

На правах рукописи

Николаев Дмитрий Петрович

Алгоритмы цветовой сегментации, применимые в условиях
сложного освещения сцены

Специальность 05.13.18 – математическое моделирование,
численные методы и комплексы программ

Автореферат диссертации на соискание ученой степени кандидата
физико-математических наук

МОСКВА, 2004 г.

Работа выполнена в Институте проблем передачи информации РАН

Научный руководитель: доктор физико-математических наук,
профессор
Чуличков Алексей Иванович

Консультант: доктор биологических наук,
профессор
Рожкова Галина Ивановна

Официальные оппоненты: доктор физико-математических наук,
профессор
Голубцов Пётр Викторович

кандидат физико-математических наук
Юрин Дмитрий Владимирович

Ведущая организация: Институт системного анализа РАН

Защита состоится «__» _____ 2004 года в _____ часов на заседании диссертационного совета К 501.001.17 в Московском государственном университете им. М.В. Ломоносова по адресу: 119992 ГСП-2, г. Москва, Воробьёвы горы, дом 1, строение 2.

С диссертацией можно ознакомиться в библиотеке физического факультета МГУ.

Автореферат разослан «__» _____ 2004 года.

Учёный секретарь диссертационного
совета К 501.001.17, д.ф.-м.н, профессор

П.А. Поляков

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Алгоритмы оконтуривания границ объектов на цветном изображении (алгоритмы объектной цветовой сегментации) являются необходимым инструментом для решения различных прикладных задач в области обработки цветных изображений, связанных с их редактированием, анализом, синтезом, восстановлением и сжатием. Хотя в настоящий момент уже разработано большое число таких алгоритмов, как для автоматической, так и для контролируемой оператором сегментации, использование большинства из них не обеспечивает удовлетворительного результата. Причина этого в первую очередь заключается в том, что в этих алгоритмах моделью изображения однородно окрашенного объекта является однородный по цвету участок изображения. Однако, из-за существенных неоднородностей мощности и цветности освещения в пространстве сцены, а также по причине сложной структуры индикатрис рассеяния поверхностей объектов в сцене, изображение однородно окрашенного объекта является, как правило, существенно неоднородным по цвету. Более того, на изображении объекта появляются и дополнительные контрастные границы (границы бликов, теней и затенений). В итоге, при сегментации изображения классическими методами объект дробится на более мелкие области, границы которых не соответствуют какому-либо скачку отражательных свойств поверхности. Поэтому очевидно, что задача объектной сегментации не может быть удовлетворительно решена без учета оптических явлений, порождающих на изображении границы различных типов. Хотя уже давно понятно, что алгоритмы цветовой сегментации можно развивать, только ориентируясь на физическую модель сцены (с ее ограничениями и приближениями), число таких алгоритмов, тем более реализованных программно, исчисляется единицами (Николаев, 1988; Klinker et al., 1990; Gevers and Smeulders, 1999). В то же время потребность в устойчиво работающих алгоритмах цветовой сегментации велика. Таким образом, актуальность создания новых методов и алгоритмов обработки цветного изображения (основывающихся на физическом подходе) вполне очевидна.

Цель и задачи работы

Цель работы состоит в комплексном решении задачи различения на изображении сложной зрительной сцены областей, соответствующих однородно окрашенным объектам, и выявления их действительных границ (задачи цветовой объектной сегментации).

В работе решались следующие научные задачи:

1. Физическое обоснование линейной модели формирования цветного изображения и определение пределов её применимости.
2. Развитие методов оценки адекватности линейной модели формирования изображения по отношению к анализируемым изображениям.
3. Создание математической и алгоритмической базы для решения задачи цветовой сегментации в пределах применимости линейной модели.
4. Создание библиотеки алгоритмов автоматической и полуавтоматической объектной сегментации цветных изображений трёхмерных сцен со сложным освещением для использования прикладными программами обработки и сжатия изображений.

Методы исследований, достоверность и обоснованность результатов

Методологическую основу работы составляют современные методы цифровой обработки сигналов, методы статистического оценивания, теории вероятностей и математической статистики, вычислительные методы. Достоверность результатов диссертации обеспечивается использованием строгих математических методов для расчета, обработки и анализа данных. Достоверность также подтверждается численным экспериментом.

Научная новизна работы

1. Впервые дано физическое обоснование положениям линейной теории формирования спектрального стимула и определены границы её применимости. Построена исчерпывающая классификация рангов (типов) сцен и объектов сцены по сложности описывающей их математической модели.
2. Впервые линейная модель формирования изображения была исследована с помощью теории надежности статистических гипотез. Впервые, основываясь

на теории проверки статистических гипотез, были построены оптимальные критерии сегментации изображений сцен со сложным освещением.

3. Впервые на основе выводов линейной теории формирования спектрального стимула разработана алгоритмическая схема сегментации изображений, не требующая информации об условиях освещения и свойствах поверхностей объектов сцены. В отличие от аналогичных разработок, также основывающихся на теоретическом рассмотрении законов светорассеяния для построения квазиинвариантных (по отношению к условиям освещения и наблюдения) функций компонент изображения, но ориентированных на анализ определенных типов сцен, предложенная в работе схема корректна и применима при существенно меньших ограничениях на сложность сцены. Универсальность схемы по отношению к задаче типизации сцен обеспечивает ее применимость к реальным изображениям.

4. На основе предложенного подхода впервые разработаны эффективные алгоритмы автоматической цветовой сегментации и полуавтоматического выделения объектов, комбинирующие зарекомендовавшие себя в практических приложениях подходы и алгоритмы (алгоритм поиска водоразделов, метод слияния областей с использованием графа соседства) с результатами линейной теории формирования спектрального стимула и теории проверки статистических гипотез. Алгоритмы предназначены для обработки изображений высокого качества в условиях сложного цветного освещения.

5. Впервые разработан алгоритм автоматической цветовой сегментации для сенсоров, калибровка которых невозможна, использующий положения теории формирования спектрального стимула. Алгоритм предназначен для обработки изображений низкого качества.

Практическая ценность и реализация результатов работы

Разработанные схемы и алгоритмы сегментации могут применяться для обработки видеопоследовательностей изображений с целью их анализа, редактирования и сжатия. В дополнение к этому они могут быть использованы для обработки статических изображений, таких как фотографии, журнальные страницы и документы. Отдельные варианты предлагаемых алгоритмов цветовой сегментации были внедрены в качестве части программного

обеспечения Института перспективных технологий Самсунг (SAIT, Южная Корея) в рамках проекта по реализации стандарта MPEG-4 и поданы на патентование. На базе предложенных алгоритмов разработан фильтр бинаризации изображений цветных документов, вошедший в ядро сканирования и распознавания печатных и рукопечатных документов “Scanify” компании Cognitive Technologies, Ltd (Россия).

