

## СОДЕРЖАНИЕ

<b>Пленарное заседание</b>	<b>6</b>
<i>Н. В. Лысенко</i>	6
Телевидение в ЛЭТИ	
<i>А. А. Умбиталиев, А. К. Цыбулин, А. А. Манцевов, И. А. Зубакин, А. В. Морозов</i>	9
Адаптация телевизионных систем как следствие максимизации	
качества информации	
<i>В. В. Хрящев</i>	17
Опыт преподавания истории телевидения в дисциплине магистратуры	
«История и методология науки»	
<i>В. В. Хрящев</i>	21
К вопросу о неэталонной оценке качества телевизионных изображений	
<i>А. В. Кузичкин, А. Ю. Аганов, А. А. Таранов</i>	25
Комплекс для управления телевизионной инфраструктурой	
сложных территориально-распределенных объектов	
<i>И. Л. Бурцев, Т. Ю. Мамаева, А. К. Рыжаков, В. Я. Яцук</i>	30
Доступная ЭКБ ИП для бортовой и наземной РЭА	
<b>Секция 1. Методы и устройства формирования</b>	<b>32</b>
<b>видеосигнала</b>	
<i>В. С. Дюмин, М. В. Четвергов, М. Г. Выдревич</i>	32
Фоточувствительный прибор с зарядовой связью	
с встроенным регистром умножения	
<i>М. В. Четвергов, А. В. Корюшкин, В. В. Петров, В. А. Локтев</i>	34
Метрологическое обеспечение измерений параметров ФПЗС	
<i>Т. С. Ларина, Е. В. Рыпакова, Д. Н. Сергеев, М. В. Четвергов</i>	36
Конструктивные особенности матричных фоточувствительных приборов	
с зарядовой связью для космических систем	
<i>Г. И. Вишневский, М. В. Четвергов, Н. С. Чепилко, М. Г. Выдревич, А. Г. Попов</i>	38
Производство фоточувствительных КМОП микросхем	
<i>Д. В. Бородин</i>	40
Параметры для оптимизации ячейки КМОП фотоматрицы	
<i>П. С. Баранов, Д. А. Белоус, В. Т. Литвин, А. А. Манцевов</i>	44
Темновые токи твердотельных фотоприемников	
при высокой рабочей температуре	
<i>Д. А. Белоус</i>	50
Оптимизация температурного режима твердотельного фотоприемника	
при работе в ближней инфракрасной области спектра	
<i>П. В. Столбов</i>	54
Высокоскоростная специализированная матричная камера	
с расширенным динамическим диапазоном и сжатием изображения	
<i>В. С. Шулика, А. А. Шведов</i>	58
Современные матрицы для получения изображения в ультрафиолетовом,	
видимом и инфракрасном спектральных диапазонах	
<i>И. В. Богданов, А. Н. Величко, А. В. Степовой, В. С. Шеленговский</i>	59
Метод физического моделирования температурного контраста	

<i>И. В. Богданов, В. В. Сгадлев, П. А. Сенченков, А. В. Степовой</i>	63
Математическая модель измерения уровня сигнала для юстировки камер	
<i>П. С. Баранов, Д. А. Белоус, Н. Н. Какушкин, А. А. Манцеветов</i>	65
Наблюдение нагретых тел кремниевыми фотоприемниками	
<i>П. С. Баранов, И. О. Позолотин</i>	70
Оптимизация спектральной характеристики телевизионной камеры в условиях переменной освещенности	
<i>Р. С. Ермолаев</i>	75
Измерение дисторсии объективов	
<b>Секция 2. Видеоаналитика</b>	77
<i>Ш. С. Фахми, Ю. И. Мукало, И. А. Гаврилов, Е. В. Костикова</i>	77
Обнаружение и классификация транспортных средств видеопотока в реальном времени	
<i>М. Альмахрук, И. А. Гаврилов, Ю. И. Мукало, Ш. С. Фахми</i>	86
Обнаружение и распознавание транспортных средств по триангуляционным сеткам	
<i>Б. С. Тимофеев, А. А. Мотыко</i>	94
Телевизионная система анализа пассажиропотока	
<i>А. Л. Приоров, М. А. Кулагин</i>	101
Применение нейронной сети в задаче детектирования железнодорожных светофоров	
<i>А. И. Бобровский, Ш. С. Фахми, Я. В. Алексеенко, Е. П. Ермаков</i>	106
Видеосистемы на кристалле селекции объектов на примере обнаружения лесных пожаров	
<i>А. И. Бобровский, Ш. С. Фахми, Ю. И. Мукало, И. А. Гаврилов</i>	114
Новый метод обнаружения объектов на звёздном фоне с использованием регулярных решёток	
<i>И. С. Фомин, Д. А. Громошинский, А. В. Бахшиев</i>	124
Обнаружение визуальных ориентиров в задачахстыковки с помощью глубокой нейронной сети	
<i>Л. И. Ивановский, В. В. Хрящев, Д. Е. Храбров</i>	130
Распознавание эмоций по изображению лица человека на основе сверточных нейронных сетей	
<i>А. А. Лебедев, В. В. Хрящев, В. А. Павлов</i>	133
Разработка алгоритма биометрической идентификации по изображению лица на основе сверточных нейронных сетей	
<i>М. В. Кипке, Н. Н. Гаврютин, Т. М. Сухов, А. А. Сорокин, Ю. А. Балошин, М. А. Сухова, А. В. Арсеньев, С. В. Василевич</i>	137
Выделение контура тела человека по цвету кожи на бытовом фоне в фотоизображении	
<i>Е. П. Петров, Н. Л. Харина, Е. Д. Ржаникова</i>	142
Метод сокращения времени передачи видео с борта микро беспилотного летательного аппарата	
<i>В. П. Кирнос, А. Л. Приоров, В. А. Антипов, В. А. Ковкина</i>	148
Алгоритм генерации базы изображений подстилающей поверхности	
<i>В. П. Кирнос, А. Л. Приоров, В. А. Антипов, В. А. Ковкина</i>	152
Проблемы картирования пространства в задаче SLAM	

<i>A. A. Чиркунова</i>	156
Оптимизация спектральной характеристики телекамеры при обнаружении малоконтрастных объектов	
<i>M. A. Занина, B. A. Павлов</i>	161
Совместное применение алгоритма TLD и фильтра частиц для задачи сопровождения объектов	
<b>Секция 3. Цифровая обработка и кодирование видеинформации</b>	<b>165</b>
<i>Д. С. Андреев</i>	165
Метод обработки разреженной карты глубины	
<i>И. В. Жигулина</i>	171
Выбор апертуры фильтра при двумерной дискретной квазиптимальной фильтрации	
<i>H. A. Обухова, A. A. Поздеев</i>	176
Методы контрастирования медицинских ТВ изображений	
<i>A. Г. Шоберг, K. A. Шоберг</i>	182
Сравнение изображений в частотной области при блочной форме выполнения преобразования	
<i>H. A. Андриянов</i>	185
Сжатие изображений на основе моделей случайных полей с изменяющимися параметрами	
<i>E. П. Петров, Н. Л. Харина, П. Н. Сухих</i>	189
Метод внутrikадрового предсказания	
<i>Нгуен Ван Чыонг, Доан Тиен Бан</i>	194
Исследование блочных алгоритмов кодирования преобразованием в стандартах сжатия мультимедийных данных	
<i>Ю. В. Кузнецов, A. A. Щаденко</i>	198
Оптимальное кодирование изображений в печати	
<i>A. И. Бобровский, M. M. Альмахрук, Ш. С. Фахми, A. Салем</i>	207
Саккадические движения глаз и концепция кодирования изображений по опорным точкам	
<i>H. A. Обухова, И. Г. Зубов</i>	212
Автоматический метод формирования области интереса в прикладных телевизионных системах медицинского назначения	
<i>C. B. Gross</i>	216
Встраиваемая система коррекции курса самолета на базе частичного декодирования сигнала H.264/AVC	
<i>B. A. Сухарев, A. B. Богословский</i>	221
Моделирование турбулентных искажений	
<i>A. Н. Бучатский, C. П. Куликов</i>	224
Сравнительный анализ характеристик измерительных приборов для контроля параметров сигнала цифрового эфирного телевидения стандарта DVB-T2	
<i>A. B. Проконич</i>	226
Методы повышения качества видеинформации в подводной среде	
<i>A. K. Цыцулин, И. А. Зубакин</i>	228
Проблема разделения информации на доминантную и фоновую	
<i>Г. М. Лабков</i>	230
Исследование областей встраивания видеостеганоконтейнера	

## Пленарное заседание

Руководитель пленарного заседания - заведующий кафедрой «Телевидение и Видеотехника» СПбГЭТУ «ЛЭТИ», профессор Н. В. Лысенко

---

**Н. В. Лысенко**  
СПбГЭТУ «ЛЭТИ»

### Телевидение в ЛЭТИ

*Приведены исторические данные о начале развития телевидения и телевизионной подготовке специалистов в старейшем электротехническом вузе Европы – ЛЭТИ. Представлены основные направления развития телевизионной техники в XXI веке в ЛЭТИ.*

#### **Телевидение, исследования в области телевизионной техники, история развития телевидения в ЛЭТИ**

В истории ЛЭТИ были предпосылки к подготовке специалистов по телевидению, как средству коммуникации еще в начале прошлого века. Так, например, в журнале «Электротехник» (1900 г., № 6) отмечалось, что в Электротехническом институте известный деятель РТО К.Д. Перский 18 и 19 марта 1900 года читал лекцию «О способах видеть на больших расстояниях при помощи электричества». Чтение лекции сопровождалось демонстрацией «световых картинок» (т.е. слайдов) с чертежами, поясняющими устройство аппаратов П.И. Бахметьева, П. Нипкова, Я. Щепаника и др. Лектор впервые публично объявил о патентовании «молодым ученым А.А. Полумордвиновым» телевизионной системы для передачи цветного изображения путем последовательной передачи сигналов трех основных цветов. В том же году инженер-технолог Полумордвинов для пополнения знаний в области электротехники поступил на 3-й курс Электротехнического института, а по окончании этого вуза работал в нем несколько лет в качестве преподавателя телефонии [1]. В краткой биографии он сообщил о получении субсидии 2000 руб. от Военного министерства на разработку изобретенного им устройства, но не указал, насколько удалось ему продвинуться в реализации изобретения.

В советские годы в ЛЭТИ читал лекции и работал со студентами основоположник электронного телевидения Б.Л. Розинг. Один из энтузиастов механического телевидения профессор ЛЭТИ В.А. Гуров, автор учебника «Основы дальновидения» (1936) поставил в институте курс «Телевидение» и организовал в 1938 г. лабораторию телевидения при кафедре телемеханики.

В ноябре 1945 г. при поддержке академика А.И.Берга в ЛЭТИ была образована кафедра телевидения – а ее заведующим назначен профессор Я.А.Рыftин.

Вначале деятельность небольшого коллектива кафедры в составе профессора Я.А. Рыftина, доцентов О.Б. Лурье и О.А. Александрова была сконцентрирована на разработке учебных планов, определении тематики научных исследований, определении содержания учебных дисциплин. Содержание курсов, раскрывающих основы телевизионной техники, совершенствовалось вместе с развитием телевидения и наполнялось содержанием из научной периодики и собственных исследований, которые отчасти базировались на

оптико-механических системах. К последним можно отнести работы Я.А. Рыфтина «К расчету параметров телевидения» (1934) и «Разрешающая сила телевизионных систем» (1947). Можно также вспомнить, что О.Б. Лурье руководил работами по созданию оптико-механической аппаратуры довоенного Киевского телецентра и не без юмора говорил о своем анти-приоритете (был последним создателем оптико-механической системы). В дальнейшем на кафедре получили развитие такие научные направления как теоретические и практические проблемы повышения эффективности телевизионных систем и качества телевизионной передачи (М.В.Антипин); разработка основ анализа и инженерных принципов расчета видеоусилителей (О.Б.Лурье); были разработаны новый метод телевизионной передачи «скользящим растром» и метод оценки разрешающей способности телевизионной системы по испытательной таблице с зонами Френеля (Я.А.Рыфтин, М.В.Антипин); разработан новый способ повышения качества телевизионной передачи путем анаморфирования раstra и изображения (Я.А.Рыфтин). За годы руководства кафедрой профессором Я.А.Рыфтиным (1945-1973 г.г.) на кафедре разработан и осуществлен новый стробоскопический прибор «свет-свет» (Г.А.Эйссенгардт, В.Н.Семенов); разработаны новые критерии оценки качества телевизионной передачи, метод оценки сложности изображения, а также теория разрешающей способности и чувствительности систем с контурным разложением (Ю.П.Куликовский, Н.Н.Степанов); решена одна из основных проблем телевизионной автоматики – опознание образа (Н.Н.Степанов, Я.Л.Штермер); обнаружен новый «эффект пульсации – адаптации» на заряженной мишени трубки с накоплением (Я.А.Рыфтин); разработан прибор, позволяющий получать электрические сигналы, пропорциональные угловым положениям зрительной оси непосредственно от глаза наблюдателя (исследования в области медицины, биологии, процессов мышления).

Ученики Я.А.Рыфтина Ю.П.Куликовский и М.В.Антипин, успешно защитившие в конце 60 – х годов докторские диссертации, стали впоследствии ректорами вузов Ленинграда: ЛЭИС им. М.А.Бонч-Бруевича и ЛИКИ.

В годы заведования Я.А.Рыфтиным кафедрой телевидения в ЛЭТИ была создана уникальная учебная лаборатория, в которой студенты реализовывали весь процесс проектирования, изготовления, настройки и эксплуатации поочередно всех основных устройств телевизионной системы: от усилителей до разверток и синхрогенератора. В наши дни этот подход аналогичен так называемой инициативе CDIO «Инженерное образование на основе жизненного цикла технологий, систем, процессов и изделий: **Conceive** (Задумай), **Design** (Спроектируй), **Implement** (Реализуй), **Operate** (Управляй)».

Сегодня можно сказать, что тогда в 60-е и 70-е годы кафедра телевидения, да и ЛЭТИ в целом готовили действительно инженеров – специалистов высочайшего класса. И в этом есть несомненная заслуга Я.А.Рыфтина.

В последующие годы развитие телевидения в ЛЭТИ связано с цифровой обработкой изображений и появлением твердотельных оптико-электронных преобразователей. Появляется новое направление, рассматривающее телевизионную систему и потребителя переданной ею информации – автомат или человека - как единую видеоИнформационную систему (Ю.П.Куликовский, Н.В.Лысенко); разрабатываются методы записи – воспроизведения сигналов изображения на «лазерных» дисках (Л.М.Штутман); проводятся работы по пространственному кодированию и компрессии видеоИнформации, получившими существенное развитие с появлением твердотельных оптико-электронных

преобразователей, а также цифровых средств обработки видеосигнала (С.Д.Егорова, А.М.Мончак); создается первая в СССР учебная замкнутая телевизионная система с индивидуальными видеопросмотровыми устройствами и системой обратной связи с сенсорным вводом информации (Ю.П.Куликовский, Н.В.Лысенко); реализуется визуализация невидимых информационных полей (Ю.Н.Хомяков); расширяются работы в области твердотельных телекамер и цифровых телевизионных систем, генерации и обработки сигналов цветного телевидения (Г.А.Эйссенгардт, А.А.Манцевов); проводятся работы по применению спектронального телевидения в целях дистанционного экологического мониторинга (Р.Е.Быков, Н.В.Игнатьева, Ю.М.Титов); создаются рентгенотелевизионные системы (Р.Е.Быков) и многоцветные фотометрические обзорные телевизионные системы для космических телескопов (А.А.Манцевов), телевизионно-компьютерные авиа- и автомобильные тренажеры (Н.В.Лысенко, В.И.Киврин); разрабатываются стандарты изображения для систем цифрового телевидения (Я.А.Рыфтин, Л.Л.Полосин).

В 1976-98 г.г. кафедрой руководил д.т.н., заслуженный профессор СПбГЭТУ, заслуженный деятель науки РФ, приглашенный профессор Пекинского университета Р.Е.Быков – основатель научно-педагогической школы в области цифровых телевизионных систем "Анализ и обработка цветных изображений".

Новый век открыл для телевидения и новые направления исследований и разработок: классификация изображений случайных полей; обнаружение и оценивание координат точечных объектов; контроль целостности видеоданных; разработка теории гетерогенных видеоинформационных систем; сверхскоростной трекинг движущихся объектов для стационарного и динамического видеонаблюдения; разработка высокочувствительных телевизионных камер на отечественных КМОП-сенсорах для перспективных направлений авиа- и космических технологий (обзорные камеры, камерыстыковки, камеры посадки и др.); создание телевизионных систем с контрастной чувствительностью, ограниченной фотонным шумом; разработка мультиспектральных телевизионных систем медицинского назначения; совершенствование структуры и алгоритмов кодирующих светофильтров монохромных фотоприемников; разработка методов решения задач декомпозиции на аппаратную и программную составляющие для устройств класса «система на кристалле»; совершенствование аппаратно-программных средств контроля целостности видеоданных.

#### Библиографический список

1. Урвалов В.А. Из истории цветного телевидения. О «контактах» Д.Л. Бэрда и А.А. Полумордвинова //Broadcasting. Телевидение и радиовещание. – 2013. – № 7. – С.50-51; № 8. – С.46-47.

#### TELEVISION IN LETI

**N. V. Lysenko**

Saint-Petersburg electrotechnical university "LETI"

*Given historical data about the early development of television and television training in the oldest electrotechnical University in Europe – LETI. The basic directions of development of television technology in the XXI century in LETI*

**Television, research in the field of television technology, history of development of television in LETI**

**А. А. Умбиталиев<sup>1</sup>, А. К. Цыцулин<sup>1</sup>, А. А. Манцевов<sup>1,2</sup>,**

**И. А. Зубакин<sup>2</sup>, А. В. Морозов<sup>1,2</sup>**

<sup>1</sup>АО «НИИ телевидения»

<sup>2</sup>СПбГЭТУ «ЛЭТИ»

## Адаптация телевизионных систем как следствие максимизации качества информации

*Показано, что адаптация накопления и квантования сигналов в телекамере при любых неблагоприятных условиях – изменении освещённости, скоростного смаза и т. п. является следствием и средством повышения качества информации при изменении статистик наблюдаемого сигнала.*

### **Качество информации, адаптация, накопление, квантование**

Научной школой НИИ телевидения на протяжении десятилетий развиваются два направления: теории информации в приложении к кодированию зашумлённых сигналов и адаптации прикладных телевизионных систем. После первой книги [1], впервые во весь голос заявившей о единстве целей этих двух направлений, важными вехами в развитии понимания этих двух направлений стали работы [2]–[7]. Вместе с тем идея первичности теории (здесь – теории информации), которая должна подсказывать практике (адаптации телевизионных систем) методы решения задач, нуждается в популяризации. Мелким, хотя и показательным, признаком такой потребности является то, что ещё редактор книги [1] потребовал разместить материал по теории информации после описания свойств новых тогда приборов с зарядовой связью и их существенно более широких возможностей по перестройке параметров разложения по сравнению с электронно-лучевыми фотоприёмниками. А глубинной причиной разрыва этих разделов является распространенное мнение о том, что теория информации (теория Шеннона) лишь устанавливает некие границы, а не является конструктивной теорией. Мы надеемся наглядно показать ущербность этого мнения и конструктивность связи теории и практики. Существует такой вид литературного творчества, как «повести, состоящие из одних названий», и по большому счету данный доклад может быть отнесен к этому жанру. Вместе с тем остановимся на основных понятиях, идеях и результатах влияния теории информации на адаптацию прикладных телевизионных систем. Так как мы намерены показать, что адаптация прикладных телевизионных систем является следствием решения оптимизационной задачи в рамках теории информации, то следует напомнить, как сформировалось понятие качества информации в рамках вероятностной теории информации.

Развиваемая нами концепция качества информации, во многом ориентированная именно на *кодирование источника* в прикладном телевидении, в первую очередь определяет структуру и оптимальные параметры телекамер (параметров разложения изображения в фотоприёмнике и квантования, которые являются первичными и главными элементами кодирования изображения в телекамере) и цифровых кодеров источника, обеспечивающих сжатие видеинформации без потерь.

Качество информации можно определить по-разному; научной школой НИИ телевидения развивается концепция качества как **отношения энтропии источника к информационному риску**:  $Q = H/R$ . Мы рассматриваем два вида информационного риска, которые при оценке качества информации риски суммируются,  $R = R_C + R_n$ :

- $R_C$ , связанного с влиянием соотношения энтропии источника и пропускной способности (наличия фоновой информации или проявления ненадёжности канала);
- $R_n$ , связанного с влиянием входного шума (наличием потери полезной информации и шумовой информацией – передаваемой через канал информацией о входном шуме).

У этого, вообще говоря, произвольного определения есть аналогия: в теории решений качество оценки определяется отношением дисперсии оцениваемого случайного процесса к среднему риску (среднеквадратической ошибке).

Информационный риск  $R_C$  включает избыточную информацию  $I_\Phi$  (характеризующую неполноту загрузки канала и определяемую «для удобства» как разность между пропускной способностью канала и энтропией источника) и потерю доминантной информации  $\Delta I_C$ . Эти две компоненты согласно основной теореме Шеннона для дискретного канала с шумом соответствуют ситуациям:

- $H > C$  (имеется потеря доминантной информации  $\Delta I_C$ , соответствующая «ненадёжности» канала);
- $H < C$  (из-за отсутствия сжатия сигнала имеется фоновая информация  $I_\Phi$  тем большая, чем больше избыточность канала);
- $H = C$  (информационное равновесие, при котором благодаря оптимальному сжатию  $\Delta I_C = 0$ ,  $I_\Phi = 0$ ).

В общем случае информационный риск, как и средний (байесовский) риск в теории решений включает весовые коэффициенты при входящих слагаемых, но в простейшем случае информационный риск имеет вид:

$$R_C = \Delta I_C + I_\Phi. \quad (1)$$

И фоновая (избыточная) информация и потеря доминантной информации (ненадёжность) определяются разностью энтропии источника и пропускной способности, поэтому:

$$Q_C = \frac{H}{R_C} = \frac{H}{|H - C|}. \quad (2)$$

При этом при всех соотношениях энтропии  $H$  доминантного сигнала и пропускной способности  $C$  количество передаваемой доминантной информации составляет

$$I(x, y) = \min\{H, C\}. \quad (3)$$

Качество информации зависит от соотношения энтропии  $H$  и пропускной способности  $C$  (рис. 1) и характеризуется разрывом в точке информационного равновесия  $H = C$ .

Этот разрыв, обусловленный лишь потенциальной достижимостью равенства энтропии и пропускной способности, имеет три проявления: в бесконечность обращаются не только отмеченные К. Шенноном сложность и задержка, но и качество информации (2).

К. Шеннон построил график достижимой области надёжности в зависимости от соотношения энтропии источника и пропускной способности, т. е. допускал их любое соотношение, но формулировка его теорем подразумевает стремление приблизиться к равенству энтропии источника и пропускной способности, которое возможно при сколь угодно малой ненадёжности. При этом он отмечал, что скорость  $C$  не может быть в действительности достигнута; к ней можно только приближаться в пределе, используя всё более и более сложные способы кодирования и при всём большем запаздывании в работе передатчика и приёмника. Это означает, что бесконечность качества в формуле

(2) при передаче неискажённых сигналов может быть лишь потенциальной, а не актуальной.

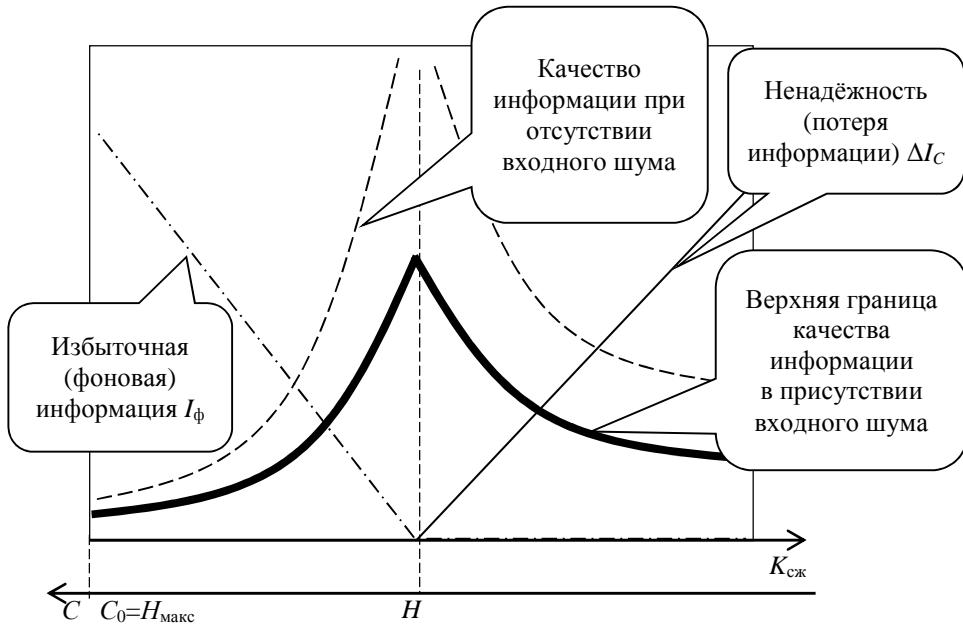


Рис. 1. Зависимость качества информации от соотношения энтропии источника  $H$  и пропускной способности канала  $C$  (нижняя шкала; максимум пропускной способности  $C_0$  соответствует максимальной энтропии источника  $H_{\max}$  при выбранном алфавите); шкала коэффициентов сжатия  $K_{\text{сж}}$  относительно  $C_0$  имеет противоположное направление

Теория кодирования зашумлённых сигналов для гауссовских процессов при использовании разложения Карунена–Лоэва и информационном равновесии (фоновая информация  $I_\phi$  равна нулю) привела к определению качества информации  $Q_n$ , лимитируемого входным шумом, которое можно записать в виде:

$$Q_n = \frac{H}{Dl + I_{\text{ш}}}. \quad (4)$$

Важным выводом теории кодирования зашумлённых сигналов, использующей формулу (4), является закон о существовании нижней границы информационного риска  $R_n$  [4]:  $\min R_n \approx 1$  бит/символ, и, как следствие, **верхней границы** (безразмерного) качества информации  $Q_n$ , численно равной энтропии  $H$ . Обобщение формул (1)–(4) позволяет определить предельное качество информации через сумму информационных рисков, обусловленных различием энтропии источника и пропускной способности  $R_C$  и влиянием входного шума  $R_n$ .

$$Q = \frac{H}{R_C + R_n} = \frac{H}{(Dl_C + I_\phi) + (Dl_n + I_{\text{ш}})}. \quad (5)$$

Формула (5) имеет следующие характерные точки:  $Q = 0$  при  $H = 0$ ;  $Q = H$  при  $H = C$ ;  $Q = 1$  при  $H/C \rightarrow \infty$ . Предельное качество  $Q = H$  потенциально достижимо при выполнении ряда условий, в том числе оптимального кодирования как с точки зрения подавления фоновой информации, т. е.  $R_C = 0$ , так и с точки зрения компромисса между потерей полезной информации и загрузкой канала связи передачей информации о входном шуме, т. е.  $R_n = 1$ .

При кодировании зашумлённых сигналов предельное значение качества информации (5) имеет конечное значение (рис. 1).

Применение концепции качества информации к сигналам, формируемым ТВ камерой, проведём на уровне синтаксической информации, измеряемой на выходе АЦП с акцентом на ключевых операциях в ТВ камере, ответственных за повышение качества информации – накоплении и квантовании. При этом традиционно воспользуемся оценкой сверху количества информации, без учёта корреляционных связей изображения, учитываемых при построении цифрового кодера сжатия сигналов, а также простой формулой информационного риска  $R_n$  без весовых коэффициентов. Это позволяет использовать простую формулу для качества информации (5), в которой все величины измеряются в битах на пикселе.

Для оценки влияния квантования на качество информации приведём (5) к виду, использующему не только число  $N$  пикселов, но и число *измеренных* элементов разрешения  $N_p$  (из-за шума, влияния объектива, смаза и т. п.  $N_p \leq N$ ), а также энтропийную мощность полезного сигнала  $D_{Hc}$ , измеренную дисперсию  $\sigma^2$  шума на изображении (приняв её гауссовой и одинаковой по всему изображению, равной дисперсии шума на «сером» участке изображения) и энтропийную мощность шума квантования  $D_{Hkb}$  (определенную числом уровней квантования):

$$Q = \frac{\frac{N_p^2}{2} \log \frac{D_{Hc}}{\sigma^2}}{\left| \frac{N_p^2}{2} \log \frac{D_{Hc}}{\sigma^2} - N^2 C \right| + \frac{N_p^2}{2} \log \frac{\infty}{\epsilon} + \frac{D_{Hkb}}{\sigma^2} \frac{\infty}{\emptyset} + \frac{N^2}{2} \log \frac{\infty}{\epsilon} + \frac{\sigma^2}{D_{Hkb}} \frac{\infty}{\emptyset}}. \quad (6)$$

Это даёт такую картину: существует максимум качества информации при оптимальных дискретизации и квантовании. При переменных условиях наблюдения такая оптимизация трансформируется в адаптацию. Логика увеличения интервалов накопления при ухудшении условий наблюдения (падения освещённости, смаз, расфокусировка и т. п.) состоит в том, что эти факторы уменьшают энтропию сигнала, и для максимизации качества информации необходимо понизить требуемую пропускную способность, сократив число отсчётов сигнала («эффективных» пикселов и/или кадровую частоту).

Измерение зависимости качества информации от числа уровней квантования (рис. 2) показали наличия оптимума, а зависимости от освещённости дало очевидный результат (рис. 3): до определённого предела (насыщение пикселя, после чего в ТВ камере включается автоматическая регулировка времени накопления, АРВН) наблюдается монотонный рост качества.

Измерения качества информации в зависимости от степени бинирования, т. е. группирования сигналов различного количества пикселов при наблюдении в «жёстких» условиях малой освещённости (рис. 4) или скоростного смаза (рис. 5) показали наличие максимума качества информации, достигаемого при тем больших размерах зоны накопления, чем хуже условия наблюдения – идея, впервые рассмотренная на начальном этапе твердотельной революции в телевидении [1].

Общность философии дискретизации по полю и квантования амплитуды проявляется в том, что их оптимальность определяется минимумом информационного риска и максимумом качества информации (5). Кроме того, при оптимальном квантовании амплитуды сигнала гистограмма шума содержит три отсчёта: так же, как и число отсчётов при оптимальной пространственной дискретизации ( $3 \times 3$ ) изображения в задаче измерения координат точечных объектов (звёзд) [3], [4].

Так как максимум качества информации при наличии скоростного смаза (рис. 5) зависящий от площади зоны накопления (количество суммируемых пикселов), во многом определяется значением скорости смещения объекта в течение времени кадра (времени накопления), то в зависимости от имеющейся априорной информации о сюжете могут

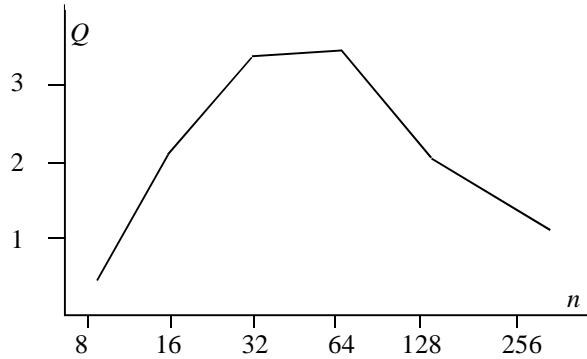


Рис. 2. Зависимость качества информации от числа  $n$  уровней квантования

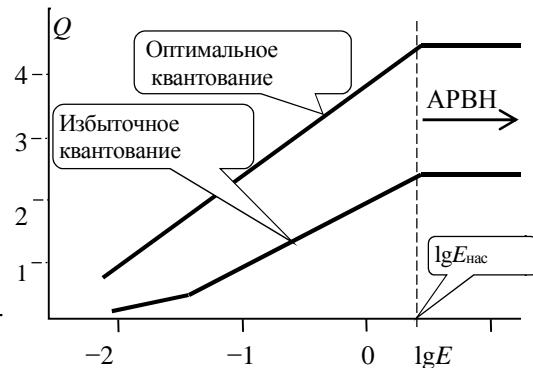


Рис. 3. Зависимость качества информации от освещённости

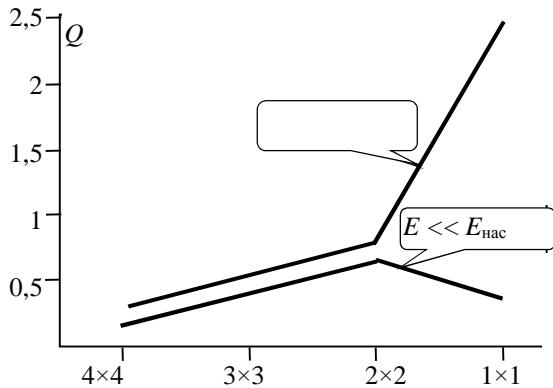


Рис. 4. Зависимость качества информации от бинирования при разных освещённостях

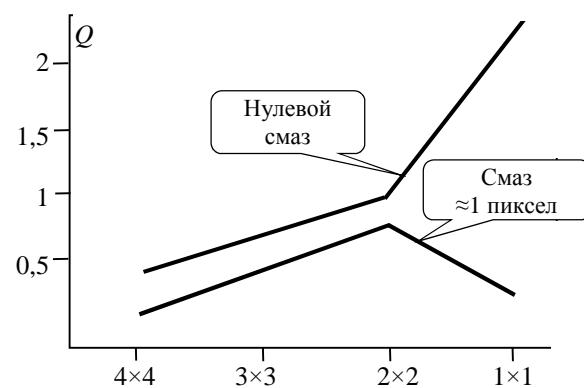


Рис.5. Зависимость качества информации от бинирования при скоростном смазе,  $E \approx E_{\text{нac}}$

быть поставлены две задачи: при постоянном скоростном смазе – оптимизация размеров зоны накопления; при переменном скоростном смазе – адаптация телевизионной системы.

Если решать задачу максимизации качества информации (например, на основе минимума динамического промаха визирной оси системы относительно наблюдаемого объекта), то задача адаптации усложняется. Может оказаться, что требуемые оптимальные значения размера пикселя и кадровой частоты *одновременно* не достижимы. Отсюда возникает идея *взаимообмена кадровой частоты и чёткости изображения*: на большой дистанции актуальна предельная чёткость изображения, на близкой – предельная кадровая частота. Идея такого изменения [4], [6], [7] во многом является следствием итерационного принципа контроля сближения истыковки космических аппаратов, выдвинутого С. П. Королёвым, и является следствием простого принципа: *для каждого сюжета существует свой оптимальный видеоряд*.

Не останавливаясь на адаптации квантования к динамике сюжета [8] (динамический диапазон сигнала монотонно увеличивается при уменьшении кадровой частоты, для чего в фотоприёмных матрицах часто делают два выходных устройства: «быстрый» и «малошумящий»), покажем влияние концепции качества информации на адаптацию пространственно-временной дискретизации изображений при наблюдении подвижных объектов.

Основная идея адаптации телевизионной системы к изменению скоростного среза состоит в том, что при ограниченной скорости считывания сигнала с фотоприёмной матрицы на численное значение среза оказывает влияние не только изменение площади зоны накопления (числа суммируемых пикселов), но возможное (или требуемое) изменение кадровой частоты. Так, если увеличить размер зоны накопления, то в силу уменьшения числа считываемых «эффективных» пикселов открывается возможность повышения кадровой частоты для уменьшения значения скоростного среза.

В связи с очень часто имеющейся недостаточно подробной дискретизацией многомерного поля облучённостей основным компонентом качества информации становятся ошибки дискретизации. Эти ошибки трактуются как ошибки оценки формы сигнала, для неслучайных моделей сигнала могут трактоваться как промахи системы наведения на движущуюся цель [4], [6], [7]. Формализация качества  $Q$  формируемой информации в таком варианте доминирования ошибок дискретизации соответствует минимизации среднеквадратической ошибки  $\epsilon$  оценивания формы сигнала при ограничении скорости  $C_0$  считывания видеинформации с фотоприёмной матрицы. При этом вклад шумовой информации в потерю качества учитывается компромиссным выбором разрядности цифрового сигнала, определяемой ёмкостью пикселя  $N_e$  в электронах [3]:

$$Q = \frac{H}{DI + I_{\text{ш}}}, \quad @ C \leq C_0, \epsilon > \epsilon_{\text{мин}} \quad Q \gg \frac{H}{\frac{1}{2} \log \frac{\epsilon}{\epsilon_{\text{мин}}}}. \quad (7)$$

Оптимизация предполагает два этапа: определение оптимальных параметров многомерной дискретизации и ключевых правил формирования управлений, затем, исходя из дополнительного ограничения дискретности изменения параметров дискретизации и условия устойчивости системы управления параметрами ТВ камеры, определяются свойства замкнутой системы управления [7].

Задача на поиск минимума суммарной ошибки оценивания случайного поля достаточно просто решается при гипотезе о симметрии пространственной и временной автокорреляционных функций. Эта задача относится к *изопериметрическим* (её классический вариант – нахождение фигуры максимальной площади при ограниченном периметре), в которой в роли переменных выступают парциальные ошибки оценивания  $\epsilon_j = D_{\Delta j}$  (дисперсии приращений сигнала), а в роли «периметра» – доступная скорость  $C_0$  считывания информации с фотоприёмной матрицы. В статистической теории связи примером изопериметрической задачи является нахождение распределение заданной ошибки по спектру передаваемого сигнала (эпсилон-энтропии). Очень важный результат теории информации сформулирован А. Н. Колмогоровым, полученный путём решения вариационной задачи (уравнение Эйлера с множителями Лагранжа) на поиск минимума информации (эпсилон-энтропии, или *rate distortion* по Шеннону): при передаче многомерного сигнала ошибки передачи должны быть распределены *равномерно по всем степеням свободы*.

Из максимизации качества информации следует наличие *оптимального (ненулевого) скоростного среза*, и вытекают три важных вопроса [7]:

- какие значения должны иметь оптимальные пространственные и временные интервалы накопления (проблематика теории информации);

• как должны формироваться управляющие сигналы для адаптации пространственной и временной чёткости (проблематика теории управления);

• какими свойствами должна обладать система для обеспечения устойчивости (переход системы из одного состояния в другое не должен вести к немедленному переходу в исходное состояние; проблематика теории управления, в частности статистической динамики радиотехнических следящих систем).

Из решения изопериметрической задачи на максимум при ограничении (7) вытекает главный теоретический результат [4], [6], [7] – **правило динамического выравнивания дисперсий приращений сигнала по всем аргументам** (межэлементных, межстрочных и межкадровой разностей видеосигнала):

$$D_{\Delta x} = D_{\Delta y} = D_{\Delta t}. \quad (8)$$

Эти статистики телевизионных сигналов иногда оказываются далёко не равны, например, при очень больших дистанциях до кооперируемого космического аппарата межкадровая разность очень мала, при очень малых – очень велика. При специальном управлении параметрами разложения они могут иметь близкие значения (6).

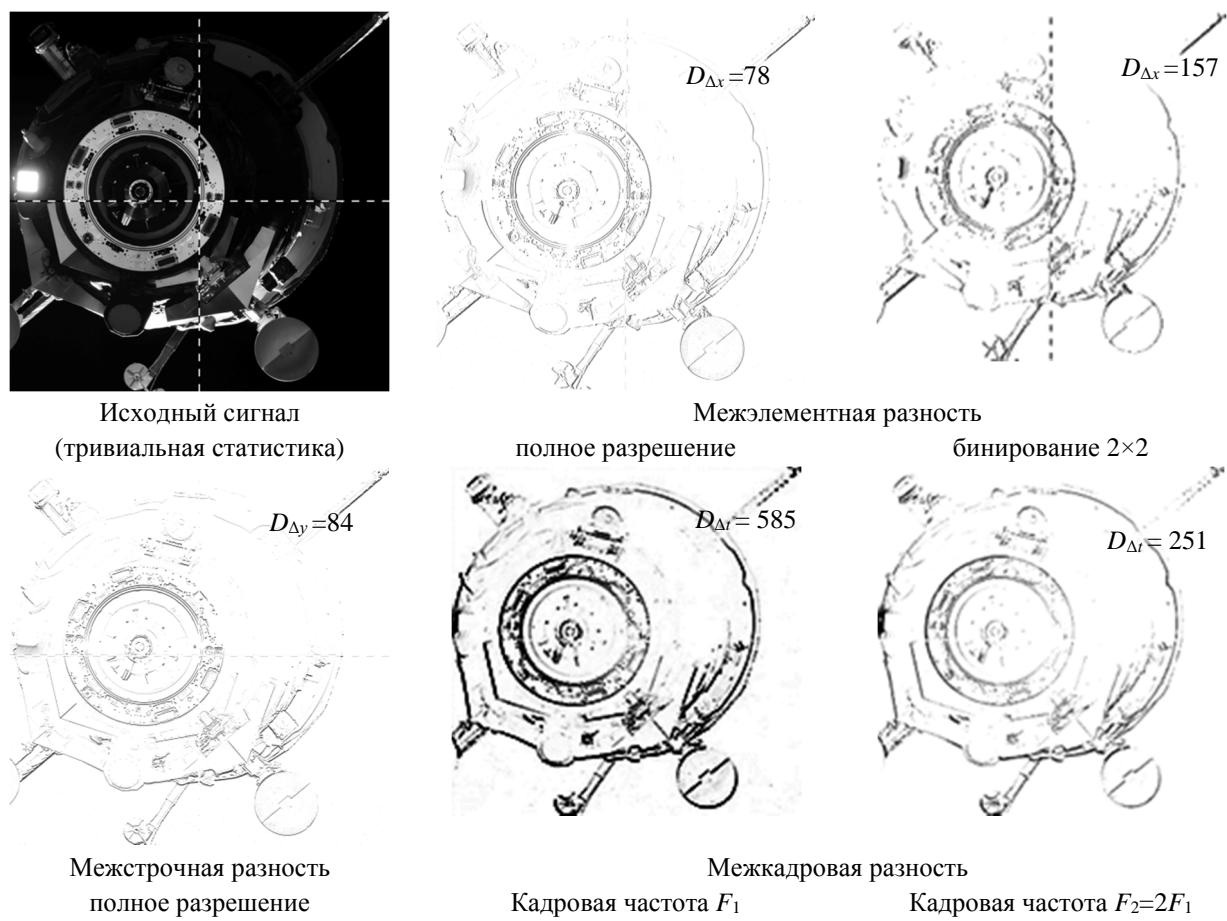


Рис. 6. Статистики телевизионного сигнала (приращения – негатив, усилено)

При переходе из одного состояния в другое возможны различные варианты стабилизации уровня сигнала, формируемого ТВ камерой. Например, когда при принятии решения о необходимости повышения кадровой частоты пиксели группируются по 4 («бининг» вида  $2 \times 2$ ) и накопленный за кадр видеосигнал возрастает в 4 раза, а кадровая частота возрастает вдвое, поэтому накопленный за кадр видеосигнал уменьшается в 2 раза и для обеспечения устойчивости системы необходимо дополнительное двукратное сокращение времени накопления [4].

