i=1}^k D_i^* = D$  ;
- 2)  $D_i^* \cap D_j^* = \emptyset$  для любых  $i \neq j$  ;
- 3)  $LP(D_i^*) = True$  для любого  $i$  ;
- 4)  $LP(D_i^* \cap D_j^*) = False$  для любых  $i \neq j$  .

Предикат  $LP$  называется предикатом однородности. В частном случае предикат  $LP$  определяется как:

$$LP(D_i^*) = \begin{cases} True - \text{если } f(x_1, y_1) = \dots = f(x_M, y_M) \\ false - \text{в противном случае} \end{cases},$$

где  $f(x_m, y_m) \in D_i^*$ ,  $m = 1, 2, \dots, M$ ;  $M$  – число точек в области, или

$$LP(D_i^*) = \begin{cases} True - \text{если } |f(x_m, y_m) - f(x_1, y_1)| \leq T \\ false - \text{в противном случае} \end{cases},$$

где:  $(x_m, y_m), (x_1, y_1)$  – произвольные точки из области  $D_i^*$ ,  $T$  – некоторый заданный порог.

Отсюда следует, что формально сегментация представляет собой оператор, задающий преобразование  $f(x, y) \rightarrow G(x, y)$ , где  $G(x, y) = L_i$  при  $(x, y) \in D_i^*$ ,  $i = 1, 2, \dots, k$ ,  $L_i$  – имя  $i$ -ой области [4, 10].

Как правило, алгоритмы сегментации изображений основываются на одном из двух базовых свойств сигнала яркости: *однородности* и *разрывности* [1, 8-17].

В системах распознавания рукописных и машинописных текстов сегментация изображений является весьма важным этапом в последовательности операций. Роль сегментации сводится к отысканию на плоскости изображения элементов, подвергаемых распознаванию, таких как строки символов, отдельные слова или символы, рисунки, таблицы и другие объекты, содержащиеся в тексте. Многие современные подходы к сегментации текста основаны на использовании метода диаграмм Вороного [17]. Выделение текстовых блоков в таких методах происходит в два этапа: первый заключается в построении диаграммы Вороного выделенной области. Второй этап использует полученную диаграмму как средство быстрого поиска так называемых «смежных» точек и производит постепенное укрупнение рассматриваемых блоков (т.е. выделение строк, затем их объединение в абзацы, параграфы и т.п.).

Наиболее эффективный алгоритм построения диаграммы Вороного на плоскости был предложил Форчун [18]. Подробное изложение деталей реализации и представления данных приводится в работе [21]. Для выделения текстовых блоков (т.е. абзацев, параграфов) и строк внутри этих блоков можно использовать обобщение диаграммы Вороного [22], когда каждый символ трактуется как объект, состоящий из множества связанных точек.

Для построения обобщённой диаграммы Вороного на основе точечных диаграмм предполагается использовать множества всех точек. Так, в случае изображения с высоким разрешением, построение диаграммы Вороного для множества всех точек является

достаточно затратным (например, при разрешении  $1\ 637 \times 1\ 481$  точек и 10% заполнении изображения чёрными точками потребуются построение порядка 200 000 ячеек диаграммы Вороного).

Учитывая вычислительную сложность построения диаграммы Вороного области, существует возможность получения новых, более эффективных алгоритмов сегментации текста, основанные на анализе взаимного расположения центров масс символов, например, алгоритм поиска «соседних» символов, принадлежащих одному слову [6]. Упомянутые методы являются новым эффективным способом выделения строк, слов и изолированных символов текста, который используется в системе распознавания рукописных и печатных текстов.

### 3. Подходы к разработке системы распознавания контента изображения

В целом, задача распознавания контента изображений может быть достоверно и легко решена человеком. Распознавание – это отнесение конкретной реализации, представленной значениями ее признаков, к одному из фиксированного перечня классов по определённому решающему правилу в соответствии с поставленной целью. То есть распознавание может осуществляться любой живой или неживой системой, выполняющей такие функции, как измерение значений признаков и производство вычислений, реализующих решающее правило. Как правило перечень информативных признаков и решающие правила либо задаются распознающей системе извне, либо формируются самой системой. Оптимальное решающее правило в свою очередь необходимо строить с помощью оценки риска потерь. Это позволяет выбрать наиболее информативную систему признаков, которые используются при распознавании, и др.