Основные результаты и положения, выдвигаемые на защиту

1. Цветное изображение однородно-окрашенного объекта при освещении источниками света различной цветности формируется в соответствии с линейной моделью цветных изображений: цветовое распределение объекта лежит в линейном подмножестве цветного пространства линейного сенсора.
2. Адекватность линейной модели формирования изображения может быть проверена с помощью теории надежности статистических гипотез.
3. Параметры линейной модели формирования изображения могут быть найдены оптимальным образом методом собственных векторов или методом обобщённого преобразования Хафа, в зависимости от модели шума.
4. Алгоритмы сегментации, разработанные на основе линейной модели формирования изображения и теории надежности статистических гипотез, позволяют эффективно решать задачу различения объектов на изображении, в том числе в условиях сложного освещения (несколько источников света различной цветности).

Апробация работы

Основные результаты диссертации докладывались на международных конференциях “Искусственные интеллектуальные системы” и “Интеллектуальные САПР” (IEEE AIS и CAD) (пос. Дивноморское, Россия, сентябрь 2002, сентябрь 2003 и сентябрь 2004), на международном симпозиуме “25th European conference on Visual Perception” (г. Глазго, Великобритания, август 2002), на международном семинаре “6th German-Russian Workshop on Pattern Recognition and Image Understanding” (OGRW-6-2003) (пос. Катунь, Россия, август 2003), а также неоднократно обсуждались на семинарах лаборатории обработки сенсорной информации Института проблем передачи

информации РАН, на семинарах отдела когнитивных и компьютерных технологий Института системного анализа РАН и мультимедийной лаборатории Института перспективных технологий Самсунг (SAIT, Южная Корея).

Публикации, личный вклад автора

По материалам диссертации опубликовано 12 научных работ, из них 2 – тезисы докладов, 2 – патентные публикации.

Изложенные в диссертации результаты получены лично автором или при его непосредственном участии. В работах, выполненных вместе с соавторами, вклад автора является определяющим.

Структура и объём диссертации

Диссертация состоит из введения, трех частей: литературного обзора (глава I), теоретической части (глава II), алгоритмической и экспериментальной части (главы III, IV и V) и заключения. Работа изложена на 125 страницах, включающих 22 рисунка и список литературы из 120 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Во введении определяется место данной работы в общем ряду исследований по цветовой сегментации. Вводится корректное с физической точки зрения определение задачи цветовой сегментации как задачи цветоразличения. Формулируется цель диссертационной работы. Приводится краткая аннотация ее содержания.

В главе I приведен обзор экспериментальных и теоретических работ, относящихся к проблеме цветовой сегментации. В конце главы ставится задача настоящего исследования.

В разделе I.1. приводятся основные понятия цветовой теории, такие как пространство спектрального распределения, цветовые пространства (ЦП), цветовые координаты и гистограммы цветового распределения, и рассматривается возможность перехода к квазиинвариантным цветовым координатам как способа избежать анализа многомерных цветовых распределений. В разделе I.2. рассматриваются наиболее часто используемые на данный момент алгоритмы цветовой сегментации изображения,

основывающиеся, как правило, на некоторых локальных свойствах самого изображения и не использующие информацию о процессах, лежащих в основе взаимодействия света с объектами сцены, – процессов, которые собственно и формируют изображение. Анализируются причины неудовлетворительности конечных результатов работы существующих алгоритмов и отмечается ценность некоторых из них как сервисных алгоритмов обработки изображения (в том числе и для разработанных в данной работе схем сегментации). В разделе I.3. кратко описаны физические явления, лежащие в основе формирования изображения, и линейная модель формирования спектрального стимула, положенная в основу настоящего подхода к проблеме цветовой сегментации. Излагается представление о ранговой классификации сцен, опирающееся на различия в форме цветовых распределений, соответствующих в ЦП сенсора объектам сцены. На основе проведенного анализа обосновывается адекватность использования геометрических особенностей кластеров цветового распределения для сегментации изображений. Раздел I.4. посвящен обзору исследований цветовой сегментации изображения в человеческом зрении.

Глава II содержит материал, полученный диссертантом с соавторами, и посвящена теоретическому рассмотрению преобразования спектра излучения при отражении от поверхности объекта и при взаимодействии с сенсорной системой для широкого многообразия типов сцен (содержащих тени, затенения, блики и интеррефлексы). На основе этого материала строится линейная модель формирования спектрального стимула, даётся обоснование ее корректности, а также строятся критерии, позволяющие проверить её адекватность для данного изображения. Показывается, что для изображений, удовлетворяющих модели, задача цветовой сегментации может быть поставлена в терминах морфологического анализа математически строго и корректно как задача аппроксимации формы изображений.

В разделе II.1. кратко приводятся определения морфологического анализа и вводится общая модель изображения. Изображение определяется как вектор-функция $\vec{c}(\vec{r})$, заданная на некотором поле зрения – подмножестве Ω плоскости \mathfrak{R}_2 , заданная в виде:

$$\vec{c}(\vec{r}) = \sum_{i=1}^N \vec{c}_i(\vec{r}) \cdot \omega_i(\vec{r}), \quad (1)$$

где N – число изображённых объектов, $\omega_i(\vec{r})$ – индикаторная функция i -го объекта, а $\vec{c}_i(\vec{r})$ – его цветовое распределение. Индикаторные функции определяются следующим образом: поле зрения Ω разбито на области $A_i \subset \Omega$, соответствующие изображениям отдельных объектов ($i=1, \dots, N$; $i \neq j \Rightarrow A_i \cap A_j = \emptyset$; $\bigcup_{i=1}^N A_i = \Omega$). Тогда $\omega_i(\vec{r}) = \begin{cases} 1, \vec{r} \in A_i \\ 0, \vec{r} \notin A_i \end{cases}$.

При этом функции цветового распределения $\vec{c}_i(\vec{r})$ считаются принадлежащими некоторому классу функций \mathbf{C} , определяемому из специфики задачи.

Раздел II.2. посвящён определению класса цветовых распределений, возникающих при регистрации изображений различных сцен. В этом разделе строится более общая, чем в других работах (Николаев, 1988; Klinker et al., 1990; Brill, 1990), модель формирования спектрального стимула, что позволяет распространить её на более широкий круг реальных сцен. Обобщенная линейная модель основывается на трёх допущениях. Первое – адекватность приближения лучевой оптики. Второе – возможность представить спектральную яркость источников света в сцене в виде произведения спектрального и геометрического сомножителей. Третье – возможность представить спектральную двухлучевую функцию отражательной способности (ДФОС) поверхности объектов сцены в виде произведения спектрального и геометрического сомножителей или суммы таких произведений. На основе анализа известных экспериментальных данных демонстрируется адекватность этих допущений.