Правило (8) динамического выравнивания дисперсий межэлементных и межкадровой разностей видеосигнала в совокупности с гипотезой об изотропности изображения по полю (дисперсии межэлементных и межстрочных разностей примерно равны, см. рис. 6) диктует построение системы управления с использованием формирователя *отношения* внутrikадровых и межкадровых статистик (дисперсий первых конечных разностей). Реализация такой адаптивной телевизионной системы показала её устойчивость и эффективность в смысле максимума качества информации при изменении статистик сигнала при наблюдении приближающегося объекта.

Таким образом, адаптация накопления и квантования в телекамере имеет теоретическую основу в виде концепции качества синтаксической информации. В частности, *адаптация* телевизионной системы с обменом чёткости на кадровую частоту не только обосновывается с позиций *теории* информации, но и является воплощением введённого С. П. Королевым *практического* принципа итерационного сближения космических аппаратов.

#### Библиографический список

1. Твердотельное телевидение /Л. И. Хромов, Н. В. Лебедев, А. К. Цыцуллин, А. Н. Куликов. М.: Радио и связь, 1986. 184 с.
2. Хромов Л. И., Цыцуллин А. К., Куликов А. Н. Видеоинформатика. М.: Радио и связь, 1991. 192 с.
3. Твердотельная революция в телевидении: Телевизионные системы на основе приборов с зарядовой связью, систем на кристалле и видеосистем на кристалле/ Под ред. А. А. Умбиталиева и А. К. Цыцуллина. – М.: Радио и связь, 2006. 312 с.
4. Твердотельные телекамеры: накопление качества информации/ Цыцуллин А. К., Адамов Д. Ю., Манцевтов А. А, Зубакин И. А./ СПб., Изд-во СПбГЭТУ «ЛЭТИ», 2014. 272 с.
5. Цыцуллин А. К. Развитие теории информации научной школой НИИ телевидения. Вопросы радиоэлектроники, серия техника телевидения, 2015, вып. 3. С. 66–87.
6. Управление режимом накопления в твердотельных фотоприёмниках / Умбиталиев А. А., Цыцуллин А. К., Манцевтов А. А. и др./Оптический журнал, 2012, № 11, с. 84–92.
7. Статистический анализ и синтез прикладных телевизионных систем/ Умбиталиев А. А., Пятков В. В., Цыцуллин А. К. и др./ Вопросы радиоэлектроники, серия техника телевидения, 2015, вып. 6. С. 4–14.
8. Видеосистема на кристалле с адаптацией к сюжету / Цыцуллин А. К., Левко Г. В., Морозов А. В. и др. Патент РФ № 2612191 от 03.03.2017. Опубл. БИ № 7.

#### ADAPTATION OF TELEVISION SYSTEMS AS A CONSEQUENCE OF MAXIMIZING THE QUALITY OF INFORMATION

**A. A. Umbitaliev<sup>1</sup>, A. K. Tsytsulin<sup>1</sup>, A. A. Mantsvetov<sup>1,2</sup>, I. A. Zubakin<sup>2</sup>, A. V. Morozov<sup>1,2</sup>**

<sup>1</sup>Television Research Institute

<sup>2</sup>Saint-Petersburg electrotechnical university "LETI"

*It is shown that the adaptation of accumulation and quantization of signals in the television camera under any unfavorable conditions-changes in illumination, speed lubrication, etc., is a consequence and means of improving the quality of information when the statistics of the observed signal change.*

**Information quality, adaptation, accumulation, quantization**

**В. В. Хрящев**  
*Ярославский государственный университет им. П. Г. Демидова*

## Опыт преподавания истории телевидения в дисциплине магистратуры «История и методология науки»

*Проводимый инновационный образовательный проект направлен на разработку и аprobацию системы смешанного (blended) обучения для магистрантов технических и компьютерных специальностей. В качестве пилотного проекта выбран гуманитарный курс «История и методология науки» для магистрантов ЯрГУ, обучающихся по направлениям подготовки «Инфокоммуникационные технологии и системы связи», «Радиотехника», «Радиофизика».*

**История науки, история телевидения, Зворыкин, Сарнов, образовательный стартап**

### Система смешанного обучения

Вопросы изучения истории телевидения рассматриваются в ЯрГУ в 1 семестре соответствующего курса обучения в магистратуре «История и методология науки» [1, 2]. Автор данной работы уже 10 лет преподаёт часть этого курса, посвященную истории науки магистрантам специальностей «Инфокоммуникационные технологии и системы связи», «Радиотехника», «Радиофизика». Наработанный практический опыт позволил сформировать следующие основные проблемы образовательного процесса по данной дисциплине:

- низкой активностью студентов магистратуры в течении семестра;
- на современном этапе общения со студентами в рамках данного курса классические методики чтения лекций под запись или озвучивание презентаций работают не эффективно;
- наблюдается крайне низкое умение студентами технических специальностей излагать свои мысли письменно;
- де-факто институт библиотеки для студентов технических специальностей не существует;
- низкий интерес студентов к поиску и приобретению печатных изданий по профилю курса.

В тоже время доступность различных электронных материалов в сети интернет за последние 10-15 лет существенно возросла. В сети интернет доступно огромное количеством качественных материалов по истории науки по профилю изучаемой специальности (фильм Леонида Парфенова «Зворыкин-Муромец», фильмы про Николу Тесла, видеоматериалы про изобретение радио, материалы про возникновение и развитие Кремниевой Долины), имеющихся в свободном доступе. Все это заставляет преподавателя обратить внимание к методологии смешанного (blended) обучения, при котором аудиторные часы тратятся в основном на обсуждение и интерпретацию материала, с которым студенты знакомятся дома самостоятельно [3]. Для внедрения системы смешанного обучения в ЯрГУ им. П.Г. Демидова совместно с малым инновационным предприятием ООО «Центр эффективных коммуникаций» была разработана соответствующая облачная образовательная система под брендом Study24Seven [4].

Для ее разработки были выбраны направления кроссплатформенной мобильной и веб-разработки. Самыми популярными на данный момент являются мобильные платформы iOS и Android. Преимуществом разработки является реализация приложения под несколько, изначально не совместимых, платформ с использованием одного кода. Это позволяет в значительной мере снизить стоимость разработки, так как достаточно одного разработчика для реализации мобильного приложения под все платформы. При наличии положительных сторон всегда появляются негативные моменты. Отрицательными моментами кроссплатформенной разработки является ограниченные возможности приложений при сравнении с нативными приложениями. Второй недостаток является, в некоторой степени, преимуществом – общий дизайн пользовательского интерфейса. Не важно, под какую платформу разрабатывается приложение, оно будет иметь одинаковый интерфейс и логику экранов.

Для разработки кроссплатформенного мобильного приложения был выбран фреймворк Phonegap. PhoneGap, называемый также Apache Callback, основан на Apache Cordova, что включает в себя множество разработанных ранее библиотек по взаимодействию со встроенными сервисами смартфонов. Данный инструмент позволяет создать приложения для мобильных устройств используя JavaScript, HTML5 и CSS3, без необходимости знания «родных» языков программирования (например, Objective-C), под все мобильные операционные системы (iOS, Android, Bada и т. д.). Готовое приложение компилируется в виде установочных пакетов для каждой мобильной операционной системы. Была разработана система быстрой авторизации с 4 вариантами: через электронную почту, facebook, vk, twitter.

### **Преподавание истории телевидения**

Одним из основных вопросов в рамках курса «История и методология науки» является вопрос об истории создания, развития и продвижения на рынок телевизионных систем. В рамках данной проблемы студентам предлагается ознакомиться самостоятельно с документальным фильмом Леонида Парфенова «Зворыкин-Муромец», дополнительными материалами, подобранными преподавателем и написать эссе на тему: «Кто внес больший вклад в развитие телевидения – Зворыкин или Сарнов?» (рис. 1). Данное эссе размещается в системе Study24Seven и проверяется преподавателем (рис. 2). Во время аудиторных занятий происходит обсуждение наиболее оригинальных и убедительных доводов студентов по указанной теме. Анализируются аналогии, например Стив Джобс и Стив Возняк. По накопленной за последние 5 лет статистике примерно 80-90% студентов в дебатах отстаивают позицию Зворыкина.

Таким образом, использование системы Study24Seven позволило повысить мотивацию студентов на изучение соответствующего раздела курса, лучше сбалансировать аудиторную нагрузку и самостоятельную работу студентов.



а)

б)

Рис.1. Творцы истории телевидения на первоначальных этапах:  
Владимир Козьмич Зворыкин (а); Давид Абрамович Сарнов (б)

Рис. 2. Web-интерфейс системы Study24Seven

### Заключение

Разработанное мобильное приложение Study24Seven является кроссплатформенным и имеет реализацию под современные версии операционных систем iOS и Android. Данний результат достигнут с помощью архитектуры приложения, а также за счет использования технологии PhoneGap iOnic. Данний подход позволил увеличить скорость разработки, а также уменьшить стоимость будущей поддержки за счет работы над одним кодом, который адаптивен к разным платформам. Разработанное мобильное приложение доступно к использованию на платформах Android и iOS и загружено в соответствующие магазины приложений Google Play и AppStore. Разработка соответствующей методики системы смешанного образования повысит эффективность при обучении студентов технических специальностей гуманитарным дисциплинам.

Работа выполнена при поддержке гранта Фонда Владимира Потанина.

### Библиографический список

1. М. А. Быховский. Развитие телекоммуникаций. На пути к информационному обществу. История телеграфа, телефона и радио до начала XX века. Либроком, 2013.
2. Kovarik, B., 2011. Revolutions in Communication: Media History from Gutenberg to the Digital Age. Bloomsbury Academic.
3. Horn, M. et al., 2014. Blended: Using Disruptive Innovation to Improve Schools. Jossey-Bass, Wiley.
4. Khryashchev V., Kasatkina N., Sokolenko D. Readiness for mobile learning: Multidisciplinary cases from Yaroslavl State University // Proceedings of the 13th International conference "Mobile learning 2017". Budapest, Hungary, 10-12 April, 2017. P. 114-118.

### **EXPERIENCE IN TEACHING TELEVISION HISTORY IN THE MASTER DISCIPLINE "HISTORY AND METHODOLOGY OF SCIENCE"**

**V. Khryashchev**

P.G. Demidov Yaroslavl State University

*The innovative educational project is aimed at the development and approbation of a mixed (blended) educational project for masters of technical and computer specialties. As a pilot project, the humanitarian course "History and Methodology of Science" was selected for master students of P.G. Demidov Yaroslavl State University who study master program "Infocommunication technologies and communication systems".*

**History of science, history of television, Zworykin, Sarnoff, education startup**

**В. В. Хрящев**  
*Ярославский государственный университет им. П. Г. Демидова*

## К вопросу о неэталонной оценке качества телевизионных изображений

*Представлены основные результаты по разработке и исследованию алгоритмов для неэталонной оценки качества изображений на основе локальных бинарных шаблонов (НОК ЛБШ). Данный неэталонный алгоритм является универсальным - для оценки не нужно предварительно знать тип искажения, которому подверглось исследуемое изображение. Проведено сравнение с известными эталонными и неэталонными алгоритмами оценки качества телевизионных изображений.*

### **Неэталонная оценка качества изображений, рандомизированные деревья, локальные бинарные шаблоны**

Алгоритмы обработки, сжатия и кодирования визуальной информации неизбежно вносят искажения различного типа, что делает актуальным вопрос об оценке параметров этих искажений и объективного оценивания качества искаженного изображения. На современном этапе развития все методы оценки качества классифицируются на две большие категории [1-4].

1. Субъективные критерии оценки качества изображений. Такая оценка в данном случае осуществляется человеком (оценки MOS по рекомендации ITU-R BT.500-11).

2. Объективные критерии оценки качества изображений. Здесь оценка производится алгоритмически. Такой алгоритм может использовать оригинал изображения для оценки качества, то есть быть эталонным, или может определять качество без наличия неискаженной копии изображения, то есть быть неэталонным.

В связи с необходимостью к унификации исследований в данной области в последние годы были предложены две общедоступные тестовые базы изображений с экспертными оценками MOS: LIVE (LIVE Image Quality Assessment Database, University of Texas at Austin, USA) и TID2008/2013 (Tampere Image Database, Tampere, Finland) [5, 6].

Задача неэталонной оценки качества изображений вызывает в последние годы повышенный интерес исследователей. Большая часть предлагаемых неэталонных алгоритмов разработана для конкретных типов априори известных искажений в изображении (блочность, размытие и др.), что не всегда удобно для использования в практических приложениях. Обзор существующих неэталонных критериев оценки качества изображений и видеопоследовательностей показал, что универсальных решений в данной области на настоящий момент не существует. Перспективным подходом представляется решение данной проблемы на основе современных алгоритмов машинного обучения. Для этого следует, во-первых, определить признаки, отвечающие за качество изображений, во-вторых, определить эффективный классификатор, способный разделять изображения разного качества на классы [7-9].

В данной работе в качестве признаков в задаче оценки качества изображений были выбраны локальные бинарные шаблоны (ЛБШ). Для описания изображения в пространстве признаков ЛБШ используем гистограмму, столбцы которой соответствуют десятичным значениям бинарных кодов шаблонов [10]. На рис. 1 представлены гистограммы равномерных инвариантных к повороту ЛБШ для изображений из базы LIVE, искаженных различными типами искажений. Как видно из анализа данных

гистограмм – ЛБШ может служить признаком, указывающим на присутствие искажений в изображении.

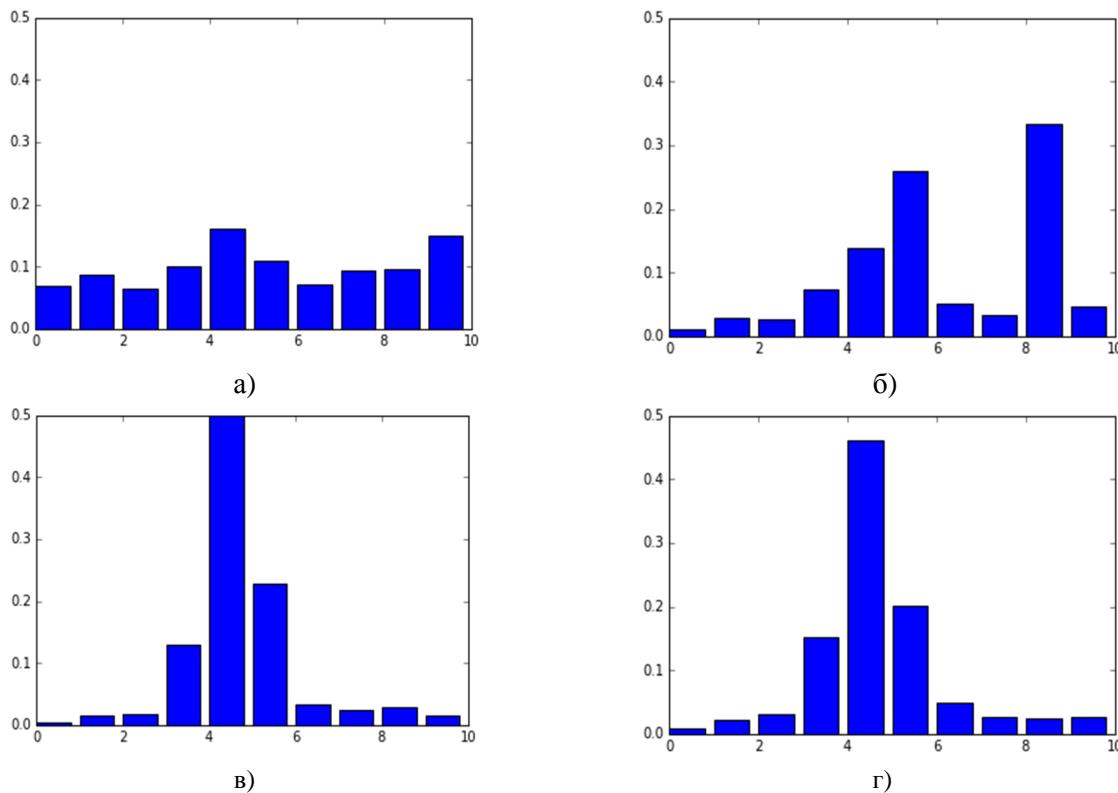


Рис. 1. Гистограмма равномерных инвариантных к сдвигу ЛБШ для искаженных изображений:  
а) оригинал; б) jpeg; в) jpeg2000; г) размытие

Таким образом, схема предлагаемого алгоритма неэталонной оценки качества с использованием локальных бинарных шаблонов и алгоритмов машинного обучения (НОК ЛБШ) выглядит следующим образом (рис. 2). В качестве вектора признаков для предложенного алгоритма выступают столбцы гистограмм многомасштабных равномерных локальных бинарных шаблонов, инвариантных к сдвигу. Гистограммы для каждого масштаба строятся отдельно, после чего формируется результирующая гистограмма путем склейки гистограмм всех масштабов. Всего для каждого изображения рассчитывается 54 признака. Выбор многомасштабных локальных бинарных шаблонов обусловлен их повышенной точностью по сравнению с одномасштабными шаблонами.

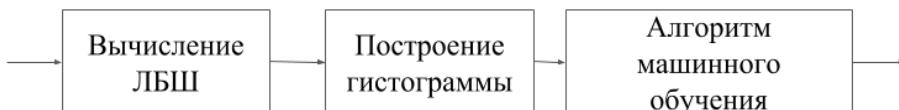


Рис. 2. Алгоритм неэталонной оценки качества изображения

Получены результаты исследований для алгоритма НОК ЛБШ с различными классификаторами, такими как бустинг (AdaBoost), машина опорных векторов, градиентный бустинг над решающими деревьями и рандомизированные деревья (табл. 1).

Полученные результаты показывают, что для алгоритма НОК ЛБШ по всем трем метрикам наилучшие результаты показывает классификатор на основе рандомизированных деревьев. Получены результаты по тестированию точности алгоритмов НОК ЛБШ-1 и его усложненной модификации НОК ЛБШ-2 [3] в сравнении с известными эталонными и неэталонными метриками на стандартных базах изображений LIVE и TID2013 (табл. 2, табл. 3).

Коэффициенты корреляции Пирсона (верхнее значение) и Спирмена (нижнее значение) при использовании многомасштабных ЛБШ для различных классификаторов на базе изображений LIVE

Таблица 1.

Тип искажения Классификатор	JPEG	JPEG 2000	Белый шум	Размытие	Битовые ошибки	Все
Бустинг	0,923	0,943	0,950	0,945	0,886	0,924
	0,901	0,932	0,967	0,942	0,926	0,926
МашинаОпорных векторов	0,809	0,884	-0,484	0,933	0,797	0,618
	0,854	0,883	-0,364	0,931	0,796	0,567
Градиентный бустинг	0,962	0,962	0,981	0,967	0,931	0,959
	0,937	0,945	0,964	0,945	0,904	0,947
Рандомизированные деревья	<b>0,970</b>	<b>0,968</b>	<b>0,987</b>	<b>0,971</b>	<b>0,936</b>	<b>0,965</b>
	<b>0,952</b>	<b>0,957</b>	<b>0,974</b>	<b>0,955</b>	<b>0,912</b>	<b>0,956</b>

Коэффициенты корреляции Пирсона (верхнее значение) Спирмена (нижнее значение) алгоритмов оценки качества изображений на тестовой базе LIVE

Таблица 2.

Название алгоритма	JPEG	JPEG 2000	Белый шум	Размытие	Битовые ошибки	Все типы
PSNR	0,903	0,876	0,917	0,780	0,880	0,859
	0,883	0,865	0,941	0,752	0,874	0,864
SSIM	0,946	0,941	0,982	0,900	<b>0,951</b>	0,907
	0,947	0,939	0,964	0,905	<b>0,939</b>	0,913
MS-SSIM	<b>0,979</b>	<b>0,975</b>	0,988	0,965	0,949	0,949
	<b>0,979</b>	0,963	0,977	0,954	0,939	0,940
LBIQ	0,935	0,910	0,976	0,910	0,838	0,909
	0,929	0,904	0,970	0,898	0,822	0,906
LD-TS	0,851	0,827	0,954	0,932	0,878	0,877
	0,833	0,820	0,957	0,925	0,886	0,883
GRNN	0,880	0,828	0,989	0,825	0,819	0,837
	0,872	0,816	0,979	0,833	0,735	0,827
НОКПО	0,973	0,923	0,985	0,951	0,903	0,942
	0,965	0,914	0,979	0,951	0,877	0,940
<b>НОК ЛБШ-1</b>	0,968	0,968	0,987	0,971	0,936	0,965
	0,952	0,957	0,974	0,955	0,912	0,956
<b>НОК ЛБШ-2</b>	0,972	0,973	<b>0,990</b>	<b>0,980</b>	<b>0,952</b>	<b>0,973</b>
	0,953	0,964	0,983	0,967	0,929	0,960

Сравнительные результаты работы алгоритмов НОК ЛБШ-1 и НОК ЛБШ-2 на тестовой базе TID2013

Таблица 3.

Алгоритм	Коэффициент корреляции Спирмена	Коэффициент корреляции Кенделла
<b>НОК ЛБШ-2</b>	<b>0,859</b>	<b>0,670</b>
FSIMx	0,851	0,667
PSNR-HA	0,819	0,643
PSNR-HMA	0,813	0,632
FSIM	0,801	0,630
НОК ЛБШ-1	0,790	0,600
MS-SIM	0,787	0,608
PSNRc	0,687	0,508
VSNR	0,681	0,508
PSNR-HVS	0,654	0,496
PSNR	0,640	0,482
SSIM	0,637	0,470

Результаты тестирования на базе LIVE показывают, что предложенный алгоритм НОК ЛБШ-2 показывает наилучшие результаты для следующих типов искажений: JPEG2000, белый шум, размытие, а также в среднем по всей базе LIVE. Результаты тестирования на базе TID2013 показывают, что алгоритм НОК ЛБШ-2 опережает по точности известные эталонные и неэталонные метрики. Алгоритм НОК ЛБШ-1 находится на одном уровне с известным алгоритмом MS-SSIM. Работа выполнена при поддержке гранта РФФИ № 15-07-08674.

### Библиографический список

1. Дворкович В.П., Дворкович А.В. Цифровые видеоИнформационные системы (теория и практика) // Техносфера, 2012. 1008 с.
2. Wang Z., Bovik A. Modern Image Quality Assessment. Morgan and Claypool Publishing Company, 2006.
3. Приоров А.Л., Ненахов И.С., Хрящев В.В. Разработка и анализ алгоритма неэталонной оценки качества изображений на основе локальных бинарных шаблонов // Успехи современной радиоэлектроники. 2016. № 7. С. 46–52.
4. Хрящев В.В., Ненахов И.С., Приоров А.Л. Использование многомасштабных локальных бинарных шаблонов в задаче оценки качества изображений // Телевидение: передача и обработка изображений: Тез. 13-й междунар. конф. – СПб., 2016. С. 119–123.
5. Sheikh H., Wang Z., Cormack L., Bovik A. LIVE image quality assessment database release 2, 2006. <http://live.ece.utexas.edu/research/quality/subjective.htm>.
6. Ponomarenko N., Jin L., Ieremeiev O., Lukin V., Egiazarian K., Astola J., Vozel B., Chehdi K., Carli M., Battisti F., Jay Kuo C., Image database TID2013: Peculiarities, results and perspectives // Signal Processing: Image Communication, 2015. V. 30, P. 57-77.
7. GaoX., LiX., TaoD. Universal no reference image quality assessment metrics based on local dependency // IEEE ACPR, 2011.
8. Saad M., Bovik A. C., Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain // IEEE Trans. Image Process., vol. 21, no. 8, pp.3339–3352, 2012.
9. Tong H., Li M., Zhang H., Zhang C. Learning no-reference quality metric by examples // Proc. 11th Int. Multimedia Modelling Conf., Jan. 2005, pp. 247–254.
10. Pietikainen M., Hadid A., Zhao G., Ahonen T., Computer vision using local binary patterns // Springer 2011, XVI, 212p.

### NO-REFERENCE QUALITY ASSESSMENT OF TELEVISION IMAGES

**V. Khryashchev**

P.G. Demidov Yaroslavl State University

*The main results of development and investigation of the algorithm for no-reference estimation of image quality on the basis of local binary patterns (NR LBP) are presented. This no-reference algorithm is universal - for estimation it is not necessary to know in advance the type of image distortion. Comparison with known reference and no-reference algorithms for quality assessment of television images is given.*

**No-reference image quality assessment, random forest, local binary patterns**

**А. В. Кузичкин, А. Ю. Аганов, А. А. Таранов**  
АО «НИИ телевидения»

## Комплекс для управления телевизионной инфраструктурой сложных территориально-распределенных объектов

*В работе рассматриваются особенности управления телевизионной инфраструктурой сложных территориально-распределенных объектов на примере космодрома «Восточный». Приведен состав и принципы построения разработанного и поставленного на космодром цифрового комплекса коммутации и распределения телевизионной информации.*

### **Телевизионная информация, коммутация и распределение ТВ-информации, Ethernet**

Телевизионная информация (ТВ-информация, ТИ) играет все более важную роль в обеспечении задач, решаемых сложными территориально-распределенными объектами. Наиболее значимыми областями использования телевизионных систем на таких объектах являются:

- контроль выполнения технологических операций;
- мониторинг состояния объектов и действий обслуживающего персонала;
- получение и обработка видеинформации от взаимодействующих систем;
- отображение наиболее важной информации на рабочих местах должностных лиц и операторов технологических расчетов и дежурных смен;
- запись, хранение и воспроизведение ТИ (по требованию);
- обеспечение охраны объектов.

В последние годы АО «НИИ телевидения» разработало и поставило заказчикам целый ряд систем, обеспечивающих управление телевизионной инфраструктурой сложных территориально-распределенных объектов. Структуру и принципы работы этих систем рассмотрим на примере одной из последних работ, выполненных для космодрома «Восточный»

Основными элементами телевизионной инфраструктуры космодрома являются:

1. Системы технологического телевидения основных комплексов космодрома (стартовых комплексов, технических комплексов, командно-измерительного комплекса, бортовых систем видеоконтроля, заводов по производству и хранению компонентов топлива, аэродромного комплекса, объектов железнодорожных и автомобильных коммуникаций и т.п.);
2. Система видеомониторинга (СВМ) [2];
3. Передвижные телевизионные станции (ПТС) [3];
4. Приемный телевизионный комплекс (для обмена ТИ с КА);
5. Беспилотные летательные аппараты с телевизионной аппаратурой;
6. Телевизионные охранные системы.

Своевременное и достоверное доведение ТИ от большого количества её источников до внешних и внутренних потребителей является важнейшей задачей телевизионной инфраструктуры космодромов.

На существующих космодромах телевизионная инфраструктура создавалась эволюционным путем, последовательно по мере развития основных комплексов и понимания возрастающей роли телевизионных

На космодроме «Восточный» для управления процессами сбора, обработки, распределения и передачи телевизионной информации заблаговременно до первого пуска была создана специальная система управления телевизионной инфраструктурой. В период

с 2013 по 2016 годы АО «НИИ телевидения» по заказу ФГУП «ЦЭНКИ» разработо и развернуло на космодроме Цифровой комплекс коммутации и распределения ТИ (ЦКК и РТИ) [3]. Комплекс состоит из автоматизированного рабочего места оператора (АРМО), аппаратуры коммутации и распределения ТИ (АКИРТИ), аппаратуры сопряжения с источниками ТИ (АСИТИ) и аппаратуры отображения ТИ (АОТИ).

ЦКК и РТИ предназначен для выполнения следующих основных задач:

- прием ТИ от всех источников ТИ на космодроме (технологическое телевидение, система видеомониторинга, приемный ТВ-комплекс, ПТС, спутниковое ТВ);
- согласование поступающей ТИ по скорости, типу и параметрам алгоритма кодирования с параметрами аппаратуры ЦКК и РТИ, аппаратуры отображения ТИ у внешних и внутренних потребителей, характеристиками каналов связи;
- коммутация поступающей ТИ внешним и внутренним потребителям;
- управление правами доступа потребителей к ТИ;
- отображение ТИ у потребителей в мульти- и моноэкранных режимах;
- мониторинг состояния и управление параметрами и режимами работы основных составных частей ЦКК и РТИ;
- автоматический контроль качества передачи ТИ в ЦКК и РТИ;
- запись, хранение и воспроизведение (по требованию) поступающей ТИ.

Состав опытного образца ЦКК и РТИ, развернутого на космодроме к первому пуску (рис. 1):

- четыре комплекта аппаратуры сопряжения с источниками ТИ;
- АКИРТИ;
- аппаратура АРМО (рис.2), состоящая из стойки с серверным оборудованием со специальным программным обеспечением (СПО), управляющим работой всего комплекса, рабочего стола оператора АРМО (из 4 моноблоков) и табло коллективного пользования (из трех большеразмерных мониторов с портативными компьютерами);
- 20 комплектов АОТИ.

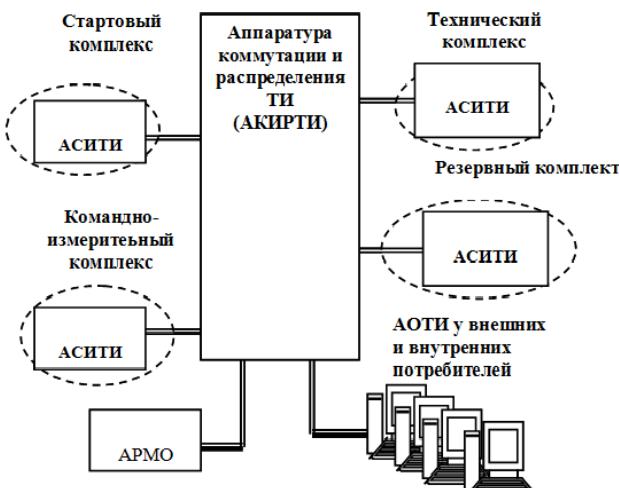


Рис. 1. Структурная схема ЦКК и РТИ

Базовыми протоколами формирования, обработки и обмена ТИ в ЦКК и РТИ являются протоколы, построенные на основе IP-протоколов. Обмен ТИ между объектами космодрома осуществляется с помощью специальной сети передачи данных (СПД). СПД космодрома строится на основе опорной волоконно-оптической сети. Топология сети – звезда.

Технология передачи ТИ всем внутренним потребителям – мультикаст: каждому частному потоку ТИ от конкретной видеокамеры или сервера присваивается свой мультикастинговый IP-адрес (из диапазона адресов, зарезервированных под мультикаст) и весь пакет рассыпается всем получателям ТИ. В зависимости от информации, поступающей из АРМО от подсистемы специального программного обеспечения (СПО),

управляющей правами доступа к ТИ, каждому получателю ТИ устанавливается перечень источников ТИ, которые он может просматривать на своей АОТИ.

Внешний вид опытного образца ЦКК и РТИ, развернутого в лабораторных условиях АО «НИИ телевидения» приведён на рис. 2 и 3.

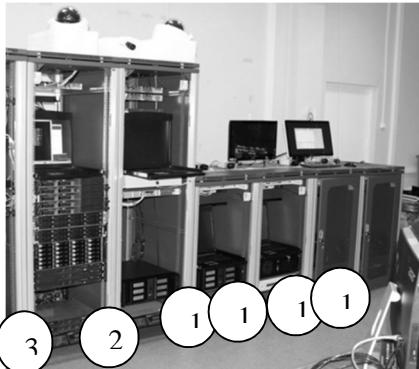


Рис. 2. Внешний вид опытного образца ЦКК и РТИ во время заводских испытаний:  
1- АСИТИ; 2- АКИРТИ; 3- стойка аппаратуры АРМО

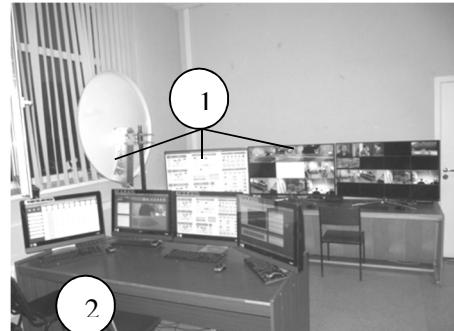


Рис.3. Внешний вид опытного образца ЦКК и РТИ во время заводских испытаний: 1 – табло коллективного пользования, 2 – рабочий стол оператора

СПО, установленное на серверном оборудовании аппаратуры АРМО, разрабатывалось совместно с ОАО "Новосибирский институт программных систем". СПО состоит из следующих основных модулей (компонентов):

- модуль управления параметрами и режимами работы основных элементов ЦКК и РТИ;
- модуль мониторинга состояние каналов и всех основных элементов ЦКК и РТИ;
- модуль управления формированием общего мультикастингового потока ТИ, рассылаемого всем пользователям;
- модуль управления правами доступа потребителей к ТИ;
- модуль управления отображением ТИ;
- модуль управления регистрацией ТИ;
- модуль управления формированием ТИ для передачи по каналам дальней связи для внешних потребителей;
- модуль сопряжения ТИ, получаемой от источников ТИ, с аппаратурой и алгоритмами обработки ТИ в ЦКК и РТИ;
- модуль сопряжения ТИ, отправляемой ЦКК и РТИ, с аппаратурой и алгоритмами обработки ТИ у потребителей ТИ;
- модуль хранения текущих настроек ЦКК и РТИ;
- модуль ведения журналов событий и нештатных ситуаций ЦКК и РТИ;
- модуль пользовательского интерфейса.

Управление режимами работы и состоянием элементов ЦКК и РТИ, осуществляющееся с АРМО, выполняется с помощью виртуальных панелей и графического интерфейса на сенсорном экране мониторов, компьютерной клавиатуры и компьютерной мыши. В режимах диагностики, мониторинга и управления элементами ЦКК и РТИ мониторы с сенсорными экранами отображают в графической форме текущую информацию о состоянии и режимах работы всех элементов ЦКК и РТИ, наложенную на мнемоническую схему космодрома. Предусмотрен вывод информации с управляющих мониторов на большие табло отображения.

Структурная схема взаимодействия основных источников и потребителей телевизионной информации с ЦКК и РТИ приведена на рис. 4.

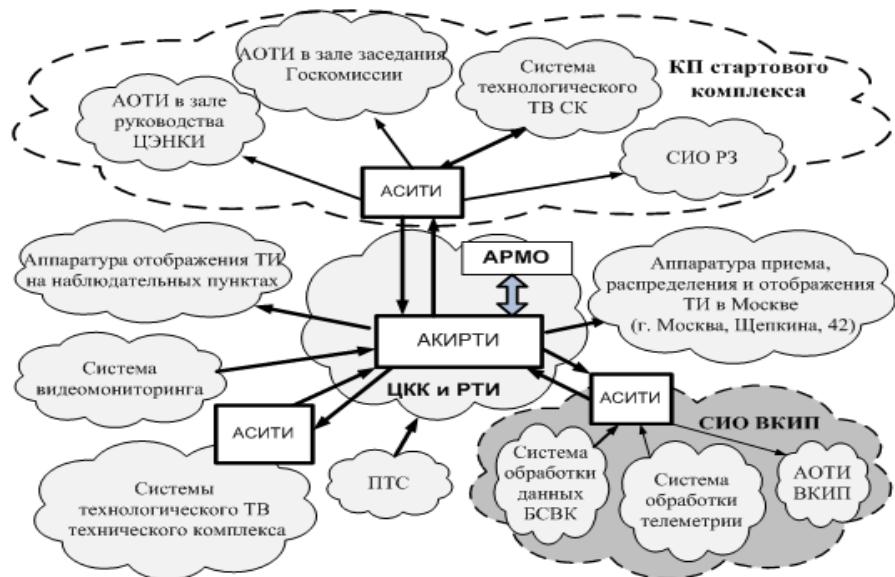


Рис.4. Структурная схема информационного взаимодействия основных источников и потребителей телевизионной информации космодрома с ЦКК и РТИ

На схеме показаны следующие основные источники ТИ: система технологического телевидения стартового комплекса: система обработки данных бортовой системы видеоконтроля (БСВК); передвижная телевизионная станция (ПТС); тестовые видео камеры из состава АСИТИ на основных объектах пускового минимума; система обработки телеметрических данных; система видеомониторинга; система технологического телевидения технического комплекса.

Основными получателями ТИ на Рис.4 являются: аппаратура отображения телевизионной информации (АОТИ), входящая в состав ЦКК и РТИ; аппаратура отображения ТИ системы технологического телевидения стартового комплекса; система информационного обеспечения восточного командно-измерительного пункта (СИО ВKİП); система информационного обеспечения руководителя запуска (СИО РЗ); аппаратура отображения ТИ в зале заседания Госкомиссии; аппаратура отображения ТИ на наблюдательных пунктах; аппаратура приема, распределения и отображения ТИ в Москве в аппаратной ФГУП «ЦЭНКИ»

Первый пуск показал правильность решений, принятых при разработке ЦКК и РТИ. ЦКК и РТИ с успехом выполнил важные и довольно непростые задачи сопряжения ЦКК и РТИ с большим количеством разнородных информационных сетей источников и пользователей ТИ, использующих различное специальное программное обеспечение и видеокамеры разных производителей с разными кодерами компрессии ТИ.

Управление распределением ТИ в процессе подготовки и выполнения первого пуска осуществлялось с пульта АРМО (Рис. 5).



Рис. 5. Аппаратная АРМО ЦКК и РТИ на временном узле связи 28 апреля 2016г. за 3 минуты до пуска

В процессе подготовки и выполнения первого пуска разработанный комплекс ЦКК и РТИ продемонстрировал устойчивую и безотказную работу, обеспечил качественную доставку и отображение ТИ на АОТИ всех потребителей ТИ на космодроме "Восточный".

На Рис. 6-8 представлены фотографии, иллюстрирующие совместную работу ЦКК и РТИ с ПТС и СВМ (также разработанные АО «НИИ телевидения») и отображение ТИ на некоторых наиболее важных объектах космодрома "Восточный" в процессе подготовки и выполнения первого пуска.



Рис. 6. Отображение ТИ от ПТС на видеостене СИО РЗ



Рис. 7. Отображение видеопотока от ПТС и СВМ средствами системы технологического ТВ стартового комплекса (видеопоток ПТС – в верхнем левом углу)



Рис. 8. Изображение, передаваемое ЦКК и РТИ от ПТС внешним потребителям

Эксплуатация комплекса ЦКК и РТИ в процессе подготовки и выполнения первого пуска выполнялась объединенным расчетом в составе специалистов ФГУП «ЦЭНКИ» и АО «НИИ телевидения».

#### Библиографический список

1. Система приёма и трансляции видеоинформации с площадок космодрома / Умбиталиев А. А., Кузичкин А. В., Севастьянов Д. А. и др. Вопросы радиоэлектроники, серия Техника телевидения, Вып.2, 2014. С. 57–61
2. Передвижная телевизионная станция по стандарту телевидения высокой четкости HD-SDI для космодрома "Восточный" / Кузичкин А. В., Баланин Л.Н., Жуков Л.И. и др. Вопросы радиоэлектроники, серия Техника телевидения, Вып.3, 2016. С. 117–123.
3. Цифровой комплекс коммутации и распределения телевизионной информации космодрома «Восточный» / Умбиталиев А. А., Кузичкин А. В., Аганов А. А. и др. Вопросы радиоэлектроники, серия Техника телевидения, Вып.2, 2015. С. 13–20.

#### **COMPLEX FOR MANAGEMENT OF TELEVISION INFRASTRUCTURE THE DIFFICULT TERRITORIAL DISTRIBUTED OBJECTS**

**A.V. Kuzichkin, A.J. Aganov, A.A.Taranov**  
JSC Scientific Research Institute of Television

*In operation consider features of control of television infrastructure of difficult geographically distributed objects on the example of the spaceport "East". The composition and the principles of creation of the digital complex of switching and distribution of television information developed and delivered on the spaceport is given.*

**TV information, management, switching, spaceport "East"**

**И. Л. Бурцев, Т. Ю. Мамаева, А. К. Рыжаков, В. Я. Яцук**  
АО "Восток", ООО "ЭФО"

## Доступная ЭКБ ИП для бортовой и наземной РЭА

*Рассмотрены возможности поставок в Россию электронной компонентной базы иностранного производства (ЭКБ ИП) для российских разработчиков бортовой и наземной аппаратуры космического назначения в условиях экспортных ограничений США и ЕС на основе практического опыта квалифицированного поставщика электронных компонентов компании «Восток».*

**Лицензии, ЭКБ, экспорт, EUS, Microchip, TI, Teledyne E2V, Isocom Limited, DDC, STMicroelectronics, «Восток»**

Известно, что режим экспортного контроля накладывает ряд ограничений на вывоз товара из страны-экспортера: необходимость получения лицензии на экспорт, запрет на экспорт в некоторые страны, контроль за конечными пользователями и за передачей технологий. Высоконадежные и радиационно-стойкие космические комплектующие изделия, как правило, подлежат специальному экспортному контролю. Они входят в группу товаров двойного назначения, перечень которых определяется списком торгового контроля страны-экспортера.