В общем случае при проектировании сложных автоматизированных систем распознавания с обучением выполняются следующие этапы:

1. **Постановка задачи распознавания.** На этом этапе осуществляется определение объектов распознавания, составление априорного алфавита классов, формулировка цели распознавания и задание критерия оценки эффективности системы распознавания.
2. **Формирование априорного словаря признаков.** Производится поиск всех существенных свойств объектов распознавания, имеющих численную интерпретацию и позволяющих достичь цели распознавания, а также производится описание объектов распознавания на основе априорного словаря признаков.
3. **Выбор математического аппарата для классификации.** Требуется определить класс системы распознавания в зависимости от типа признаков априорного словаря (признаки могут быть детерминированные, логические, структурные или вероятностные [28]), а также алгоритма распознавания в зависимости от полноты априорной информации.
4. **Нормирование признаков** априорного словаря производится с целью приведения значений признаков к определённому интервалу изменений и преобразовании их в безразмерные величины. Данный этап необходим при формировании систем распознавания, использующих детерминированные и вероятностные признаки.
5. **Формирование рабочего словаря признаков** из априорного словаря, с помощью которого классифицируемые объекты можно с высокой степенью достоверности разделить на заданное число классов.

6. **Формирование обучающей выборки** (обучающего набора) объектов всех распознаваемых классов. На данном этапе важно, чтобы количество объектов для обучения было достаточным для эффективного обучения системы распознавания.
7. **Обучение системы распознавания.** Системе распознавания предоставляются вектора признаков объектов из обучающей выборки и информация о принадлежности каждого объекта к определенному классу. В результате обучения система распознавания формирует разделяющее правило по заданному критерию, например минимизирующего ошибку классификации.
8. **Оценка эффективности** системы распознавания. Осуществляется формирование экзаменационной выборки и оценка результатов ее распознавания по заданному правилу.

В решении задачи определения полного перечня признаков – главное найти все признаки, характеризующие существо распознаваемых объектов (явлений). Любые ограничения, любая неполнота могут привести к ошибкам или полной невозможности правильной классификации объектов (явлений).

Реально даже целая группа признаков может оказаться неэффективной. Поэтому для решения задачи создания системы распознавания необходимо найти все возможные признаки, описывающие объекты распознавания, с тем, чтобы при оценке эффективности решений системы не возвращаться к этой задаче, обнаружив ограниченность выбранных признаков на последующих этапах разработки.

При наличии двух или более числа классов цель выбора признаков состоит в выборе таких признаков, которые являются наиболее эффективными с точки зрения разделимости классов. Разделимость классов не зависит от системы координат, и этим критерии разделимости отличаются от критериев для одного распределения. Разделимость классов зависит не только от распределений объектов в классах, но также от используемого классификатора. Например, оптимальный набор признаков для классификатора с линейным решающим правилом может не быть оптимальным для других классификаторов при тех же распределениях. Для того чтобы избежать этого дополнительного усложнения, ищется оптимальный набор признаков для байесовского классификатора; это позволяет минимизировать ошибку классификации. Тогда разделимость классов будет эквивалентна вероятности ошибки байесовского классификатора, который является оптимальным с точки зрения минимизации риска принятия неверного решения.

Следовательно, с теоретической точки зрения вероятность ошибки является наилучшим критерием эффективности признаков. Кроме того, на практике одним из наиболее распространенных критериев является вероятность ошибки, полученная экспериментально; а именно, интуитивно выбрав набор признаков, строят байесовский классификатор и экспериментально подсчитывают число ошибок классификаций. Эта процедура является гибкой, не зависит от вида распределения и теоретически позволяет найти оптимальное решение.

Самый естественный и простой способ определения состоит в подсчете неверных классификаций при распознавании достаточно большого числа реализаций сигналов. Вычислением отношения числа ошибочных ответов к общему числу предъявленных системе распознавания реализаций сигналов можно составить представление о величине вероятности ошибочного распознавания:

$$P_{ош} = \frac{n_{ош}}{n_p}$$

где:  $n_{ou}$  – число ошибочных решений,  $n_p$  – общее число объектов предъявленных для распознавания.