В работе показано, что при таких допущениях спектральный стимул $F(\lambda)$ от каждой точки объекта является элементом линейного подмножества размерности n пространства спектральных функций. Действительно, световой поток, порождающий изображение сцены, формируется в результате двух процессов – распространения света в прозрачной среде и его рассеяния поверхностями объектов. В общем случае источник может быть задан спектральной яркостью излучения:

$$B_\lambda(\lambda, \vec{r}, \vec{n}_{out}) = \frac{\delta P}{(\vec{n}_{out} \cdot \vec{n}_{surf}) \cdot \delta\lambda \cdot \delta\omega \cdot \delta S}, \quad |\vec{n}_{out}| = 1, \quad |\vec{n}_{surf}| = 1.$$

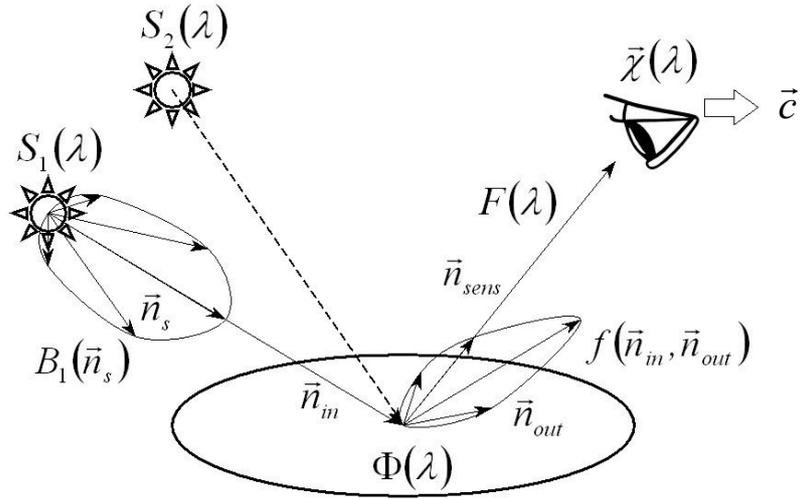


Рис 1. Схема формирования изображения сцены. Точечные источники света S_i со спектральной яркостью излучения $B_{\lambda i}(\lambda, \vec{r}, \vec{n}_s) = S_i(\lambda) \cdot B_i(\vec{n}_s) \cdot \delta(\vec{r} - \vec{r}_{0i})$ (здесь \vec{r}_{0i} – координаты i -го источника) освещают плоскость Φ с СДФОС $f_\lambda(\lambda, \vec{r}, \vec{n}_{in}, \vec{n}_{out}) = \Phi(\lambda) \cdot f(\vec{n}_{in}, \vec{n}_{out}) \cdot \delta(\vec{r} \cdot \vec{e}_z)$ (здесь \vec{e}_z – орт вертикальной оси). Сенсор со спектральной чувствительностью датчиков $\vec{\chi}(\lambda)$ регистрирует в точке светочувствительной матрицы некоторый спектральный стимул $F(\lambda)$, преобразуя его в вектор-стимул \vec{c} , соответствующий данному элементу изображения.

Здесь δP – интегральная мощность излучения в диапазоне длин волн $(\lambda, \lambda + \delta\lambda)$, излучаемая площадкой δS источника из точки \vec{r} внутри телесного угла $\delta\omega$ в направлении, заданном вектором \vec{n}_{out} ; \vec{n}_{surf} – нормаль к площадке δS . В большинстве случаев распределение спектральной яркости может быть переписано в виде:

$$B_\lambda(\lambda, \vec{r}, \vec{n}_{out}) = S(\lambda) \cdot B(\vec{r}, \vec{n}_{out}), \quad |\vec{n}_{out}| = 1, \quad \int_0^\infty S(\lambda) d\lambda = 1.$$

Здесь $B(\vec{r}, \vec{n}_{out})$ – интегральная яркость излучения, а $S(\lambda)$ – относительное спектральное распределение излучения (Рис. 1).

Если не рассматривать нелинейно-оптических эффектов, то отражательные свойства тел, образующих сцену, можно охарактеризовать спектральной ДФОС (СДФОС) $f_\lambda(\lambda, \vec{r}, \vec{n}_{in}, \vec{n}_{out})$, задающей отношение спектральной яркости поверхности к её освещённости в диапазоне длин волн $(\lambda, \lambda + d\lambda)$ в каждой точке \vec{r} поверхности тела при освещении в направлении \vec{n}_{in} и наблюдении в направлении \vec{n}_{out} в системе координат, связанной с элементарной площадкой поверхности в точке \vec{r} (Рис. 1). При рассмотренных

ограничениях ДФОС представима в виде суммы следующего вида:

$$f_{\lambda}(\lambda, \vec{r}, \vec{n}_{in}, \vec{n}_{out}) = \sum_k \Phi_k(\lambda) \cdot f_k(\vec{r}, \vec{n}_{in}, \vec{n}_{out}),$$

где $|\vec{n}_{in}|=1$, $|\vec{n}_{out}|=1$, $\int_0^{\infty} \Phi_k(\lambda) d\lambda = 1$, и $\Phi_k(\lambda)$ – линейно независимы.

Здесь $f_k(\vec{r}, \vec{n}_{in}, \vec{n}_{out})$ – интегральные ДФОС, а $\Phi_k(\lambda)$ – относительные спектральные распределения коэффициента отражения. Количество значимых членов этой суммы зависит от свойств материала поверхности.

При этих условиях, спектральный стимул $F(\lambda, \vec{r})$ от сцены, содержащей несколько источников с различными относительными спектральными распределениями излучения и несколько окрашенных поверхностей с различными СДФОС, с учетом интеррефлексов может быть записан в следующем виде:

$$F(\lambda, \vec{r}) = \sum_i g_i(\vec{r}) \cdot S_i(\lambda) + \sum_{i,n,k} g_{i,n}(\vec{r}) \cdot S_i(\lambda) \cdot \Phi_{n,k}(\lambda) + \sum_{i,n_1,n_2,k_1,k_2} g_{i,n_1,n_2,k_1,k_2}(\vec{r}) \cdot S_i(\lambda) \cdot \Phi_{n_1,k_1}(\lambda) \cdot \Phi_{n_2,k_2}(\lambda) + \dots \quad (2)$$

где $S_i(\lambda)$ – относительное спектральное распределение излучения i -того источника света; $\Phi_{n,k}(\lambda)$ – k -тое относительное спектральное распределение коэффициента отражения в разложении СДФОС n -ной поверхности, а $g_{i,n_1,n_2,\dots,n_m,k_1,k_2,\dots,k_m}(\vec{r})$ – геометрический фактор для луча света от i -того источника света, последовательно отражённого от m поверхностей n_1, n_2, \dots, n_m .