Требуют обязательного оформления экспортных лицензий на поставку в РФ изделия с кодами 9A515.e и 9A515.d. (радиационно-стойкие электронные компоненты, специально разработанные для космических применений), 3A001A5 (микроконтроллеры, процессоры, АЦП, ЦАП, ПЛИС с производительностью, превышающей определённые границы), 3A001A2 (микроконтроллеры, процессоры, АЦП, ЦАП, ПЛИС, имеющие по документации расширенные температурные диапазоны). Электронным компонентам, формально подпадающим под описание лицензируемых групп ЭКБ, но с меньшими технологическими характеристиками, присваиваются номера, содержащие в своем составе код 99 (EAR99, 3A991, 5A991G и т.д.). Экспорт таких компонентов не требует оформления лицензии. Но и в этом случае необходима идентификация конечного пользователя и изделия, в котором будет применяться выбранный электронный компонент.

Для оформления лицензии в Службу экспортного контроля страны-экспортера должен быть предоставлен специальный пакет документов. Обязательным пунктом в этом пакете является форма EUS (заявление конечного пользователя) /EUC (сертификат конечного пользователя), содержащая информацию о конечном пользователе, об изделии, в котором будет применяться электронный компонент, о логистической цепочке для поставляемой партии. Заполненный документ проходит несколько итераций согласования с производителем ЭКБ и конечным пользователем. Окончательно согласованный вариант формы распечатывается на бланке организации конечного пользователя, подписывается, заверяется печатью и отправляется на стадию лицензирования. Экспортная лицензия выдается на ограниченное время и на определенное количество микросхем.

Вопросы доступности ЭКБ иностранного производства рассматриваются на примере продукции европейских и американских производителей, входящих в программу поставок АО «Восток» и имеющих сертифицированные производственные линии для изготовления комплектующих космического назначения. Основу перечня составляют высоконадежные и радиационно-стойкие электронные компоненты, выпускаемые

компаниями Microchip (ранее Atmel), Teledyne E2V&Peregrine, Isocom, Kemet, DDC (ранее Maxwell), Pulse, Smiths Interconnect (EMC, Hypertac, IDI, Lorch, Millitech, RF Labs, Sabritec, TECOM, TRAK), STMicroelectronics, Texas Instruments. Здесь мы хотели бы обратить внимание на то, что у представленных производителей имеется достаточно большое количество компонентов космического назначения, не требующих специального разрешения на экспорт в РФ.

Для случая, когда экспорт высоконадежных и радиационно-стойких компонентов по каким-либо причинам является заблокированным, мы проработали и предлагаем вашему вниманию возможные варианты замещения попавшей под экспортные ограничения ЭКБ. Ниже приводятся рекомендуемые варианты замещения заблокированной ЭКБ не связанными экспортными ограничениями функциональными аналогами. Так, например, номенклатура E2V в определенной степени перекликается с производственной программой Texas Instruments. Радиационно-стойкие MOSFET-транзисторы от английской компании Isocom Limited могут заменить продукцию американской фирмы IR. Изделия компаний Texas Instruments и STMicroelectronics, не подпадающие под ограничения, могут применяться взамен собственных лицензионных или изделий других производителей.

Примерами подобных замещений служат результаты технического сопровождения компаниями АО "Восток" и ООО "ЭФО" проектов, реализуемых на предприятиях Роскосмоса и ОПК: произведена замена микросхемы памяти EEPROM производства компании DDC на функциональный аналог от Microchip с некоторым снижением требований по радиационной стойкости; осуществлена замена лицензионного АЦП от Texas Instruments на аналогичный QML-V вариант от того же производителя, свободный от экспортных ограничений; предложена замена герметичных оптронов производства Micropac полными аналогами, выпускаемыми английской компанией ISOCOM Limited.

Основываясь на опыте нашей практической работы в условиях введенных ограничений, мы можем сделать следующие выводы относительно доступности ЭКБ иностранного производства на текущий момент: лицензии на экспорт в РФ высоконадежной и радиационно-стойкой ЭКБ производства США выдаются только для международных космических программ при условии подписания конечным пользователем формы EUS; лицензии на экспорт в РФ высоконадежной и радиационно-стойкой ЭКБ производства стран ЕС выдаются для международных космических программ и российских космических программ гражданского назначения при условии подписания конечным пользователем формы EUC; высоконадежные и радиационно-стойкие компоненты категорий NEC (No Export Control) и EAR99 не подлежат специальному контролю и могут свободно поставляться в РФ гражданским заказчикам ЭКБ.

## Секция 1. Методы и устройства формирования видеосигнала

Руководитель секции – д.т.н., проф. А. К. Цыцуллин (АО «НИИ телевидения»)

Зам. руководителя – к.т.н., доц. А. А. Манцевов (СПбГЭТУ «ЛЭТИ»)

Секретарь секции – инженер Д. А. Белоус (СПбГЭТУ «ЛЭТИ»)

---

**В. С. Дюмин, М. В. Четвергов, М. Г. Выдревич**  
ЗАО «НПП «ЭЛАР»

### Фоточувствительный прибор с зарядовой связью с встроенным регистром умножения

*Данная статья посвящена фоточувствительному прибору с зарядовой связью с электронным умножением (ФПЗС с ЭУ), который позволяет решить проблему регистрации изображения при низких уровнях освещенности. Показано влияние интеграции регистра умножения в ФПЗС. Рассмотрены основные виды шумов для данного прибора.*

**ФПЗС с ЭУ, регистр умножения, электронное умножение, шум считывания, шум-фактор, ударная ионизация, низкие уровни освещенности**

Одним из основных параметров фотодетектора является чувствительность, характеризующая способность реагировать на малые сигналы. В условиях слабого освещения чувствительность ФПЗС ограничивается шумовым сигналом, и получение необходимого изображения становится затруднительным. Решением данной проблемы стало создание ФПЗС с ЭУ (electron multiplying charge-coupled device – EMCCD), способного работать при чрезвычайно низких уровнях освещенности порядка  $10^{-4}$ - $10^{-5}$  люкса.

Для стандартных ФПЗС матриц в слабых световых потоках характерны малые соотношения сигнал-шум из-за значительного влияния шумов, в особенности, шума считывания. В EMCCD сигнальные электроны еще до процесса считывания выходным устройством умножаются в специальной части горизонтального сдвигового регистра - регистре умножения за счет лавинных процессов. Таким образом, шум считывания отходит на второй план и перестает играть доминирующую роль, что позволяет увеличить соотношение сигнал-шум при низких уровнях освещенности.

Следует отметить, что при использовании умножающего регистра возникает дополнительный источник шума благодаря вероятностной природе процесса ударной ионизации. Данный вид шума может быть оценен при помощи шум-фактора (шум умножения)  $F$ . В рамках данной работы было показано, что шум-фактор при большом числе каскадов умножения  $n$  и малой вероятности рождения одного нового носителя  $\alpha$  слабо зависит от усиления, а квадрат его значения не превышает двух, тем самым доказана применимость EMCCD.

Основные источники шума ФПЗС с ЭУ: дробовый (фотонный) шум, шум считывания, темновой ток, шум умножения [2]. Было показано, что с шумом считывания порядка 100 электронов на 1 зарядовый пакет наиболее эффективно справляется регистр с коэффициентом умножения  $G$  от 50 и до 100. При охлаждении кристалла до минус 15°C темновой ток понижается до значения порядка 0,02 электрона/пиксель/кадр при площасти

пикселя  $3,6 \cdot 10^{-6} \text{ см}^2$  и времени нахождения зарядового пакета в структуре ФПЗС 77 мс. Из всего вышесказанного можно сделать вывод, что доминирующим видом шума в EMCCD становится фотонный шум, который следует из квантовой природы света и определяется как корень квадратный из числа фотонов, падающих на поверхность фоточувствительной матрицы.

Таким образом, в ФПЗС с ЭУ соотношение сигнал-шум приближается к теоретическому пределу.

#### Библиографический список

1. Казначеев С.А. Разработка телевизионного приёмника, использующего эффект внутреннего электронного умножения для повышения эффективности регистрации и исследований малофотонных изображений: автореф. дис. на соиск. учен. степ. канд. технических наук (05.11.07) / Казначеев Сергей Александрович; Национальный исследовательский университет «МЭИ» (ФГБОУ ВО «НИУ «МЭИ») – Москва, 2015. – 20 с.
2. Robbins M.S. The Noise Performance of Electron Multiplying Charge Coupled Devices / M. S. Robbins, B. J. Hadwen // IEEE Trans. Electron Devices. – 2003. – Vol.12. – P. 1227–1232.

### CHARGE-COUPLED DEVICE WITH BUILT-IN MULTIPLICATION REGISTER

**V. S. Dyumin, M. V. Chetvergov, M. G. Vydrevich**  
CJSC "ELAR"

*This article is devoted to a charge-coupled device with electron's multiplication (EMCCD), which allows to solve the problem of image receiving at low light levels. The influence of applying multiplication register in CCD was shown. The main types of noise for this device were considered.*

**EMCCD, multiplication register, multiplication of electrons, read noise, excess noise factor, impact ionization, low light levels.**

**М. В. Четвергов, А. В. Корюшкин, В. В. Петров, В. А. Локтев**  
ЗАО «НПП «ЭЛАР»

## Метрологическое обеспечение измерений параметров ФПЗС

*В статье приведены тезисы доклада, посвященного вопросам метрологического обеспечения измерений параметров ФПЗС в ЗАО «НПП «ЭЛАР». В частности рассмотрено стендовое измерительное оборудование и методики измерений.*

### **ФПЗС, метрология, измерения**

В докладе рассмотрены вопросы метрологического обеспечения измерений и испытаний параметров матричных фоточувствительных приборов с зарядовой связью (ФПЗС).

ЗАО «НПП «ЭЛАР» занимается разработкой и изготовлением матричных фоточувствительных приборов (ФПЗС) специального назначения уже десятки лет. На протяжении всего времени выпуска ФПЗС стоял один из главных вопросов – обеспечение полноты и точности измерения параметров. Причем, так как ФПЗС относится к элементам функциональной электроники и имеет сложную внутреннюю структуру, скорее можно говорить даже не об измерении параметров, а об их исследовании. Необходимость именно в исследовании параметров также обусловлена различиями оптимальных режимов работы ФПЗС в разной аппаратуре заказчика.

В докладе рассмотрены две основные составляющие метрологического обеспечения измерений – стендовое оборудование и методики измерений. Также дается оценка необходимости участия производства в исследованиях характеристик (контроль параметров в технологическом процессе, изготовление специальных приборов-свидетелей и другое).

Ключевой частью стендового оборудования является его разработка и изготовление специально для целей измерения ФПЗС. Костяк оборудования составляют несколько моделей осветителей и блоки управления, разрабатываемые индивидуально для каждого наименования изделий. Задачей осветителей является создание равномерного пятна засветки, регулировки в широких пределах облученности и задания спектрального состава облучения. Перед блоком управления ставятся следующие задачи: формирование электрического режима и временных диаграмм управляющих напряжений, а также усиление, преобразование в цифровой код и передача выходного сигнала в компьютер, где с помощью специального программного обеспечения производится его обработка и вычисление фотоэлектрических параметров Изделия.

Основной же составляющей измерений являются методики измерений, при этом именно с ними имеется ряд трудностей. Существующие нормативные документы (ГОСТ 28953-91, ОСТ 11 0259-85, ОСТ 11 0290.0-86 и др.) частично устарели или не всегда применимы к ФПЗС с рядом конструктивных особенностей. Примером этого могут быть ФПЗС, работающие в режиме временной задержки и накопления (ВЗН), которые, в отличие от обычных кадровых ФПЗС, предназначены для считывания движущегося изображения. Ввиду этого ЗАО «НПП «ЭЛАР», по согласованию с заказчиками и на основе имеющейся нормативной документации, разрабатывает собственные методики измерений, основные из которых рассматриваются в данном докладе.

В частности, на основе ISO 12233:2000(E) разработана методика измерения ФПМ с помощью «оптического ножа», который наносится непосредственно на поверхность ФПЗС в технологическом процессе. При этом ЧКХ может быть рассчитана по результатам измерения ФПМ или измерена с помощью штриховой меры, когда формирование оптического ножа на поверхности ФПЗС невозможно.

Другим примером необходимости разных подходов к измерениям приборов разных конструкций может быть измерение спектральной характеристики квантового выхода. Разработано три метода таких измерений. Первый основан на использовании монохроматора и измерении тока выходного устройства. Второй метод базируется на использовании набора узкополосных светофильтров. Третий метод совмещает измерение нормированной зависимости с помощью монохроматора с уточнением по измерениям за узкополосными светофильтрами. Каждый из этих методов имеет свои преимущества и недостатки и находит применение для ФПЗС разных конструкций.

Кроме того ЗАО «НПП «ЭЛАР» разработаны методы измерения многих других параметров, а именно: крутизна преобразования выходного устройства, напряжение насыщения, зарядовая вместимость, динамический диапазон, шум считывания, пиковая неравномерность выходного сигнала, интегральная чувствительность, темнового тока (и его зависимость от температуры), эффективности переноса по горизонтали и вертикали, число дефектов и другие.

#### Библиографический список

1. ГОСТ 28953-91 Приборы фоточувствительные с переносом заряда. Методы измерения параметров.
2. ГОСТ 11 0259-85 Приборы фоточувствительные с переносом заряда. Система параметров.
3. ГОСТ 11 0290.0-86 Приборы фоточувствительные с переносом заряда. Общие требования при измерении параметров.

### METROLOGICAL SUPPORT OF MEASUREMENTS OF CCD PARAMETERS

***M. V. Chetvergov, A. V. Koriushkin, V. V. Petrov, V. A. Loktev.***  
CJSC «ELAR»

*Theses of the report devoted to questions of metrological support of measurements of the CCD parameters in CJSC «ELAR» are provided in article. Measuring equipment and measurement techniques is considered.*

**CCD, metrology, measurements**

**Т. С. Ларина, Е. В. Рыпакова, Д. Н. Сергеев, М. В. Четвергов**  
ЗАО «НПП «ЭЛАР»

## Конструктивные особенности матричных фоточувствительных приборов с зарядовой связью для космических систем

*В работе рассматриваются наиболее востребованные для Российского космоса фоточувствительные приборы с зарядовой связью (ФПЗС), работающие в режиме временной задержки и накопления (ВЗН), разработанные и выпускаемые ЗАО «НПП «ЭЛАР». Рассматриваются перспективные направления развития матричных ФПЗС ВЗН.*

**Фоточувствительный прибор с зарядовой связью, режим временной задержки и накопления, ЗАО «НПП «Элар», широкоформатные ФПЗС ВЗН, многокристальные гибридные сборки**

В докладе приведен обзор разработанных и выпускаемых ЗАО «НПП «ЭЛАР» фоточувствительных приборов с зарядовой связью (ФПЗС), работающих в режиме временной задержки и накопления (ВЗН), применяемых в космических аппаратах (КА) дистанционного зондирования земли (ДЗЗ), а также рассмотрены перспективные направления развития таких ФПЗС.

На данный момент для Российского космоса наиболее востребованными фотоприемниками являются: ФПЗС «Круз-6» и ФЧМ «Круз-Ц-Б». Данные датчики предназначены для регистрации движущегося изображения со спутника.

Микросхема ФПЗС «Круз-6» представляет собой бескорпусной прибор с форматом фотозоны – 1536’192 элемента, размером ячейки 6’6 мкм, и обеспечивает получение панхроматического изображения в видимом диапазоне с высоким разрешением.

Фоточувствительный модуль «Круз-Ц-Б» представляет собой бескорпусной фотоприемник с тремя фотозонами форматом – 768’64 элемента, размером ячейки 18’18 мкм, и обеспечивает получение мультиспектрального изображения за счет интегрированного с микросхемой трехзонного светофильтра. ФЧМ изготавливается различных типов (в настоящее время 5 типов ФЧМ), соответствующих типам используемых светофильтров.

Данные ФПЗС предназначены для монтажа в большие фокальные плоскости, размер которых может превышать полметра, а формат по горизонтали составлять нескольких десятков тысяч элементов.

На данный момент ФПЗС «Круз-6» совместно с ФЧМ «Круз-Ц-Б» установлены на следующих КА: «Ресурс-П» № 1, «Ресурс-П» № 2, «Ресурс-П» № 3 и еще на нескольких спутниках специального назначения.

Указанные ФПЗС были разработаны на базе предшественника - ФПЗС ВЗН «Круз». Этот тип фотоприемников успешно проработал в аппаратуре КА «Ресурс-ДК» весь срок его эксплуатации.

В настоящий момент ЗАО «НПП «ЭЛАР» прорабатывает два перспективных направления развития матричных приборов ФПЗС ВЗН.

Первое направление связано с использованием однокристальных широкоформатных ФПЗС ВЗН («Прибор-ОЭК-9», «Прибор-ОЭК-Ц», «Комби-ФПЗС») для создания больших фокальных плоскостей. За счет того, что для сборки фокальной плоскости потребуется

всего несколько таких приборов, процесс монтажа их во многом облегчается и ускоряется во времени. Однако, данные фотоприемники не обеспечивают минимальных расстояний между фотозонами в составе гибридной сборки.

Второе направление связано с разработкой многокристальной гибридной ФПЗС ВЗН сборки в корпусном исполнении, что позволит использовать такой модуль как самостоятельное изделие или как компонент широкоформатной фокальной плоскости. В ЗАО «НПП «ЭЛАР» разрабатывается изделие подобного типа – ФПЗС «Текстура» – на базе ФПЗС ВЗН «Круиз-б». Также ведется проработка многокристальных сборок, представляющих собой совокупность кристаллов ФПЗС ВЗН «Круиз-Ц-Б» и светофильтров, расположенных в одном корпусе с совмещенной светозащитной блендой, и позволяющих получать мультиспектральное изображение. Очевидно, монтаж таких сборок в фокальную плоскость не требует наличия чистых помещений. Такие сборки будут интересны относительно новому направлению в космической промышленности – производству малых космических аппаратов.

#### **MAIN DIRECTIONS OF DEVELOPMENT OF MATRIX CCD TDI FOR SYSTEMS OF REMOTE SENSING OF THE EARTH**

**T. S. Larina, E. V. Rypakova, D. N. Sergeev, M. V. Chetvergov**  
CJSC «ELAR», Saint-Petersburg

*The most popular of the Russian space CCD TDI designed and produced by ZAO "NPP "ELAR" is described. There are considered the perspective directions of development of these CCD-TDI.*

**Photosensitive charge-coupled device, a mode time delay and accumulation, ZAO "NPP "ELAR", large-format CCD TDI, multichip hybrid modules.**

**Г. И. Вишневский<sup>1</sup>, М. В. Четвергов<sup>1</sup>, Н. С. Чепилко<sup>1</sup>,**

**М. Г. Выдревич<sup>2</sup>, А. Г. Попов<sup>2</sup>**

<sup>1</sup>ЗАО «НПП «ЭЛАР»

<sup>2</sup>АО «НПП «СИЛАР»

## Производство фоточувствительных КМОП микросхем

*В статье приведены тезисы доклада, посвященного вопросам производства фоточувствительных КМОП микросхем предприятиями ЗАО «НПП «ЭЛАР» и АО «НПП «СИЛАР».*

### **КМОП, фоточувствительная микросхема, производство**

В докладе рассмотрены вопросы производства фоточувствительных КМОП микросхем.

ЗАО «НПП «ЭЛАР» занимается разработкой и изготовлением матричных фоточувствительных приборов (ФПЗС) специального назначения уже десятки лет. В 2014 году совместно с АО «НПП «СИЛАР» было принято решение начать разработку фоточувствительных КМОП микросхем для различных задач. При этом разработкой кристаллов занимается АО «НПП «СИЛАР», а сборкой, тестированием, испытаниями – ЗАО «НПП «ЭЛАР».

В качестве экспериментального образца было решено разработать и изготовить фоточувствительную матрицу форматом 660x484 элемента, из которых фоточувствительными являются 640x480 элементов.

На момент начала работы над матрицей выбранная фабрика не имела стандартного процесса для производства фоточувствительных матриц, так как ни в одном из существующих не было отработано производство p-in-диодов. В результате анализа было принято решение использовать процесс с технологическими нормами 350 нм; ожидалось, что это позволит создать фоточувствительную ячейку с прогнозируемыми параметрами.

Разработанная микросхема имеет следующие особенности:

- Размер ячейки 14x14 мкм;
- Тип затвора – GLOBAL SHUTTER;
- Дополнительные элементы, для калибровки и настройки;
- 12-ти разрядный АЦП;
- Аналоговый и цифровой выходы;
- Аналоговый сигнал выводится в соответствии со стандартом VGA.

На первой версии кристалла были измерены: спектральная характеристика чувствительности (квантовая эффективность в максимуме составила 30% на длине волны 540 нм); зарядовая вместимость (~150 тыс. электронов), крутизна преобразования (~7 мкВ/эл).

После получения первой версии кристалла были выявлены и исправлены ошибки проектирования. В частности, исправлена ошибка топологии, в результате которой в цифровом режиме были видны наводки практически по всему полю прибора; исправлены недочеты ячеек для калибровки.

Были внесены другие изменения: в ячейке заменен p-n переход хранения на емкость затвора n-mos транзистора; добавлен умножитель частоты с ФАПЧ; добавлен

технологический модуль водородного отжига; разработаны две версии кристалла: с обычной ячейкой и ячейкой увеличенной зарядовой вместимости.

Внесенные изменения показали, что в новой версии кристалла:

- Отсутствуют яркие белые дефекты;
- Зарядовая вместимость обычной ячейки составляет ~ 150 тыс. электронов;
- Зарядовая вместимость ячейки с увеличенной емкостью составляет ~ 1 млн. электронов;
- Умножитель частоты с ФАПЧ работает в штатном режиме.

Полученные результаты показывают, что, используя выбранную технологию, можно производить фоточувствительные КМОП кристаллы с ожидаемыми параметрами. В данный момент разработан и отлаживается кристалл форматом 1320x1024 элементов для сопряжения с ЭОП.

На данный момент в АО «НПП «СИЛАР» и ЗАО «НПП «ЭЛАР»: согласуются ТЗ и договор на разработку нового фоточувствительного кристалла для комплектования многоканальной спектральной системы, прорабатывается эскизный проект для создания кристалла для комплектования гиперспектральной аппаратуры, прорабатываются основные узлы зарядо-чувствительной микросхемы для масс-спектрометра.

## **PRODUCTION OF IMAGE CMOS SENSORS**

**G.I. Vishnevsky<sup>1</sup>, M.V. Chetvergov<sup>1</sup>, N.S. Chepilko<sup>1</sup>, M.G. Vydrevich<sup>2</sup>, A.G. Popov<sup>2</sup>**

<sup>1</sup>CJSC «ELAR», Saint-Petersburg

<sup>2</sup>JSC «SILAR», Saint-Petersburg

*These of the report devoted to questions of production of image CMOS sensors by CJSC «ELAR» and CJSC «SILAR». First results and current developments.*

**CMOS, photosensitive chip, production**

**Д. В. Бородин**  
ООО "РТК Инпекс", Мытищи, РФ  
ОАО "НПП "Пульсар", Москва, РФ

## Параметры для оптимизации ячейки КМОП фотоматрицы

Для технологии, используемой при изготовлении отечественных КМОП фотоприемников видимого диапазона форматов 1024' 1024 ячеек размером 15' 15 мкм и 1280' 1024 ячеек размером 13' 13 мкм, определены: темновой ток донной части фотодиода; темновой ток периметра фотодиода; площадь сбора фотозаряда, что позволяет оптимизировать конструкцию фотоячейки.

### КМОП, фотоприемник, темновой ток, площадь сбора фотозаряда

Предельные параметры фотоприемных матриц зависят в значительной степени от темновых токов  $I_{\text{темн}}$  расположенных в ячейках фотодиодов. Именно этот ток определяет темновой сигнал и его шум, а тем самым и динамический диапазон КМОП фотоматрицы, пороговую экспозицию. Типовое значение  $I_{\text{темн}}$  составляет фемтоамперы [1] и достаточно затруднено для прямых измерений. Ток  $I_{\text{темн}}$  складывается из темнового тока донной части р-п-перехода и темнового тока, связанного с периметром фотодиода. Поскольку фотодиод может иметь достаточно произвольный размер и форму, для проектирования необходимо знать оба тока, однако фабрики-изготовители пластин их не указывают в соответствующей документации. Также важным параметром является соотношение площади фотодиода и площади сбора фотозаряда фотодиодом, которые могут существенно отличаться, хотя иногда высказываются предположения об их равенстве [2].

Для темнового сигнала на выходе фотоприемника  $U_{\text{тем}}$  при времени накопления  $T_h$ , можно написать соотношение

$$U_{\text{тем}} / (k \times T_h) = (A \times I_{\text{темн\_S}} + P \times I_{\text{темн\_P}}) / C_{\text{фд}},$$

где  $k$  – коэффициент передачи напряжения от фотодиода до выхода микросхемы (в данном случае  $k \approx 1$ ),  $A$  и  $P$  – площадь донной части и периметр фотодиода,  $C_A$  и  $C_P$  – удельные емкости донной и боковой частей фотодиода,  $I_{\text{темн\_S}}$  и  $I_{\text{темн\_P}}$  – удельные темновые токи донной и боковой частей фотодиода,  $C_{\text{фд}} = (A \times C_A + P \times C_P)$ . Соотношение удобно переписать в виде:

$$I_{\text{темн\_S}} = (U_{\text{тем}} / (k \times T_h) - X_P \times I_{\text{темн\_P}}) / X_S, \quad (1)$$

где  $X_P = P / C_{\text{фд}}$ ,  $X_S = A / C_{\text{фд}}$  – топологические и технологические параметры фотодиода, известные для конкретного процесса и конструкции, а  $U_{\text{тем}} / (k \times T_h)$  – измеряемая величина. На рис.1 для трех, расположенных на одном кристалле, ячеек с фотодиодами размером 8×8, 16×16 и 25×25 мкм, приведены измеренные зависимости темнового сигнала от времени накопления, дающие для каждого фотодиода значение  $U_{\text{тем}} / (k \times T_h)$ . В результате может быть построено семейство зависимостей  $I_{\text{темн\_S}} = f(I_{\text{темн\_P}})$  по уравнению (1) для каждого фотодиода (рис. 2 и 3). Поскольку все структуры созданы в одном процессе, значения  $I_{\text{темн\_S}}$  и  $I_{\text{темн\_P}}$  для каждого из диодов должно быть одинаковыми, чему соответствует точка пересечения всех трех зависимостей, что и дает  $I_{\text{темн\_S}} \approx 0,08...0,14$  фА/мкм<sup>2</sup> и  $I_{\text{темн\_P}} \approx 0,20...0,25$  фА/мкм. Учет зависимостей емкостей от напряжения

сигнала вероятно дал бы несколько более точные значения и меньшее отклонение при переходе от  $T_h = 0,5$  сек к 1 сек.

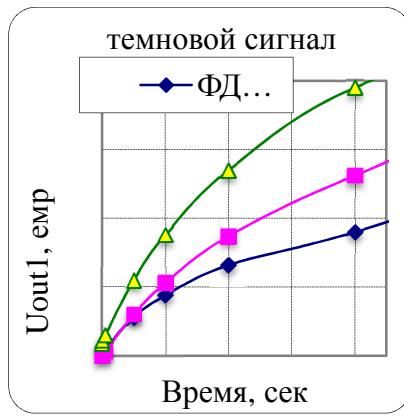


Рис.1. Темновой сигнал

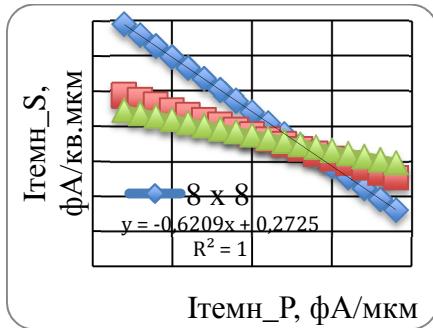


Рис.2.  $T_h = 0,5$  сек

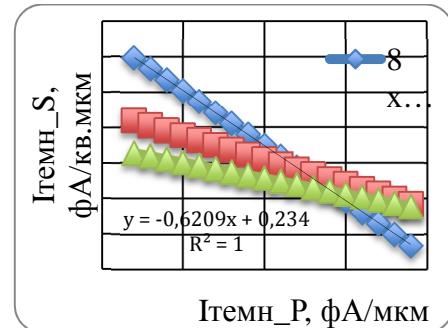


Рис.3.  $T_h = 1$  сек

При изменении формы фотодиода, то есть отношения периметра к площади, изменяются вклады в суммарную емкость и темновой ток от периметра и донной частей фотодиода, что приводит к изменению темнового сигнала. На рис.4 приведено семейство зависимостей темнового сигнала от площади фотодиода при его фиксированном периметре от 8 до 40 мкм. Уменьшение площади фотодиода при заданном периметре приводит к увеличению темнового сигнала. Особенно это существенно для ячеек малых размеров. Максимальной же площади (и минимальному темновому сигналу) соответствует форма окружности.

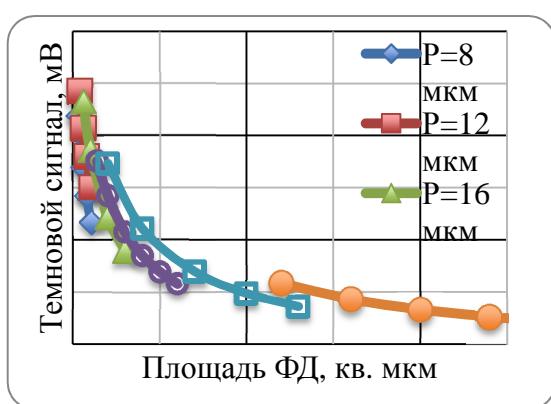


Рис.4. Зависимости темнового сигнала от площади фотодиода при фиксированном периметре от 8 до 40 мкм.

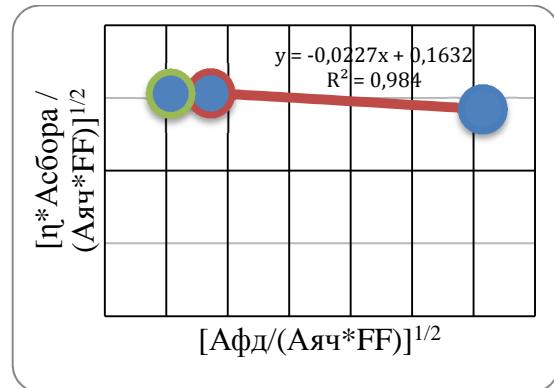


Рис.5. Площадь сбора фотозаряда не изменяется для фотодиодов размером (кв. корень из площади) 3,8; 17 и 33 мкм для ячеек 13' 13, 40' 40 и 40' 40 мкм с филл-факторами 0,51; 0,81 и 0,81 соответственно.

Известно, что поскольку длина диффузии неосновных носителей как правило, превышает размеры современных фотоячеек, то фотозаряд коллективизируется не только из области встроенного поля фотодиода, а со значительно большего объема. Если площадь, с которой фотодиодом собирается фотозаряд составляет  $A_{\text{сбора}}$ , то для освещенности  $E$  можно записать выражение

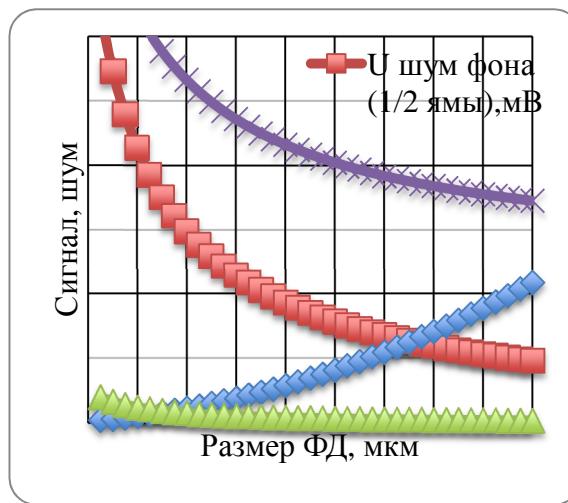
$$A_{\text{сбора}} \cdot E \cdot T_h \cdot \eta = C_{\text{фд}} \cdot U_c,$$

где  $U_c$  – напряжение сигнала,  $\eta$  – константа, связанная со спектральными характеристиками излучателя и приемника, фотометрическим коэффициентом и др. Показательно отношение площади сбора к площади ячейки, открытой для облучения  $A_{\text{яч}} \cdot FF$ , где  $FF$  – филл-фактор, поэтому выражение удобно записать в виде

$$\eta \cdot A_{\text{сбора}} / (A_{\text{яч}} \cdot FF) = C_{\text{фд}} \cdot S_v / (A_{\text{яч}} \cdot FF), \quad (2)$$

где  $S_v = U_c / (E \cdot T_h)$  – интегральная чувствительность. На рис.5 приведено отношение (2) в зависимости от  $A_{\text{фд}} / (A_{\text{яч}} \cdot FF)$ , где  $A_{\text{фд}}$  – площадь фотодиода, для трех фотодиодов, одинаковой конструкции и изготовленных на одной пластине, с филл-факторами от 0,5 до 0,8, расположенных в ячейках размером от 13 × 13 до 40 × 40 мкм. Видно, что при значительном изменении размеров фотодиода, площадь сбора не изменяется, что возможно, если она ограничивается площадью ячейки, открытой для облучения ( $A_{\text{яч}} \cdot FF$ ).

В итоге могут быть оценены (рис.6) темновой сигнал и его шум, максимальная



зарядовая емкость в зависимости от размера фотодиода.

Тем самым можно заключить, что для данной технологии оптимальен фотодиод с максимальной площадью при заданном периметре (например, по форме близкий к окружности или восьмиугольнику). Поскольку показано, что фотозаряд собирается со всей открытой для облучения площади ячейки, площадь фотодиода определяется требованием, в первую очередь, к максимальной зарядовой емкости, а также, в меньшей степени, в зависимости от предполагаемого режима работы прибора, требованием к коэффициенту преобразования заряда в напряжение, поскольку это требование может быть реализовано и другим способом, например за счет накопления не на емкости фотодиода [3].

#### Библиографический список

Рис.6. Зависимости темнового сигнала, его шума, шума фона (при заполнении фоновым сигналом половины зарядовой емкости), максимальной зарядовой емкости от размера квадратного фотодиода со стороной от 2 до 20 мкм при  $T_h = 20$  мс.

1. I.Brouk, K.Alameh, Y.Nemirovsky. Design and characterization of CMOS/SOI image sensors. IEEE trans. of electron devices. 2007, vol. 54, № 3, p. 468-475.
2. А.А.Манцоветов. Чувствительность КМОП-фотоприемников с активным пикселом. Вопросы радиоэлектроники, серия "Техника телевидения", 2014 г., № 2, стр. 18-24.
3. Д.В.Бородин, Ю.В.Осипов, В.В.Васильев. Эффективное использование динамического диапазона КМОП фотоматриц и проблемы его расширения. Вопросы радиоэлектроники, серия "Техника телевидения", 2016 г., № 4, стр. 40-47.

## THE PARAMETERS TO OPTIMIZE CELL OF CMOS IMAGE SENSOR

**Borodin D. V.**

«RTC Inpex, Itd», Mytischy, Russia

JSC "S&P Enterprise "Pulsar", Moscow, Russia

*For the technology used in the manufacture of domestic CIS in the visible range of formats 1024' 1024 cells in size 15' 15 microns and 1280' 1024 cells in size 13' 13  $\mu$ m, defined: the dark current of the bottom part of the photodiode; the dark current of the photodiode perimeter; area of a photo charge collection, which allows to optimize the design of the photo cell.*

**CMOS image sensor, the dark current, the area of the photo charge collection**

**П. С. Баранов<sup>1</sup>, Д. А. Белоус<sup>1</sup>, В. Т. Литвин<sup>2</sup>, А. А. Манцеветов<sup>1</sup>**

<sup>1</sup>СПбГЭТУ «ЛЭТИ»

<sup>2</sup>Санкт-Петербургский горный университет

## Темновые токи твердотельных фотоприемников при высокой рабочей температуре

Получены выражения для темновых токов кремниевых фотоприемников, выполненных по технологиям ПЗС и КМОП при рабочих температурах до +130 °C. Показано, что темновые токи КМОП-сенсоров существенно меньше, чем у матричных ПЗС. Режим электронного затвора позволяет уменьшить количеством темновых электронов и минимизировать их влияние на качество формируемого изображения. Проведенные экспериментальные исследования для различных типов фотоприемников подтвердили полученные теоретические результаты.

**Твердотельный фотоприемник, темновой ток, электронный затвор, матричный ПЗС, КМОП-сенсор**

В настоящее время высокотемпературные телевизионные системы (ТВС) оказываются востребованными в таких областях металлургия, химическая промышленность, варка стекла, физические эксперименты нефтегазодобычи и т. д. Существующая высокотемпературная элементная база позволяет работать до температур +150 °C и более.

Однако, основой современной ТВС является твердотельный кремниевый фотоприемник (ФП), выполненный по технологии ПЗС или КМОП, темновые токи которого экспоненциально растут с ростом температуры. В связи с этим становится актуальным анализ темновых токов твердотельных ФП при высокой рабочей температуре.

Число электронов, генерированных темновым током [1]

$$n_{\text{TT}} = j_{\text{TT}} T A / q,$$

где  $q = 1,6 \times 10^{-19}$  Кл – заряд электрона;  $A = \Delta x \cdot \Delta y$  – площадь светочувствительного элемента;  $T$  – время накопления;  $j_{\text{TT}}$  – плотность темнового тока

$$j_{\text{TT}} = \frac{qn_c W}{2t} + \frac{qv_0 n_c}{2} + \frac{qD_d n_c^2}{L_d N_a}, \quad (1)$$

где первое слагаемое – ток, обусловленный объёмной генерацией в обедненном слое, второе – ток поверхностной генерации, третье – диффузионный ток генерации неосновных носителей в электронейтральной части подложки, поступающий через область обеднения в инверсионный слой;  $n_c$  – собственная концентрация носителей;  $W$  – глубина обедненного слоя;  $\tau$  – время жизни неосновных носителей;  $v_0$  – скорость поверхностной рекомбинации;  $D_d$  – коэффициент диффузии;  $L_d$  – диффузионная длина;  $N_a$  – концентрация акцепторной примеси.

Предположим, что электроны появились вследствие термогенерации. Этот процесс случаен, и число электронов, генерированных за время  $T$ , описывается распределением плотности вероятности Пуассона [2]. Для распределения Пуассона дисперсия численно равна среднему значению.

Накопленному темновому заряду с числом электронов  $n_{\text{TT}}$  также свойственны дробовые шумы [3], определяемые выражением

$$\bar{n}_{\text{ш}_{\text{TT}}} = \sqrt{n_{\text{TT}}} = \sqrt{\frac{j_{\text{TT}}TA}{q}},$$

где  $\bar{n}_{\text{ш}_{\text{TT}}}$  – среднеквадратическое отклонение (СКО) числа шумовых электронов.

Таким образом, СКО результирующего шума ФП при наличии темнового заряда

$$\bar{n}_{\text{ш}} = \sqrt{\bar{n}_{\text{ш}_{\text{TT}}}^2 + \bar{n}_{\text{ш}_{\text{сч}}}^2} = \sqrt{\bar{n}_{\text{ш}_{\text{TT}}} + \bar{n}_{\text{ш}_{\text{сч}}}^2},$$

где  $\bar{n}_{\text{ш}}$  – среднее значение результирующего шума ФП при наличии темнового заряда,  $\bar{n}_{\text{ш}_{\text{сч}}}$  – шумы считывания ФП при отсутствии темнового заряда.

Основной проблемой, возникающей при работе ФП при высокой температуре, является то, что число темновых электронов экспоненциально возрастает с увеличением температуры [4, 5]. Так, плотность темнового тока увеличивается в два раза при повышении температуры на 7...9 °C [1].

В матричных ПЗС со строчным переносом темновые токи генерируются в фотодиодах и в регистрах переноса. В фотодиодах вторая и третья компоненты темнового тока отсутствуют и, к тому же, темновые токи могут быть уменьшены применением режима электронного затвора. В тоже время темновые токи регистров переноса на порядок больше, чем в фотодиодах и не снижаются в режиме электронного затвора.

В КМОП-сенсорах с бегущим затвором темновые токи генерируются только в фотодиодах, а в КМОП-сенсорах с глобальным затвором ни могут возникать и в емкостях хранения, если в качестве таковых используются обратносмещенные диоды [6].

На рис. 1 представлены результаты расчетов значений уровня темнового тока черно-белых твердотельных ФП фирмы Sony в зависимости от температуры.

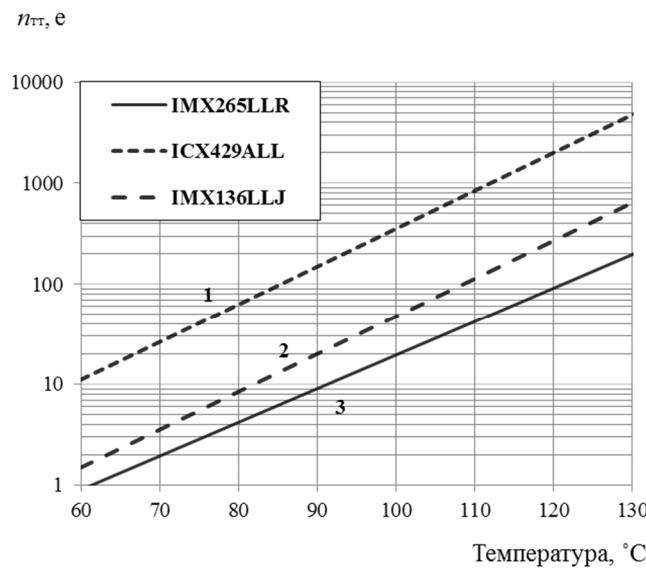


Рис.1. Результаты расчетов уровня темновых токов твердотельных ФП фирмы Sony в зависимости от температуры

График 1 показывает зависимость суммарного темнового тока фотодиодов и регистров МПЗС со строчным переносом ICX429ALL от температуры, график 2 зависимость темнового тока для КМОП-сенсора с бегущим затвором IMX136LLJ, а график 3 – для КМОП-сенсора глобальным затвором IMX265LLR.