Такой метод определения качества распознавания пригоден для любых систем. Однако, он требует значительных экономических затрат и очень большого числа актов распознавания, особенно в условиях малой величины вероятности ошибочного распознавания. Поэтому на практике стремятся использовать аналитические соотношения для определения качества распознавания.

Наряду с величиной вероятности ошибочного распознавания на практике широко используется и другой показатель – величина вероятности правильного распознавания:

$$P_{np} = 1 - P_{ou}$$

Обе эти величины характеризуют качество распознавания в среднем, т.е. по всей совокупности распознаваемых классов. В то же время, каждый из классов может быть распознан с различным качеством, определяемым соответствующей величиной вероятности ошибочного или правильного распознавания  $P_{ou.i}$  ( $P_{np.i}$ ). Между средней величиной ошибочного (правильного) распознавания и величинами вероятностей ошибочного (правильного) распознавания отдельных классов существует очевидная связь:

$$P_{ou} = \frac{1}{K} \sum_{i=1}^K P_{ou.i}, \quad P_{np} = \frac{1}{K} \sum_{i=1}^K P_{np.i}$$

где  $K$  – количество распознаваемых классов.

Помимо вероятностных оценок качества распознавания на практике часто используются и информационные показатели. Одним из таких показателей является количество информации, которое получает система в процессе распознавания в результате измерения значений признаков классов.

Из теории информации известно, что количество информации, которое будет получено о классах при распознавании, равно

$$I_n = H(K) - H\left(\frac{K}{G_n}\right)$$

где:  $H(K) = -P(x) \log P(x)$  – мера неопределенности (априорная энтропия) для

случайной величины  $x$ , появляющейся с вероятностью  $P(x)$ ,  $H\left(\frac{K}{G_n}\right)$  – мера неопределенности (апостериорная энтропия) для случайной величины  $x$ , после измерения значения соответствующего признака.

Очевидно, что качество распознавания определяется как степень полезности (информативностью) признаков, так и их количеством. Поэтому, отбирая признаки для распознавания, следует учитывать, в первую очередь, их влияние на качество распознавания. В принципе признаки могут выбираться произвольно с экспериментальной оценкой качества распознавания системы. Однако, такой подход к выбору признаков связан с большими экономическими затратами и практически нецелесообразен. Более привлекательным является подход, основанный на изучении аналитических зависимостей между качеством распознавания и информативностью признаков и их количеством.

**Выводы.** Несмотря на то, что успехи в области создания систем распознавания цифровых изображений впечатляющие, до сих пор такие системы не могут конкурировать с человеком при решении сложных задач. Задача распознавания образов является основной в большинстве интеллектуальных систем. Целью распознавания образов является классификация объектов (образов) по нескольким категориям или классам. Другими словами,

суть задачи распознавания – установить, обладают ли изучаемые объекты фиксированным конечным набором признаков, позволяющим отнести их к определенному классу. Если говорить об эффективности различных методов и алгоритмов цифровой обработки изображений, выполняемой с целью распознавания объектов и оценки их параметров, то она в значительной степени зависит от результатов выполнения процедуры сегментации контента в регистрируемых изображениях. Наиболее эффективным современным методом сегментации текста на изображении можно выделить метод, основанный на применении диаграмм Вороного, на что и следует опираться в процессе разработки и создания так называемой системы понимания изображений.

В области автоматического распознавания объектов на изображении существует ряд общих проблем, для решения которых прикладываются усилия ученых и разработчиков всего мира. Одна из главных – это то, что задачи данного направления необходимо решать в условиях ограниченного ресурса времени.