Геометрические факторы зависят от мощности источника света, интегральных ДФОС поверхностей и взаимного расположения источников, объектов и сенсора в сцене.

Ранг объекта на изображении определяется через рассмотрение совокупности спектральных функций L , элементами которой являются разности между спектральными стимулами от каждой точки объекта и спектрального стимула от произвольной фиксированной точки объекта \vec{r}_0 : $L = \{F(\lambda, \vec{r}) - F(\lambda, \vec{r}_0)\}$. Пусть ранг L в линейном пространстве относительных спектральных функций равен n , тогда (2) можно переписать в виде:

$$F(\lambda, \vec{r}) = g_c \cdot M_c(\lambda) + \sum_{i=1}^n g_i(\vec{r}) \cdot M_i(\lambda), \quad (3)$$

где $\{M_i(\lambda)\}$ – ортонормированный базис L , а $M_c(\lambda)$ ортогональна этому базису. Таким образом, спектральный стимул $F(\lambda)$ от каждой точки объекта является элементом линейного подмногообразия размерности n пространства спектральных функций. Эта размерность и называется рангом данного объекта. Ранг объекта зависит и от типа его поверхности, и от условий его освещения и наблюдения.

Сенсор преобразует поступающий сигнал в каждой точке (x, y) своей светочувствительной матрицы из спектрального стимула $F(\lambda)$ в т.н. цветовой вектор-стимул $\vec{c}(x, y)$:

$$\vec{c}(x, y) = \int_0^{\infty} F(\lambda, \vec{r}(x, y)) \cdot \vec{\chi}(\lambda) d\lambda, \quad (4)$$

где $\vec{\chi}(\lambda)$ – вектор спектральных чувствительностей сенсора (трёхмерный в случае RGB-камеры или, в определённых условиях, глаза человека). Таким образом, сенсор проецирует вектор-стимул из пространства функций в конечномерное ЦП, сохраняя линейные свойства распределения для случаев рангов меньших размерности ЦП. Из (4) также следует, что на формирование изображения не влияют спектральные особенности стимула $F(\lambda, \vec{r})$ в тех областях спектра, где все компоненты вектора $\vec{\chi}(\lambda)$ близки к нулю. Таким образом, все требования линейной теории, сформулированные выше, должны выполняться только в видимой для сенсора части спектра.

Для изображений, удовлетворяющих линейной модели, класс \mathbf{C} функций цветового распределения $\vec{c}_i(\vec{r})$ представления (1) можно записать в явном виде:

$$\vec{c}(\vec{r}) = \sum_{i=1}^N \vec{c}_i(\vec{r}) \cdot \omega_i(\vec{r}) = \sum_{i=1}^N (g_{i,0} \cdot \vec{c}_{i,0} + g_{i,1}(\vec{r}) \cdot \vec{c}_{i,1} + g_{i,2}(\vec{r}) \cdot \vec{c}_{i,2}) \cdot \omega_i(\vec{r}), \quad (5)$$

где $(\vec{c}_{i,j} \cdot \vec{c}_{i,k}) = \delta_{j,k}$. Тем самым получено определение формы цветного изображения трёхмерной статичной сцены при сложных условиях освещения. Форма (5) является конусом в пространстве изображений, а задача цветовой объектной сегментации заключается в построении проекции на этот конус.

На рис. 2 показано, как выглядят уединённые кластеры цветовых распределений реальных сцен.

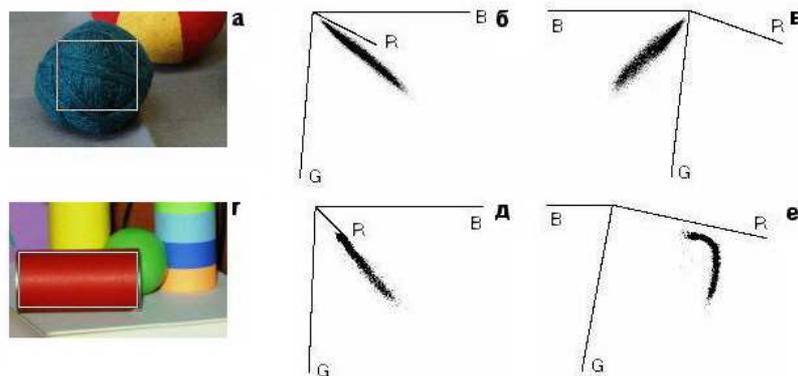


Рис. 2. Примеры цветовых гистограмм реальных изображений в RGB-пространстве. а) Объект ранга 1 (клубок синей шерсти при дневном освещении). б), в) Проекция кластера ранга 1 в ЦП. г) Объект ранга 2 (глянцевый красный цилиндр при освещении лампой накаливания). д), е) Проекция кластера ранга 2 в ЦП.

В разделе II.3. приводятся примеры сцен, создающих, согласно теории, в ЦП сенсора кластеры различных рангов (минимальных размерностей натянутого линейного подмногообразия). Показывается, что для некоторых классов сцен перепроецирование цветового распределения из ЦП в пространства пониженной размерности сохраняет свойство вырожденности кластеров распределения, что позволяет существенно ускорить сегментацию при использовании одних и тех же алгоритмов. Чтобы обозначить классы сцен, позволяющих редукцию ЦП, была уточнена систематика рангов сцен с учетом расположения линейных подмногообразий относительно главной диагонали нормализованного ЦП – ахроматической оси. Сведены в таблицу примеры сцен различных рангов расширенной систематики. Показано, для каких практических случаев возможно упрощение анализа цветовых распределений за счет проекции трёхмерного ЦП на плоскость цветности, а для каких вообще достаточно рассмотреть распределение цветового тона.

В разделе II.4 вводится параметризация моделей линейных кластеров различных рангов и сравниваются алгоритмы определения параметров их распределений – обобщённое преобразование Хафа и метод собственных векторов (или, что то же самое, метод наименьших квадратов). В качестве оптимального выбора предлагается использовать преобразование Хафа, если это позволяют вычислительные мощности системы. В противном случае предлагается использовать метод пристрелки с начальным приближением, вычисляемым с помощью метода собственных векторов. Дополнительно

рассматриваются особенности применения этих методов для анализа цветовых распределений на циклической оси цветового тона.