Для проверки полученных теоретических результатов был разработан экспериментальный стенд, рис. 2. Стенд включал термостат, в котором располагалась четыре черно-белых ТВ камеры TBK1–TBK4. TBK1 выполнена на КМОП-сенсоре с бегущим затвором IMX136LLJ, TBK2 выполнена на КМОП-сенсоре с глобальным затвором AR0134CS фирмы ON Semiconductor, TBK3 выполнена на КМОП-сенсоре с глобальным затвором IMX265LLR фирмы Sony, TBK4 выполнена на МПЗС ICX429ALL фирмы Sony.

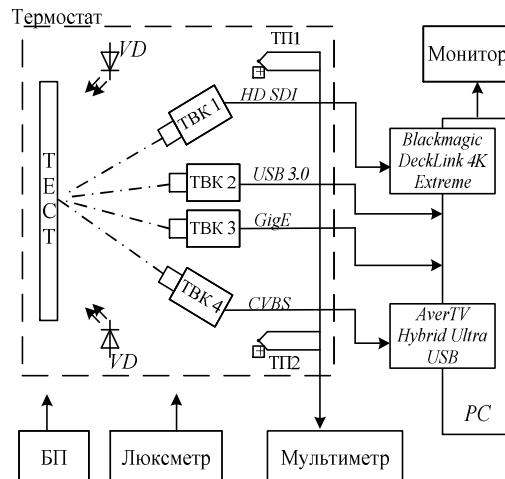


Рис. 2. Структурная схема экспериментального стенда

Для каждой ТВК и выбранных значений температуры в диапазоне от +50 °C до +130 °C снималось 16 последовательных изображений в формате \*.bmp при отсутствии освещения. Усредненное значение цифрового видеосигнала в одном и том же пикселе в последовательности изображений принималось как оценка математического ожидания сигнала, сумма квадратов разностей математического ожидания, поделенная на 15, как оценка несмещенной дисперсии. Отношение сигнал/шум вычислялось как отношение оценки математического ожидания к СКО, равному корню из оценки дисперсии.

На рис. 3 представлены результаты вычисления значения оценки числа темновых электронов для всех перечисленных выше ФП, снятые при двух временах накопления – 20 мс и 2,5 мс в зависимости от температуры. Все зависимости носят экспоненциальный характер.

Для МПЗС со строчным переносом ICX429ALL, рис. 3а, при температуре +130 °C и времени накопления 20 мс темновые электроны занимают до трети емкости фотодиода, резко ухудшая отношение сигнал/шум и динамический диапазон ФП. Уменьшение времени накопления до 2,5 мс приводит к незначительному снижению, что объясняется тем, что превалирующими являются темновой ток вертикальных регистров, на которые режим электронного затвора не оказывает влияния. Картина темновых токов вертикальных регистров имеет вертикальную структуру, т. к. их среднее значение усредняется вдоль вертикального регистра.

Для КМОП-сенсора с глобальным затвором AR0134CS, рис. 3б, темновые электроны при температуре +130 °C занимают до 40% максимальной ёмкости фотодиода. При уменьшении времени накопления среднее число темновых электронов падает, но не пропорционально времени накопления. Это объясняется тем, что помимо темновых электронов, генерируемых в фотодиодах, они также образуются в емкостях хранения. Так

как время хранения различно для разных строк, образуется градиент среднего уровня темнового заряда от первых строк к последним.

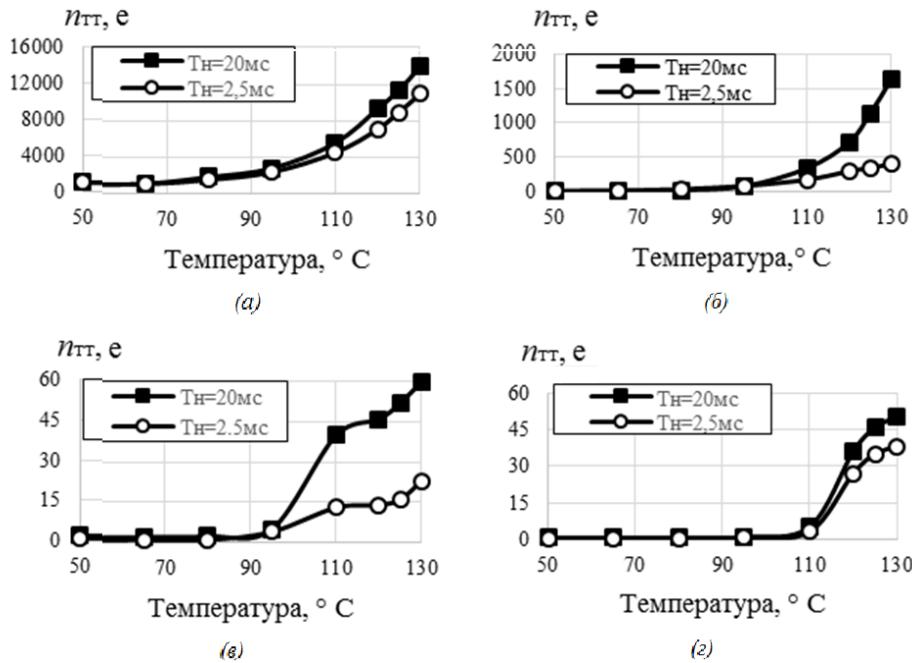


Рис. 3. Результаты измерений значений числа темновых электронов, снятые при двух временах накопления – 20 мс и 2,5 мс в зависимости от температуры: МПЗС со строчным переносом *ICX429ALL* – *а*, КМОП-сенсор с глобальным затвором *AR0134CS* – *б*, КМОП-сенсор с бегущим затвором *IMX136LLJ* – *в*, КМОП-сенсор с глобальным затвором *IMX265LLR* – *г*

На рис. 4 представлены распределения среднего значения числа темновых электронов вдоль столбца КМОП-сенсора с глобальным затвором AR0134CS при различных значениях температуры. Начиная с температуры +120 °C по мере увеличения номера строки появляется описанный выше градиент темнового заряда, который становится практически двукратным при температуре +135 °C.

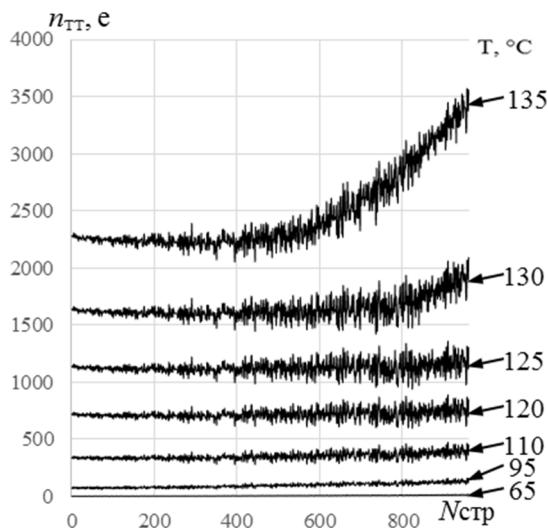


Рис.4. Распределения среднего значения числа темновых электронов вдоль столбца КМОП-сенсора с глобальным затвором AR0134CS при различных значениях температуры

Из графиков для КМОП-сенсора с бегущим затвором IMX136LLJ, рис. 3в, и КМОП-сенсора с глобальным затвором IMX265LLR, рис. 3г, среднее значение числа темновых

электронов составило 40...50 при температуре +130 °С и времени накопления 20 мс в центре кадра. При уменьшении времени накопления до 2,5 мс их число уменьшилось до 1,5...2,5 электронов. Полученные значения числа темновых электронов гораздо ниже, чем указано в технической документации на матрицы.

На рис. 5 приведены значения оценок СКО шума в зависимости от температуры для КМОП-сенсора с бегущим затвором IMX136LLJ и КМОП-сенсора с глобальным затвором IMX265LLR.

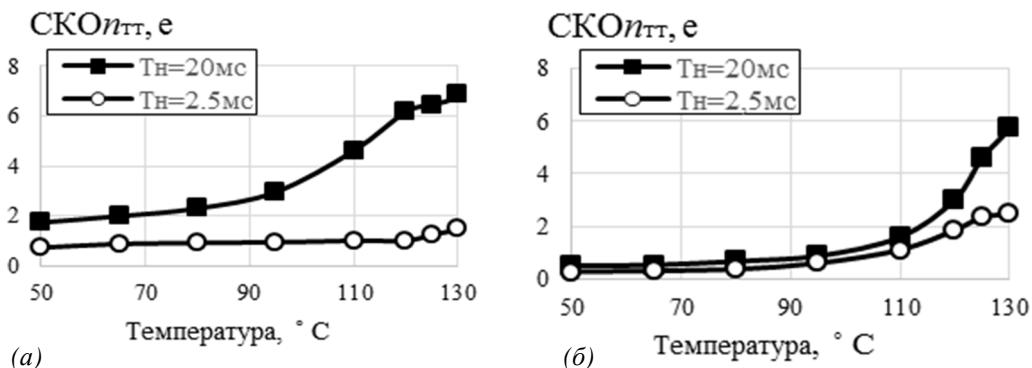


Рис. 5. Оценки СКО шума в зависимости от температуры:

КМОП-сенсор с бегущим затвором IMX136LLJ – (а);  
КМОП-сенсор с глобальным затвором IMX265LLR – (б)

Значения СКО возрастают, в первую очередь, за счет увеличения дробовых шумов темнового заряда, однако следует учитывать, что такие компоненты как шум сброса, тепловые шумы транзисторов и др. также возрастают с температурой.

### Выводы

1. Проведённые расчёты показывают, что в МПЗС основной вклад в темновой ток вносят регистры переноса, и при температуре +130 °С его уровень занимает значительную часть динамического диапазона. Для КМОП-сенсоров фирмы Sony как с бегущим затвором, так и с глобальным, расчётный уровень числа темновых электронов при высокой температуре, составляет не более нескольких сотен и может быть снижен использованием режима работы с электронным затвором.

2. Проведённые экспериментальные исследования подтвердили экспоненциальный характер возрастания числа темновых электронов с повышением температуры.

3. Для МПЗС со строчным переносом ICX429ALL при температуре +130 °С и времени накопления 20 мс значение числа темновых электронов составило 14000. Уменьшение времени накопления до 2,5 мс приводит к незначительному снижению их количества, что объясняется тем, что превалирующими являются темновые токи вертикальных регистров.

4. Для КМОП-сенсора с глобальным затвором AR0134CS при температуре +130 °С и времени накопления 20 мс число темновых электронов составило 1700 электронов. При уменьшении времени накопления до 2,5 мс среднее число темновых электронов снизилось, но не пропорционально времени накопления. Это объясняется тем, что темновые электроны образуются также и в емкостях хранения. Последнее приводит к возрастанию темнового заряда по мере увеличения номера строки, т. к. время хранения сигнала на этих емкостях увеличивается.

5. Для КМОП-сенсоров фирмы Sony как с бегущим, так и глобальным затвором при температуре +130°C и времени накопления 20 мс значения уровня темновых зарядов оказались аномально малы и составили менее 50 электронов, что значительно ниже паспортных данных. Необходимо отметить, что в последних реализациях данного КМОП-сенсора его рабочая температура повышенена до +75 °C и на самом кристалле установлен датчик температуры.

#### Библиографический список

1. Физика полупроводниковых приборов: В 2-х книгах. Кн. 1. Пер. с англ. – 2-е перераб. и доп. изд. – М.: Мир, 1984. – 456 с.
2. Приборы с зарядовой связью / Под ред. М. Хоувза и Д. Моргана: Пер. с англ. – М.: Энергоиздат, 1981. – 376 с.
3. Цифровое преобразование изображений / Быков Р. Е., Фрайер Р., Иванов К. В., Манцветов А. А. Под ред. Профессора Быкова Р. Е. – М.: Горячая линия – Телеком, 2003. – 228 с.
4. Durini D. High Performance Silicon Imaging. Fundamentals and Applications of CMOS and CCD Image Sensor. Woodhead Pub. 2014. 450 p.
5. Image sensors and signal processing for digital still cameras / Ed. by J. Nakamura CRC Press, Inc. Boca Raton, FL, USA 2006. 322 p.
6. Dark current of the solid-state imagers at high temperature. P. Baranov, D. Belous, V. Litvin, A. Mantsvetov // Proceedings of 2017 ElConRus IEEE, Saint-Petersburg, Russia, 1-3 February 2017.

#### DARK CURRENTS OF SOLID-STATE IMAGE SENSORS AT HIGH OPERATING TEMPERATURE

**P. S. Baranov<sup>1</sup>, D. A. Belous<sup>1</sup>, V. T. Litvin<sup>2</sup>, A. A. Mantsvetov<sup>1</sup>**

<sup>1</sup>Saint-Petersburg state electrotechnical University "LETI"

<sup>2</sup>National Mineral Resources University

*Expressions are obtained for dark currents of silicon photodetectors made using CCD and CMOS technology at operating temperatures up to +130 °C. It is shown that the dark currents of CMOS sensors are much smaller than for the matrix CCDs. The electronic shutter mode allows reducing the number of dark electrons and minimizing their effect on the quality of the image being formed. The carried out experimental researches for various types of photodetectors have confirmed the received theoretical*

**Solid-state image sensor, dark current, electronic shutter, matrix CCD, CMOS sensor**

**Д. А. Белоус**  
СПбГЭТУ «ЛЭТИ»

# Оптимизация температурного режима твердотельного фотоприемника при работе в ближней инфракрасной области спектра

*В работе исследована возможность оптимизации температурного режима КМОП-сенсора с целью повышения отношения сигнал/шум при работе в ближней ИК-области спектра, в частности, на длине волны 1064 нм.*

**КМОП-сенсор, матричный ПЗС, инфракрасный участок спектра, квантовая эффективность, высокая температура, лидар**

Использование ближней инфракрасной (ИК) области спектра для работы телевизионно-лазерных систем обусловлено наличием окна прозрачности атмосферы и меньшим рассеянием ИК-излучения на аэрозолях. К тому же существуют непрерывные и импульсные лазеры на алюмоиттриевом гранате (YAG) с длиной волны излучения 1064 нм, способные генерировать наносекундные импульсы с частотой до сотен герц и энергией в единицы и более джоулей.

В современных лидарных системах часто используются лазеры с длиной волны излучения 1064 нм. В своем составе такие системы, как правило, имеют стробируемые электронно-оптические преобразователи (ЭОП), сочлененные с твердотельным фотоприемником (ФП). Использование ЭОП приводит к излишнему усложнению системы, резкому повышению ее стоимости, снижению надежности и долговечности.

Современные твердотельные ФП, работающие в режиме электронного затвора, позволяют достичь времени экспозиции, соизмеримого со стробируемым ЭОП.

Современные матричные ПЗС и КМОП-сенсоры фирмы Sony позволяют работать с временами накопления десятки нс [1, 2]. Однако квантовая эффективность (quantum efficiency – QE) и тех и других типов ФП на длине волны 1064 нм невелика и составляет, как правило, единицы процентов.

Современные КМОП-сенсоры характеризуются незначительным уровнем темновых токов [3], в связи с чем могут функционировать при высоких температурах.

В твердотельных ФП значение QE определяется многими факторами, такими, как коэффициент отражения поверхности кремния, наличие и материал микролинзы, глубина обедненной области кремния, объемная плотность легирующих примесей и т. д. Кроме того, в диапазоне длин волн 900...1100 нм значение QE во многом определяется шириной ЗЗ.

Зависимость ширины ЗЗ полупроводника описывается формулой Варшни [4]:

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta},$$

где Т – абсолютная температура;  $\alpha$  – температурный коэффициент изменения ширины ЗЗ;  $\beta$  – константа смещения шкалы температуры.

С увеличением температуры кристалла кремния ширина его запрещенной зоны (ЗЗ) уменьшается, рис.1 [5].

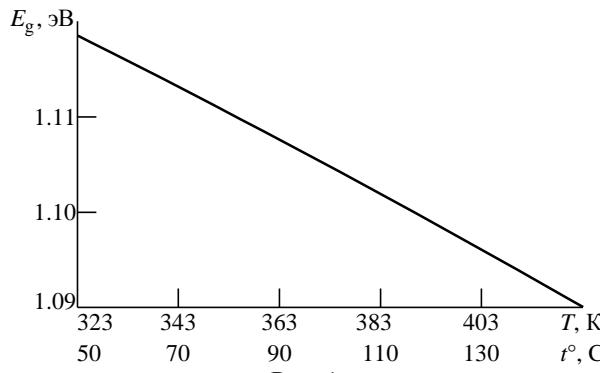


Рис. 1

Для оценки вероятности того, что фотон определенной длины волны сгенерирует электрон, который преодолеет ЗЗ кремния, найдена зависимость отношения энергии фотона к ширине ЗЗ кремния (табл. 1) [5]. Расчеты выполнены для температуры +50 °C. Энергия фотона определена в соответствии с формулой Планка

$$E_{\text{ph}}(\lambda) = hc/\lambda,$$

где  $h = 4.315 \times 10^{-15}$  эВ·с – постоянная Планка;  $c = 3.0 \times 10^8$  м/с – скорость света в вакууме.

Зависимость отношения энергии фотона к ширине ЗЗ кремния

Таблица 1

$\lambda$ , нм	900	940	980	1020	1060	1100
$E_{\text{ph}}(1)/E_g(323)$	1.223	1.181	1.156	1.132	1.110	1.009

Таким образом, с ростом температуры ширина ЗЗ кремния уменьшается, а отношение  $E_{\text{ph}}(1)/E_g(T)$  при фиксированной длине волны – увеличивается, что приводит к росту значения QE.

Результаты расчета температурных зависимостей квантовой эффективности при работе на длинах волн 940 нм и 1064 нм [5] представлены на рис. 2. Кривые 1 построены для матричного ПЗС со строчным переносом ICX429ALL фирмы Sony [6], содержащего  $752 \times 582$  пикселя размером  $8.3 \times 8.6$  мкм<sup>2</sup>; кривые 2 – для КМОП-сенсора с бегущим затвором IMX136LLJ той же фирмы [7] с числом пикселов  $1920 \times 1200$  и размером пикселя  $2.8 \times 2.8$  мкм<sup>2</sup>.

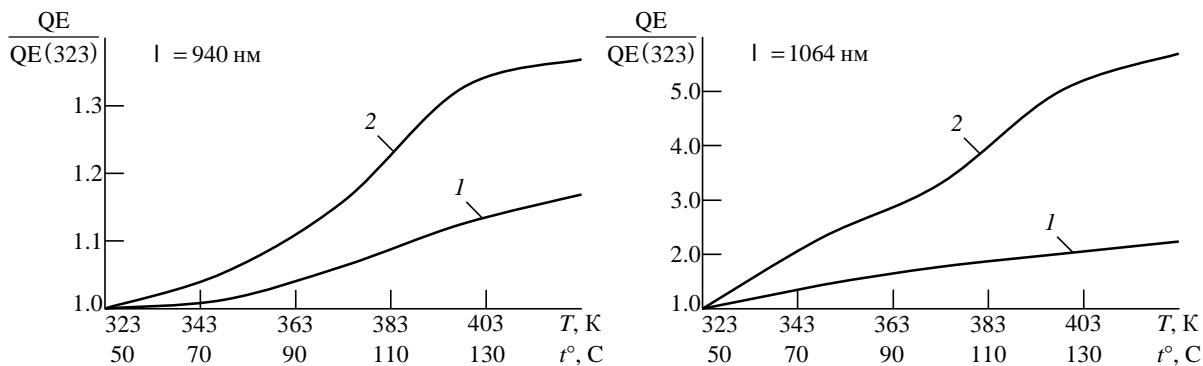


Рис. 2

На длине волны 1064 нм при температуре +130 °C выигрыш QE по отношению к температуре 323 К (+50 °C) составляет до 2 раз для матричного ПЗС и до 5 раз для КМОП-сенсора.

Для проверки результатов теоретического расчета был разработан экспериментальный стенд (рис. 3). Стенд содержал термостат, в котором для реализации черно-белого перепада, по которому измерялось значение видеосигнала, располагались модель абсолютно черного тела (АЧТ) и диффузный отражатель (ДО), освещаемые рядом ИК-светодиодов VD с максимумами длины волны излучения 940 нм и 1050 нм, а также непрерывным лазерным излучением с длиной волны 1064 нм. В термостате располагалась черно-белая телевизионная камера (ТВК), выполненная на КМОП-сенсоре с бегущим затвором IMX136LLJ фирмы Sony с интерфейсом HD-SDI. У ТВК отключались все автоматические регулировки, устанавливавшиеся режим  $g=1$ . Время накопления составляло 20 мс или 2.5 мс. Выходной сигнал ТВК передавался на плату захвата видео, установленную в персональном компьютере (PC). Температура в термостате изменялась в диапазоне  $+50^{\circ}\text{C}...+130^{\circ}\text{C}$  и контролировалась термопарами ТП1 и ТП2, подключенными к мультиметру.

Питание всех устройств осуществлялось от блока питания БП.

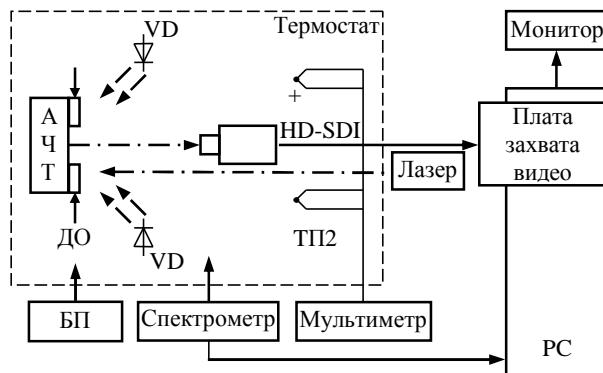


Рис. 3

На рис. 4 приведены результаты эксперимента, которые показывают рост амплитуды выходного сигнала ФП относительно сигнала при температуре 323 К ( $+50^{\circ}\text{C}$ ), обусловленный увеличением QE с повышением температуры при работе в ближнем ИК-диапазоне для длин волн 940, 1050 и 1064 нм.

При работе на длине волны 1064 нм и температуре  $+130^{\circ}\text{C}$  увеличение QE составило 4 раза, что близко к теоретически полученному значению.

Таким образом, влияние уменьшения ширины ЗЗ при повышении температуры на увеличение значения QE достаточно велико и может быть использовано для оптимизации температурного режима работы ФП.

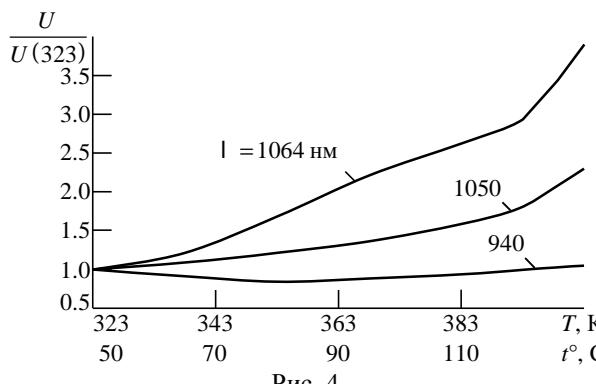


Рис. 4

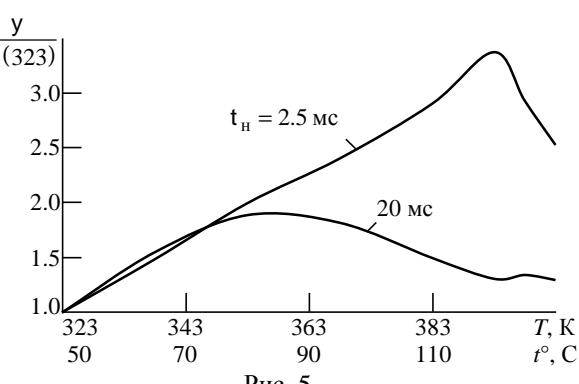


Рис. 5

Анализируя совместно температурные зависимости роста QE и СКО шума из [3] можно сделать вывод о том, что существует оптимальное значение температуры кристалла ФП по критерию отношения сигнала/шум  $\psi$ .

На рис. 5 представлены зависимости относительного увеличения этого отношения при увеличении температуры для двух времен накопления КМОП-сенсора ICX136LLJ фирмы Sony. Из них следует, что для времени накопления 20 мс оптимальным является температура кристалла +80 °C, а при включении электронного затвора ( $t_h = 2.5$  мс) это значение возрастает до +120 °C. При этом выигрыш в отношении сигнал/шум в первом режиме составляет 1.9 раз, во втором – 3.4 раза.

Таким образом, оптимизация температурного режима работы твердотельного ФП приводит к повышению пороговой чувствительности при малых значениях полезного сигнала. При больших значениях сигнала определяющим компонентом будет фотонный шум и выигрыш будет пропорционален росту QE при рабочей температуре.

#### Библиографический список

1. Управление режимом накопления в твердотельных фотоприемниках / А. А. Умбитаев, А. К. Цыцулин, А. А. Манцветов, В. В. Козлов, А. Е. Рычажников, П. С. Баранов, А. В. Иванова // Опт. Журнал. 2012. Т. 79, № 11. С. 84–92.
2. URL: [http://www.sony-semicon.co.jp/products\\_en/new\\_pro/february\\_2016/imx253\\_255\\_e.html](http://www.sony-semicon.co.jp/products_en/new_pro/february_2016/imx253_255_e.html) (дата обращения 06.05.2017)
3. Стрижнев К. В., Белоус Д. А., Баранов П. С., Литвин В. Т., Манцветов А. А., Михайлов В. А. Анализ количества темновых электронов твердотельных фотоприемников при высокой рабочей температуре // Вопросы радиоэлектроники. Сер. Техника телевидения. 2017. Вып. 2. С. 31–40.
4. Varshni Y. P. Temperature dependence of the energy gap in semiconductors // Physica. 1967. Vol. 34, no 1, pp. 149–154.
5. Белоус Д. А. Чувствительность твердотельных фотоприемников в ближней инфракрасной области спектра при высокой температуре // Вопросы радиоэлектроники. Сер. Техника телевидения. 2017. Вып. 2. С. 41–47.
6. URL: [www.npk-photonica.ru/images/icx429all2.pdf](http://www.npk-photonica.ru/images/icx429all2.pdf) (дата обращения 06.05.2017)
7. URL: [http://www.sony-semicon.co.jp/products\\_en/IS/sensor0/img/product/cmos/imx136lqj\\_llj.pdf](http://www.sony-semicon.co.jp/products_en/IS/sensor0/img/product/cmos/imx136lqj_llj.pdf) (дата обращения 06.05.2017)

#### OPTIMIZATION OF THE TEMPERATURE MODE OF A SOLID-STATE IMAGE SENSOR IN A SHORT-WAVE INFRARED REGION

*Belous Darya A.*

Saint Petersburg Electrotechnical University “LETI”

*In operation the possibility of optimization of temperature condition of a CMOS sensor for the purpose of increase in the relation signal/noise by operation in short-range IR-area of a range, in particular, on wavelength of 1064 nanometers is probed.*

**CMOS sensor, matrix CCD, infrared section of a range, quantum efficiency, high temperature, lidar**

**П. В. Столбов**  
СПбГЭТУ каф. РЭС

## Высокоскоростная специализированная матричная камера с расширенным динамическим диапазоном и сжатием изображения

*Рассмотрена аппаратная модернизация стереотелевизионной системы обнаружения контактного провода путем создания высокоскоростной матричной камеры с расширенным динамическим диапазоном и сжатием изображения. Представлена структура разрабатываемой камеры и разработанная в ПЛИС архитектура сжатия изображения.*

**Высокоскоростная матричная камера, расширение динамического диапазона, сжатие по алгоритму JPEG, обработка изображения в ПЛИС**

Одним из способов реализации контроля инфраструктуры контактных сетей железных дорог является использование специального вагона-лаборатории испытаний контактной сети (ВИКС), который осуществляет регулярные объезды главных путей электрифицированных железнодорожных магистралей. Для обнаружения и бесконтактного измерения положения (высота, смещение) контактного провода (КП) в составе ВИКС используется стереотелевизионная система (СТС).

При обнаружении КП в существующей СТС существуют проблемы связанные с трудностью выделения сигнала о КП на фоне помех. Основными причинами, вызывающими данную проблему являются:

- ограниченный динамический диапазон используемого линейного ПЗС сенсора;
- неоптимальные алгоритмы обнаружения с точки зрения помехоустойчивости.

Для повышения эффективности работы СТС требуется ее усовершенствование. Усовершенствование идет по двум направлениям:

- разработка новых алгоритмов обработки изображения, устойчивых к изменяющимся внешним условиям;
- аппаратная модернизация СТС.

В рамках данной работы будет уделено внимание аппаратной модернизации, а именно замены линейного ПЗС сенсора высокоскоростным матричным сенсором с расширенным динамическим диапазоном и создании новой высокоскоростной специализированной матричной камеры на его основе.

Это позволит:

- добиться лучших характеристик обнаружения в условиях меняющегося шумового фона за счет расширения динамического диапазона по сравнению с линейным сенсором;
- обеспечить повышение информативности системы за счет более высокого быстродействия;
- применить новые алгоритмы обнаружения (в частности возможна внутrikадровая обработка изображения).

Так как необходимо реализовать видеосистему реального времени, встраиваемую в существующий программно-вычислительный комплекс, то, исходя из пропускной способности канала передачи данных и высокого требования к частоте кадров получаемых с камер, необходимо реализовать сжатие изображения непосредственно внутри камеры.

Представленные на рынке коммерчески оправданные варианты обладают следующими скоростными характеристиками: около 30 кадров/с FHD изображения в формате MJPEG, что не удовлетворяет скоростным требованиям видеосистемы. Отметим, что они также требуют затраты на встраивание в программно вычислительный комплекс ВИКС. Это обуславливает целесообразность разработки новой камеры, удовлетворяющей техническим требованиям.

В рамках модернизации была осуществлена замена линейного ПЗС сенсора высокоскоростным матричным сенсором CMV2000 с глобальным затвором и расширенным динамическим диапазоном. Расширение динамического диапазона до 90 дБ достигается за счет использования нелинейной свет - сигнальной характеристики светочувствительных элементов сенсора, представленной на рисунке 1.

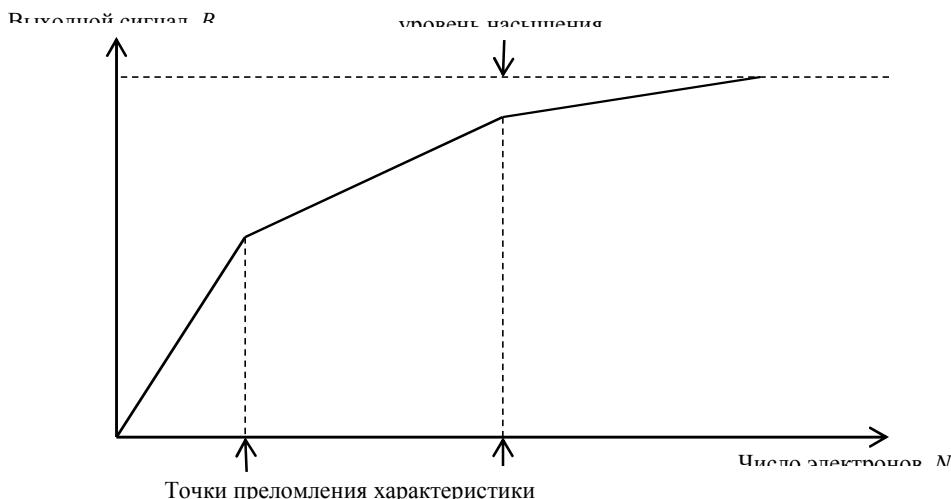


Рис. 1. Нелинейная свет - сигнальная характеристика светочувствительного элемента

Как видно из рисунка 2, это позволит обнаруживать объекты, расположенные на фоне широко варьируемой световой обстановки.

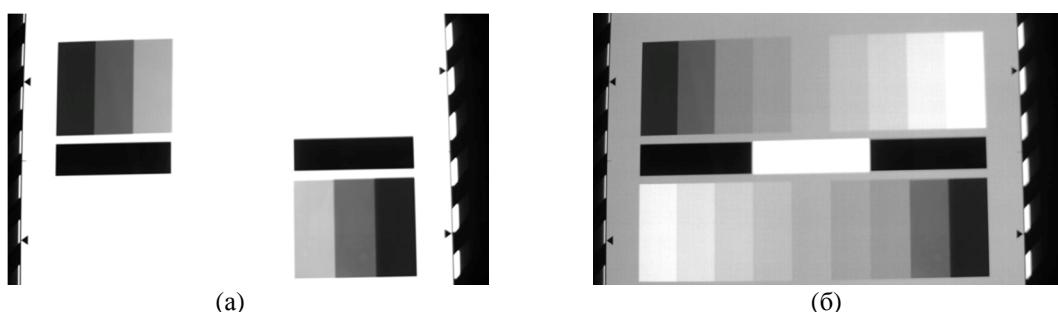


Рис. 2. (а) – снимок, сделанный при линейном отклике. (б) – снимок, сделанный при кусочно-линейном отклике

В матрице используется 16 LVDS - каналов для вывода изображения, а также 1 управляющий LVDS - канал, обеспечивающий правильное получение кадра на схеме захвата изображения. Для управления режимом работы матрицы в ней реализован последовательный SPI – интерфейс, по которому производится программирование таких параметров как: выбор области интереса, управление затвором, управление свет - сигнальной характеристикой, выбор используемых LVDS каналов, определение разрядности АЦП и др. Функциональная схема, представленная на рисунке 3, отображает основные блоки разрабатываемой камеры. Сигнальный процессор осуществляет обработку захваченного изображения и передачу полученной информации на ПК, на котором идет сохранение и вывод результатов. Также он обеспечивает передачу

управляющих сигналов для матрицы в соответствии с выбранным режимом работы камеры.

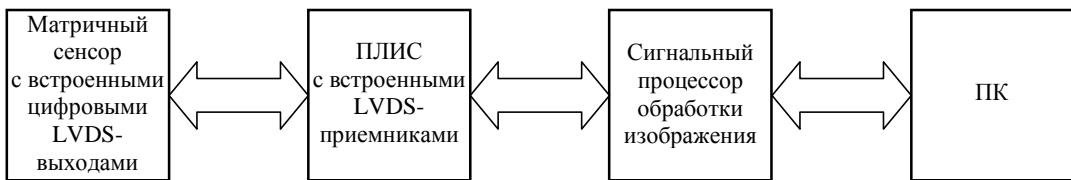


Рис. 3. Структурная схема разрабатываемой камеры

В качестве схемы захвата и сжатия изображения с сенсора используется ПЛИС фирмы Intel семейства Cyclone V. Особенностью данной ПЛИС является наличие встроенных приемников LVDS сигналов, обеспечивающих согласование волновых сопротивлений, соответствующих стандарту LVDS. С алгоритмом захвата изображения с сенсора можно ознакомиться в [1].

В ПЛИС было реализовано ядро сжатия по стандарту JPEG, позволяющее варьировать коэффициент сжатия изображения. При использовании стандарта сжатия JPEG возможно использовать стандартные, представленные в [2] таблицы кодов Хаффмана. Это дает возможность эффективно использовать ограниченный объем высокоскоростной встроенной в ПЛИС динамической памяти, однократно обращаясь к блокам пикселов 8\*8 для их кодирования и последующей передачи, освобождая область памяти для сохранения пикселов следующего кадра. Также стандарт предоставляет возможность параллельного кодирования различных областей одного кадра. Таким образом, благодаря применению конвейерной архитектуры вычислений, а также распараллеливания процессов обработки потенциальная скорость получения сжатого изображения формата 2048x1088 в разработанной в ПЛИС схеме достигает 70 кадров/с. При этом, изменяя коэффициент сжатия, возможно обеспечить удовлетворяющую пропускной способности канала приема видеоданных скорость передачи. Разработанное устройство также осуществляет передачу на сигнальный процессор и несжатых строк напрямую из ОЗУ. Структурная схема сжатия по стандарту JPEG в ПЛИС представлена на рис 4.

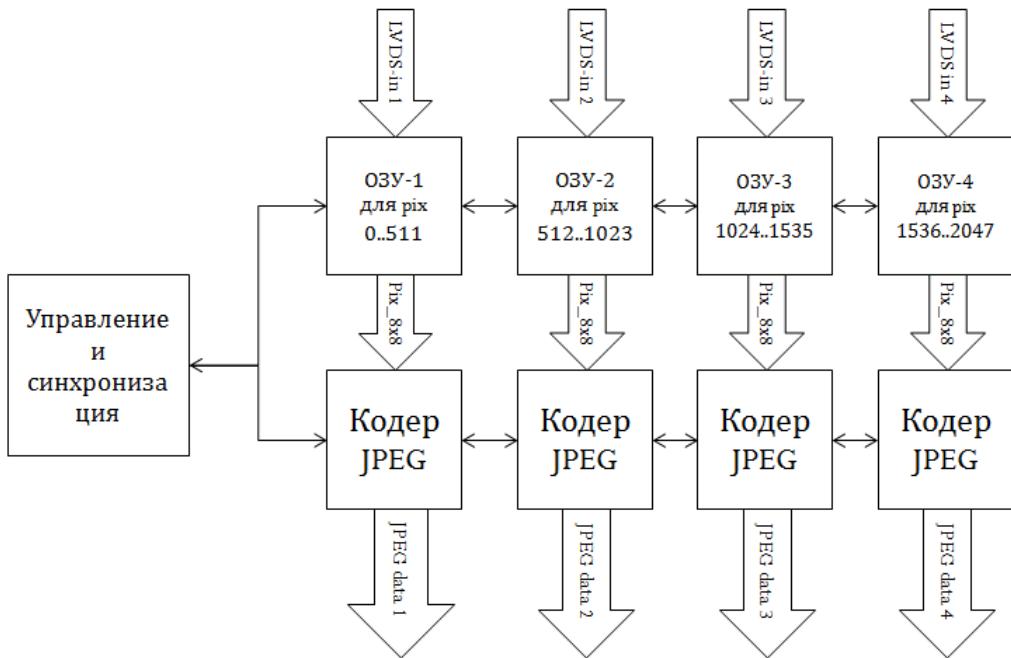


Рис. 4. Структурная схема сжатия изображения

Таким образом, разрабатываемая видеосистема удовлетворяет скоростным требованиям и дает возможность повысить эффективность обнаружения искомых объектов.

#### Библиографический список

1. Столбов П. В., Сиротинин В. И., Воронин А. В. Матричная камера в составе телевизионной системы для обнаружения контактного провода / П. В. Столбов, В. И. Сиротинин, А. В. Воронин // Известия высших учебных заведений России. Радиоэлектроника. – 2016. - №4. - с. 50 - 55.
2. Information technology – digital compression and coding of continuous-tone still images – Requirements and guidelines. [Электронный ресурс] // URL: <https://www.w3.org/graphics/jpeg/itu-t81.pdf> (дата обращения 15.05.2017)

### **HIGH-SPEED SPECIALIZED MATRIX CAMERA WITH EXTENDED DYNAMIC RANGE AND IMAGE COMPRESSION**

**Stolbov P.V.**

Saint Petersburg Electrotechnical University "LETI"

*The hardware upgrade of the stereo-TV system for detecting the contact wire is presented in this paper*

**High-speed matrix camera, dynamic range expansion, JPEG compression, image processing in FPGA.**

**В. С. Шулика, А. А. Шведов**  
ООО «НПК «Фотоника», С-Петербург

# Современные матрицы для получения изображения в ультрафиолетовом, видимом и инфракрасном спектральных диапазонах

*В представленном докладе рассмотрены существующие на сегодняшний день современные и перспективные матрицы для получения изображения в различных спектральных диапазонах с указанием области применения таких приборов. Приведено сравнение характеристик современных КМОП- и ПЗС- матриц, выделены основные преимущества каждой из технологий.*

**КМОП, ПЗС, ультрафиолетовый диапазон, видимый диапазон, инфракрасный диапазон, матрицы, гиперспектральные системы**

Рассматриваемые спектральные диапазоны включают в себя УФ область спектра (UV), видимый диапазон, ближний ИК-диапазон, SWIR (Коротковолновая ИК-область спектра) и LWIR – длинноволновая ИК-область спектра.

Особый интерес вызывают сегодня так называемые гиперспектральные системы, позволяющие получать спектральный слепок изучаемого объекта или группы объектов. Такие приборы находят все больше областей для их применения: от систем спасения до интеллектуальных промышленных комплексов.

Будут также указаны основные особенности ближнего и дальнего ИК-диапазонов, озвучены перспективы развития направления и основные параметры, представляющие наибольший интерес для современных видеосистем.

## Библиографический список

1. Ермаков О. Н. Прикладная оптоэлектроника. М. Техносфера, 2004. – 416 с.
2. CCD vs CMOS [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://www.teledynedalsa.com/imaging/knowledgecenter/appnotes/ccd-vs-cmos/>. – Загл. с экрана. – (12.10.2013).
3. Сысоева С.В. Актуальные технологии и применения датчиков автомобильных систем активной безопасности // Компоненты и технологии. – 2006. – №11. (–С. 1 -10).
4. WuXingxing, LiuJinguo, ZhouHuaide, ZhangBoyan (Чанчуньский институт оптики, точной механики и физики, Китайская академия наук, Чанчунь, Китай). Космические видеосистемы на основе EMCCD и CMOS для работы в условиях сверхнизкой освещённости. Публикация на русском языке на сайте «НПК Фотоника» <http://www.npk-photonica.ru/images/kosmicheskie-videosistemy-na-osnove-emccd-i-cmos-dlya-raboty-v-usloviyah-sverhnizkoy-osveschyonnosti-pdf123526.pdf>

## PRESENT-DAY IMAGE SENSORS IN UV, VIS AND IR SPECTRAL RANGES

**Vladislav Shulika, Alexander Shvedov**  
NPK Fotonika, Saint Petersburg

*The spectral ranges under consideration include the UV spectral region (UV), the visible range, the near infrared range, the SWIR (Short-wave IR spectral region), and the LWIR-long-wave IR spectral region. The characteristics of CMOS and CCD matrices will be compared, and the main advantages of each of the technologies are outlined.*

**CMOS, CCD, UV spectral range, VIS spectral range, IR spectral**

**И. В. Богданов, А. Н. Величко, А. В. Степовой,  
В. С. Шеленговский**  
АО «КБточмаш им. А. Э. Нудельмана»

## Метод физического моделирования температурного контраста

*В работе рассмотрен метод физического моделирования температурного контраста в лабораторных условиях для оценки алгоритмов работы оптико-электронных систем космического назначения. Описан экспериментальный стенд и представлены результаты испытаний. Обоснованы состав и требования к данному типу аппаратуры физического моделирования.*

**Абсолютно чёрное тело, оптико-электронные приборы, стенд, температурный контраст, фоно-целевая обстановка**

В данной работе рассматривается метод физического моделирования температурного контраста для оценки алгоритмов, основанных на сравнении сигналов в разных спектральных диапазонах, оптико-электронных приборов (ОЭП) космического назначения [1, 2]. Подобные алгоритмы используются космическими аппаратами (КА) для наблюдения за объектами искусственного естественного происхождения.