## ЛИТЕРАТУРА

1. Форсайт Д.А., Понс Ж. Компьютерное зрение. Современный подход: – М.: Издательский дом «Вильямс», 2004. – 928 с.: ил. – Парал. Тит. Англ.
2. Розенфельд А., Дейвис Л.С. Сегментация и модели изображений. ТИИЭР. 1979. Т.67, №5. С. 71-82.
3. Пытьев Ю.П., Чуликов А.И. Методы морфологического анализа изображений. –М.: ФИЗМАТЛИТ, 2010. – 336 с.
4. Системы технического зрения (принципиальные основы, аппаратное и математическое обеспечение) / А.Н. Писаревский [и др.]; Под общ. ред. А.Н. Писаревского, А.Ф. Чернявского. Л.: Машиностроение. (Ленингр. отделение), 1988. 424 с.: ил.
5. Фурман Я.А., Юрьев А.Н., Яншин В.В. Цифровые методы обработки и распознавания бинарных изображений. – Красноярск: Изд-во Краснояр. ун-та, 1992. – 248 с.
6. Запрягаев С.А., Сорокин А.И. Сегментация рукописных и машинописных текстов методом диаграмм Вороного // Вестник ВГУ, серия: Системный анализ и информационные технологии, 2010, №1, с. 160.
7. Фу К. Последовательные методы в распознавании образов и обучении машин: перев. с англ., изд-во «Наука», Главная редакция физико-математической литературы. – М., 1971. 256 с.
8. Путьтин Е.П., Аверин С.И. Обработка изображений в робототехнике. М.: Машиностроение, 1990. 320 с.
9. Ким Н.В. Анализ и обработка изображений в системах технического зрения: Учебное пособие. Москва: МАИ, 2001. 164 с.
10. Техническое зрение роботов / В.И. Мошкин [и др.]. М.: Машиностроение, 1990. 272 с.
11. Красильников Н.Н. Цифровая обработка 2D- и 3D-изображений: Учебное пособие. СПб.: БХВ - Петербург, 2011. 608 с.: ил.
12. Шапиро Л., Стокман Дж. Компьютерное зрение. М.: БИНОМ. Лаборатория знаний, 2006. 752 с.

13. Абламейко С.В., Лагуновский Д.М. Обработка изображений: технология, методы, применение: Учебное пособие. М.: Амалфей, 2000. 304 с.
14. Canny J. A. Computational Approach for Edge Detection // IEEE Trans. Pattern Anal. Machine Intel. 1986. Vol. 8, N. 6. P. 679-698.
15. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде МАТЛАВ. М.: Техносфера, 2006. 616 с.
16. Колючкин В.Я., Нгуен К.М., Чан Т.Х. Алгоритмы обработки изображения в системах машинного зрения роботизированных производственных линий // Нейрокомпьютеры: разработка, применение. – М.: Издательство «Радиотехника», 2014. №3. С. 44 - 51.
17. Kise K. Segmentation of page images using the area voronoi diagram / K. Kise, A. Sato, M. Iwata // Computer Vision and Image Understanding. –1998. – Vol. 3, no. 70. – P. 370–382.
18. Fortune S. A sweepline algorithm for Voronoi diagrams / S. Fortune // Proceedings of the second annual symposium on Computational geometry. – 1986. – P. 313 – 322.
19. Shafait F. Performance Comparison of Six Algorithms for Page Segmentation / F. Shafait, D. Keysers, T. Breuel // Image Understanding and Pattern Recognition (IUPR) research group. – 2006. – 12 pp.
20. Препарата Ф. Вычислительная геометрия / Ф. Препарата, М. Шеймос – М.: Мир, 1989. – 295 с.
21. Computational Geometry Algorithms and Applications / [Edited by M. Berg]. – 3rd Edition. –Berlin.: Springer-Verlag, 2008 – 386 pp.
22. Wang Z. Word Extraction Using Area Voronoi Diagram / Z. Wang, Y. Lu, C. Lim // CVPRW '03. – 2003. – P. 31 – 36.
23. Мухамедияров Р.М. Машинное зрение: понятия, задачи и области применения // [http://www.rusnauka.com/25\\_NPM\\_2009/Informatica/50975.doc.htm](http://www.rusnauka.com/25_NPM_2009/Informatica/50975.doc.htm).
24. Местецкий Л.М. Математические методы распознавания образов (курс лекций) / Московский государственный университет имени М.В. Ломоносова. – 2004 г.
25. Мазуров В.Д. Комитеты систем неравенств и задача распознавания / В.Д. Мазуров // Кибернетика, 2004, №2. С. 140-146.
26. Жук С.В. Обзор современных методов сегментации растровых изображений / С.В. Жук // – Известия ВолГТУ. – 2009, №6. С. 115-118.
27. Дорогов А.Ю. Быстродействующий алгоритм семантической классификации JPEG-изображений / А.Ю. Дорогов, Р.Г. Курбанов, В.В. Разин // Санкт-Петербургский государственный электротехнический университет (СПбГЭТУ) "ЛЭТИ", 2006.
28. Горелик А.Л., Скрипкин В.А. Методы распознавания: Учебное пособие для вузов. Изд. 4. – М: Букинист. –2004. 262 с.