Метод наименьших квадратов решает задачу аппроксимации формы одного сегмента оптимально в квадратичной метрике при наличии нормального шума:

$$\int_{A_i} \left\| \bar{c}(\bar{r}) - \left(g_{i,0} \cdot \bar{c}_{i,0} + \sum_{j=1}^R g_{i,j}(\bar{r}) \cdot \bar{c}_{i,j} \right) \right\|^2 d\bar{r} \equiv \int_{A_i} \rho_R^2(\bar{r}) d\bar{r} \rightarrow \min_{\bar{c}_{i,j}, g_{i,j}},$$

где $(\bar{c}_{i,j} \cdot \bar{c}_{i,k}) = \delta_{j,k}$, а R – ранг линейного подмногообразия. В свою очередь, метод ОПХ решает ту же задачу, оптимизируя другой функционал, что обеспечивает методу устойчивость по отношению к импульсному (выбросовому) шуму:

$$\int_{A_i} \frac{\exp(-\rho_R^2(\bar{r})/2 \cdot \sigma^2)}{\sqrt{2 \cdot \pi \cdot \sigma}} d\bar{r} \rightarrow \max_{\bar{c}_{i,j}, g_{i,j}},$$

где σ – параметр сглаживания. Сглаживание исходного распределения требуется для подавления нормальной компоненты шума, что особенно важно при малых размерах сегмента A_i .

В разделе показано, что преобразование Хафа в некотором смысле обратимо с гауссовским сглаживанием, что позволяет производить подгонку параметра σ без перевычисления преобразования. Если обозначить преобразование Хафа для k -мерных линейных подмногообразий в n -мерном пространстве как $H_{n,k}$, то этот факт можно записать следующим образом:

$$H_{n,k} \{G_\sigma(\bar{c}) * f(\bar{c})\}(\bar{p}_\varphi, \bar{p}_\rho) = G_\sigma(\bar{p}_\rho) * H_{n,k} \{f(\bar{c})\}(\bar{p}_\varphi, \bar{p}_\rho),$$

где $G_\sigma(\bar{p}) = \frac{\exp(-|\bar{p}|^2/2 \cdot \sigma^2)}{(\sqrt{2 \cdot \pi \cdot \sigma})^{\dim(\bar{p})}}$, \bar{p}_φ – угловые координаты подмногообразия, а \bar{p}_ρ

– его пространственные координаты. При этом $\dim(\bar{p}_\varphi) = k \cdot (n - k)$, $\dim(\bar{p}_\rho) = n - k$.

Кроме того, в разделе предложена простая быстрая схема суммирования для преобразования Хафа ($O(n^2 \cdot \log n)$ операций для $H_{2,1}$), не использующая преобразование Фурье. В массиве с линейными размерами $n \times n = 2^D \times 2^D$ суммирование вдоль прямой, соединяющей точки (x, y) и $(x + shift, y + 2^{\text{deg}} - 1)$ в

горизонтальную полосу $n \times 2^{\text{deg}}$ ($0 < \text{deg} \leq D$) с вертикальным смещением $y = k \cdot 2^{\text{deg}}$ ($0 \leq k < 2^{D-\text{deg}}$), проводится по рекуррентной формуле

$$\begin{aligned} \left[(x, y), (x + \text{shift}, y + 2^{\text{deg}} - 1) \right] = & \left[(x, y), (x + \lfloor \text{shift} / 2 \rfloor, y + 2^{\text{deg}-1} - 1) \right] \cup \\ & \left[(x + \lceil \text{shift} / 2 \rceil, y + 2^{\text{deg}-1}), (x + \text{shift}, y + 2^{\text{deg}} - 1) \right] \end{aligned} \quad (6)$$

что и обеспечивает (при порядке суммирования по deg от 0 до $D-1$) увеличение быстродействия за счёт отсутствия повторного суммирования для фрагментов дискретного представления прямых, входящих в несколько сумм.

В главе III описан алгоритм быстрого выделения маски объекта на цветном изображении сцены со сложными условиями освещения под частичным контролем пользователя. Алгоритм позволяет определять искомую однородно окрашенную область по указанному малому участку области. Предлагаемый алгоритм выделения объекта построен на процедуре итерационного уточнения его модели по мере роста области выделения. В качестве модели цветового распределения объекта выбрана линейная модель ранга 1, поскольку, как было показано, большинство однородно окрашенных объектов при ахроматическом (либо малонасыщенном) освещении удовлетворяют модели ранга 1, – их цветовое распределение образует в ЦП отрезок прямой.

Кластер ранга 1 задаётся уравнением прямой в канонической форме. Параметры модели – координаты точки $g_0 \cdot \vec{c}_0$, через которую проходит эта прямая, и вектора \vec{c}_1 , задающего ее направление ($\vec{c}_0^2 = 1$, $\vec{c}_1^2 = 1$, $(\vec{c}_0 \cdot \vec{c}_1) = 0$). Квадрат невязки $\rho^2(x, y)$ между пикселем изображения с координатами (x, y) и моделью равен:

$$\rho^2(x, y) = \|\vec{c}(x, y) - g_0 \cdot \vec{c}_0\|^2 - (\vec{c}(x, y) \cdot \vec{c}_1)^2.$$

Собственно алгоритм состоит в следующем:

1. Методом обобщённого преобразования Хафа найти приближение параметров модели $\{g_{0,1}, \vec{c}_{0,1}, \vec{c}_{1,1}\}$ для начального сегмента S_0 , заданного на входе. Задать ρ_1^{\max} , как среднюю невязку по области S_0 .

2. Построить карту невязок $\rho_i(x, y)$, которая представляет собой массив расхождений между моделью с параметрами $\{g_{0,i}, \bar{c}_{0,i}, \bar{c}_{1,i}\}$ и пикселями исходного изображения $\bar{c}(x, y)$ (для первой итерации $i = 1$).
3. Произвести заливку карты $\rho_i(x, y)$ по уровню ρ_i^{\max} из области S_0 . Результат заливки – новое приближение рассматриваемой области, S_i . Если S_i покрывает все изображение, перейти к (7).
4. Вычислить среднее отношение невязок $G_i = \tilde{\rho}_i^{ext} / \tilde{\rho}_i^{int}$ для оценки “качества границы”. Здесь $\tilde{\rho}_i^{int}$ – средняя невязка среди пикселей, входящих в S_i , а $\tilde{\rho}_i^{ext}$ – средняя невязка среди пограничных с S_i пикселей.
5. Методом пристрелки вычислить новые параметры модели $\{\bar{p}_{i+1}, \bar{v}_{i+1}\}$ для области S_i .
6. Увеличить порог заливки $\rho_{i+1}^{\max} = \rho_i^{\max} + \Delta\rho$ и перейти к (2) для следующего i .
7. Проанализировать результат выделения объекта: выделить в последовательности G_i локальные максимумы: для каждого i , удовлетворяющего условию $(G_{i-1} \geq G_i < G_{i+1})$, пометить S_i в качестве кандидата на искомое выделение.