Широкое распространение для условий космического пространства получили ОЭП работающие в средне- и дальневолном инфракрасном (ИК) диапазоне. Одной из задач таких приборов является поиск на слабоизлучающем фоне космоса требуемых малоразмерных объектов среди точечных источников с достаточно близкими энергетическими параметрами [3]. Для оценки эффективности алгоритмов поиска применяются лабораторные стенды формирования фоно-целевой обстановки (ФЦО) в угловом поле ОЭП. Воспроизводимая ФЦО должна отвечать требованиям таким как однородный и низкотемпературный фон и наличие определённого температурного контраста объектов регистрации.

В АО «КБточмаш им. А. Э. Нудельмана» были проведены исследования по разработке метода физического моделирования температурного контраста, которые могут быть использованы для создания условий ФЦО по отработке алгоритмов поиска ОЭП КА.

Экспериментальный стенд (рис. 1) собран на базе оптической скамьи ОСК-2ЦЛ, тепловизионной камеры (ТПВК), двух коллиматоров, плоского зеркала, «зеркальной» диафрагмы с блендоид и двух моделей абсолютно чёрного тела (АЧТ).

Оптическая схема стенда состоит из двух каналов – I и II. Канал I служит для формирования излучения от заданного объекта регистрации и включает в себя: модель АЧТ поз. 6 ограниченную диафрагмой поз. 5, плоское зеркало поз. 3 и коллиматор поз. 2. Канал II необходим для формирования однородного и низкотемпературного фона вокруг объекта регистрации и содержит: модель АЧТ поз. 8, коллиматор поз. 7, плоское зеркало поз. 3 и диафрагму поз. 5, на которую проецируется излучение модели АЧТ поз. 8.

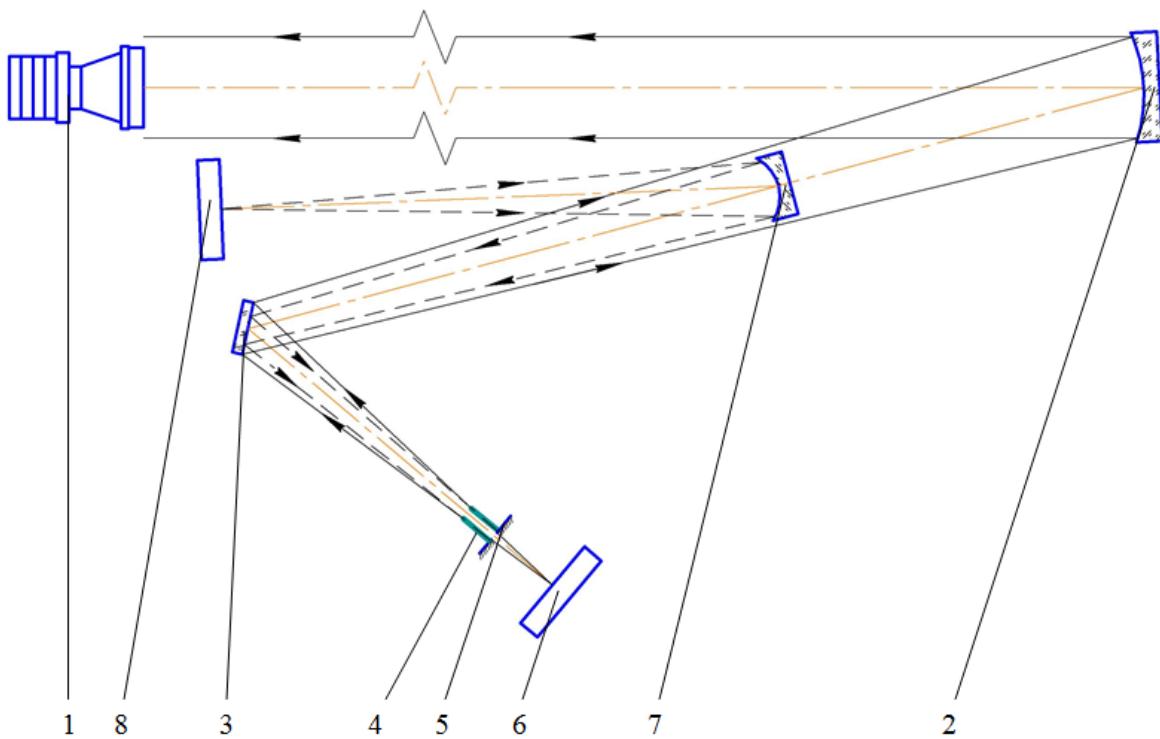


Рис. 1. Схема экспериментального стенда. 1 – ТПВК; 2 – коллиматор (канал I); 3 – плоское зеркало (каналы I и II); 4 – бленда; 5 – «зеркальная» диафрагма (каналы I и II); 6 – модель АЧТ (канал I); 7 – коллиматор (канал II); 8 – модель АЧТ (канал II)

Выходной зрачок длиннофокусного коллиматора поз. 2 создаёт изображение объекта регистрации и сопряжен с входным зрачком объектива ТПВК поз. 1. Объектом регистрации служит точечный источник, в качестве которого выступает модель АЧТ поз. 6 с диафрагмой поз. 5. На «зеркальной» поверхности диафрагмы поз. 5 дополнительным коллиматором поз. 7 проецируется низкотемпературное излучение от модели АЧТ поз. 8. За счёт выбранной геометрии оптической системы (каналов I и II) удаётся сформировать два независимых друг от друга сигнала, создавая, таким образом, заданный температурный контраст.

Для уменьшения габаритных размеров стенда использовано плоское зеркало поз. 3, поворачивающее оптическую ось системы. Бленда поз. 4 служит для уменьшения телесного угла засветки и, как следствие, уменьшения вредного (рассеянного) света попадающего на диафрагму поз. 5. Передняя и задняя поверхности диафрагмы отполированы по ГОСТ 11141-84 и по своим оптическим свойствам близки к зеркальному покрытию. Это также позволяет уменьшить время прогрева диафрагмы поз. 5 от модели АЧТ поз. 6.

Испытания экспериментального стенда проводились в лабораторных условиях. Цель испытаний – сформировать и обосновать метод физического моделирования точечного изображения объекта регистрации со значениями температурного контраста по отношению к фону, которые бы удовлетворяли условиям работы ОЭП КА. В процессе отработки стенда были использованы несколько типов латунных диафрагм. Диаметр отверстия диафрагмы  $D_{диаф} = 0,4$  мм, а температуры моделей АЧТ отвечающих за объект регистрации и фон, соответственно,  $T_{об.рег.} = 297$  К и  $T_{фон} = 260$  К.

Ниже приведены полученные в результате испытаний графики распределения интенсивности излучения в градациях серого по чувствительным элементам матричного фотоприёмного устройства ТПВК (рис. 2 и 3) полученные из кадров изображений.

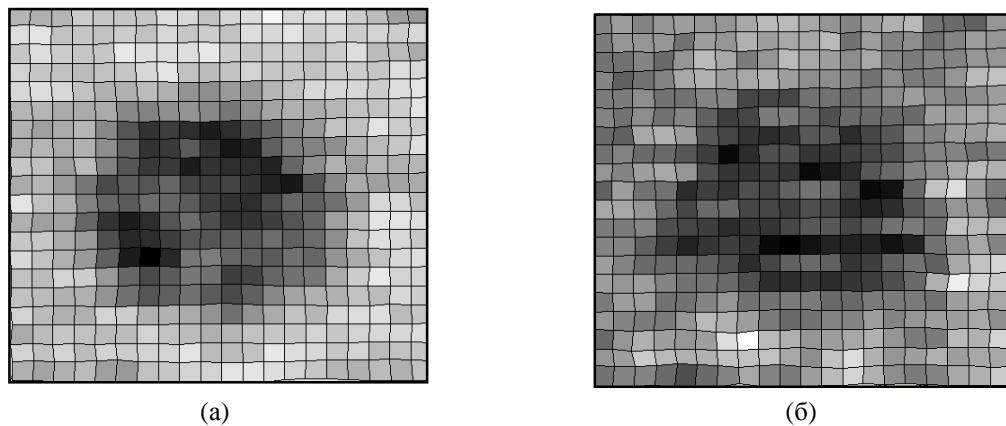


Рис. 2. Диафрагма из неотполированной латуни (а) и латуни покрытой чёрной эмалью (б)

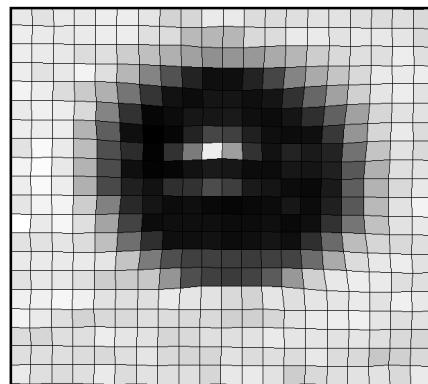


Рис. 3. Диафрагма из латуни с полированными поверхностями

Проведённые испытания показали работоспособность метода физического моделирования температурного контраста с «зеркальной» диафрагмой (полированной, в соответствии с ГОСТ 11141-84, латунью). Для дальнейшего развития метода будут проведены исследования по определению наиболее эффективной конструкции диафрагм (форм, материалов и оптических покрытий).

## Библиографический список

1. Корнышев Н.П. Оценка порогового контраста слабо люминесцирующих объектов // Вопросы радиоэлектроники, серия Техника телевидения, 2010, вып. 2, с. 56-65.
  2. Ковин С.Д., Сагдуллаев Ю.С. Спектрゾональное телевидение и тепловидение в задачах селекции и идентификации объектов // Вопросы радиоэлектроники, серия Техника телевидения, 2013, вып. 2, с. 123-136.
  3. Люхин А.В., Умбиталиев А.А. Задачи космических оборонных видеоинформационных систем // Вопросы радиоэлектроники, серия Техника телевидения, 2013, вып. 2, с. 3-14.

## METHOD OF SIMULATION OF THERMAL CONTRAST

**I.V. Bogdanov, A.N. Velichko, A.V. Stepovoy, V.S. Shelengovskii**  
JSC "Nudelman Precision Engineering Design Bureau"

In work the method of physical modeling of temperature contrast in laboratory environment for an assessment of algorithms of selection of optical-electronic systems of space appointment is considered. The experimental stand is described and results of tests are presented. The structure and requirements to this type of the equipment of physical modeling are proved.

**Blackbody, optoelectronic, stand, thermal contrast, background and target situation**

**И. В. Богданов, В. В. Сгадлев, П. А. Сенченков, А. В. Степовой**  
АО «КБточмаш им. А. Э. Нудельмана»

## Математическая модель измерения уровня сигнала для юстировки камер

*В данной статье рассмотрена разработанная математическая модель измерения уровня сигнала для юстировки камер. Модель предназначена для отображения видеоданных с камер в режиме реального времени, воспроизведения ранее записанных видеоданных для их анализа, а также для отображения обработанных видеоданных в виде графической трёхмерной структуры сигнала. Полученные результаты могут быть использованы в процессе разработки оптико-электронной аппаратуры или для проведения оптико-электронных измерений.*

### **Измерения, камеры, математическая модель, программное обеспечение, сигнал**

Разработка оптико-электронной аппаратуры на современных предприятиях сопровождается проведением большого числа измерений и экспериментов, в том числе, с применением вычислительной техники [1, 2]. Текущий уровень её развития позволяет использовать высокоскоростные интерфейсы (например, Ethernet и PCI) и многопоточность [3] для решения многих проблем в информационных системах.

Разработанная математическая модель представлена в виде программного обеспечения (ПО) написанного на языке Pascal в среде Delphi. Для работы с видеоданными используется дополнительный компонент DelphiX. Данный компонент позволяет напрямую использовать аппаратное ускорение графической платы персонального компьютера с помощью возможностей DirectX, для отображения видеоданных в режиме реального времени без эффекта мерцания кадров.

Математическая модель измерения уровня сигнала применяется для настройки, калибровки или юстировки видеокамер, проведения прецизионных измерений с помощью этих камер, анализа ранее записанных видеоданных, а также может использоваться для отработки алгоритмов поиска. В процессе работы ПО принимаемый видеопоток раскладывается по чувствительным элементам матричного фотоприёмного устройства камеры. На основе этого формируется соответствующий трёхмерный график с полным отображением его характеристик (в градациях серого) в режиме реального времени:

- уровней сигнала, шума и фона;
- математическое ожидание;
- среднеквадратическая ошибка;
- определение количества занимаемых объектом чувствительных элементов, т.е. его размер;
- график распределения уровня сигнала по чувствительным элементам камеры.

При необходимости, можно производить выбор анализируемой области (строб) на кадре, осуществлять программное переключение между несколькими параллельно подключёнными камерами и загружать для работы, ранее записанные видеоданные. График распределения уровня сигнала в градациях серого и текущий кадр изображения сохраняются в формате JPEG. Скриншот ПО представлен на рисунке 1.

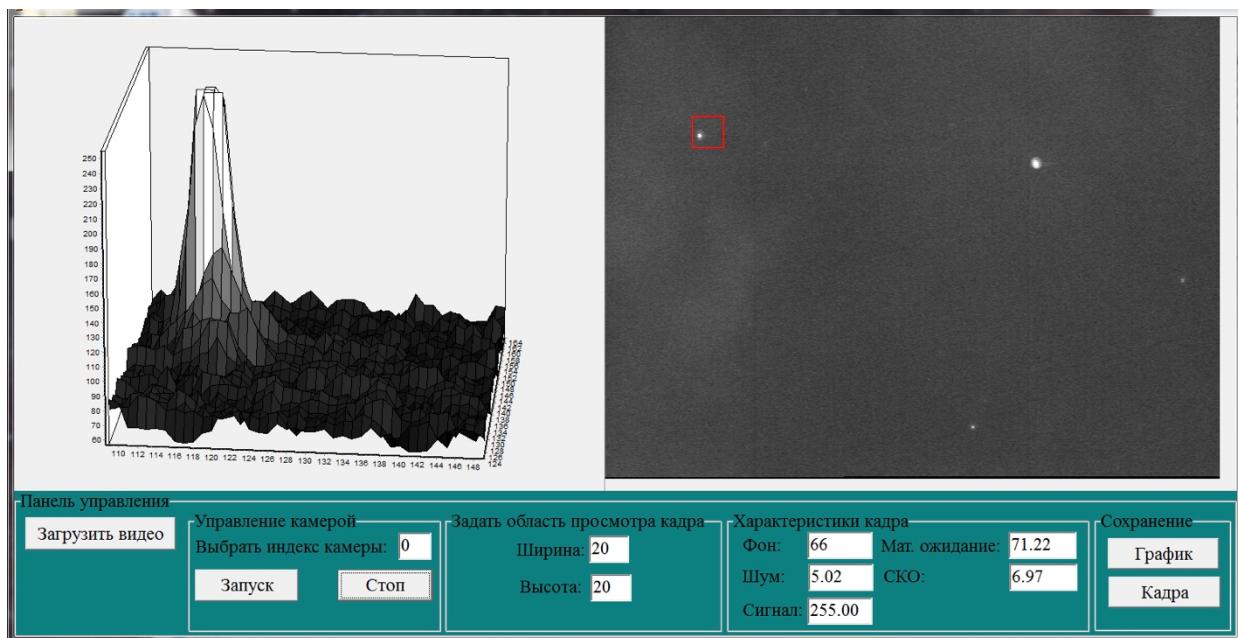


Рис. 1. Скриншот интерфейса ПО

### Библиографический список

1. Алмахрук М., Салем А., Еид М. и Фахми Ш.С. Алгоритмы и устройства обработки изображений по опорным точкам // Тезисы 13-ой международной конференции “Телевидение: передача и обработка изображений”, 2016, с. 89-96.
2. Жигулина И.В. Энергетические характеристики изображений и видеопоследовательностей // Тезисы 13-ой международной конференции “Телевидение: передача и обработка изображений”, 2016, с. 128-131.
3. Земсков А.И., Фащев В.В. Применение многопоточности при разработке производительных информационных систем // Сборник докладов Шестой научно-технической конференции молодых учёных и специалистов, посвящённой 80-летию со дня рождения А.А. Леманского, 2015, с. 605-610.

### MATHEMATICAL MODEL FOR MEASUREMENT OF THE LEVEL OF A SIGNAL FOR CALIBRATION OF CAMERAS

**I.V. Bogdanov, V.V. Sgadlev, P.A. Senchenkov and A.V. Stepovoy**  
JSC “Nudelman Precision Engineering Design Bureau”

*In this article the developed mathematical model of measurement of level of a signal for calibration of cameras is considered. The model is intended for “real-time” display of videodata from cameras, reproduction of earlier written down videodata for their analysis, and also for display of the processed videodata in the form of graphic 3D structure of a signal. The received results can be used in the course of development of the optical-electronic equipment or for carrying out optical-electronic measurements.*

**Measurement, cameras, mathematical model, software, signal**

**П. С. Баранов, Д. А. Белоус, Н. Н. Какушкин, А. А. Манцевов**  
СПбГЭТУ «ЛЭТИ»

## Наблюдение нагретых тел кремниевыми фотоприемниками

*Теоретически и экспериментально исследована возможность получения тепловых изображений в диапазоне температур от 200°C до 350°C в ближнем инфракрасном диапазоне спектра при помощи кремниевых фотоприемников. Установлено, что при помощи кремниевого фотоприемника можно наблюдать нагретые тела с температурой выше 260°C.*

**Тепловое излучение, абсолютно черное тело, тепловизор, твердотельный фотоприемник, КМОП-сENSOR, матричный ПЗС, чувствительность фотоприемника**

На сегодняшний день тепловизионные камеры используются во многих сферах деятельности человека. Основные элементы тепловизора – микроболометрическая матрица и объектив, обычно выполненный из германия, весьма сложны в производстве, а их стоимость составляет более 90% от стоимости тепловизора. В тоже время существуют задачи наблюдения нагретых тел в диапазоне температур от +200 °C до +350 °C и более, которые еще не видны в видимом диапазоне длин волн, но уже излучают оптическое излучение в ближнем инфракрасном (ИК) диапазоне 800...1100 нм.

Все тела, температура которых превышает температуру абсолютного нуля излучают электромагнитное тепловое излучение в соответствии с законом Планка [1]

$$R_{\lambda T} = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{hc/k\lambda T} - 1}, \frac{\text{Вт}}{\text{см}^3}$$

где  $h$  – постоянная Планка;  $c$  – скорость света в вакууме;  $\lambda$  - длина волны излучения;  $k$  – постоянная Больцмана,  $T$  – абсолютная температура.

При помощи пакета математического моделирования Wolfram Mathematica был произведен расчет излучательной способности нагретого тела. Излучение нагретого тела с высокой точностью моделируется излучением абсолютно черного тела (АЧТ), которое описывает уравнение Планка.

Для наблюдения излучения нагретого тела был выбран КМОП-сенсор фирмы Sony IMX265LLR, имеющий  $2064 \times 1544$  пикселов с размером  $3,45 \times 3,45$  мкм<sup>2</sup>. Данный КМОП-сенсор работает в режиме глобального затвора, имеет малые темновые токи, малые шумы считывания и высокую квантовую эффективность, в том числе и в ближней ИК области спектра.

Необходимо отметить, что на сегодняшний день в КМОП-технологиях появились матрицы с повышенной чувствительностью в ближнем ИК-диапазоне с длиной волны до 1100 нм. Это стало возможным благодаря утолщенной подложке [2].

Малые темновые токи КМОП-сенсора фирмы Sony позволяют повысить чувствительность камеры на его основы путем увеличения времени накопления, которое для данного фотоприемника может составлять до 32 с.

К тому же, при повышении температуры кристалла чувствительность кремниевых ФП в ближнем ИК диапазоне длин волн возрастает, причем большего повышения удается достигнуть при использовании КМОП-сенсоров [3].

Чтобы распознать нагретые тела при определенной температуре, требуется определенное минимальное количество фотоэлектронов, образованных в пикселе. Зная минимальное количество фотоэлектронов, участвующих в формировании видеосигнала, можно определить минимальную температуру излучения нагретого тела, которую способен распознать КМОП-сенсор с заданным отношением сигнал/шум.

Собственные шумы КМОП-сенсора составляют 2,5 электрона. Тогда, с учетом фотонного шума сигнала, для получения отношения сигнал/шум равного 2,5 должно быть накоплено 10 фотоэлектронов. Для расчетов использовалась спектральная характеристика КМОП-сенсора IMX265LLR, свойства оптического тракта телевизионной камеры, а именно коэффициент светопередачи объектива, а также излучательной способность нагретого тела.

В результате расчетов в пакете Wolfram Mathematica, были получены результаты, представленные на рис.1.

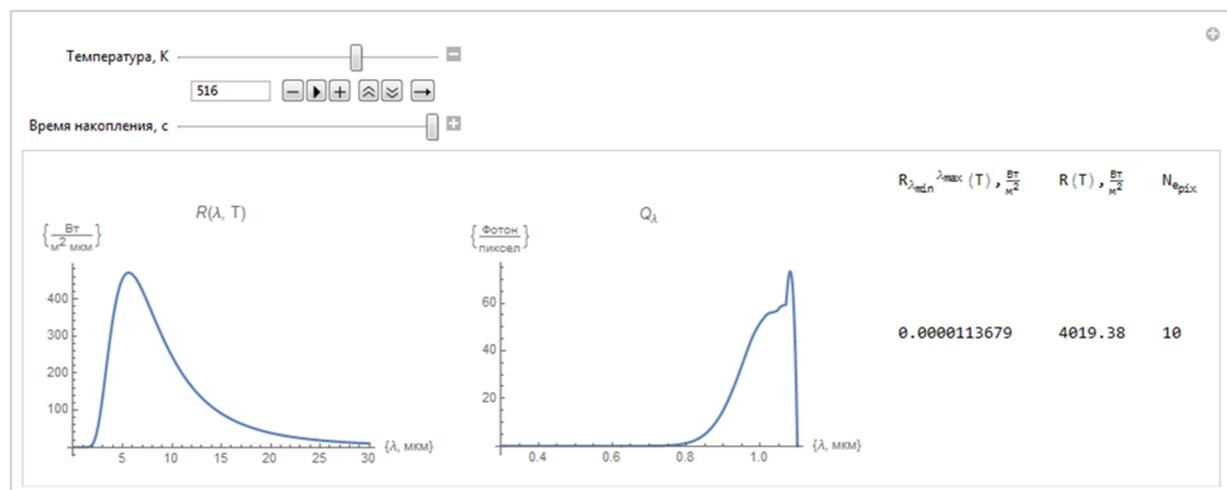


Рис.1. Результат выполнения вычислений в Wolfram Mathematica

На рисунке 1 изображены результаты, полученные при температуре 516 К или +243 °C. На рисунке 1 слева направо представлены: спектральная плотность потока излучения АЧТ, спектральная плотность потока фотоэлектронов, образовавшихся в одном пикселе КМОП-сенсора, значение потока излучения на КМОП-сенсоре, значение потока излучения АЧТ во всем спектральном диапазоне, количество фотонов, при котором распознается нагретое тело.

Зависимость отношения сигнал/шум от температуры

$$S_n(T) = \frac{N_{e_n}(T)}{\sqrt{N_{e_n}(T) + N_{c_q}^2}},$$

где  $N_{e_n}(T)$  - количество фотоэлектронов, участвующих в формировании сигнала при различных температурах;  $N_{c_q}$  – шум считывания КМОП-сенсора.

На рис. 2 представлена расчетная зависимость отношения сигнала/шум от температуры.

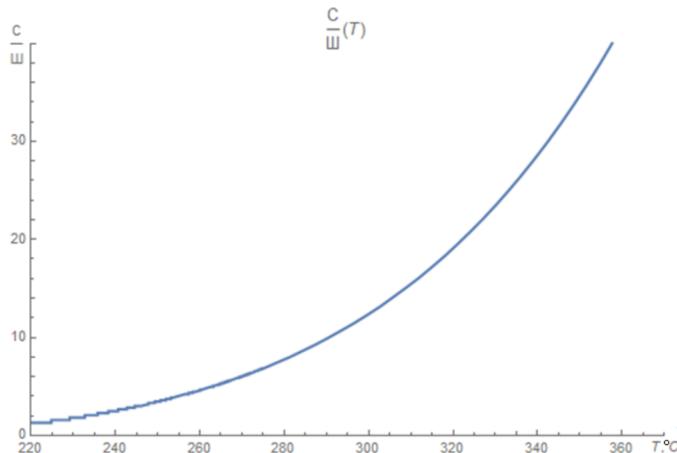


Рис. 2. Зависимость отношения сигнал/шум от температуры

Для проверки результатов теоретического расчета был разработан экспериментальный стенд, структурная схема которого представлена на рис. 3.

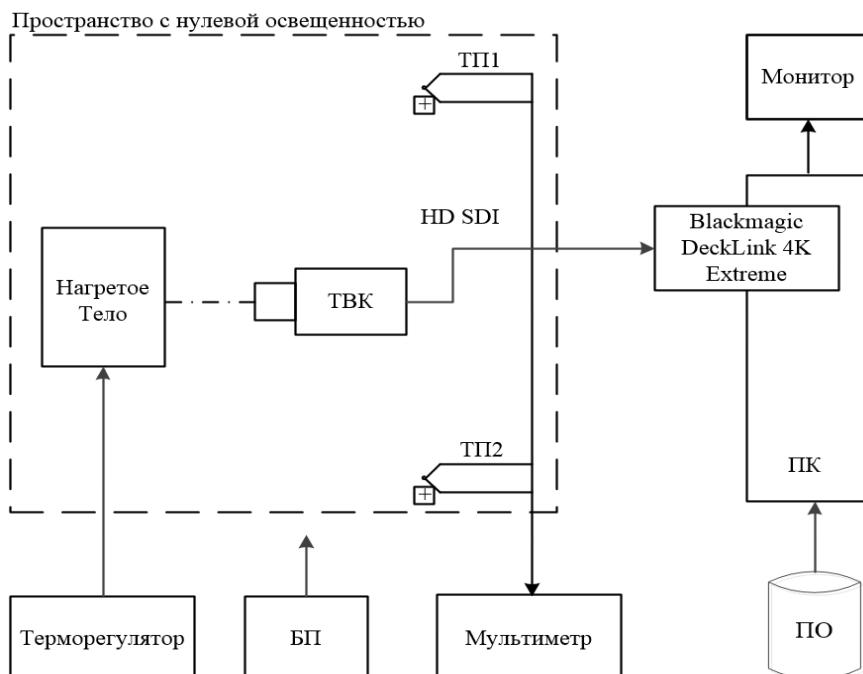


Рис. 3. Структурная схема экспериментального стенда

Стенд включал в свой состав: пространство с нулевой освещенностью, нагреваемый объект, черно-белую телевизионную камеру (ТВК), две термопары (ТП), персональный компьютер (ПК) с платой захвата видео по интерфейсу HD-SDI фирмы Blackmagic, монитор, мультиметр и блок питания (БП).

Пространство с нулевой освещенностью использовалось для того, чтобы исключить из эксперимента внешнее излучение.

В качестве нагреваемого тела использовался паяльник с медным жалом, а в качестве терморегулятора паяльная станция фирмы Lukey.

Температура нагреваемого тела измерялась термопарами ТП1 и ТП2, закрепленных на конце и основании жала паяльника и подключенных к мультиметрам фирмы S-Line.

Питание всех устройств осуществлялось от БП.

Использовалась камера фирмы Point Grey, выполненная на КМОП-сенсоре с глобальным затвором IMX265LLJ с интерфейсом HD-SDI.

Во время эксперимента в камере были отключены все автоматические регулировки, аналоговое усиление выставлено на максимум +48 дБ (256 раз), а время накопления установлено равным 1000 мс.

С помощью камеры при заданной температуре захватывались последовательности изображений 16 последовательных изображений в формате BMP. Температура изменялась в диапазоне от +220 °C до +260 °C шагом 5 °C.

Изображения, полученные при двух значениях температуры, приведены на рис. 4.

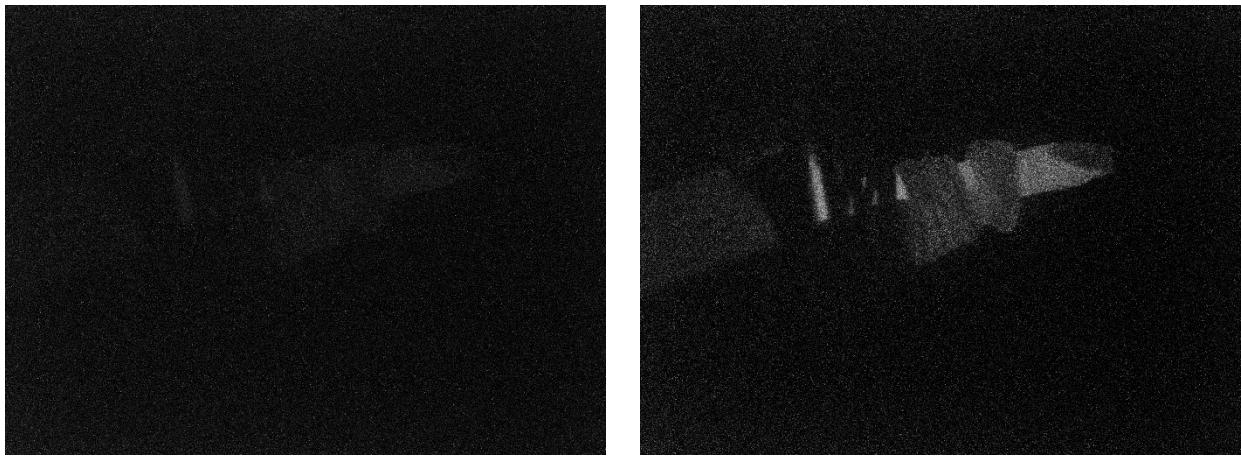


Рис. 4. Изображение нагретого тела при: а) 288,5 °C, б) 351 °C

В точках наиболее белого и черного по последовательности кадров путем статистической обработки вычислялись оценки математических ожиданий и среднеквадратического отклонения.

Значения температуры нагретого тела и соответствующего ей отношения сигнал/шум сведены в таблицу 1.

Значения отношения сигнала/шума в зависимости от температуры

Таблица 1

$T, {}^{\circ}\text{C}$	264	278,5	288,5	302,5	315	326,5	338,5	351	363,5
$S_n$	2,6	3,9	4,4	4,6	4,9	5,2	5,8	8,3	13,1

Значения температур выбирались, как среднеарифметическое значения термопар.

Зависимость соотношения сигнала/шум от температуры представлены на рис. 5.

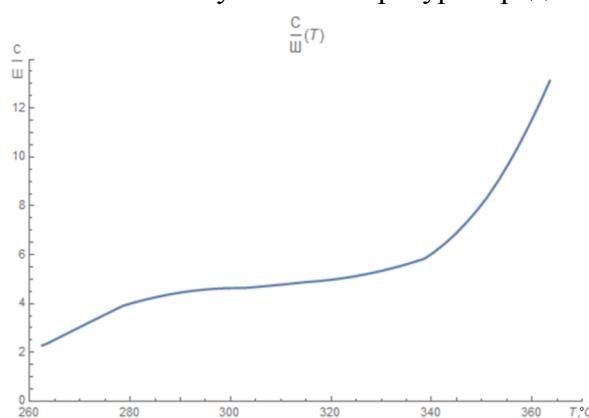


Рис. 5. Зависимость отношения сигнала/шум от температуры

Результаты теоретических и экспериментальных данных отличаются в среднем на 8%, что связано с тем, что в расчетах использовалась модель АЧТ, а в экспериментальной

модели его аппроксимировало медное жало паяльника, характеристики которого отличаются от характеристик идеального АЧТ.

Таким образом установлено, что кремниевые фотоприемники, в частности КМОП-сensоры могут быть использованы для наблюдения нагретых тел, температура которых выше +250 °C...+260 °C.

### Библиографический список

1. Ллойд Дж. Системы тепловидения: Перевод с английского под редакцией А.И. Горячева. - М.: Мир, 1978. - 414 с.
2. URL: <https://www.baslerweb.com/ru/vision-campus/tehnologii> - kamer/kamery-blizhnego-ik-diapazona/ (дата обращения: 20.04.2017)
3. Белоус Д. А. Чувствительность твердотельных фотоприемников в ближней инфракрасной области спектра при высокой температуре // Вопросы радиоэлектроники. Сер. Техника телевидения. 2017. № 2.

## OBSERVATION OF HEATED BODIES BY SILICON IMAGE SENSOR

**P. S. Baranov, D. A. Belous, N. N. Kakushkin, A. A. Mantsvetov**  
Saint-Petersburg state electrotechnical University "LETI"

*Theoretically and experimentally the possibility of obtaining thermal images in the range of temperatures from 200 °C to 350 °C in a short wave infrared range by means of silicon image sensor is probed. It is set that by means of the silicon image sensor it is possible to watch heated bodies with a temperature above 260 °C.*

**Thermal radiation, absolutely black body, thermal imager, solid state image sensor, CMOS image sensor, area CCD, image sensor sensitivity**

**П. С. Баранов, И. О. Позолотин**  
СПбГЭТУ «ЛЭТИ»

## Оптимизация спектральной характеристики телевизионной камеры в условиях переменной освещенности

*В работе разработан алгоритм оптимизации спектральной характеристики телевизионной камеры в условиях переменной освещенности. Построена математическая модель в среде моделирования Wolfram Mathematica. Проведено моделирование, которое показало работоспособность разработанного метода оптимизации.*

**Спектральные характеристики телевизионной камеры, оптимальный спектральный фильтр, переменная освещенность**

Задаче оптимизации характеристики спектральной чувствительности фотоприемника посвящен значительный объем публикаций. В работе [1] показано, что при известных спектральных зависимостях светового потока от фона и интересующего объекта, характеристика пропускания оптимального светофильтра будет иметь прямоугольный вид. В работах [2-4] приведены решения для целого ряда моделей спектральных характеристик фона и объекта, в том числе при учете собственных и изменяющимся фоне.

Однако для максимально обобщенной задачи оптимизация спектральной характеристики телевизионной (ТВ) камеры не решена. В данной работе задача оптимизации решается численными методами, а не аналитическими, как например в [5].

В качестве переменных параметров выступают:  $F_s(l)$  - спектральная характеристика объекта,  $F_b(l)$  - спектральная характеристика фона,  $q(l)$  - спектральная эффективность фотоприемника (ФП),  $WC$  – глубина потенциальной ямы ФП,  $D$  – дисперсия собственных шумов ФП,  $E_s$  - диапазон освещенностей от фона,  $E_b$  - диапазон освещенностей от объекта.

В качестве критерия оптимизации в ряде работе [2-4] выбрано максимум отношения сигнал/шум ( $y$ ), а в работе [6] максимум произведения отношения сигнала/шум на контраст ( $K$ ). В данной работе будем использовать последний критерий.

Отношение сигнал/шум определим как:

$$y = \frac{\int_0^{\infty} N_s(l) dl}{\sqrt{\int_0^{\infty} N_s(l) dl + \int_0^{\infty} N_b(l) dl + D}} \quad (1)$$

где  $N_s$  – количество сигнальных фотоэлектронов,  $N_b$  – количество фотоэлектронов образованных от фона,  $D$  – дисперсия собственных шумов фотоприемника, также приведенная к числу шумовых электронов.

Интегральный контраст определим в соответствии с [4]:

$$K_I = \frac{\int_0^{\infty} N_S(\lambda) d\lambda}{\int_0^{\infty} N_B(\lambda) d\lambda}.$$

Спектральные плотности количества сигнальных и фоновых фотоэлектронов  $N_S(\lambda)$  и  $N_B(\lambda)$  равны, соответственно

$$N_S(\lambda) = C_1 \propto q(\lambda) \tau(\lambda) F_S(\lambda),$$

$$N_B(\lambda) = C_2 \propto q(\lambda) \tau(\lambda) F_B(\lambda),$$

где  $C_1, C_2$  – константы,  $\tau(\lambda)$  – спектральная характеристика оптимального светофильтра (которая является искомой при решении данной задачи), рис.2. Отметим, что в соответствии с результатами [1, 2] функция  $\tau(\lambda)$  может принимать значения, равные либо 0, либо 1.

На рисунке 1 представлен предлагаемый алгоритм оптимизации спектральной характеристики телевизионной камеры в условиях переменной освещенности.

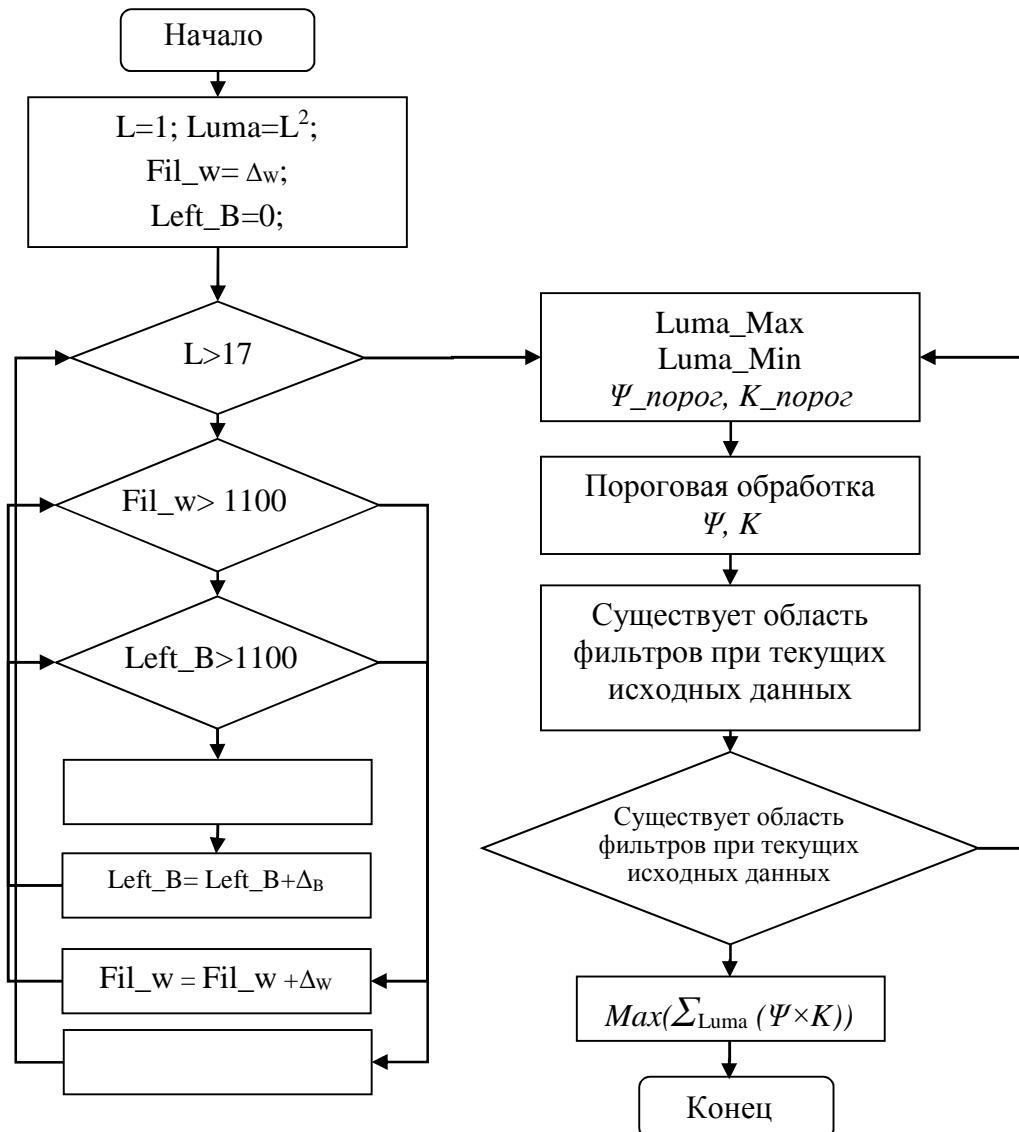


Рис. 1. Алгоритм оптимизации спектральной характеристики телевизионной камеры в условиях переменной освещенности

Продемонстрируем работу алгоритма на примере оптимизации спектральной ТВ камеры для обнаружения танка НАТО на фоне лесной поросли (спектральные характеристики взяты из открытых источников [7]) при условии, что наблюдение ведется при освещенностях от 1 лк до 131072 лк.

На первом этапе задаются исходные данные для работы алгоритма. На рисунке 2, а представлены спектральные характеристики ФП, фона (лесная поросль) и объекта (Танк). Далее необходимо задать начальные данные для поиска фильтра: Left\_B - левая граница фильтра на текущем шаге (начальное значение 400 нм),  $\Delta_B$  – шаг изменения левой границы фильтра (в данном примере 25 нм), Fil\_w - ширина фильтра на текущем шаге (начальное значение 25 нм)  $\Delta_w$  – шаг изменения ширины фильтра (в данном примере 25 нм), Luma – яркость фона ( $2^L$ , где L от 1 до 17). На рисунке 2, б приведены примеры фильтров на некоторых этапах работы алгоритма.