**Revyakin Andrey Mikhailovich**

The academy of federal security guard service of the Russian Federation, Russia, Orel  
E-mail: [dronrev@mail.ru](mailto:dronrev@mail.ru)

**Skurnovich Aleksey Valentinovich**

The academy of federal security guard service of the Russian Federation, Russia, Orel  
E-mail: [alexeymail2010@mail.ru](mailto:alexeymail2010@mail.ru)

## **Approaches to the development of a recognition system to solve the problem of determining the content of digital images**

**Abstract.** This article describes a variety of methods of pattern recognition for digital images (content of the images). The grounding on development of image context recognition systems is provided as they are used in most intelligent systems. The essence of the problem of recognition is also represented. There is a detailed and generalized of methods of digital image processing that highlights the most effective once. There is also a analysis of approaches to the construction of recognition systems and a formation of feature space while developing of the complex automated systems with training.

In particular, the article presents the analysis of the main methods of image segmentation. The article reveals the content of the most effective and modern methods.

Basing on the conducted research the authors conclude that the information generalized in the article can serve as a theoretical basis for the development of recognition systems while solving the problem of determining the content of digital images. The most effective modern method of text segmentation on the image based on application of Voronoi diagrams. This method can be relied on while the development of the image understanding systems, but it requires adaptation to modern conditions, because of the limited resource of time to solve the such type of information problems.

**Keywords:** computer vision; image content; system pattern recognition; image segmentation; morphological image analysis methods; methods of digital image processing

### **REFERENCES**

1. Forsayt D.A., Pons Zh. Komp'yuternoe zrenie. Sovremennyy podkhod: – M.: Izdatel'skiy dom «Vil'yams», 2004. – 928 s.: il. – Paral. Tit. Angl.
2. Rozenfel'd A., Deyvis L.S. Segmentatsiya i modeli izobrazheniy. TIHER. 1979. T.67, №5. S. 71-82.
3. Pyt'ev Yu.P., Chulikov A.I. Metody morfologicheskogo analiza izobrazheniy. –M.: FIZMATLIT, 2010. – 336 s.
4. Sistemy tekhnicheskogo zreniya (printsipial'nye osnovy, apparatnoe i matematicheskoe obespechenie) / A.N. Pisarevkiy [i dr.]; Pod obshch. red. A.N. Pisarevskogo, A.F. Chernyavskogo. L.: Mashinostroenie. (Leningr. otdelenie), 1988. 424 s.: il.
5. Furman Ya.A., Yur'ev A.N., Yanshin V.V. Tsifrovye metody obrabotki i raspoznavaniya binarnykh izobrazheniy. – Krasnoyarsk: Izd-vo Krasnoyar. un-ta, 1992. – 248 s.
6. Zapryagaev S.A., Sorokin A.I. Segmentatsiya rukopisnykh i mashinopisnykh tekstov metodom diagramm Voronogo // Vestnik VGU, seriya: Sistemnyy analiz i informatsionnye tekhnologii, 2010, №1, s. 160.
7. Fu K. Posledovatel'nye metody v raspoznavanii obrazov i obuchenii mashin: perev. s angl., izd-vo «Nauka», Glavnaya redaktsiya fiziko-matematicheskoy literatury. – M., 1971. 256 s.