В результате работы данного алгоритма определяется несколько гипотез о границах объекта. Пользователь выбирает из них искомый. Пример работы алгоритма выделения объекта приведен на Рис. 3.

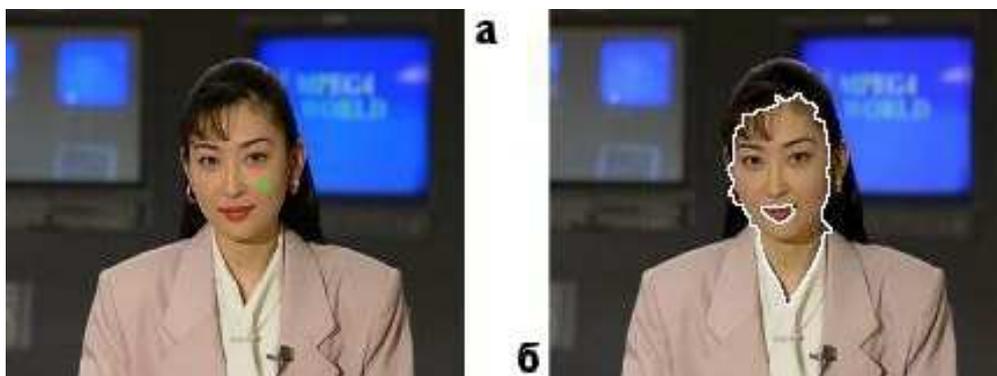


Рис. 3. Результаты применения алгоритма выделения объекта. а) Исходное изображение, точка на правой щеке маркирует начальную точку выделения. б) Результат выделения (гипотеза 2 из 7).

В главе IV диссертации описывается оригинальный алгоритм автоматической сегментации изображений, полученных с линейного сенсора. Информация о геометрических особенностях изображений объектов учитывается алгоритмом в минимальной степени, что обеспечивает высокое качество сегментации для объектов произвольной сложной формы.

В разделе IV.1. описывается алгоритм предварительной сегментации, позволяющий существенно сократить время обработки изображения основным алгоритмом. Задача предварительной сегментации – разделить изображение на области площадью хотя бы в несколько пикселей так, чтобы истинные границы объектов не пересекали этих областей. В работе используется обобщённый на случай цветных изображений алгоритм поиска водоразделов (ранее он применялся только для случая монохромных изображений). Основная посылка алгоритма поиска водоразделов состоит в том, что каждый сегмент должен содержать точно один минимум значения некоторого детектора границ, и градиентный спуск от любого пикселя сегмента по карте детектора границ должен приводить к этому минимуму. Это условие реализуется проведением градиентного спуска из каждой точки изображения и группированием точек по их принадлежности к минимумам. В результате границы сегментов (групп пикселей) проходят вдоль «хребтов» значений детектора краев. Поскольку классический оператор градиента может быть применен только для скалярного, но не для векторного поля, в цветовой версии алгоритма используется векторный аналог градиента (т.н. цветовой градиент). Его модуль записывается следующим образом:

$$G = \frac{\Sigma_{11} + \Sigma_{22} + \sqrt{(\Sigma_{11} - \Sigma_{22})^2 + 4 \cdot \Sigma_{12}^2}}{2}, \text{ где } \Sigma_{jk}(x_1, x_2) = \sum_{i=1}^n \frac{\partial C_i(x_1, x_2)}{\partial x_j} \cdot \frac{\partial C_i(x_1, x_2)}{\partial x_k}.$$

Здесь $C_i(x_1, x_2)$ – цветовые компоненты изображения в линейном ЦП, n – размерность ЦП (для RGB-пространства – 3), а G – модуль цветового градиента (максимум модуля изменения вектора C_i по всем направлениям).

В разделе IV.2. изложена техника сегментации слиянием областей с помощью графа соседства областей (ГСО). Каждая вершина ГСО представляет сегмент изображения. Две вершины связаны ребром тогда и только тогда, когда соответствующие сегменты имеют общую границу. Каждому ребру приписывается вес, характеризующий вероятность того, что эти сегменты

имеют несовпадающие параметры модели. Граф последовательно модифицируется слиянием пар вершин, соединённых ребром с наименьшим весом. Слияние производится до тех пор, пока наилучшее ребро не будет иметь вес выше некоторого порога. Качество сегментации, очевидно, определяется выбором весовой функции и уровнем порога. Кроме того, метод ГСО требует в качестве входных данных некоторую карту предварительной сегментации, которая будет только «огрубляться» в процессе обработки.

В разделе строятся весовые функции ГСО, основанные на оптимальных критериях проверки статистических гипотез. Для моделей ранга 0 задача ставится как проверка гипотезы, что изображение $\vec{c}(\vec{r})$ представимо в форме

$$\vec{c}(\vec{r}) = \sum_{i=1}^2 \vec{c}_i \cdot \omega_i(\vec{r}), \quad \vec{c}_1 \neq \vec{c}_2 \text{ против альтернативы } \vec{c}_1 = \vec{c}_2.$$

Показано, что в этом случае оптимальным является следующий критерий:

$$j = \frac{\|(P_B - P_0)\vec{c}(\vec{r})\|^2}{\|(I - P_B)\vec{c}(\vec{r})\|^2},$$

где I – единичный оператор, P_B – проектор на подпространство изображений

изображений $\vec{c}(\vec{r}) = \sum_{i=1}^2 \vec{c}_i \cdot \omega_i(\vec{r})$, $\vec{c}_1 \neq \vec{c}_2$, а P_0 – на подпространство пустых

изображений.

В разделе IV.3. изложен основной алгоритм сегментации, опирающийся на метод ГСО. В предложенном алгоритме используется семейство весовых функций, соответствующих моделям кластеров различных рангов. В общем виде эти весовые функции записываются следующим образом:

$$d_j[k, l] = \sqrt{\frac{\sum_{i=1}^{n_k} \rho_j^2(\vec{p}_{k,i})}{n_l} + \frac{\sum_{i=1}^{n_l} \rho_j^2(\vec{p}_{l,i})}{n_k}},$$

где $d_j[k, l]$ – вес ребра, связывающего узлы, соответствующие сегментам k и l ,

j – ранг модельного распределения, $\rho_j(\vec{p})$ – отклонение пикселей \vec{p} от идеальной модели для области, объединяющей сегменты k и l , $\vec{p}_{k,i}$ – пиксель i сегмента k , а n_k – кол-во пикселей в сегменте k .