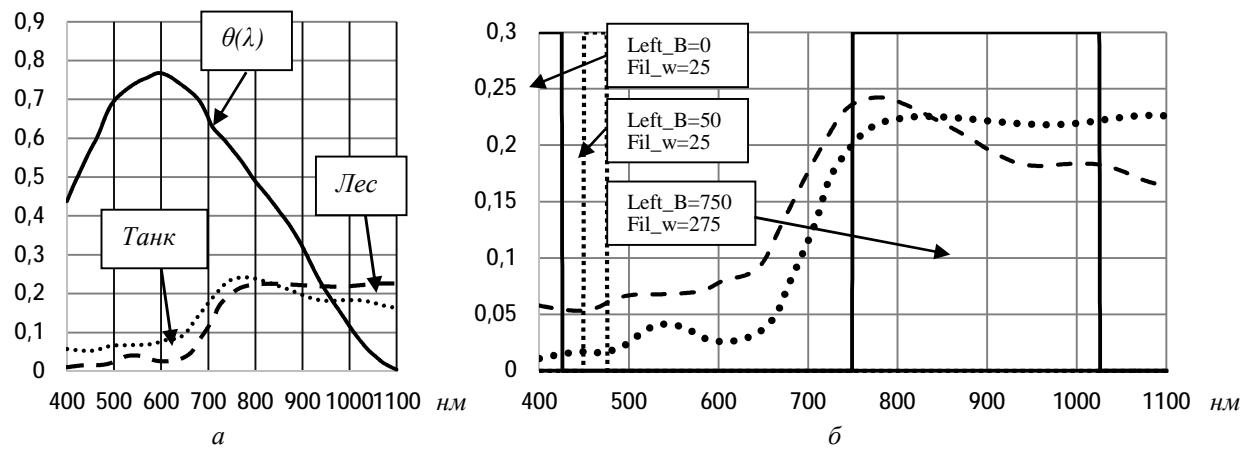


Рис. 2. Алгоритм оптимизации спектральной характеристики телевизионной камеры в условиях переменной освещенности

После прохождения работы алгоритма для одной освещенности формируются распределения отношения сигнала/шума  $\Psi$  и контраста  $K$  для всех комбинаций фильтра (рисунок 3). Допустимые варианты фильтра образуют треугольную матрицу. При заданных начальных параметрах существует 406 допустимых вариантов фильтров. После заполнения матрицы для сигнала/шума  $\Psi$  и контраста  $K$  при накально освещенности процедура повторяется для всех трубеумых освещенностей.

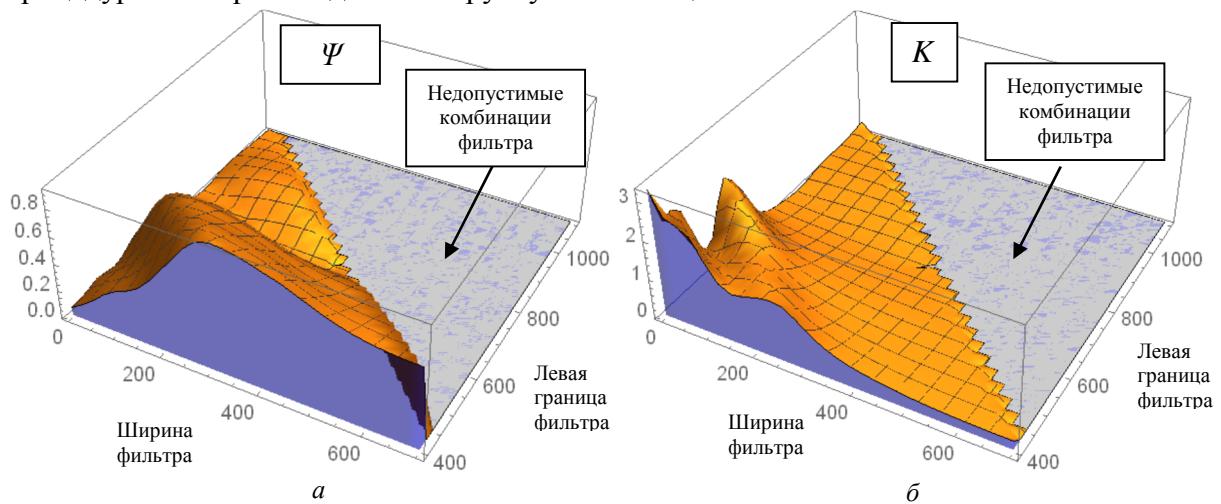


Рис. 3. Распределения отношения сигнала/шум  $\Psi$  и контраста  $K$  для  $Luma = 1$  лк

Для поиска фильтра необходимо задать минимально допустимое отношение сигнала/шум  $\Psi$  и минимально допустимый контраст  $K$  ( $\Psi_{\text{порог}}, K_{\text{порог}}$ ). Для каждой

освещенности производится пороговая обработка распределения отношения сигнал/шум  $\Psi$  в соответствии с выбранным порогом (рисунок 4). Можно отметить, что для заданного порога, при освещенности 1 лк не существует такого спектрального фильтра, который бы обеспечил минимально допустимое значение  $\Psi$ . При увеличении освещенности до 2 лк появляется 82 различных комбинаций спектрального фильтра, что отражается на рисунке 4 в виде белой области. При увеличении освещенности заданному условию  $\Psi_{\text{порог}}=1$  удовлетворяет все большее количество фильтров, что отражено на рисунке 4.

Аналогичная пороговая обработка производится и для распределения контраста  $K$  при  $K_{\text{порог}}=1$  (рисунок 5, а). На искомый фильтр должен удовлетворять совместной оценке  $(K \times \Psi)$ , которая представлена на рисунке 5, б. На рисунке 5, в представлена  $\sum_{\text{Luma}} (\Psi \times K)$ , а также указан фильтр удовлетворяющий условию максимума, который имеет параметры  $\text{Fil\_w}=275$  нм,  $\text{Left\_B}=0$  нм.

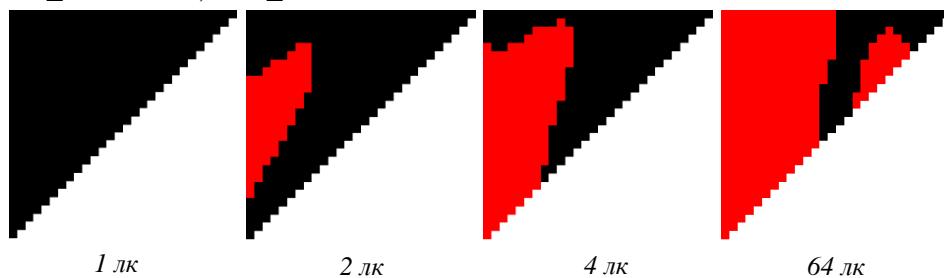


Рис. 4. Пример пороговой обработки распределения отношения сигнал/шум  $\Psi$  при пороге 1

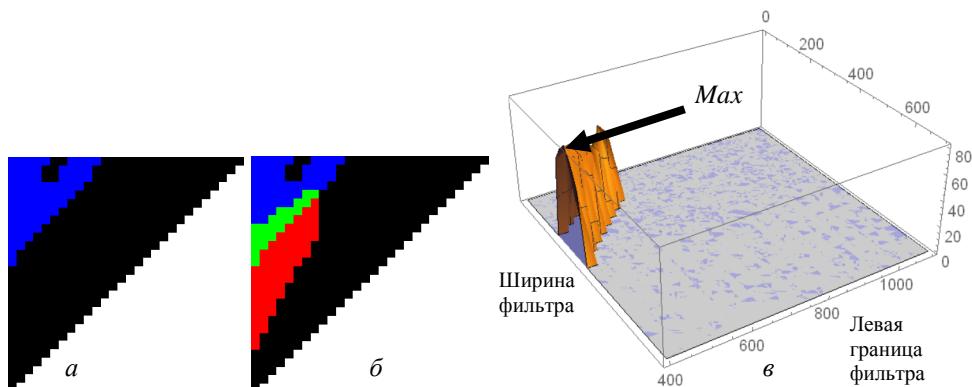


Рис. 5. Результаты пороговой обработки

В результате работы разработан алгоритм оптимизации спектральной характеристики телевизионной камеры в условиях переменной освещенности. Можно отметить, что благодаря использованию среды моделирования Wolfram Mathematica возможно практически в реальном времени производить подстройку всех интересующих параметров.

#### Библиографический список

1. Kleinhans W.A. Optimum spectral filtering for background-limited infrared system. «J. Opt. Soc. Amer.», 1965, vol.55, No 1, p.104
2. Сущев Г.А., Резников Б.И. Выбор спектральной характеристики оптимального светофильтра в прикладном телевидении. – Вопросы радиоэлектроники, сер. Техника телевидения. – 1973, №1.
3. Брацлавец П.Ф., Росселевич И.А., Хромов Л.И. Космическое телевидение. Изд. второе, дополненное. М., «Связь», 1973. 247 с.
4. Хромов Л.И., Бринкен Н.О., Литвинчук Л.А., Сущев Г.А. Специфика теории обнаружения в телевидении. – Вопросы радиоэлектроники, сер. Техника телевидения. – 1973, №4.

5. Иванова А. В., Манцевов А. А, Цыцуллин А. К. Оптимизация характеристики спектральной чувствительности фотоприемника при наблюдении малоконтрастных изображений // Вопросы радиоэлектроники. Сер. Техника телевидения. 2012. Вып. 1. С. 90–98.
5. Васильев Ю.Н., Гольдштейн Ю.А., Контрович В.Я., Короленко В.Н. Оптимизация спектральной характеристики фотоэлектрического преобразователя лучистой энергии. – Вопросы радиоэлектроники, сер. Техника телевидения. – 1969, №3.
6. Чиркунов А.А. Методы и аппаратно-программные средства оптимизации и адаптации телевизионных систем обнаружения малоконтрастных объектов: дис. ... канд. техн. наук: 05.12.04 / Санкт–Петербургский электротехнический университет. Санкт–Петербург, 2017. - 177 с.
7. Гуревич Ю. К. Оптико-электронные приборы ночного видения. – М.: ФИЗМАТ, 2014. -276 с. – ISBN 978-5-9221-1550-6.

## **OPTIMIZATION OF A SPECTRAL CHARACTERISTIC OF THE TV CAMERA IN THE CONDITIONS OF VARIABLE ILLUMINANCE**

**P.S. Baranov, I.O. Pozolotin**  
SPbETU

*In operation the algorithm of optimization of a spectral characteristic of the TV camera in the conditions of variable illuminance is developed. The mathematical model in the environment of Wolfram Mathematica simulation is constructed. Simulation which showed operability of the developed optimization method is carried out.*

**Spectral characteristics of the TV camera, optimum spectral filter, variable illuminance**

**P. С. Ермолаев**  
АО «НИИ телевидения»

## Измерение дисторсии объективов

*В работе исследуется метод измерения дисторсии объективов, основанный на вычислении координат малоразмерных объектов по методу центра тяжести. После этого производится центрирование полученных массивов координат и вычисление значения дисторсии исследуемого объектива с учетом значения линейного увеличения.*

### Дисторсия, метод центра тяжести

При оценке качества объективов важным критерием является наличие оптических aberrаций, которые могут привести к значительному снижению качества получаемого изображения.

Одним из основных типов aberrаций является дисторсия, при которой коэффициент линейного увеличения изменяется по полю зрения объектива. При этом нарушается геометрическое подобие между объектом и его изображением [1]. Дисторсия проявляется в том, что правильный квадрат изображается в виде подушки или бочки или еще более сложной фигуры.

В данной работе рассматривается метод измерения дисторсии объективов, основанный на измерении координат малоразмерных объектов методом центра тяжести.

С помощью исследуемого объектива на фотоприемнике телевизионной камеры получают изображение тестового объекта. В качестве тестового объекта используется Full HD AMOLED дисплей с разрешением  $1920 \times 1080$  и глубиной цвета – 16 млн. цветов. На дисплей выводится тестовое изображение, представляющее собой черный фон с сеткой  $10 \times 10$  из красных точек (рис. 1).

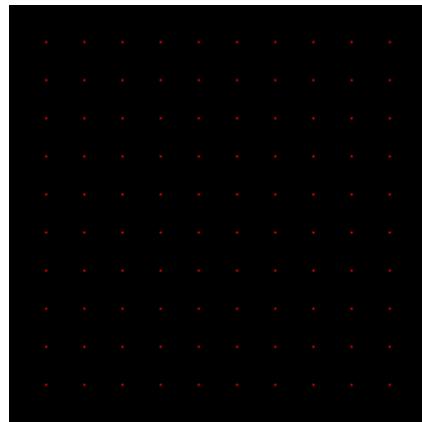


Рис. 1. Тестовое изображение

После этого определяются координаты каждой точки на тестовом изображении ( $x_i$ ,  $y_i$ ) и исследуемом изображении ( $x'_i$ ,  $y'_i$ ). Для этой задачи используется метод центра тяжести, поскольку он обладает устойчивостью к изменению форме сигнала [2]. Затем производится центрирование полученных массивов координат

$$x_{0i} = x_i - \overline{\mathbf{x}}_i / N,$$

$$y_{0i} = y_i - \overline{\mathbf{y}}_i / N,$$

$$x'_{0i} = x'_i - \frac{\sum_{i=1}^N x'_i}{N},$$
$$y'_{0i} = y'_i - \frac{\sum_{i=1}^N y'_i}{N},$$

где  $N$  – количество точек на изображении [3].

Значение дисторсии определяется с помощью следующих выражений:

$$\Delta x'_i = x'_{0i} - V_0 x_{0i},$$

$$\Delta y'_i = y'_{0i} - V_0 y_{0i},$$

где  $V_0$  – номинальное значение линейного увеличения исследуемого объектива.

Для экспериментального исследования описанного метода использовался объектив LM8XC фирмы Kowa с фокусным расстоянием 8,5 мм и относительным отверстием F2,8. Результаты эксперимента показали, что значение дисторсии используемого объектива составляет минус 0,17 %, что с допустимой погрешностью соответствует заявленному паспортному значению – минус 0,12 %.

#### Библиографический список

1. Иофис, Е.А. Фотокинотехника / Е.А. Иофис. – М.: Советская энциклопедия, 1981. – 447 с.
2. Твердотельная революция в телевидении: Телевизионные системы на основе приборов с зарядовой связью, систем на кристалле и видеосистем на кристалле / В.В. Березин, А.А. Умбаталиев, Ш.С. Фахми, А.К. Цыцулин, Н.Н. Шипилов; под ред. А.А. Умбаталиева и А.К. Цыцулина. – М.: Радио и связь, 2006. – 310 с.
3. Родионов, С.А. Обработка результатов измерения дисторсии проекционных объективов / С.А. Родионов, Н.Б. Вознесенский, Э.М. Шекольян // Известия высших учебных заведений. Приборостроение. – 1991. – № 7. – С. 61 – 68.

#### MEASUREMENT OF THE DISTORTION OF LENSES

*Ermolaev R.S.*

*Television Research Institute*

*This work investigates the method of measuring the distortion of lenses based on the calculation of coordinates of small objects using the Center of Gravity Method. After that, obtained arrays of coordinates are centered and value of the distortion of the investigated lens calculates taking into account the value of the linear magnification.*

**Distortion, Center of Gravity Method**

## Секция 2. Видеоаналитика

Руководитель секции – д.т.н., проф. Н. А. Обухова (СПбГЭТУ «ЛЭТИ»)

Зам. руководителя – к.т.н., доц. А. А. Мотыко (СПбГЭТУ «ЛЭТИ»)

Секретарь секции – асс. А. А. Чиркунова (СПбГЭТУ «ЛЭТИ»)

---

**Ш. С. Фахми<sup>1,2</sup>, Ю. И. Мукало<sup>3</sup>, И. А. Гаврилов<sup>3</sup>, Е. В. Костикова<sup>4</sup>**

<sup>1</sup>Институт проблем транспорта им. Н.С. Соломенко ФАНО

<sup>2</sup>СПбГЭТУ «ЛЭТИ»

<sup>3</sup>Научно-исследовательский институт прикладных проблем

<sup>4</sup>Государственный университет морского и речного флота имени адмирала С.О. Макарова

## Обнаружение и классификация транспортных средств видеопотока в реальном времени

*Предложен комбинированный метод обнаружения и распознавания машин на основе пространственно-рекурсивного метода. Рассмотрены все этапы обучения, обнаружения и распознавания машин на изображениях.*

**Обнаружение, распознавание машин, Метод Виолы-Джонса, выборка, каскад**

### Введение

Одна из наиболее распространенных задач компьютерного зрения на транспорте – задача об обнаружении транспортных средств (далее машин) и классификации. В рамках этой задачи требуется определить, принадлежит ли некоторый объект на изображении или объекты к определенному классу. Примером такой задачи является задача об обнаружении и распознавании машин на изображениях и классификации их на три группы: класс машин массой до трёх тон, класс машин от трёх до восьми тон и машины более восьми тон.

Интерес к применению регулярной пирамиды в задаче сегментации изображения объясняется тем, что подобная иерархическая структура представляет собой набор разбиений исходного изображения, в котором форма областей является контекстно-зависимой. Это означает, что структура данных такой пирамиды обеспечивает выделение деталей изображения, для чего каждое из разбиений, т.е. каждый уровень пирамиды представляется как графом, так и набором рецептивных полей [1].

Один из методов применения машинного обучения в задачах компьютерного зрения – метод Виолы–Джонса [2], являющийся одним из наиболее известных методов. Суть, которого заключается в следующем. Вместо одного сложного классификатора использовать каскад сильных классификаторов, построенных из слабых классификаторов, представляющих собой детекторы признаков (особенностей), характеризующих требуемый объект классификации.

Слабые классификаторы составляют основу получаемого сильного классификатора и во многом определяют его эффективность. Заметим, что прямоугольные особенности не всегда достаточно хорошо могут описать характерные признаки искомого класса либо для этого требуется линейная комбинация достаточно большого числа таких особенностей. Таким образом, несмотря на высокую в среднем скорость работы получаемых

классификаторов, в худшем случае требуется вычислять большое число особенностей. В работе Виолы и Джонса [2-4] для задачи детекции машин был получен классификатор, в котором использовалось 4385 детекторов прямоугольных особенностей.

В данной статье поставлена задача исследования комбинации алгоритмов пространственно-рекурсивного метода (ПРМ) разбиения и поиска опорных точек и метода Виолы – Джонса при решении задачи обучения, обнаружения и распознавания машин.

Синтезированный при помощи комбинированного метода классификатор должен распознавать признаки не хуже классификатора, синтезированного методом Виолы–Джонса.

Существует два способа повышения быстродействия методов компьютерного зрения:

- алгоритмическое улучшение с применением различных способов описания признаков объектов;

- аппаратная реализация на программируемой логической интегральной схеме (ПЛИС) с использованием современных САПР систем на кристалле.

Важно отметить, при семантическом анализе изображений, как правило, необходимо реализовать следующие функции:

- автоматическую локализацию объектов с заданными свойствами;

- автоматический поиск объектов на изображениях, аналогично воспринимаемых и выделяемых человеком;

- выделение фрагментов связанного множества пикселов на изображении, несущих семантическую нагрузку.

Локализация и идентификация объектов на изображении традиционно сводится к описанию набора характеристик искомого объекта. Наиболее сложной проблемой при этом является выявление и локализация границ фрагментов, несущих семантическое содержание. В предлагаемом методе рассматривается возможность построения фрагментов в виде древовидной связанной динамической структуры данных.

Использование детекторов признаков на основе ПРМ может дать преимущество в эффективности и позволит уменьшить число слабых классификаторов в каскаде, поскольку прямоугольные особенности ограничены формой, а алгоритмы ПРМ разбиения изображения на полигоны различной формы и размера, могут распознавать не только прямоугольные элементы. При этом рекурсия легко реализуется на ПЛИС [3]. Главное их преимущество – отсутствие сложных операций. Все, что необходимо – это разбиение и поиск опорных точек в пределах полигона. Поэтому использование ПРМ в детекторах особенностей гарантирует их легкую реализацию на ПЛИС с минимальными затратами ресурсов.

*Метод Виолы-Джонса.* Метод Виолы-Джонса имеет огромные преимущества в скорости распознавания машин на изображении, но недостатками этого алгоритма являются долгое время обучения классификатора и невозможность распознавания машин с углом наклона более 30 градусов. Также можно сказать, что данный метод является одним из лучших по соотношению показателей эффективности распознавания и скорость работы.

Задача, стоящая перед разработчиками алгоритмов обнаружения машин, заключается в следующем: имеется *изображение*, на котором есть искомые объекты (т.е.

машины). Алгоритм должен в результате своей работы определить найденные машины и пометить их, выделив рамкой - область нахождения объекта (рис. 1).

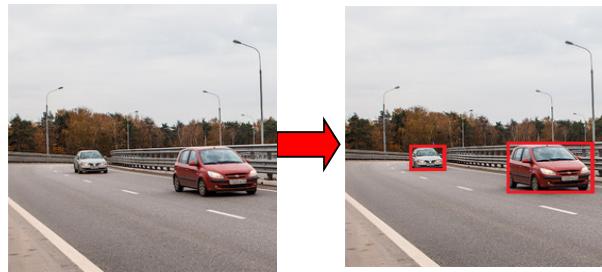


Рис. 1. Этап обнаружения

*Интегральное представление изображений.* Интегральное представление изображения – это матрица (т.е. сканирующее окно), совпадающая по размерам с исходным изображением, каждый элемент которой равен сумме пикселей, находящихся выше и левее данного элемента:  $L(x,y) = \sum I(i,j)$ , где  $I(i,j)$  - значение пикселя  $(i,j)$  исходного изображения,  $L(x,y)$  - значение элемента  $(x,y)$  интегральной матрицы [2,3].

Расчет матрицы занимает линейное время, пропорциональное числу пикселей в изображении, поэтому интегральное изображение просчитывается за один проход (рис. 2).

Метод Виолы-Джонса имеет линейную зависимость скорости сканирования и количества применяемых, этапе обнаружения, признаков Хаара (рис. 2). Использование интегрального представления изображений, позволяет с помощью признаков объектов обнаружить их с очень большой скоростью независимо от величины и положения объектов на изображениях.

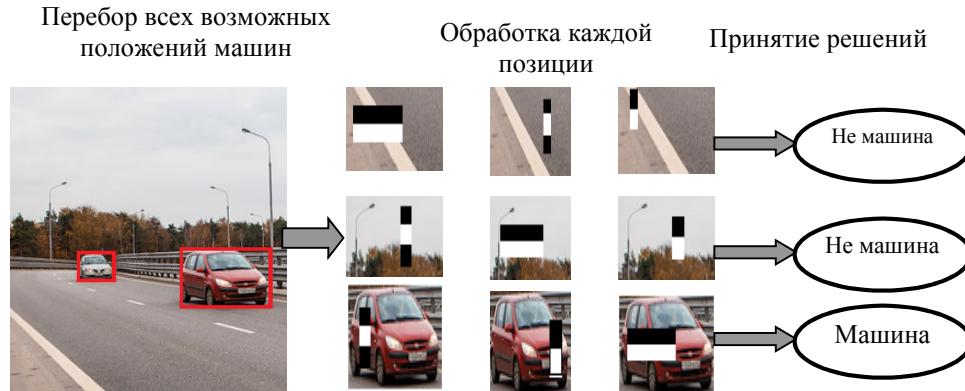


Рис. 2. Схема работы сканирующего окна

Интегральная матрица позволяет найти суммарную яркость произвольного прямоугольника изображения:

$$S(ABCD) = L(x_3, y_3) + L(x_1-1, y_1-1) - L(x_4-1, y_4) - L(x_2, y_2-1)$$

Приняв площадь прямоугольника равной значению яркости одного пикселя  $(x,y)$ , получим другую формулу, позволяющую рассчитать элементы интегральной матрицы:  $L(x,y) = I(x,y) - L(x-1,y-1) + L(x,y-1) + L(x-1,y)$ .

*Признаки Хаара.* Признаки цифрового изображения, используемые в распознавании образов. Своим названием они обязаны интуитивным сходством с вейвлетами Хаара (рис. 2).

Признаки Хаара дают значение перепада яркости между темной и светлой частью признака. Так, например, в следующих изображениях значение признака будет

положительным, так как области светлой части признака светлее, чем темной части, а, следовательно, А-область черного цвета больше В-область белого цвета. (значение белого цвета – 255, черного цвета – 0).

*Принцип сканирующего окна.* Это квадратная область, которая перемещается по изображению и масштабируется, тем самым, перебирая все возможные положения и размеры искомого объекта (рис. 2).

Схема работы сканирующего окна:

- 1) задается начальный размер окна, например, 24x24;
- 2) выбираются используемые признаки;
- 3) окно перемещается с заданным шагом по всему изображению;
- 4) в каждом положении вычисляются признаки, используя интегральное представление изображения;
- 5) значения всех признаков попадают к классификатору, который решает, присутствует ли машина в данной области;
- 6) предыдущие шаги повторяются для всех масштабов сканирующего окна.

В методе Виолы-Джонса используются *каскады признаков* (рис. 3), позволяющие быстро отбрасывать области с малой вероятностью нахождения объекта, что сильно влияет на скорость работы обнаружения. *Каскад признаков* состоит из нескольких уровней, каждый из которых включает в себя совокупность признаков и значение *порога уровня*. Причем каждый признак уровня также имеет свой собственный *порог признака*. Каскад определяется на этапе обучения. Первые уровни каскада построены так, чтобы убрать огромное количество изображений, не являющихся машинами ( $E_i$ ). Уровни становятся все сложнее и сложнее, поэтому ошибок становится гораздо меньше.

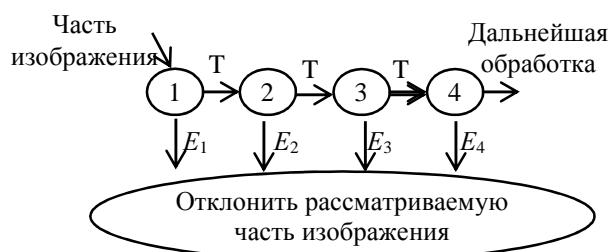


Рис. 3. Каскад формирования классификатора

*Алгоритм обнаружения на основе сканирующего окна.* Для каждого масштаба и положения сканирующего окна вычисляется значения  $j$ -признака и  $i$ -го уровня. Если значение больше порога признака, то добавляем к значению уровня число  $a$  иначе отнимаем от значения уровня число  $a_{ij}$  ( $a$  определяется на этапе обучения). А если значение уровня меньше порога  $i$ -го уровня, то необходимо переходить к следующему положению сканирующего окна. (т.е. в данной области нет искомого объекта). По завершению работы (если  $i = N$ ), то отмечается текущее положение окна как область машины квадратом и на этом заканчивается этап обнаружения машины (рис. 3).

*Обучение классификатора в методе Виолы-Джонса.* Говорят, что компьютерная программа обучается на основе опыта  $E$  по отношению к некоторому классу задач  $T$  и мере качества  $P$ , если качество решения задач из  $T$ , измеренное на основе  $P$ , улучшается с приобретением опыта  $E$ . Различают два типа обучения (рис. 3):

- 1) *Обучение по прецедентам*, или индуктивное обучение, основано на выявлении закономерностей в эмпирических данных.
- 2) *Дедуктивное обучение* предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний.

В работе используется алгоритм обучения по прецедентам, а именно алгоритм *обучения с учителем*.

Алгоритм обучения с учителем (*Supervised learning*) решает следующую задачу: Имеется множество объектов (ситуаций) и множество возможных ответов (откликов, реакций). Существует некоторая зависимость между ответами и объектами, но она неизвестна. Известна только конечная совокупность прецедентов — пар «объект, ответ», называемая *обучающей выборкой*. На основе этих данных требуется восстановить зависимость, то есть построить алгоритм, способный для любого объекта выдать достаточно точный ответ. Для измерения точности ответов определённым образом вводится функционал качества.

*Обучающая выборка*. Набор объектов, для каждого из которых априорно известно, к которому из нескольких классов он принадлежит.

В статье используются два класса объектов:

- объекты, где машины присутствуют (*положительная выборка*);
- объекты, где машины отсутствуют (*отрицательная выборка*) (рис. 2).

В обучающей выборке должны присутствовать изображения содержащие и несодержащие искомый объект (машино).

*Задача алгоритма обучения*. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется выборкой. Классовая принадлежность остальных объектов неизвестна. Требуется построить алгоритм, способный *классифицировать* произвольный объект из исходного множества.

Классифицировать объект означает, указать номер (или наименование) класса, к которому относится данный объект.

Известен алгоритм AdaBoost, который вызывает слабый классификатор в цикле. После каждого вызова обновляется распределение весов, которые отвечают важности каждого из объектов обучающего множества для классификации. На каждой итерации веса каждого неверно классифицированного объекта возрастают, таким образом, новый классификатор «фокусирует своё внимание» на этих объектах [4,5].

Рассматривается классификация объектов на два класса: *принадлежит* (1) и *не принадлежит* (-1). Применение алгоритма AdaBoost в методе Виолы-Джонса. Алгоритм бустинга в методе Виолы-Джонса имеет две ключевые особенности [2,3]:

- 1) Каскадная модель сильных классификаторов.
- 2) Определение минимальной ошибки с учетом изменения порога признака.

Каскадная модель сильных классификаторов необходима для быстрого принятия решения в областях изображения, где вероятность присутствия объекта мала. Первые уровни содержат признаки, которые не пропускают дальше большую часть изображений, не содержащих объекта обнаружения. При этом количество признаков, используемых на начальных уровнях мало, по сравнению с конечными уровнями. Определение

минимальной ошибки имеет особенность в том, что каждому признаку можно установить порог, который еще больше уменьшит ошибку классификации в алгоритме AdaBoost.

Минимальная ошибка для признака с порогом  $P$  вычисляется по следующей формуле:  $e = \min(S^+ + (T^- - S^-), S^- + (T^+ - S^+))$ , где  $T^+$  - сумма всех весов положительной выборки,  $T^-$  - сумма всех весов отрицательной выборки,  $S^+$  - сумма весов положительной выборки, значение признака которых ниже порога  $P$ ,  $S^-$  - сумма весов отрицательной выборки, значение признака которых выше порога  $P$ .

Разработано приложение для распознавания машин (рис. 4) позволяет загружать положительной ( $P$ ) и отрицательной ( $N$ ) выборки изображений, затем вычислить значений интегральной матрицы для каждого изображения выборки. Далее с помощью признаков Хаара, использующихся в обучении можно задать: максимально допустимый уровень ложных обнаружений на слой (f), минимально допустимый уровень верных обнаружений на слой (d) и целевой общий уровень ложных обнаружений (Ftarget) [2,3].

### **Комбинированный метод на основе ПРМ**

В данной работе используется преобразование изображения как совокупности двумерных пикселов, предварительно представленного в компьютере в виде одного из стандартных форматов хранения изображений, во внутреннее компьютерное представление одномерной динамической структурой данных.

Такое представление изображения назовем полигонально-рекурсивной моделью. Ее базовым элементом является полигон, определяемый как лист древовидной структуры [6] пирамиды изображения. Его атрибутом является идентификационный адрес (число), ассоциируемый с цветовыми и яркостными составляющими некоторой области изображения.

Организация полигонов изображения в виде древовидной структуры предполагает возможности описания (задания) вложенных иерархических отношений между полигонами, выделенными на различных уровнях (слоях) представления изображений.

Структура фрагментов строится итеративно начиная с «нулевого» уровня (исходное изображение) объединением пикселов с неразличимыми параметрами, например освещенности, цветности и др. Последующие уровни являются объединением фрагментов предыдущих, исходя из заданного критерия их эквивалентности.

Итеративное объединение фрагментов приводит к образованию многоуровневой иерархии, которая и представляет собой динамическое дерево. Получаемая адаптивно-динамическая структура позволяет эффективно выявлять фрагменты для семантической локализации и идентификации объектов.

На основе выше изложенного разработано Win-приложение, позволяющее в реальном времени обнаруживать и классифицировать машин по массе на три типа (рис. 4):

- а) до трёх тон, б) от трёх до восьми тон и г) от восьми и более тон.

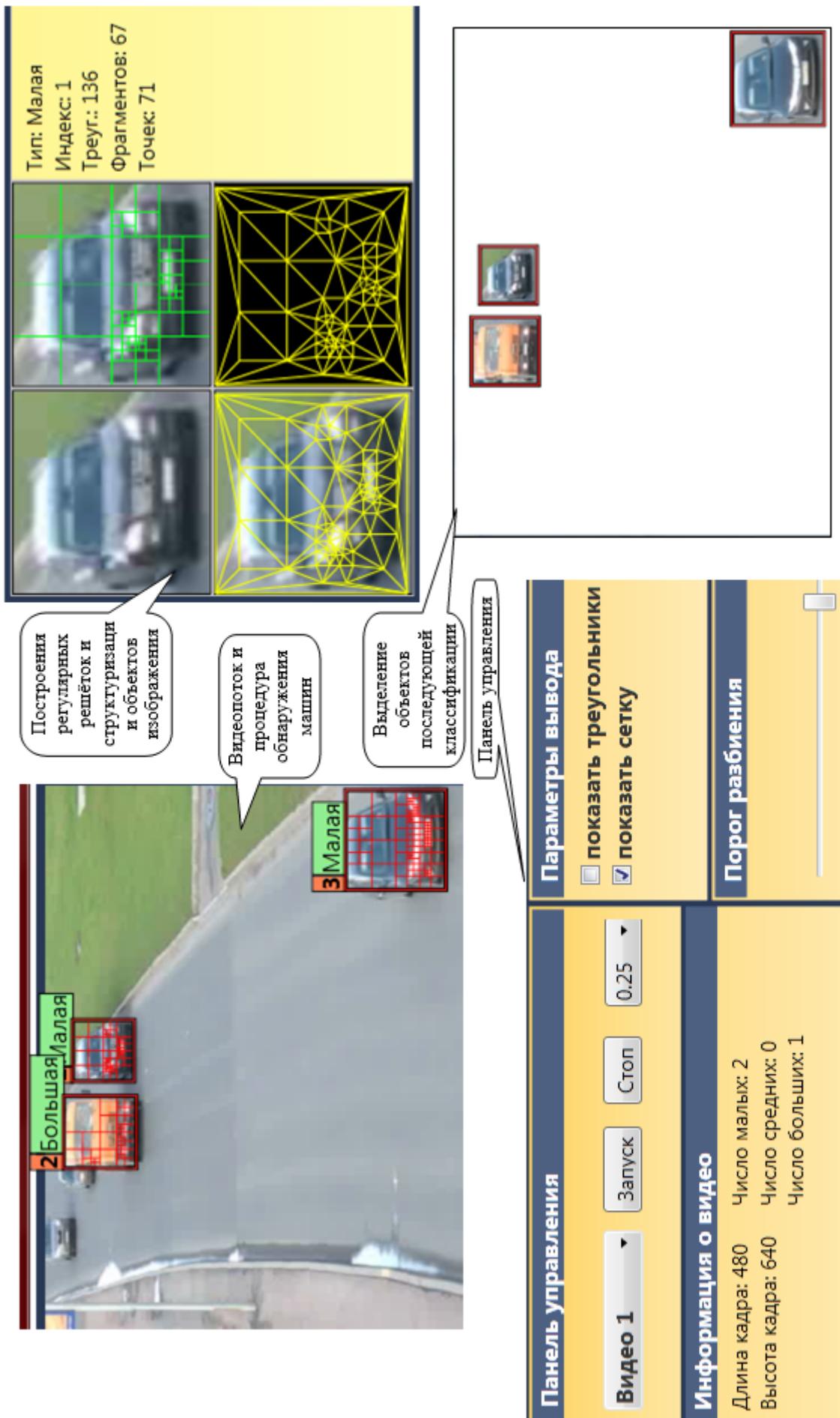


Рис. 4. Окно Win-приложения обнаружения и классификации машин

Разработанное приложение включает известный алгоритм Виолы-Джонса для решения задачи обнаружения, а для классификации и увеличения процента и точности распознавания приложение позволяет выполнить следующие процедуры (рис. 4):

- 1) на первом этапе выделения машин из общей сцены выполнить пространственно-рекурсивное разбиение с формированием неоднородной сетки изображения (рис. 4);
- 2) на втором этапе нормализации параллельно находятся опорные точки машины (максимально характеризующие выделенную область);
- 3) на третьем этапе формирования вектора признаков построить адаптивную триангуляционную сетку (маску) машины;
- 4) на последнем этапе применения классификатора сравнивать с масками по критерию минимума ошибки отклонения при идентификации машин [7,8].

Авторами предлагается дополнить метод Виолы-Джонса путём использования предварительного алгоритма рекурсивного разбиения обнаруженного машины на изображении на полигоны с последующим поиском опорных точек машины (рис. 6).

### **Результаты распознавания на основе ПРМ**

Проведены эксперименты на специально подготовленных наборах изображений. С целью проведения подробных исследований в качестве первой тестовой базы изображений выбраны фотографии различных машин, где изображения машин представлены на однородном фоне в разрешении  $64 \times 64$  пикселей. База содержит изображения более 200-х машин в 10 различных ракурсах. Изображения получены с различной освещенностью и положением.

Второй тестовой базой изображений выбрана коллекция из различных сайтов автомобилей, которая состоит из 150 изображений с разрешением  $64 \times 64$  в одном ракурсе каждое. При этом основное её отличие от базы первой группы является наличие фоновых объектов, т.е. представлены не только машины, но и обзор дороги и других объектов - неоднородный задний фон.

Результаты распознавания

Таблица 1.

Эксперимент	База машин	Комбинированный метод		Метод Виолы-Джонса	
		P, %	N, %	P, %	N, %
Эксперимент № 1	Группа из 200 машин	92, 2	7, 8	85, 2	12, 8
Эксперимент № 2	Группа из 150 машин	94, 4	6, 6	89, 3	10, 7

Численные значения правильного и ошибочного распознавания машин в экспериментах 1 и 2 представлены в табл. 1 ( $P$ ,  $N$  – оценки вероятности правильного и неверного распознавания машин, выраженные в процентах).

### **Заключение**

Предложен оригинальный метод автоматизированного распознавания машин, позволяющий существенно расширить решение задачи идентификации благодаря дополнительному разбиению области машины и представлению опорных точек машины в виде триангуляционной сетки (маски), позволяющей улучшить результат распознавания при незначительном увеличении сложности алгоритма. Использование предложенного метода позволяет структурировать информацию при обработке

изображений машин и тем самим имеет перспективу создания социализированных устройств распознавания образов в условиях шума и помех.

### Библиографический список

1. Агаронян О.С. Статистические аспекты применения диаграмм Вороного в задачах обработки изображений // Автоматика и телемеханика. 1999. № 4. С.58-64.
2. P. Viola and M.J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
3. P. Viola and M.J. Jones, «Robust real-time face detection», International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154.
4. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. CVPR, 2001.
- Freund Y., Schapire R. E. Experiments with a new boosting algorithm //International Conference on Machine Learning.\_ 1996.\_ Pp. 148–156.
5. Александров В.В., Кулешов С.В., Цветков О.В.Цифровая технология инфокоммуникации М.: Наука, 2008 с. 244.
6. Бобровский А. И., Еид М.М., Альмахрук М., Салем А., Фахми Ш. С., Березин В.В. Метод распознавания лиц на основе пространственно-рекурсивных триангуляционных сеток // Вопросы радиоэлектроники, сер. Техника телевидения, 2016, вып. 3, с. 86–91.
7. Фахми Ш. С., Левко Г. В., Гаврилов И.А., Мукало Ю.И. Пространственно-рекурсивный метод обнаружения объектов на звёздном фоне// Вопросы радиоэлектроники, сер. Техника телевидения, 2016, вып. 5, с. 67–77.

### DETECTION AND CLASSIFICATION OF VEHICLES OF THE VIDEO STREAM IN REAL TIME

**Fahmi<sup>1,3</sup> Sh.S., Gavrilov<sup>3</sup> I.A., Mukalo<sup>3</sup> Y.I. , Kostikova<sup>4</sup> E.V.**

<sup>1</sup>Institute of problems of transport after. N. With. Solomenko FANO

<sup>2</sup>Saint Petersburg electrotechnical University "LETI"

<sup>3</sup>research Institute of applied problems

<sup>4</sup>State University of sea and river fleet named after Admiral S. O. Makarov

*Proposed combined method of detection and recognition of vehicles based on spatial-recursive method. Considers the stages of learning, detection and recognition of vehicles on images.*

**Detection, identification, machines, Method viola-Jones, sampling, cascade**

**М. Альмахрук<sup>1</sup>, И. А. Гаврилов<sup>2</sup>, Ю. И. Мукало<sup>2</sup>, Ш. С. Фахми<sup>1,3,4</sup>**

<sup>1</sup>СПбГЭТУ «ЛЭТИ»

<sup>2</sup>Научно-исследовательский институт прикладных проблем

<sup>3</sup>Научно-исследовательский институт телевидения

<sup>4</sup>Институт проблем транспорта им. Н.С. Соломенко ФАНО

## Обнаружение и распознавание транспортных средств по триангуляционным сеткам

*Предлагаются метод и технологию построения триангуляционных сеток по опорным точкам объектов изображения для решения задач анализа и передачи видеонформации в реальном времени на базе платформы с встроенным ПЛИС Cyclone V и высокоскоростного USB 3 контроллера.*

**Видеосистема на кристалле, распознавание, триангуляционные сетки, ПЛИС**

### Введение

Системы компьютерного зрения все больше и больше проникают в различные области науки и техники: сельское хозяйство, метрологию, производство изделий для электронной промышленности и т.д. Основными задачами систем компьютерного зрения являются: анализ двумерного, трёхмерного изображения или видео потока, а также поиск и выделение объектов.

На протяжении последних десятилетий задаче поиска объектов на изображении уделяется большое внимание из-за того, что это позволяет производить обнаружение и классификацию объектов. На данный момент выполнено огромное количество исследований и опубликовано большое количество работ, однако, с нашей точки зрения, данная область остаётся ещё не просто не до конца изученной, а всего лишь слегка тронутой исследователями.