8. Putyatin E.P., Averin S.I. Obrabotka izobrazheniy v robototekhnike. M.: Mashinostroenie, 1990. 320 s.
9. Kim N.V. Analiz i obrabotka izobrazheniy v sistemakh tekhnicheskogo zreniya: Uchebnoe posobie. Moskva: MAI, 2001. 164 s.
10. Tekhnicheskoe zrenie robotov / V.I. Moshkin [i dr.]. M.: Mashinostroenie, 1990. 272 s.
11. Krasil'nikov N.N. Tsifrovaya obrabotka 2D- i 3D-izobrazheniy: Uchebnoe posobie. SPb.: BKhV - Peterburg, 2011. 608 s.: il.
12. Shapiro L., Stokman Dzh. Komp'yuternoe zrenie. M.: BINOM. Laboratoriya znaniy, 2006. 752 s.
13. Ablameyko S.V., Lagunovskiy D.M. Obrabotka izobrazheniy: tekhnologiya, metody, primeneniye: Uchebnoe posobie. M.: Amalfey, 2000. 304 s.
14. Canny J. A. Computational Approach for Edge Detection // IEEE Trans. Pattern Anal. Machine Intel. 1986. Vol. 8, N. 6. P. 679-698.
15. Gonsales R., Vuds R., Eddins S. Tsifrovaya obrabotka izobrazheniy v srede MATLAB. M.: Tekhnosfera, 2006. 616 s.
16. Kolyuchkin V.Ya., Nguen K.M., Chan T.Kh. Algoritmy obrabotki izobrazheniya v sistemakh mashinnogo zreniya robotizirovannykh proizvodstvennykh liniy // Neyrokom'pyutery: razrabotka, primeneniye. – M.: Izdatel'stvo «Radiotekhnika», 2014. №3. S. 44 - 51.
17. Kise K. Segmentation of page images using the area voronoi diagram / K. Kise, A. Sato, M. Iwata // Computer Vision and Image Understanding. –1998. – Vol. 3, no. 70. – P. 370–382.
18. Fortune S. A sweepline algorithm for Voronoi diagrams / S. Fortune // Proceedings of the second annual symposium on Computational geometry. – 1986. – P. 313 – 322.
19. Shafait F. Performance Comparison of Six Algorithms for Page Segmentation / F. Shafait, D. Keysers, T. Breuel // Image Understanding and Pattern Recognition (IUPR) research group. – 2006. – 12 pp.
20. Preparata F. Vychislitel'naya geometriya / F. Preparata, M. Sheymos – M.: Mir, 1989. – 295 s.
21. Computational Geometry Algorithms and Applications / [Edited by M. Berg]. – 3rd Edition. –Berlin.: Springer-Verlag, 2008 – 386 pp.
22. Wang Z. Word Extraction Using Area Voronoi Diagram / Z. Wang, Y. Lu, C. Lim // CVPRW '03. – 2003. – P. 31 – 36.
23. Mukhamediyarov R.M. Mashinnoe zrenie: ponyatiya, zadachi i oblasti primeneniya // [http://www.rusnauka.com/25\\_NPM\\_2009/Informatica/50975.doc.htm](http://www.rusnauka.com/25_NPM_2009/Informatica/50975.doc.htm).
24. Mestetskiy L.M. Matematicheskie metody raspoznavaniya obrazov (kurs lektsiy) / Moskovskiy gosudarstvennyy universitet imeni M.V. Lomonosova. – 2004 g.
25. Mazurov V.D. Komitety sistem neravenstv i zadacha raspoznavaniya / V.D. Mazurov // Kibernetika, 2004, №2. S. 140-146.
26. Zhuk S.V. Obzor sovremennykh metodov segmentatsii rastrovyykh izobrazheniy / S.V. Zhuk // – Izvestiya VolGTU. – 2009, №6. S. 115-118.
27. Dorogov A.Yu. Bystrodeystvuyushchiy algoritm semanticheskoy klassifikatsii JPEG-izobrazheniy / A.Yu. Dorogov, R.G. Kurbanov, V.V. Razin // Sankt-Peterburgskiy gosudarstvennyy elektrotekhnicheskii universitet (SPbGETU) "LETI", 2006.
28. Gorelik A.L., Skripkin V.A. Metody raspoznavaniya: Uchebnoe posobie dlya vuzov. Izd. 4. – M: Bukinist. –2004. 262 s.