Собственно алгоритм включает следующие этапы:

1. Применение техники слияния областей с весовой функцией $d_0[k,l]$ и порогом σ_0 , соответствующим по порядку уровню шума на изображении.
2. Маркирование изолированных сегментов ранга 0, то есть тех, для которых минимальное d_{min} для всех прилегающих рёбер сильно больше порога: $d_{min} > \sigma_G$, где $\sigma_G \gg \sigma_0$.
3. Применение техники слияния областей с весовой функцией $d_1[k,l]$ и порогом σ_0 , с исключением из рассмотрения рёбер, ведущих к изолированным сегментам.
4. Применение техники слияния областей с весовой функцией $d_1[k,l]$ и порогом σ_0 , с исключением из рассмотрения рёбер, соединяющих два изолированных сегмента.
5. Дополнительное маркирование изолированных сегментов ранга 1 с порогом σ_G .
6. Применение техники слияния областей с весовой функцией $d_2[k,l]$ и порогом σ_0 , с исключением из рассмотрения рёбер, ведущих к изолированным сегментам.
7. Применение техники слияния областей с весовой функцией $d_2[k,l]$ и порогом σ_0 , с исключением из рассмотрения рёбер, соединяющих два изолированных сегмента.

В работе показано, что подобная схема при правильно выбранных порогах σ_0 и σ_G позволяет получить верную сегментацию, не решая задачу определения ранга конкретного сегмента изображения. Идея состоит в том, что слияние с весовой функцией ранга i не разрушает границ зоны ранга j , если $i < j$. Таким образом, ранние этапы обработки не вносят искажений в работу более поздних (Рис. 4). Однако, этого недостаточно. Через любые две точки можно провести прямую, через любую точку и прямую – плоскость. Таким образом, на более поздних этапах обработки мы имеем шанс разрушить правильно построенные участки сегментационной карты. Чтобы этого не произошло, на этапах 2 и 5 происходит изоляция сегментов, которые могут оказаться верно найденными зонами объектов младших рангов. С другой стороны, такие сегменты могут быть порождены условиями освещения.

Например: тень, отброшенная на матовый объект, освещённый двумя источниками, – изолированный сегмент ранга 1. Но теневые сегменты имеют отличительный признак: это сегменты, соседствующие с сегментом более высокого ранга, объединение с которым не повышает ранг этого последнего. Именно такие сегменты и удаляются на этапах 4 и 7. К сожалению, до получения оценки параметров освещения не представляется возможным решить, нужно ли сливать два изолированных сегмента ранга 0. Граница между ними может оказаться как границей окраски, так и ребром однородно окрашенного объекта. Предложенный алгоритм в таких случаях слияния не производит. Пример работы алгоритма приведен на рис. 4.

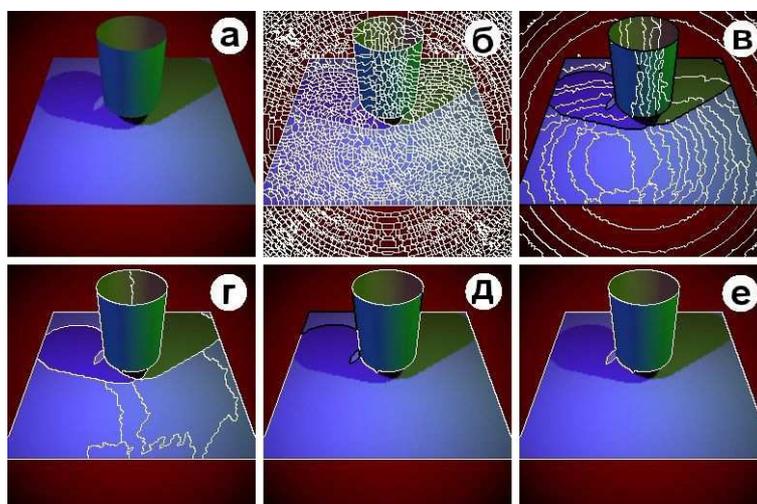


Рис. 4. Пример сегментации линейным методом. а) Исходное изображение. б) Результат предварительной сегментации. в), г), д) Промежуточные этапы сегментации (после шагов 2, 4, и 6 алгоритма, соответственно). Границы с весом, большим σ_G , помечены чёрным. е) Окончательный результат.

В главе V описан метод сегментации изображения, полученного при помощи нелинейного сенсора с неизвестными характеристиками. В ЦП таких сенсоров объекты будут проецироваться с искажениями, в связи с чем использование алгоритмов детектирования линейных подмногообразий может привести к ошибочному результату. Анализ сцен, в которых присутствуют лишь ахроматические источники (источники с относительным спектральным распределением излучения $S(\lambda) \approx const$ в видимой области спектра), показал, что каждый однородно окрашенный объект при таком освещении, в безрефлексном приближении проецируется в нелинейное ЦП как точечный кластер или кластер в форме отрезка кривой с преимущественной ориентацией

вдоль оси яркости. Это подтверждается экспериментальным исследованием изображений, полученных с реальных камер. Для таких случаев был разработан простой и эффективный алгоритм сегментации, основывающийся на детектировании в цветовом пространстве сенсора кластеров типа криволинейных отрезков.

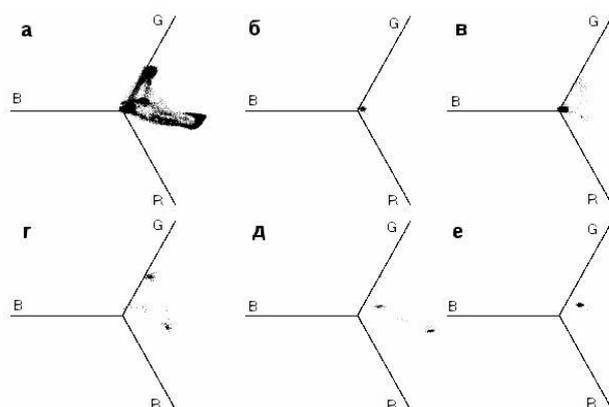


Рис. 5. Примеры сечения цветовой гистограммы плоскостями $L = (R + G + B)/3 = const$. а) Исходная гистограмма. б) – е) Сечения гистограммы ($L = 16$, $L = 48$, $L = 80$, $L = 130$, $L = 170$, соответственно).