- Методы распознавания лиц Виолы Джонса и Майкла Джонса [1,2] до сих пор являются основными в распознавании не только лиц, но и изображений в целом. В основу их методов положен метод скользящего окна, который используется во всех бюджетных фотоаппаратах для выделения лиц.

- Проблема выделения объектов от фона в естественной или искусственной обстановке и последующего распознавания всегда находилась в ряду самых приоритетных задач для исследователей, работающих в области систем машинного зрения и искусственного интеллекта. Тем не менее, множество исследований, проводящихся в ведущих научных центрах всего мира в течение нескольких десятилетий, так и не привело к созданию реально работающих систем компьютерного зрения, способных обнаруживать и распознавать объектов в любых условиях [3,4].

Серьезной проблемой, стоящей перед проектировщиками систем компьютерного зрения, является большая изменчивость визуальных образов, связанная с изменениями освещенности, окраски, масштабов, ракурсов наблюдения. Кроме того, люди имеют привычку ходить по улицам и в помещении одетыми, что приводит к существенной изменчивости изображений одного и того же человека. Однако наиболее сложной задачей компьютерного зрения является проблема устранения неоднозначности, возникающей при проектировании трехмерных объектов реального мира на плоские

изображения. Цвет и яркость отдельных пикселов на изображении также зависит от большого количества трудно прогнозируемых факторов. В число этих факторов входят [3]:

- 1) число и расположение источников света;
- 2) цвет и интенсивность излучения;
- 3) тени или отражение от окружающих объектов.

Задача обнаружения объектов на изображении осложняется также огромным объемом данных, содержащихся в изображении. Изображение может содержать тысячи пикселов, каждый из которых может иметь важное значение. Полное использование информации, содержащейся в изображении, требует наличия современной технологий с гибко-настраиваемой конфигурацией и анализа каждого пикселя на принадлежность его объекту или фону с учетом возможной изменчивости объектов. Такой анализ может потребовать высоких затрат как в требуемой памяти, так и в количестве вентилей ПЛИС.

Решение этой проблемы лежит в правильном выборе описания объектов, для обнаружения и распознавания которых создается система. Описание объекта должно его учитывать наиболее характерные особенности в виде опорных точек, отличающие его от остальных элементов окружающей сцены. Чтобы избежать субъективности при выборе нужного описания, необходимо использовать методы автоматического выбора подходящих характеристик объекта, которые реализуются на этапе обучения в комбинированных алгоритмах. В то же время существует ряд параметров в описании объекта, которые в настоящее время разработчик должен учитывать при создании систем обнаружения и распознавания. К таким параметрам относятся [4-6]:

1. 2D и 3D-представления фона и объекта. При этом 3D-алгоритмы, в отличие от 2D требуют большого числа различных описаний, соответствующих представлению объекта в различных условиях наблюдения;
2. взаимосвязанность элементов, описывающих объект и фон;
3. признаки объекта (форма, размер, объем, цвет и.т.д.), описывающие специфику объекта и его характеристики.

В самом общем случае алгоритм решения задачи обнаружения и идентификации требуемого объекта состоит из следующих шагов:

1. обнаружение факта присутствия требуемого объекта на анализируемой сцене;
2. выделение опорных точек объекта;
3. определение и выделение характерных признаков объекта;
4. определение ракурса наблюдения объекта (анфас, профиль);
5. формирование сжатого описания опорных точек объекта;
6. сравнение с эталонами и идентификация.

В зависимости от конкретных условий структура и реализация отдельных шагов алгоритма могут различаться. Важным является использование максимально доступной информации для достижения удовлетворительных результатов в сложных ситуациях изменения обстановки с большим потоком входных видеоданных.

Алгоритмы должны уметь эффективно отсеять статические и медленно изменяющиеся элементы сцены, работать в различных условиях освещенности, опознавать объект под различными ракурсами, отслеживать передвижение множества

людей и автоматически выбирать момент, подходящий для выполнения идентификации данного объекта (например, когда можно получить фронтальное изображение лица с достаточным разрешением). Для обеспечения таких возможностей алгоритма необходимо иметь:

- а) многопроцессорные реконфигурируемые сложно-функциональные блоки, включающие многокамерный обзор и анализ сцены с возможностью выделения двумерных и трёхмерных опорных точек объекта;
- б) скоростной ввод видеопотока для фильтрации элементов сцены по параметрам движения;
- в) использование цвета для выделения элементов сцены;
- г) соответствующие аппаратные коммуникационные компоненты для реализации быстрых алгоритмов преобразования и хранения признаков объекта;
- д) наличие блоков памяти в виде набора SD-карт с последовательным и параллельным обменом.

Кроме того, необходимо наличия видеокамеры с параллельным вводом изображений, с высоким разрешением и хорошей оптикой для обеспечения работы системы с большой дальностью.

Выбор метода, используемого для обнаружения и идентификации объекта на изображениях зависит от конкретных условий его применения. Например, с задачей распознавания лица человека в строго ограниченном коллективе легко справляется многослойная нейронная сеть. В то же время проблема обнаружения конкретного человека в толпе (с неопределенным составом) требует применения комбинированных методов, отсекающих случаи ложных тревог. В этом случае потребуется многоуровневая система, содержащая множество анализаторов, работающих в разных признаковых пространствах, с подсистемой принятия решений.

Ниже представлен обзор по существующим методам обнаружения и идентификации объектов изображения. В обзор включались только те методы, которые, по мнению авторов, наиболее широко используются в современных алгоритмах обнаружения [7-9].

При всем многообразии различных алгоритмов и методов распознавания изображений, типичный метод распознавания состоит из трех основных компонент:

1. преобразование исходного изображения в стандартное представление;
2. выделение основных характеристик;
3. классификация.

Кроме этого, построение метода распознавания опирается на априорную информацию о предметной области (в данном случае - характеристики объекта) и корректируется экспериментальной информацией, появляющейся по ходу разработки алгоритмов.

В связи с этим особую значимость и научно-практический интерес представляет развитие мобильных средств экспресс-обнаружения и анализа чрезвычайных ситуаций с помощью интеллектуальных методов регистрации, обработки и анализа видеинформации, обеспечивающих получение достоверной информации об требуемых объектах с использованием современных систем на кристалле и САПР на их основе.

## **Метод распознавания на основе триангуляционных сеток опорных точек**

Метод обнаружения и распознавания объектов изображений на основе адаптивных триангуляционных сеток базируется на использовании пирамидально-рекурсивного поиска опорных точек.

*Алгоритм распознавания.* Предложенный алгоритм состоит из трех процессов: обучения, сохранения шаблонов в базе видеоданных.

Процесс формирования и хранения признаков объектов изображений происходит следующим образом (рис. 1):

Шаг 1. Преобразование изображения видео фрейма в полутоновое изображение.

Шаг 2. Применение к полутоновому изображению метода Виолы–Джонса для поиска области объекта.

Шаг 3. Уменьшение размера ТС до  $64 \times 64$  пикселей (нормализация).

Шаг 4. формирование к полученному на шаге 3 изображению триангуляционные сетки (регулярных и нерегулярных) признаков объектов (рис. 2).

Шаг 5. Сохранение извлеченных признаков в базе данных.

В процессе распознавания осуществляются шаги 1–4, затем на основе применения метода главных компонент происходит сокращение числа признаков и их сравнение с признаками, хранящимися в базе данных.

Численные эксперименты выполнены на основе базы изображений сотрудников кафедры САПР СПбГЭТУ «ЛЭТИ». База содержит 23 изображений сотрудников, по 20 изображений каждого лица. При формировании базы данных размер изображений и условия съемки были разными.

В процессе распознавания выполнены шаги 1–4 для обнаружения и распознавания лиц, затем на основе применения комбинированного метода главных компонент и триангуляционных сеток опорных точек полученные результаты показали увеличения процента правильных.

*Математическая постановка задачи распознавания образов.* Дано множество  $M$  объектов  $\omega$ . Объекты задаются значениями некоторых признаков  $x_i$ ,  $i=1, \dots, N$ , наборы которых одинаковы для всех объектов. Совокупность признаков объекта  $\omega$  определяет некоторым образом его описание  $I(\omega) = (x_1(\omega), x_2(\omega), \dots, x_N(\omega))$ .

Признаки могут выражаться в терминах да/нет, да/нет/неизвестно, числовыми значениями, значениями из набора возможных вариантов и т.д.

На всём множестве  $M$  существует разбиение на подмножества (классы объектов):

$$M = \bigcup_{i=0}^m W_i.$$

Разбиение на классы может быть задано полностью или определяться некоторой априорной информацией  $I_0$  о классах  $W_i$  - например, характеристическим описанием входящих в них объектов.

Задача распознавания состоит в том, чтобы для каждого данного объекта  $\omega$  по его описанию  $I(\omega)$  и априорной (обучающей) информации  $I_0$  вычислить значения предикатов  $P_i = (\omega \in W_i)$ ,  $i=1, \dots, m$ .

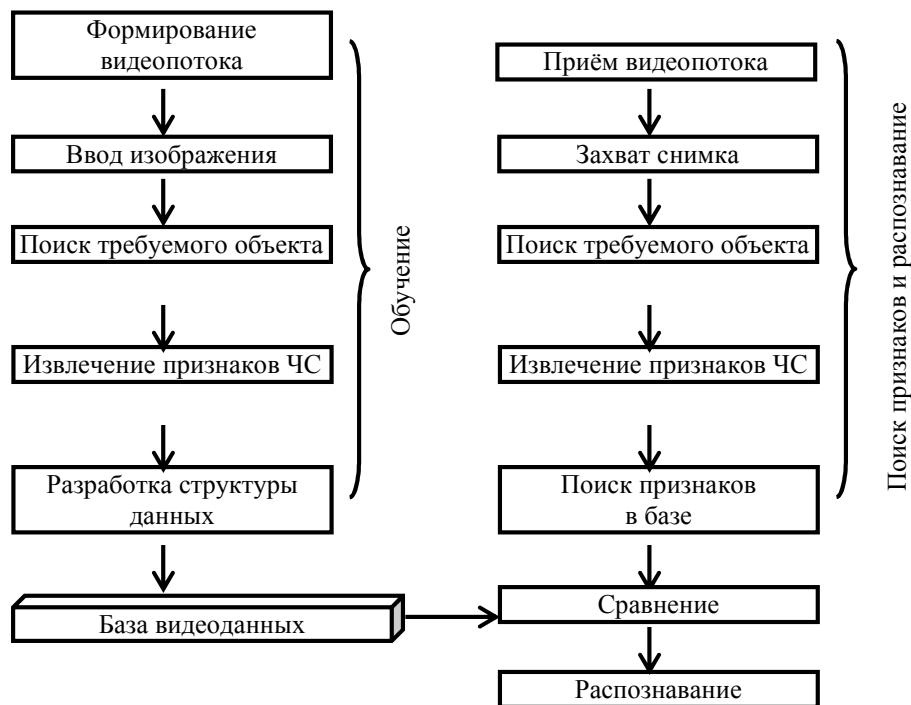


Рис. 1. Процессы обучения, поиска и распознавания МТС

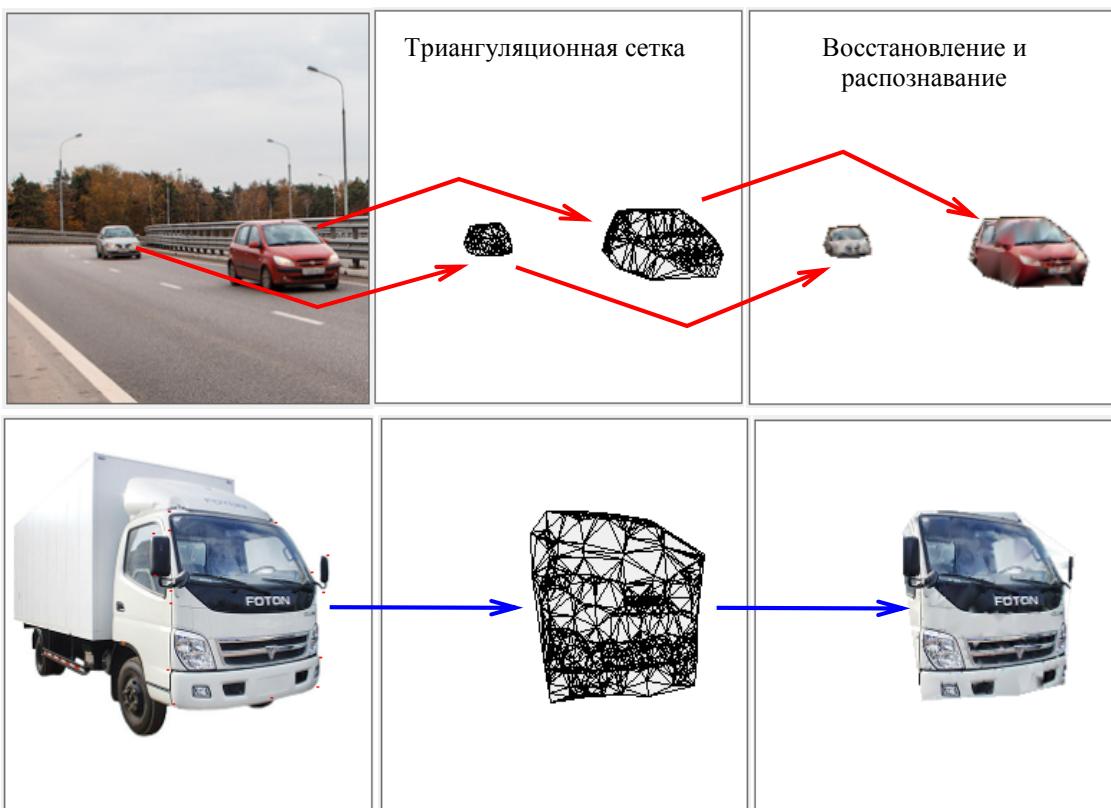


Рис. 2. Выделение ОТ объектов и восстановление по триангуляционной сетке

Для описания невозможности распознавания объектов предикаты  $P_i$  заменяются величинами  $a_i\{0(\omega \mid W_i), 1(\omega \mid W_i), D(\text{неизвестно})\}$ .

Таким образом, для рассматриваемого объекта  $\omega$  необходимо вычислить его информационный вектор  $a(\omega) = (a_1(\omega), \dots, a_m(\omega))$ .

Процедура, строящая информационный вектор  $a(\omega)$  в данном случае выражает алгоритм принятия решения об отнесении объекта  $\omega$  к тому или иному классу и называется «решающей функцией» [3].

### Технология распознавания с использованием современных ПЛИС

На кафедре САПР совместно с ОА «НИИТ» разработан новый макет на основе высокопроизводительной системы на кристалле, включающей быстродействующее процессорное ядро (*Cyclone V*) и ПЛИС, содержащее более 6-и миллионов вентилей на одном кристалле [10-12].

Платформа на базе *Cyclone V* предназначена осуществления проектов и приложений, работающих с большими объемами видеоданных и требующие высокую пропускную способность обмена. Платформа имеет расширение с интегрированными трансиверами и контроллерами памяти и подходят для применения в промышленных, проводных и беспроводных системах анализа и распознавания объектов изображений.

Платформа включает два модуля (рис. 3): *Cyclone V* и Контроллер *CYB USB3* позволяющие в реальном времени приём, анализа и передачи видеоинформации. Исследование и реализация СФ-блоков в качестве компонентов многопроцессорных систем на кристалле возможно благодаря наличию:

- **ПЛИС FPGA:** *Cyclone V*, 49 K программируемых логических элементов, 3080 Kbits встроенной памяти, 1 контроллер Hard Memory; последовательная конфигурационная память EPROM64, на плате установлен USB Blaster (разъем USB3), поддержка JTAG и AS режимов конфигурирования;
- КМОП камера (С разрешением 1024x1024 пикселей изображения), которая имеет разрешение 1Mps;
- модуль предварительной обработки видеинформации для решения задач распознавания:

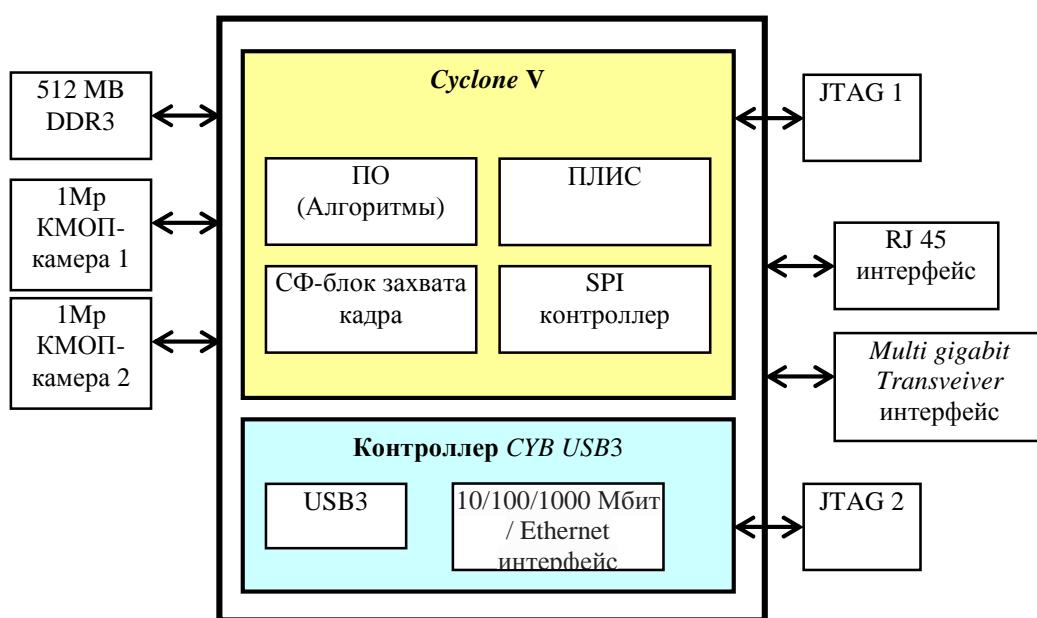


Рис. 3. Платформа проектирования видеосистем на кристалле

- встроенный комплект с ПЛИС *Cyclone V* общего назначения для разработки различных приложений, таких как кодирование и декодирование видео, обнаружение и, распознавание объектов, синтез и анализ комбинированных оптических (в т.ч. биомедицинских) сигналов нестационарных источников и т.д.;
- высокоскоростной интерфейс USB3 приемо-передачи видеинформации и т. д.

### **Заключение**

Современный подход к развитию интеллектуальных транспортных технологий распознавания образов должен обеспечивать выполнение полного маршрута системного проектирования. При этом необходимо параллельно создавать собственные библиотеки модулей с возможностью повторного использования СФ-блоков, специализированных для решения конкретных задач обучения, обнаружения и распознавания.

В заключение можно выделить следующие основные выводы:

- 1) решение проблем анализа и распознавания требуемых объектов изображений целесообразно выполнить на основе пирамидально-рекурсивного метода поиска и хранения опорных точек с использованием ПЛИС и современных САПР на их основе;
- 2) преимуществом библиотек СФ-блоков в составе видеосистем на кристалле является возможность автоматизированной разработки различных специализированных систем технического зрения на базе технологии «система на кристалле».

В заключение, следует отметить, что перечисленные методы и алгоритмы решения задач распознавания базируется на парадигме утверждении об априори качественном формируемом цифровом изображении в вычислителе, а в статье, на основе полученных результатов исследований можно выдвигать новый подход к решению этих задач, не в вычислителе а на этапе формирования видеинформации в самом источнике базируясь во-первых на основах и принципах синтеза видеосистем на кристалле, предложенных в [8], во-вторых с учётом новой аксиоматики теории информации, обозначенной в книге [9].

### **Библиографический список**

1. P. Viola and M.J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001
2. P. Viola and M.J. Jones, «Robust real-time face detection», International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154.
3. Р.Гонсалес, Р.Вудс, «Цифровая обработка изображений», ISBN 5-94836-028-8, изд-во: Техносфера, Москва, 2005. – 1072 с.
4. Gabriel Sikorjak. HVC-P: Human Vision Component. Sales presentation. Electronic& Mechanical Components Company. 2014.
5. Ramya .Rani .R, Tejaswini .M .L. Real Time, High Performance, RCE-NN Based Face Recognition System Using FPGA. Rani et al., International Journal of Advanced Research in Computer Science and Software Engineering 5(1), January - 2015, pp. 486-489.
6. Santu Sardar, Gaurav Tewari, K. A. Babu, “A Hardware/Software Co-design model for Face Recognition using Cognimem Neural Network chip ” 2011 International Conference on Image Information Processing (ICIIP 2011).
- Фахми Ш. С., Гаврилов И.А., Мукало Ю.И. Алгоритмы и устройства обнаружения объектов на звездном фоне с использованием полигонально-рекурсивных решёток// Вопросы радиоэлектроники, сер. Техника телевидения, 2017, вып. 1, с. 63–74.
7. Hripunov S.P., Blagodarjashhev I.V., Chirov D.S. [Military Robotics: modern trends and development vectors]. *Trendy i upravlenie – Trend and management*. 2015. vol. 4. pp. 410–422.

8. Lapshov V.S. et al. [Prospects for the development of autonomous robotic systems ground special military]. *Izvestija JuFU. Tehnicheskie nauki – Izvestiya SFedU. Engineering Sciences.* 2016. vol. 1(174). pp. 156–168.
9. Адамов Д. Ю. Основные вопросы проектирования видеосистем на кристалле. // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2014. – Вып. 2.. с. 3–17.
10. Твёрдотельные телекамеры: накопление качества информации / А. К. Цыцуллин, Д. Ю. Адамов, А. А. Манцевов и др. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2014. 272 с.
11. Цыцуллин А. К., Фахми Ш. С., Зубакин И. А. Начальный этап проектирования кодера источника непрерывного сигнала // Вопросы радиоэлектроники. Сер. Техника телевидения. 2010. Вып. 2. С. 17–34.

## DETECTION AND RECOGNITION OF VEHICLES BY TRIANGULATED MESHES

**Almahruk<sup>1</sup> M., Gavrilov<sup>2</sup> I.A., Mukalo<sup>2</sup> Y.I., V.V.Berezin<sup>1,2</sup>, Fahmi<sup>1,3,4</sup> Sh.S.**

<sup>1</sup>Saint-Petersburg state electrotechnical University "LETI"

<sup>2</sup>research Institute of applied problems

<sup>3</sup>research Institute of television

<sup>4</sup>Institute of problems of transport after. N. With. Solomenko FANO

*Proposed method and technology of construction the triangulated mesh on the reference image feature points to solve the problems of analysis and transmission of video information in real time on the platform with integrated Cyclone V FPGA and high speed USB 3 controller.*

### Video system on chip, recognition of a triangulated mesh, FPGA

**Б. С. Тимофеев<sup>1</sup>, А. А. Мотыко<sup>2</sup>**

<sup>1</sup>СПбГУАП

<sup>2</sup>СПбГЭТУ «ЛЭТИ»

## Телевизионная система анализа пассажиропотока

*Рассмотрены подходы к решению задачи автоматического учета потока пассажиров на транспорте. Приведены проблемы и недостатки существующих решений. Предложен новый алгоритм учета пассажиров на основании анализа видеоряда.*

**Видеоаналитика, цифровая обработка изображений, телевизионные системы на транспорте**

Вместе с повсеместным распространением относительно недорогих HD камер для систем видеонаблюдения большой интерес вызывают технологии видеоаналитики, позволяющие на основании анализа видеоряда получать новые невизуальные данные. Традиционными сферами применения видеоаналитики являются трекинг и распознавание объектов, охрана периметра, классификация объектов.

Технологии обработки и анализа видеоданных активно внедряются на транспорте. Например, комплекс технологий «умный автомобиль», включает автоматические помощники парковки, системы автономного управления и навигации, системы предупреждения о наличии опасности на пути движения и так далее. В то же время популярной задачей видеоаналитики в сфере коммерции является автоматический счетчик людей (например, посетителей магазина). Для транспорта данная задача формулируется как автоматический счетчик пассажиров, входящих и выходящих из салона транспортного средства. Польза от технологии очевидна – финансовый контроль, определение параметров пассажиропотока, загруженности линий.

Востребованность задачи сказывается на многообразии предлагаемых решений, которые можно разделить на две группы:

- контактные;
- бесконтактные.

К контактным следует отнести турникеты, устанавливаемые в проходах и так называемые «контактные ступеньки» (рис. 1) – специальные датчики, устанавливаемые под покрытие пола [1].

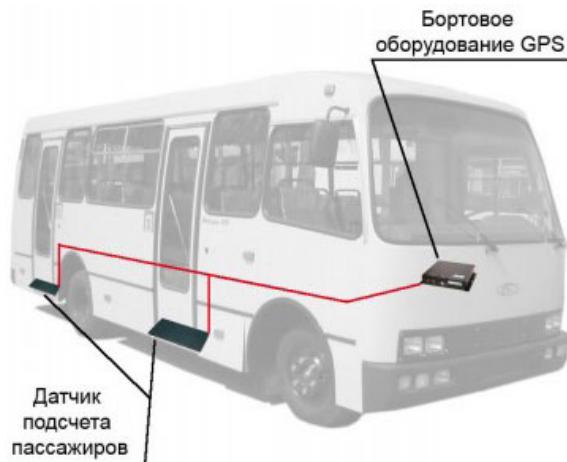


Рис. 1. Вариант контактного счетчика пассажиров

Недостатками данных подходов является, во-первых, высокий уровень погрешности (связанный, например, с неразрешимостью ситуаций многократного нажатия на датчик одним пассажиром, одновременное нажатие несколькими людьми), который не позволяет точно оценивать параметры пассажиропотока и во-вторых низкая эргономичность подобных систем. Например, внедрение турнкета значительно увеличивает время посадки и высадки пассажиров.

Применение контактных счетчиков в настоящее время сокращается и замещается новыми бесконтактными технологиями.

Бесконтактные методы имеют различные реализации, в частности использующие датчики теплового излучения [2], но преимущественно основаны на использовании камер, закрепленных над дверными проходами, и анализе видеоряда. Использование видеокамер помимо счета пассажиров позволяет вести постоянную регистрацию происходящего, что несомненно является полезным свойством.

С точки зрения типа камер, бесконтактные методы можно разделить на использующие информацию о глубине (иногда такие системы называют 3D, или RGBD камерами), и обычные видеокамеры. Предметом анализа в любом случае является последовательность кадров, но, разумеется, имея третью координату произвести подсчет пассажиров гораздо проще.

Для получения информации о глубине в системах подсчета пассажиров используют стереопару камер [3], времяпролетные (рис. 2) камеры [4], или решения, основанные на технологии Kinect от Microsoft.

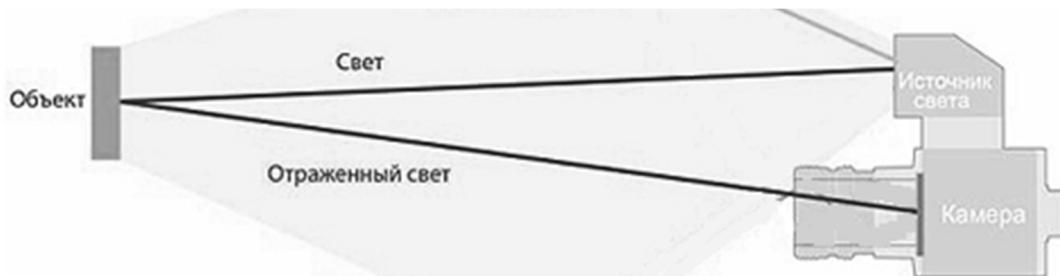


Рис. 2. Времяпролетная камера

К недостаткам подобных систем следует отнести их относительную сложность, которая оказывается при юстировке и монтаже, а также высокую цену (в среднем в три раза превышающую систему с обычной камерой), что при массовой установке значительно снижает конкурентоспособность систем такого рода.

Системы, основанные на обычных камерах, являются самыми привлекательными в плане эксплуатации и цены, однако для обеспечения точности, сравнимой с RGBD камерами, вынуждены компенсировать отсутствие информации о глубине более сложными алгоритмами анализа данных. На данный момент подобных решений не наблюдается.

Для детектирования людей, при установленной над дверным проемом камере, логичным кажется имплементация преобразования Хафа [5] для выделения окружностей. Однако, исследование показало практическую непригодность данного подхода. Во-первых, за счет головных уборов, причесок, геометрических искажений растра изображения голов людей могут иметь форму, сильно отличающуюся от окружности. Во-вторых, преобразование Хафа неустойчиво к шумам, что создает дополнительные трудности и требует сложной предобработки.

Применение соответствующих алгоритмов машинного обучения (нейронных сетей, лесов решающий деревьев) затруднено в силу большой вариативности целевых классов для детектирования, необходимостью сбора обширной базы данных (как правило, с учетом специфики используемой камеры) и требованием к последующей нетривиальной обработки результатов (как правило, требуется процедура сопровождения детектированных объектов). В таблице 1 приведено сравнение описанных подходов.

Сравнение реализаций счетчика пассажиров

Таблица 1

Описание	тип	преимущества	недостатки
турникет	контактный	точность	низкая эргономика
контактные датчики (ступени)	контактный	простота эксплуатации	низкая точность
тепловые датчики	бесконтактный	простота эксплуатации	низкая точность
3D-камеры	бесконтактный	высокая точность	цена, сложность системы
камеры наблюдения	бесконтактный	цена, простота эксплуатации	недостаточная точность

Из анализа таблицы видно, что значительными конкурентными преимуществами потенциально обладает решение, основанное на использовании камер наблюдения, но обеспечивающее высокую точность измерения. В данном материале предложен алгоритм, удовлетворяющий данным требованиям.

При входе или выходе пассажиры пересекают некоторую условную границу, отделяющую внутреннее пространство транспортного средства от окружающего мира. Для автоматического подсчета пассажиров необходимо установить факты пересечения этой границы и направления движения: на вход или выход. Целесообразно установить видеокамеру сверху, что дает возможность фиксировать перемещения пассажиров в горизонтальной плоскости. Изображение человека представляет собой нежесткий объект, для которого характерно видоизменение формы. Круглая форма головы во многих случаях изменена за счет головного убора и не может служить надежным признаком для сегментации объекта. Для подсчета числа пассажиров предложено использовать признак движения, который позволяет установить, как наличие объекта в контролируемой зоне, так и направление его перемещения.

Входными данными в системе являются видео файлы, записанные в ходе рабочего функционирования транспортного средства. Для экономии памяти и времени анализа видеозаписей целесообразно фиксировать движения пассажиров только в промежутки времени между открытием и закрытием дверей, которые суммарно составляют всего несколько процентов от общей длительности рабочего дня. Следует различать, также, движения дверей при их открытии-закрытии от перемещения пассажиров. Поэтому нужно автоматически определять, закрыты или открыты двери, и исключать из анализа отрезки времени, относящиеся к их движению.

Рассмотрим случай, когда транспортное средство имеет одностворчатые двери. Расположим на площади кадра две ловушки, первая из которых предназначена для подсчета пассажиров, а вторая – для анализа состояния двери (рис. 3). Разделим ловушки на блоки, и будем определять для них векторы движения методом сопоставления блоков. Рационально искать векторы оптического потока в тех областях

ловушки, где обнаружена энергия движения, представленная в виде суммы модулей разности яркостей пикселей в пределах блока

$$MAD = \sum_{i=1}^m \sum_{j=1}^n |L(x_j, y_i, t) - L(x_j, y_i, t-1)|,$$

в текущем  $t$  и предыдущем ( $t-1$ ) кадрах, где  $x_j, y_j$  – координаты пикселя,  $\{m, n\}$  – число пикселей в блоке [6].



Рис. 3. Расположение ловушек на кадре

Чтобы учесть направление и интенсивность движения в пределах ловушки будем использовать суммарные вектора движения по ловушкам:

$$\mathbf{v}_i = \sum_{j=1}^{p_i} \mathbf{v}_j,$$

где  $\mathbf{v}_j$  – вектор движения для  $j$ -го блока,  $i$  – номер ловушки,  $p_i$  – число активных блоков в ловушке.

Для определения отрезков времени, соответствующих перемещению дверей

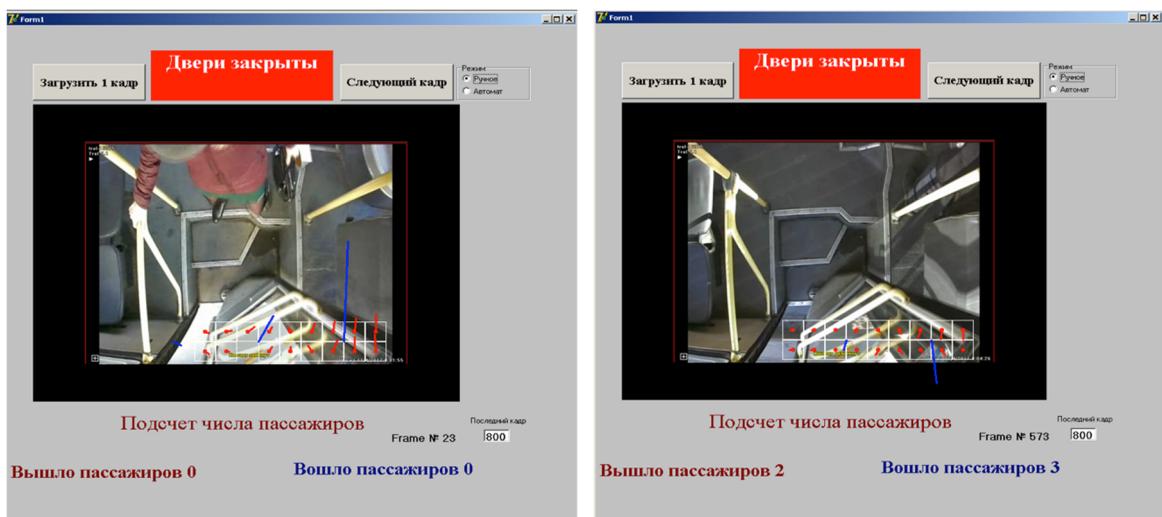


Рис. 4. Варианты движения ловушек: слева – открытие, справа – закрытие  
рассмотрим вектор движения во второй ловушке (см. рис. 3). Направление и модуль этого вектора дают однозначный ответ о состоянии дверей (рис. 4).

Знак горизонтальной составляющей вектора определяет направление движения дверей (рис. 5). На графике знаком плюс отмечены моменты времени, когда двери открыты, а знаком минус – закрыты.

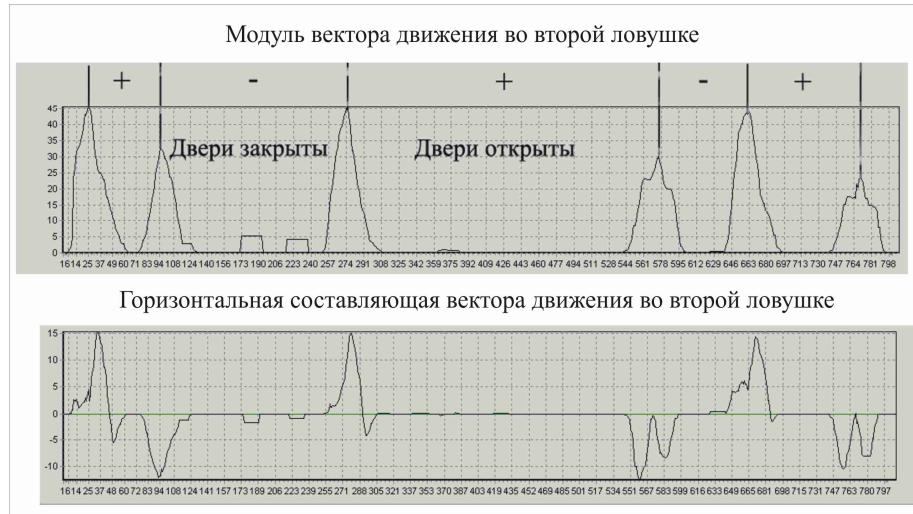


Рис. 5. Графики состояния двери

После оценки состояния дверей проводят анализ графика изменения вертикальной составляющей вектора движения в первой ловушке (рис. 6). Моменты выхода пассажиров соответствуют экстремумам максимумам на графике, а входа – минимумам. Участки кривой, относящиеся к периодам движения дверей, исключают из рассмотрения.

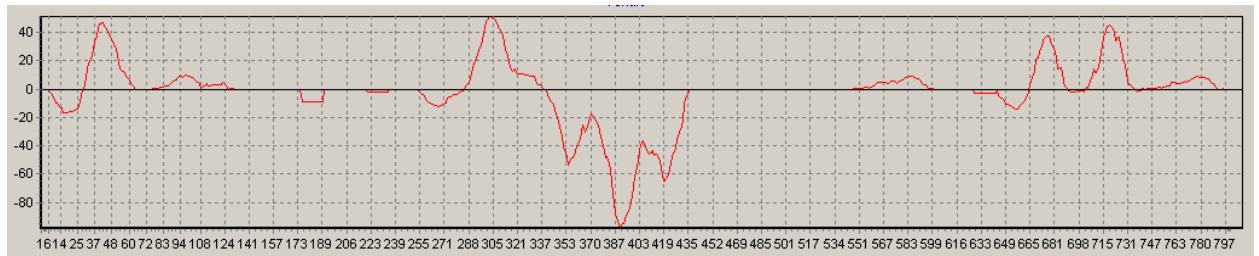


Рис. 6. График вертикальной составляющей  $v_1$ : четыре пассажира вышли и три вошли

Рисунок 7 иллюстрирует состояние системы при входе и выходе пассажира.

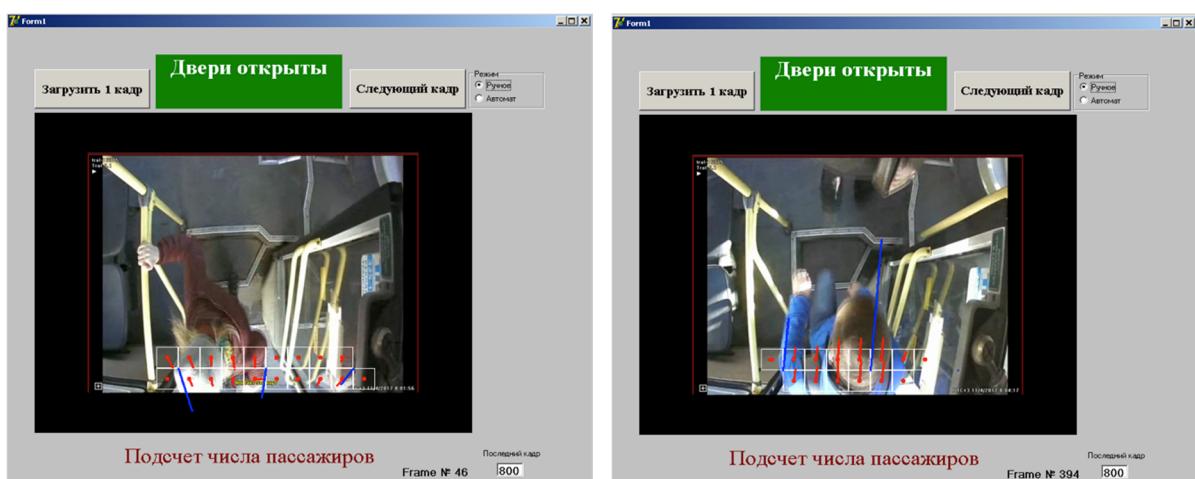


Рис. 7. Состояния системы при выходе (слева) и входе (справа) пассажиров

На всех исследованных фрагментах записи камер видеонаблюдения результаты подсчета числа вошедших и вышедших пассажиров оказались совершенно точными.

Существенно более сложной задачей является мониторинг пассажиропотока в транспортных средствах, снабженных двухстворчатыми дверями (рис. 8)



Рис. 8. Двухстворчатые двери

В этом случае возможны ситуации, когда, используя ширину дверного проема, одновременно входят два пассажира, или один входит, а другой выходит. Для решения этих задач применяют три ловушки. Подсчет числа пассажиров производят путем анализа экстремумов на графиках вертикальных составляющих векторов движения всех трех ловушек с учетом горизонтальных составляющих (рис. 9).

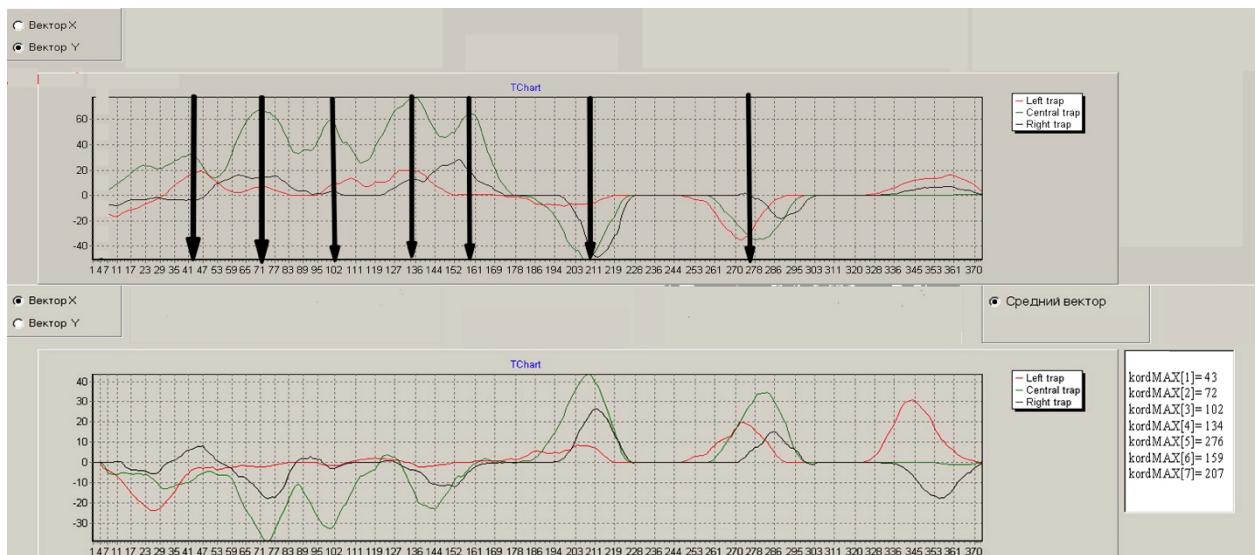


Рис. 9. Графики вертикальной (сверху) и горизонтальной (снизу) составляющих  $v_{1,2,3}$  в трех ловушках. Пять пассажиров вышли и два вошли в течение анализируемого отрезка видеопоследовательности.