Предлагаемый алгоритм состоит из этапа разбиения цветового пространства на слои семейством плоскостей, нормальных к главной диагонали цветового пространства, послойной кластеризации с помощью алгоритма поиска водоразделов, сборки кластеров и обратного проецирования на плоскость изображения с целью получения карты сегментации. Такой алгоритм показывает хорошие результаты при обработке сцен с малым количеством объектов, так как в каждом слое сечение кластера хорошо локализовано на поверхности сечения (Рис. 5). Однако, на сложных изображениях при использовании приведенного выше алгоритма не исключена вероятность ошибок сегментации, если два объекта, не имеющие общей границы, порождают в цветовом пространстве перекрывающиеся распределения. Для таких изображений предложена версия алгоритма, учитывающая геометрическое соседство точек. В ней используется разбиение изображения на небольшие, плотно покрывающие изображение, прямоугольные области, каждая из которых сегментируется независимо. При этом при сегментации текущей области анализируется цветовое распределение, взятое по большей области. При этом объект, далёкий от текущей области, уже не зашумляет её

локальное цветовое распределение, а близкие объекты вносят существенный вклад, даже если их вклад по площади в сегментируемую область мал.

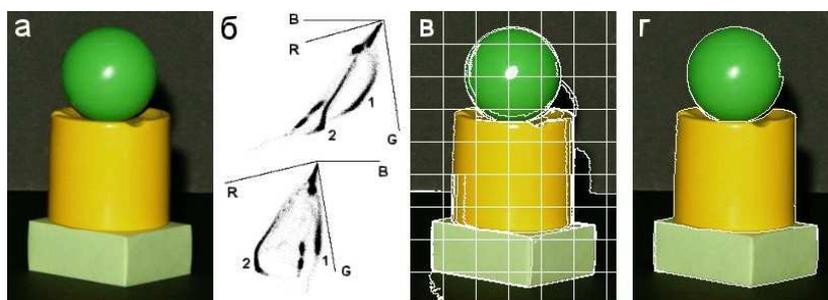


Рис. 6. Пример сегментации изображения, полученного с некалиброванного сенсора. а) Исходное изображение. б) Две проекции цветового распределения исходного изображения. Цифрой 1 помечен кластер, соответствующий на изображении шару, 2 – цилиндру. в) Результат сегментации с разбиением на зоны. г) результат сегментации после слияния областей.

Для устранения границ между «областями внимания» в алгоритме используется процедура слияния областей с помощью ГСО. В качестве меры несходства $d_{i,j}$ при этом взята следующая функция:

$$d_{i,j} = \sum_{k=1}^3 \sqrt{\frac{(\mu_{i,k} - \mu_{j,k})^2 \cdot n_i \cdot n_j}{n_i + n_j}},$$

где n_i и n_j – количество пикселей в сравниваемых сегментах, а $\mu_{i,k}$ и $\mu_{j,k}$ – средние значения k -той цветовой компоненты по соответствующим сегментам. Результат применения этого варианта алгоритма приведен на рис. 6.

Заключение подытоживает основные результаты проведенных исследований, экспериментов и практической реализации.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертации предложена информационная технология анализа цветных изображений сцен со сложным освещением. Для реализации этой технологии:

1. Разработана математическая модель цветных изображений как набора областей, цветовое распределение каждого из которых образует линейное подмногообразие цветового пространства. В частности: развита, формализована и обоснована линейная модель формирования цветного изображения; формально поставлена задача линейной цветовой сегментации; построена полная ранговая классификация сцен, различающая их по форме цветового

распределения; предложен критерий, определяющий допустимость использования цветовых пространств пониженной размерности.

2. На основе теории надежности статистических гипотез разработаны математические методы проверки адекватности линейной модели формирования изображения.

3. Предложены и обоснованы математические методы сегментации изображений, удовлетворяющих линейной модели, основанные на оптимальных критериях проверки статистических гипотез.

4. Предложены новые алгоритмы обработки изображений, в частности: алгоритм пресегментации методом водоразделов на цветном изображении, алгоритм кластеризации цветовой гистограммы путём сегментации двумерных сечений и быстрый безартефактный алгоритм преобразования Хафа.

5. Разработан комплекс программ, предназначенный для автоматической и полуавтоматической сегментации цветных изображений. Комплекс состоит из следующих модулей: модуль быстрого выделения объектов на цветном изображении (под частичным контролем пользователя); модуль автоматической сегментации изображений высокого качества в условиях сложного цветного освещения; модуль сегментации изображений, регистрируемых сенсором с неизвестными нелинейными характеристиками.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Николаев П.П., Николаев Д.П. Модели константного зрительного восприятия. III. Спектральные и перцептивные инварианты в процедурах зрительной обработки // Сенсорные системы. 1997. Т. 11. N 2. С. 181-204.
2. Kim S.K., Nikolayev D.P. Method and apparatus for sectioning image into plurality of regions // US patent application publication. Application No. 09/983032. Publication No. US 2002/0102017 A1. 2002. 25 p.
3. Nikolaev D.P., Nikolayev P.P. Linear color segmentation and its implementation // Computer Vision and Image Understanding. 2004. V. 94 (Special issue on colour for image indexing and retrieval). P. 115-139.
4. Kim S.G., Nikolayev D.P. Method and device for classifying areas of image // Korean patent abstracts. Application No. 1020000069490. Publication No. 1020020039721. KIPO. 2002. 1 p.

5. Николаев Д.П., Николаев П.П. Быстрый алгоритм выделения объектов, основанный на линейной модели формирования спектрального стимула // Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'02 и CAD-2002. М.: Изд-во Физико-математической литературы. 2002. С. 410-416.
6. Николаев Д.П., Божкова В.П., Николаев П.П. Кластеризация в цветовом пространстве как метод сегментации изображения, полученного с нелинейного сенсора // Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'03 и CAD-2003. М.: Изд-во Физико-математической литературы. 2003. С. 314-321.
7. Николаев Д.П., Николаев П.П. Гауссовская спектральная модель и её особенности в задаче цветовой константности // Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'03 и CAD-2003. М.: Изд-во Физико-математической литературы. 2003. С. 321-327.
8. Nikolaev D.P. Segmentation-based binarization method for color document images // Pattern recognition and image understanding. Proceedings of 6th German-Russian Workshop (OGRW-6). Novosibirsk 2003. P. 190-193.
9. Николаев П.П., Николаев Д.П. Сравнительный анализ гауссовской и линейных спектральных моделей в задаче оценки окраски // Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'04 и CAD-2004. М.: Изд-во Физматлит. 2004. Т. 2. С. 323-328.
10. Карпенко С.М., Николаев Д.П., Николаев П.П., Постников В.В. Быстрое преобразование Хафа с управляемой робастностью // Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'04 и CAD-2004. М.: Изд-во Физматлит. 2004. Т. 2. С. 303-309.
11. Nikolaev D.P., Bozhkova V.P., Nikolayev P.P. Linear color segmentation and its implementation // Perception. 2002. V. 31 (Supplement). P. 67-68.
12. Nikolaev D.P., Nikolayev P.P. Estimation of reflectance properties following color segmentation (Colour constancy model using colour segmentation data) // Perception. 2002. V. 31 (Supplement). P. 138.