Одновременно входящего и выходящего пассажиров распознают по знаку вертикальных составляющих векторов движения в крайних ловушках. Труднее различить одновременно входящих или выходящих пассажиров от одного, занимающего весь дверной проем, например, за счет поклажи. Такие случаи весьма редко встречаются на практике, но исключить их совсем не представляется возможным. При этом ошибка может быть в обе стороны, примерно с одинаковой вероятностью.

Мониторинг пассажиропотока целесообразно проводить по окончании дневного рабочего цикла транспортного средства путем обработки записей с камер

видеонаблюдения. Это минимизирует состав аппаратуры, установленной непосредственно на транспорте. Тестирование системы проводилось на видео файлах, полученных со стандартного видеорегистратора с разрешением 640 x 480 пикселей.

Сопоставление результатов подсчета пассажиров с метаданными (дата, время, координаты системы GPS) позволяет сделать выводы о загруженности транспортных средств на различных участках маршрута в определенные промежутки времени. Это дает возможность рационального планирования трафика движения; точный учет числа перевезенных пассажиров – повысить долю оплаченных поездок.

#### Библиографический список

1. <http://autovision.com.ua/blogcategory/sistema-teletrack/sistema-podscheta-passazhirov>
2. <http://globalv.ru/katalog/datchiki/datchik-ucheta-passazhiropotoka>
3. А. Пятницких, Технология стереоскопического видения для точного подсчёта количества людей, Современные технологии автоматизации, 2015г., выпуск 1, стр. 70-72.
4. [http://www.vektor-it.ru/katalog-gotovyh-reshenij/sistema\\_passajiropotoka\\_irma\\_matrix](http://www.vektor-it.ru/katalog-gotovyh-reshenij/sistema_passajiropotoka_irma_matrix)
5. Дегтярева А., Вежневец В. Преобразование Хафа (Hough transform). Компьютерная графика и мультимедиа, выпуск №1(1)/2003
6. Обухова Н. А., Тимофеев Б. С. Сегментация и сопровождение объектов в сложных условиях видеонаблюдения, Информационно-управляющие системы. 2008. № 6. С. 9-15.

### TELEVISION SYSTEM FOR AUTOMATIC ANALYSIS OF PASSENGER FLOW

**B. S. Timofeev<sup>1</sup>, A. A. Motyko<sup>2</sup>**

<sup>1</sup>Saint-Petersburg State University of aerospace instrumentation

<sup>2</sup>Saint Petersburg Electrotechnical University "LETI"

*Different approaches for the task of passenger flow analysis are observed and their features and limitations are described. The new algorithm is proposed for TV systems on the transport for automatic passengers counting.*

**Videoanalytics, digital image processing, television systems for transport**

**А. Л. Приоров, М. А. Кулагин**  
ФГБОУ ВО Ярославский государственный университет им. П.Г. Демидова

## Применение нейронной сети в задаче детектирования железнодорожных светофоров

*Представлен алгоритм детектирования железнодорожных светофоров с использованием нейронных сетей. Для решения поставленной задачи собрана обучающая и тестовая база изображений, спроектирована нейронная сеть. Тестирование проходило с использованием различных модификаций сети, а именно – использование разных целевых функций, необходимых для обучения по методу градиентного спуска. Так же было проведено сравнение работы нейронной сети с алгоритмом Виола-Джонса на той же базе изображений. В будущем алгоритм, построенный на основе нейронных сетей, может составить часть системы детектирования и распознавания объектов железнодорожной инфраструктуры на изображении, а также будет полезен инженерам в проектировании систем видеонаблюдения.*

**Нейронные сети, алгоритм Виола-Джонса, детектирование, целевая функция.**

### Введение

В последнее время большое внимание уделяется созданию интеллектуальных транспортных средств. Уже сейчас существуют автомобили, движущиеся без помощи человека, беспилотные самолеты и другие транспортные средства. Для того чтобы автомобиль мог двигаться в пространстве по всем правилам дорожного движения и принимать решения за человека, требуется видеть и анализировать окружающее пространство. Поэтому сейчас одним из самых популярных приложений машинного обучения является техническое зрение [1, 2].

В работе затронута тема детектирования железнодорожных светофоров с применением нейронной сети [3] и алгоритма Виола-Джонса. Задача является актуальной, так как алгоритм поиска области интереса на изображении в будущем будет лежать в основе системы, позволяющей локомотиву двигаться без участия машиниста. Создание цифровой железной дороги является на данный момент одним из приоритетных направлений развития данной области, поэтому решение задачи детектирования объектов железнодорожной инфраструктуры позволит ускорить процесс развития интеллектуального транспорта.

Применение алгоритмов машинного обучения и технического зрения для решения научно-технических задач требуют создание большой обучающей и тестовой базы данных. На данный момент уже существуют готовые базы изображений различных объектов, но задача детектирования железнодорожных светофоров требует создание собственной базы, необходимой для получения высоких результатов.

### Формирование базы изображений

Для формирования соответствующей базы использовался набор из 4305 изображений размером 64x64 пикселя (рис. 1).

Алгоритмы, спроектированные с применением нейронных сетей, требуют гораздо большего объема данных, поэтому первоначальное количество изображений в базе было увеличено с применением: аффинных преобразований (рис. 2а), зеркального отражения (рис. 2б), выравнивания гистограммы (рис. 2в).



Рис. 1. Примеры тестовых изображений из базы: а) изображения с объектом интереса (положительная выборка); б) изображения без объекта интереса (отрицательная выборка)

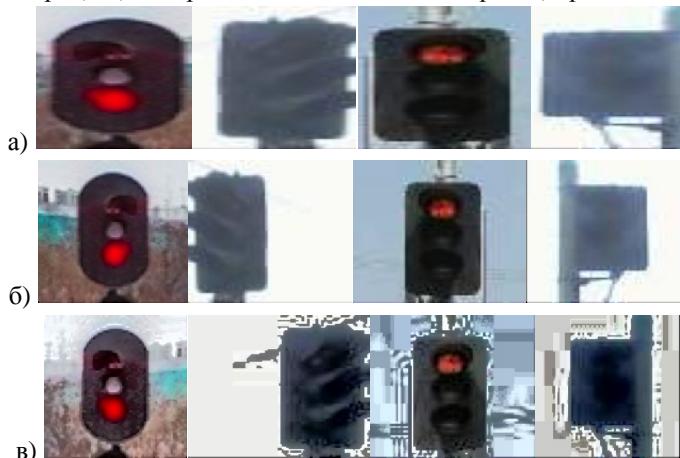


Рис. 2. Примеры полученных изображений: а) аффинные преобразования – растяжение изображения; б) зеркальное отражение изображения; в) выравнивание гистограммы

Первым вариантом расширения базы данных изображений являются аффинные преобразования, а именно – растяжение. Изображения были изменены с применением функции `wrapAffine()` из библиотеки OpenCV. Второй вариант увеличения количества изображений является зеркальное отражение. Третий вариант – выравнивание гистограммы, которое осуществлялось с использованием стандартной функции `equalizeHist` из библиотеки OpenCV.

В итоге, после расширения базы изображений, их конечное количество для обучения и тестирования составило 10000 изображений. Вся база изображений была разбита на три части: обучающая база – 7500 изображений (3007 положительная выборка и 4493 отрицательная выборка); база для валидации – 1500 изображения (750 положительная выборка и 750 отрицательная выборка); база для тестирования – 1000 изображений (500 положительная выборка и 500 отрицательная выборка). Все изображения, используемые для обучения, были размечены. Метка “1” означает присутствие на изображении железнодорожного светофора, метка “0” – его отсутствие.

В работе сравниваются алгоритмы обучения, используемые для детектирования области интереса на изображении, с применением нейронной сети и алгоритма Виола-Джонса.

### **Алгоритм Виола-Джонса**

Тестирование алгоритма Виола-Джонса проходило с использованием готовых функций из библиотеки OpenCV, предназначенной для компьютерного зрения. Более подробное описание алгоритма Виола-Джонса представлено в [4]. В этой работе, благодаря существующим программным решениям, то есть готовым библиотекам и наборам функций в них, обучение проходило в два этапа. На первом этапе применялась

функция `opencv_createsamples()`, которая приводила весь положительный набор изображений, то есть изображения с железнодорожным светофором, в общий формат и единый файл `sample.vec`. Второй этап заключался в создании итогового каскада классификаторов с использованием функции `opencv_traincascade()`. В итоге, результаты работы алгоритма показали в среднем 91% верного детектирования объекта на изображении. Работа алгоритма, протестированного в задаче детектирования железнодорожных светофоров, представлена на блок схеме (рис. 3б).



Рис. 3. Детектирование с применением: а) нейронной сети; б) алгоритма Виола-Джонса

## Нейронная сеть

Так же в данной работе, помимо алгоритма Виола-Джонса, проектируется и используется нейронная сеть с одним скрытым слоем (рис. 4). Ее обучение проводится с помощью метода градиентного спуска. На основании алгоритмов машинного обучения и компьютерного зрения, построенная нейронная сеть обучается, затем тестируется (рис. 3а). Реализовался алгоритм на языке программирования Python с использованием базовых библиотек OpenCV и NumPy. Примеры реализации других алгоритмов детектирования, но только автомобильных светофоров представлены в [5–6].

Самый левый слой нейронной сети является входным слоем, состоящим из 4096 нейронов. Крайний правый – является выходным слоем и состоит из 2 нейронов, средний слой называют скрытым и состоит из 50 нейронов (рис. 4).

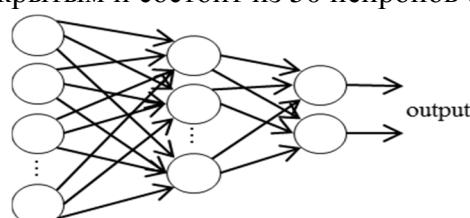


Рис. 4. Схематичная структура нейронной сети

Каждое изображение преобразуется в размер 4096x1, для того чтобы на каждый нейрон подать значение только одного пикселя для последующего обучения сети. На выходе нейронной сети активируется один из двух нейронов, демонстрирующий присутствие или отсутствие на изображении заданного объекта.

Обучения основывается на методе градиентного спуска [7], для подбора весов и смещений нейронной сети. Для количественной оценки того, насколько хорошо подобраны веса используется целевая функция (cost function). Существуют различные целевые функции. В данной работе приводится сравнение двух из них: квадратичной и функции кросс-энтропии:

$$C(w, b) = \frac{1}{2n} \sum_x (y(x) - a)^2$$

$$C(w, b) = -\frac{1}{n} \sum_x [y(x) \ln a + (1 - y(x)) \ln(1 - a)],$$

где  $w$  – совокупность всех весов нейронной сети;  $b$  – все смещения;  $n$  – совокупность всех учебных примеров;  $a$  – вектор выходов сети, когда  $x$  действует на входе сети. Задача состоит в подборе таких весов и смещений, чтобы функция затрат была минимальной. Метод градиентного спуска чаще всего используется, когда параметры (веса и смещения) функции не могут быть рассчитаны и должны быть найдены с помощью алгоритма оптимизации. Изначально подбираются произвольные значения параметров, после чего с каждым шагом коэффициенты меняются таким образом, чтобы значение функции уменьшалось. Изменение функции затрат относительно предыдущего положения вычисляется с помощью градиента функции.

В итоге нейронная сеть была обучена с использованием двух целевых функций, по очереди. В табл. 1 представлены результаты обучения на различных эпохах (эпоха – это целый цикл обучения и тестирования сети на всей обучающей и тестовой базе изображений). Видно, что функция кросс-энтропии увеличивает скорость обучения нейронной сети и результаты точности работы алгоритма на тестовой базе.

Обучение нейронной сети с использованием  
двух целевых функций

Таблица 1.

Целевая функция	Эпоха обучения	Точность работы алгоритма на тестовой базе
Квадратичная функция	10	54
	50	78
	80	93
Кросс-энтропия	10	77
	50	85
	80	94

### Выводы

С помощью алгоритмов Виола-Джонса и нейронной сети, реализованы и протестированы алгоритмы поиска объекта на изображении. Выяснено, что использование кросс-энтропии как целевой функции для обучения нейронной сети значительно сокращает время на ее обучение и увеличивает точность работы алгоритма. Алгоритм Виола-Джонса показал несколько худшие результаты, чем нейронная сеть.

Для дальнейшего улучшения результатов требуется увеличение базы данных изображений, а также применение глубоких нейронных сетей. Описанные алгоритмы в

будущем могут составить часть системы детектирования и распознавания объектов железнодорожной инфраструктуры на изображении.

#### Библиографический список

1. Шапиро Л., Стокман Дж. Компьютерное зрение. М. БИНОМ, Лаборатория знаний, 2006.
2. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход. М.: «Вильямс», 2004.
3. Осовский С. Нейронные сети для обработки информации. М. Финансы и статистика, 2002.
4. Viola P., Jones M. Rapid Object Detection using a Boosted Cascade of Simple Feature. Mitsubishi Electric Research Labs, Cambridge, 2001.
5. Zhou X., Yuan J., Liu H. A traffic light recognition algorithm based on compressive tracking. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, 2015.
6. Hwang T.H., Joo I.H., Cho S.I. Detection of Traffic Lights for Vision-Based Car Navigation System. Lecture Notes in Computer Science (PSIVT '06), Vol.4319, Dec. 2006. pp. 682–691.
7. Michael N. Neural Networks and Deep Learning . 2015. Chapter № 1.

#### APPLICATION OF A NEURAL NETWORK IN THE TASK OF DETECTING RAILWAY TRAFFIC LIGHTS

**Priorov A. L., Kulagin M. A.**

P.G. Demidov Yaroslavl State University

*This article presents an algorithm for detecting railway traffic light. This is algorithm based on using neural network. In this task was assembled a training and test image database, the neural network is designed. The testing was carried out using various cost functions required for gradient descent training. Also, a comparison was made of the work of a neural network with the algorithm Viola-Jones.*

*In the future, the algorithm will form part of the system for detecting and recognizing railway infrastructure objects on the image, and will also be useful to engineers in the design of CCTV.*

**Neural network, algorithm Viola-Jones, detection, cost function.**

**А. И. Бобровский<sup>4</sup>, Ш. С. Фахми<sup>1,2,3</sup>, Я. В. Алексеенко<sup>2</sup>,  
Е. П. Ермаков<sup>2</sup>**

<sup>1</sup> Институт проблем транспорта им. Н.С. Соломенко ФАНО

<sup>2</sup>Санкт-Петербургский университет Государственной противопожарной  
службы МЧС России (СП УГПС МЧС России)

<sup>3</sup>СПбГЭТУ «ЛЭТИ»

<sup>4</sup>Государственный НИИ прикладных проблем

## Видеосистемы на кристалле селекции объектов на примере обнаружения лесных пожаров

*Предложена обобщенная схема глобальной системы наблюдения за лесными пожарами из космоса на основе технологии «система на кристалле». Рассмотрены спектральный метод высокоскоростной передачи видеинформации и пространственный метод анализа и обнаружения лесных пожаров.*

### Видеосистема на кристалле, обнаружение пожаров, космический мониторинг

С 1997 года в МЧС России функционирует Система космического мониторинга чрезвычайных ситуаций России<sup>1</sup>, одной из главных задач её является мониторинг природных пожаров на территории Российской Федерации и приграничных районах.

Под лесными пожарами понимается неуправляемое стихийно распространяющееся по лесной территории горение. Отличие природных пожаров от других чрезвычайных ситуаций природного характера заключается в их регулярности и большой территории распространения. Последствия катастрофических лесных пожаров сказываются на протяжении длительного времени и часто приводят к необратимым последствиям [1,2].

Для своевременного доведения информации об обнаруженных термических точках СКМ ЧС России информация передается как в текстовом виде (в виде таблиц), так и размещается в географической информационной системе «Каскад<sup>1</sup>» (ГИС «Каскад») в виде векторных слоев, а также размещаются изображения (космические снимки) опасных очагов природных пожаров. Для более эффективного принятия управлеченческих решений, направленных на предотвращение человеческих жертв и материального ущерба, в ГИС «Каскад» добавлена различная необходимая информация (такая как торфяники, места массового отдыха людей, пожарные подразделения и т. п.).

Существующий визуальный контроль, осуществляемый воздушными и наземными патрулями, имеет малую эффективность из-за больших объёмов спутниковых видеоданных и сложности передачи их в реальном времени. Внедрение технологий «система на кристалле» и создание видеосистем на кристалле мониторинга (ВСнКМ), на основе эффективных методов обнаружения и передачи позволят специальным службам принять своевременные меры по предупреждению катастрофы и взять развитие ситуации под свой контроль [3].

В последние десятилетия произошел значительный рост производительности вычислительных средств и появилась новая элементная база класса систем на кристалле, которые позволили создание новых ВСнКМ и решение следующих проблем:

---

<sup>1</sup> <http://ukmmchs.ru/>

- а) внедрение современных схем с программируемой логикой в части создания высокоскоростных видеосистем приёма и передачи изображений;
- б) использование пространственно-рекурсивных методов анализа космических снимков для обнаружения лесных пожаров.

Важно отметить, повышенный интерес в решении задач мониторинга получили беспилотные летательные аппараты (БЛА), которые за счет простоты в конструкции и легкости в эксплуатации находят широкое применение в различных областях человеческой деятельности. В частности, применение БЛА при решении задач обнаружения лесных пожаров позволит существенно повысить оперативность нахождения очагов пожаров на ранней стадии возгорания [4].

Указанные обстоятельства определяют актуальность и важность рассматриваемой в данной статье задачи, посвященной исследованию и разработке ВСнКМ для обнаружения лесных пожаров и поддержки диспетчерского центра МЧС при слежении за процессом ликвидации и контроля пожаров в реальном времени.

Целью работы является исследование и разработка ВСнКМ для обнаружения и передачи информации о лесных пожарах, для этого в работе предлагается решение двух основных задач, которое должно обеспечить достижение указанной цели:

- а) спектральное преобразование для высокоскоростной передачи видеинформации в реальном времени;
- б) пространственное преобразование для анализа и обнаружения лесных пожаров и поддержки диспетчерской центра МЧС.

### **Обобщенная структурная глобальной системы наблюдения за лесными пожарами**

Очевидна тенденция одного видеонаблюдения для эффективной защиты объекта все чаще недостаточно. Российским бизнес- и госструктурам нужна сегодня не столько фиксация отдельных событий, сколько вовлечение видеонаблюдения в процессы контроля, управления и аналитики. Обеспечение комплексной безопасности также требует интеграции различных технологий и средств проектирования [5,6].

На рис. 1. предложена обобщенная структура ВСнКМ, в состав которой входит следующие ВСнК:

- 1) ВСнК I – видеосистема на кристалле оперативного обнаружения ЛС и передачи соответствующей информации диспетчерскому центру, ВСнК I б – видеосистема на кристалле приёма, анализа и выдачи соответствующих сигналов предупреждения опасности;
- 2) ВСнК II – видеосистема на кристалле приёма, анализа и передачи видеинформации;
- 3) ВСнК III – высокоскоростная параллельная видеосистема на кристалле кодирования и декодирования космических снимков изображений.

Стоит отметить, что важнейшим этапом проектирования интеллектуальных систем мониторинга является определение признаков лесных пожаров (оптических и геометрических) для обучения, обнаружения и оперативного распознавания в реальном времени.

Информация о лесных пожарах включает следующее:

- место и координаты очагов и лесных пожаров (ЛП),
- площади и стадии возгорания,
- прогнозы и степени риска,

- информацию о ближайших населенных пунктах и т. п.

Рассмотрим методы решения проблем обнаружения и передачи информации о лесных пожарах в указанных выше видеонформационных системах (I-III).

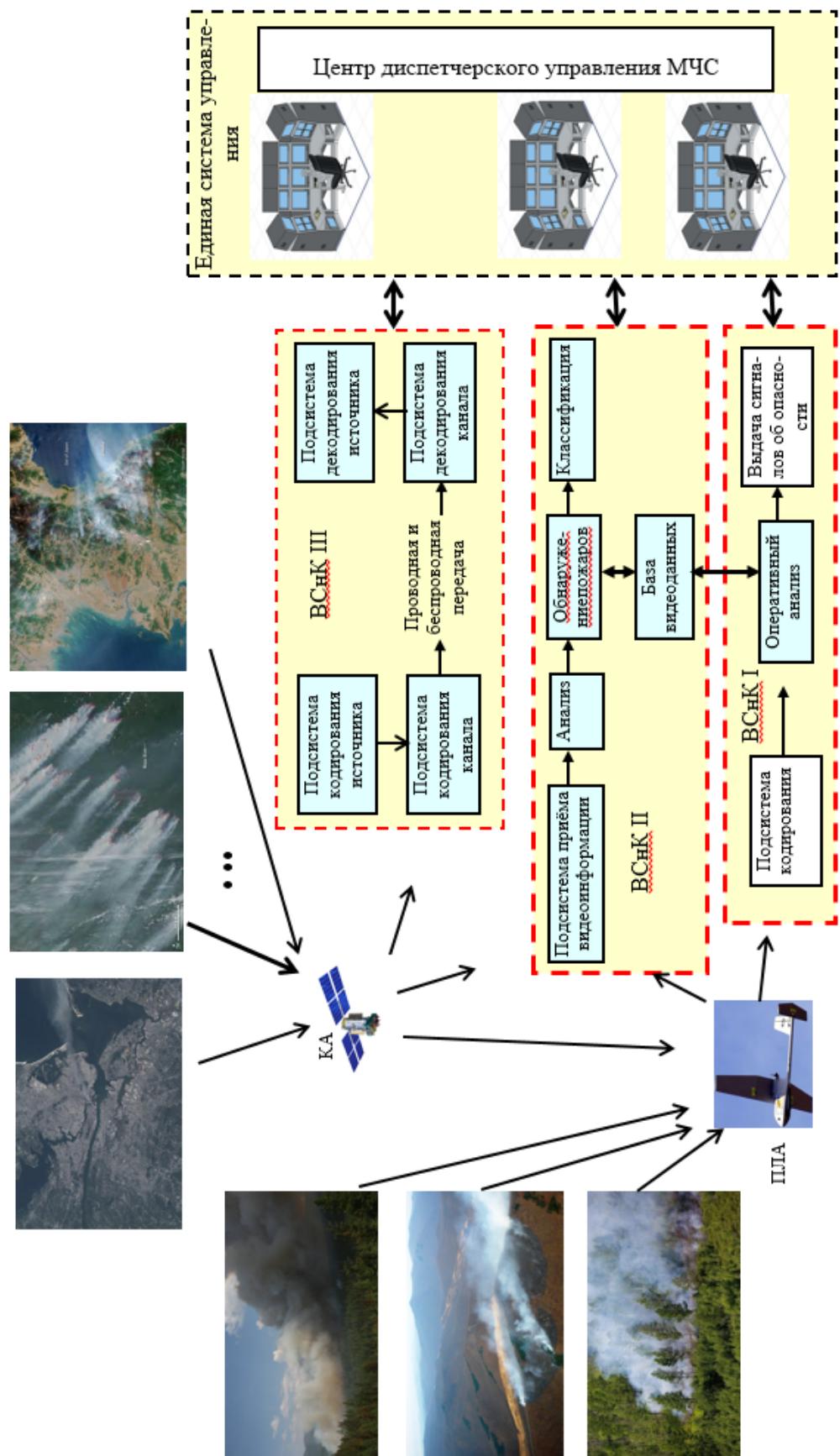


Рис. 1. Обобщенная структура модели видеонформационных систем наблюдения и контроля

## Пространственный метод обнаружения пожаров

*Видеосистемы на кристалле I и II.* Одним из перспективных направлений поиска и обнаружения лесных пожаров на изображениях является использование цветовую гамму на этапе поиска и выделения области пожаров на этапе обнаружения и вычисления площади пожара на любых его стадиях развития (рис. 2) [7].

Поиск области пожаров осуществляется рекурсивным разбиением изображения на определенное число полигонов. Далее оценивается точность аппроксимации

полигонами при условии, что плоскость полигона определяется линейной интерполяцией яркостей в пределах полигона. Те полигоны, точность аппроксимации которых не удовлетворительна, подлежат дальнейшему разбиению и анализу до тех пор, пока все изображение не окажется покрытым неравномерной полигональной сеткой. Данный метод существенно упрощает построение триангуляции, так как требуемая область жестко связана определенной структурой.

Введение регулярности явилось базой для создания параллельных алгоритмов структурирования областей пожаров на изображениях с целью определения характеристик и параметров (площадь, интенсивность, периметр, степени риска, относительный объём ущерба и т.п.) лесных пожаров по космическим снимкам. Обращено внимание на возможность деления изображения на полигоны различной формы (треугольники, квадраты и прямоугольники).

Рис. 2. Упрощенный алгоритм обнаружения лесных пожаров по космическим снимкам

При этом получены сравнительные оценки эффективности их использования при решении задачи обнаружения лесных пожаров в реальном времени.

Применение в работе рекурсивного разбиения изображения на три полигона и последующее обнаружение с использованием триангуляции объясняется следующими основными причинами:

- аналитически и экспериментально доказана эффективность трихотомии с точки зрения скорости обнаружения с минимальной ошибкой в реальном времени;
- треугольник является простейшим полигоном, вершины которого однозначно задают грани, аппроксимирующие любую область на изображениях.



## Спектральный метод передачи видеоинформации о пожарах

*Видеосистема на кристалле III.* В связи с необходимостью высокоскоростной передачи видеоинформации, видеосигнал рассмотрен, как многомерная функция, которая имеет оптимальное разложение по всем своим аргументам. Поэтому для устранения временной корреляции нет существенных формальных оснований использовать средства, отличающиеся от применяемых для пространственной декорреляции, т.е. дискретное косинусное преобразование (ДКП) [8].

В соответствии с этим в НИИТ совместно с ИПТ РАН разработан новый метод адаптивного трехмерного дискретного косинусного преобразования (ДКП-3D) для кодирования и декодирования видеоинформации, так как для большинства практически значимых классов кусочно-стационарных изображений по критерию минимума,

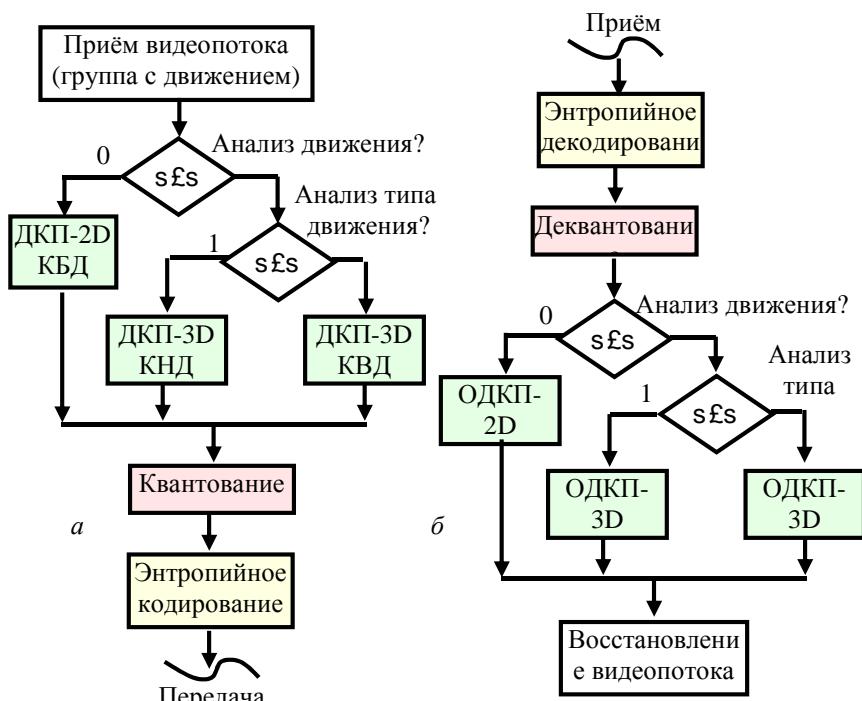


Рис. 3. Упрощенные адаптивные алгоритмы спектрального: а) кодирования и б) декодирования видеоинформации

требуемого для передачи количества информации при заданном уровне качества изображения, этот метод является оптимальным (асимптотически оптимальным).

С точки зрения вычислительной сложности алгоритм кодирования и декодирования на основе ДКП-3D характеризуется примерно в 10 раз меньшими временными затратами в сравнении с известными стандартами (MPEG). Это достигается в основном за счет временной межкадровой корреляции и отсутствия механизмов компенсации движений.

Кроме того, предложены алгоритмы кодирования изображений, адаптированные к различным динамическим сюжетам – без движения и с движением (рис. 3). При этом для групп кадров видеопотока с движением осуществляется разбиение потока на видеокубы различного размера (при интенсивном движении размер видеокуба делается меньше).

Для групп видеопотока с движением на этапе кодирования (рис. 3а), вычисляется среднеквадратическое отклонение ( $s$ ) между первым сегментом видеокуба и последним, далее в зависимости от заданных порогов ( $s_1$  и  $s_2$ ) формируются кубы: без движения (при  $s \leq s_1$  - КБД), с низкой (при  $s_1 \leq s \leq s_2$  - КНД) и высокой скоростью движения (при

$s > s_2$ - КВД). А на этапе декодирования (рис. 3б) восстановление видеопотока выполняется в обратном порядке.

Предварительный анализ движения в пределах одной группы и задание соответствующих значений порогов ( $s_1=4$  и  $s_2 =24$ ) позволили, при сохранении качества видеонформации, существенно уменьшить скорости передачи по сравнению с известными кодерами (при размерах куба 8 и 16 в 1,5...2 раза, а при размерах 4–16 в 2...3 раза).

На основе разработанного метода и алгоритмов ДКП-3D предложен вариант СФ-блока высокоскоростного преобразования элементарного ДКП-3D-видеопотока в элементарный поток MPEG-2 без трудоемких многомерных преобразований отсчетов, позволяющего: а) снижать трафик в глобальных сетях распределения данных; б) перераспределять трафик между двумя видами сетей передачи видеоданных, в которых требуется экономия трафика (глобальные сети Интернет, видео по требованию, спутниковое телевещание, IP-телевидение), и в которых этого ограничения не требуется; в) повышать качество видео за счет модифицированного обратного ДКП по времени трансформант и формирования из них Intra-кадров (с наименьшими потерями) в соответствии со стандартом MPEG-2.

### **Результаты моделирования алгоритмов обнаружения**

В результате реализации предложенных алгоритмов дистанционного зондирования и обработки космических снимков и обнаружения лесных пожаров были получены характеристики и необходимые информации, позволяющие своевременно принять меры по ликвидации и предупреждению соответствующих органов управления МЧС (рис. 4).

ВСнКМ, с точки зрения гибкости, точности, стоимости и безопасности обеспечивает значительные преимущества перед системами видеонаблюдения, характеризующими отсутствием аналитики визуальных данных, принимая визуальные и тепловые изображения из зоны пожара в реальном времени.

Следует отметить на основе вышеописанных алгоритмов анализа и передачи видеонформации и технологий в НИИТ и в ИПТ РАН созданы теоретическая база и технологии проектирования сложно-функциональных блоков для реализации различных специализированных ВСнКМ [10]. Некоторые из этих блоков эксплуатируются более 10 лет, поэтому в вышеуказанных институтах накоплен опыт, позволяющий оценивать тенденции развития ВСнКМ, которые в основном были изложены во работах [11-13].

В заключение работы следует еще раз остановиться на одной из наиболее важных возможностей развития современных ВСнКМ, а также проиллюстрировать, насколько активно она начинает использоваться уже сегодня. Такой возможностью, появившейся и быстро развивающейся в последние годы, стало появление технологий, позволяющих эффективно использовать в работе конкретной ВСнК информационные и вычислительные ресурсы различных центров представления спутниковой информации и других ВСнК. Это позволяет существенно сократить ресурсы, необходимые для создания специализированных ВСнКМ.

Фактически современные ВСнКМ могут создаваться и эксплуатироваться на основе использования уже имеющихся ресурсов, которые могут предоставляться ВСнКМ как сервисы. При таком подходе в интересах конкретных ВСнКМ необходимо лишь создавать специализированные процедуры и интерфейсы работы с данными, которые могут функционировать, используя не принадлежащие ВСнК информационные и

вычислительные ресурсы, фактически получая их в аренду (подписываясь на сервисы). Это, в конечном итоге, должно позволить существенно сократить расходы как на создание, так и на эксплуатацию многих специализированных ВСнК.

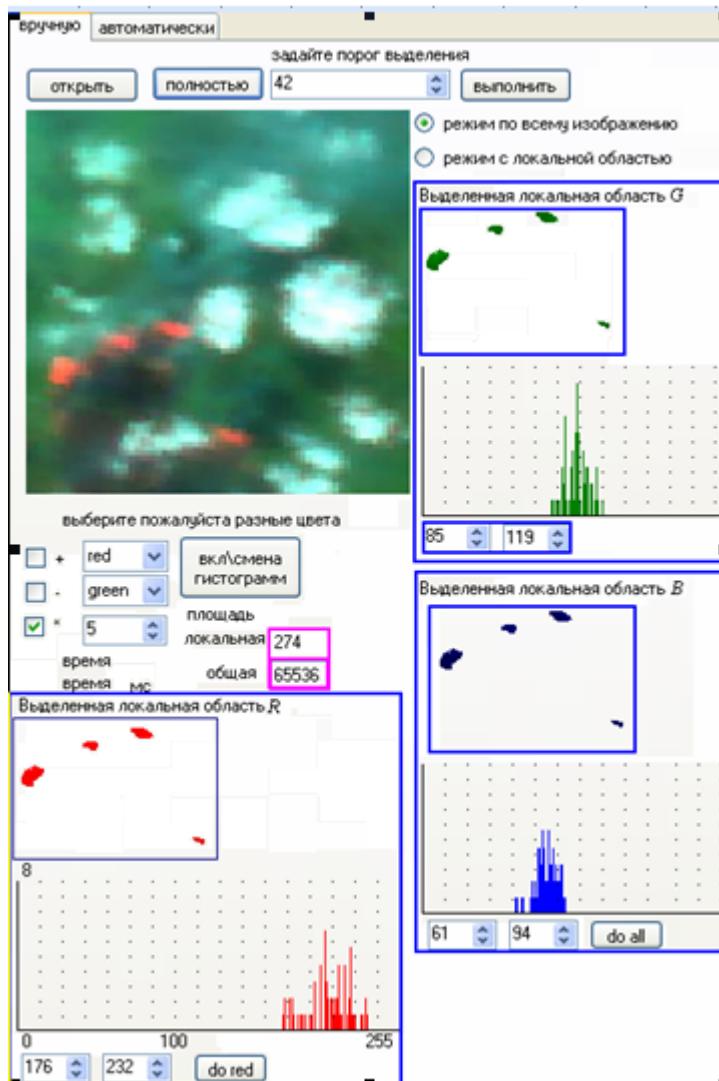


Рис. 4. результаты обработки космического снимка и выделения областей пожаров

## Заключение

В заключение можно сделать следующие важные выводы:

- 1) Лучшая стратегия тушение пожара – это предотвращение пожара сообщением о возникновении дыма. Тушение пожара начинается до возникновения пожара путём наблюдения за поступающим видеопотоком в реальном времени и с учётом признаков, характеризующих нормальное состояние лесов.
- 2) Данная работа является частью более комплексного интегрированного проекта охраны лесов от пожаров. Целостная система охраны лесов от пожаров будет базироваться на комплексной информационной системе, где будут интегрированы различные датчики и все виды деятельности, связанные с ранним обнаружением пожара в течение суток и обеспечивать диспетчерский центр МЧС необходимым видеинформацией мониторинга.
- 3) Рост объемов спутниковой информации и развитие требований и стоимости видеосистем на кристалле привели к переходу от использования локальных систем приема к глагольным системам на основе современной вычислительной базе класса система на

кристалле. В ближайшие годы системы дистанционного зондирования земли полностью перейдут на специализированные видеосистемы на кристалле с программируемой.

### Библиографический список

1. Е.А. Лупян, И.В. Балашов и др. Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 53–75.
2. Gordeev E.I., Girina O.A., Loupian E.A., Efremov V.Yu., Sorokin A.A., Mel'nikov D.V., Manevich A.G., Romanova I.M., Korolev S.P., Kramareva L.S. Vozmozhnosti ispol'zovaniya dannykh giperspektral'nykh sputnikovykh nablyudenii dlya izucheniya aktivnosti vulkanov Kamchatki s pomoshch'yu geoportala VolSatView (Using satellite hyperspectral data to study the activity of Kamchatka volcanoes on the basis of the VolSatView geoportal), Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa, 2014, Vol. 11, No. 1, pp.267–284.
3. Васильев А.С., Краснящих А.В., Коротаев В.В., Лашманов О.Ю., Ненарокомов О.Н., Лысенко Д.Ю., Широков А.С., Ярышев С.Н. Разработка программно-аппаратного комплекса обнаружения лесных пожаров методом совмещения изображений // Изв. вузов. Приборостроение. - СПб, 2012. - Т. 55. -№ 12 –С.50-56.
4. Студеникин А.В., Михалин В.А., Иванов Р.В., Магаршак С.И. Практика применения перспективных беспилотных летательных аппаратов для мониторинга и аэрофотосъемка. Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 4. С 102-106.
5. Фахми Ш.С. Концепция проектирования интеллектуальных транспортных видеосистем на основе технологии «система на кристалле»// Журнал университета водных коммуникаций. 2013. Вып. II (XVIII). С. 79–88.
6. Фахми Ш. С., Альмахрук М., Салем А., Бобровский А. И., Еид М.М., Костикова Е.В. Выделение и хранение опорных точек изображений на основе динамических рекурсивных структур // Вопросы радиоэлектроники, сер. Техника телевидения, 2016, вып. 3, с. 99–107.
7. Фахми Ш.С., Альмахрук М. М., Соколов Ю. М., Бобровский А. И., Еид М. М., Салем А. Точность, скорость и сложность устройств кодирования изображений по опорным точкам// Научно-технический вестник информационных технологий, механики и оптики. 2016. Том 16. № 4. С. 678–689.
8. Фахми Ш. С., Цыцуллин А. К. Видеосистемы на кристалле: новые архитектурные решения в задачах обработки видеинформации. Датчики и системы, 2011 г., № 4 – С.58–62.
9. Фахми Ш. С., Левко Г. В., Ибатуллин С. М., Морозов А. В., Чиркунова А. А., Чепелев А. Г. Измерение качества видеинформации на выходе цифрового кодера источника и его эффективности// Вопросы радиоэлектроники, сер. Техника телевидения, 2016, вып. 4, с. 33–40.
10. Малыгин И.Г., Комашинский В.И., Катцын Д.В. Некоторые проблемы построения когнитивных транспортных систем и сетей. В сборнике: Транспорт России: проблемы и перспективы - 2015 Материалы Юбилейной Международной научно-практической конференции. 2015. С. 3-8.
11. Адамов Д.Ю. Основные вопросы проектирования видеосистем на кристалле Вопросы радиоэлектроники. Серия: Техника телевидения. 2014. № 2. С. 3-17.
12. Умбиталиев А.А., Цыцуллин А.К., Пятков В.В., Рогачёв В.А., Левко Г.В., Морозов А.В. Статистический синтез управления телевизионной системой, адаптивной к динамике сюжета Вопросы радиоэлектроники. Серия: Техника телевидения. 2016. № 1 (24). С. 3-11.
13. Berezin V.V., Fakhmi Sh.S., Tsytulin A.K. INITIAL DESIGN STAGE OF VIDEO SYSTEMS ON A CHIP Journal of Optical Technology. 2012. Т. 79. № 11. С. 733-737.

### VIDEOSYSTEMS ON CHIP DETECTION FOREST FIRES

**A.I. Bobrovsky<sup>4</sup>, Sh. S. Fahmy<sup>1,2,3</sup>, Y. V. Alekseenko<sup>3</sup>, E. P. Ermakov<sup>3</sup>**

<sup>1</sup> Institute of problems of transport after. N. With. Solomenko FANO

<sup>2</sup> Saint-Petersburg University of State fire service of EMERCOM of Russia (SP UGPS the Ministry of emergency situations of Russia)

<sup>3</sup> Saint-Petersburg state electrotechnical University "LETI"

<sup>4</sup> State Research Institute of Applied Problems

*A generalized scheme for the global observing system for forest fires from space-based technology "system on chip". Considered spectral method of high-speed transmission of video data and spatial analysis method and the detection of forest fires.*

**Video system-on-chip, detection of fires space monitoring**

**А. И. Бобровский<sup>4</sup>, Ш. С. Фахми<sup>1,2,3</sup>, Ю. И. Мукало<sup>4</sup>, И. А. Гаврилов<sup>4</sup>**  
<sup>1</sup>СПбГЭТУ «ЛЭТИ»

<sup>2</sup>Научно-исследовательский институт телевидения

<sup>3</sup>Институт проблем транспорта им. Н.С. Соломенко ФАНО

<sup>4</sup>Научно-исследовательский институт прикладных проблем

## Новый метод обнаружения объектов на звёздном фоне с использованием регулярных решёток

*Рассмотрены алгоритмы решения задачи обнаружения объектов на звёздном фоне с использованием регулярных решёток, полученных пирамидально-рекурсивным разбиением изображений. Приведены результаты обнаружения объектов в двух режимах движения относительно звёздного фона с оценкой координат и яркости объекта.*

**Обнаружение, разбиение, структуризация, рекурсия, решётка, полигон**

### Введение

В задачах наблюдения за космическим пространством с целью обнаружения и слежения за космическими аппаратами, системах предупреждения о кометно-метеоритной опасности и системах обнаружения космического мусора стоит задача обнаружения целей, движение которых отлично от видимого движения звёздного фона [1]–[3]. Задача поиска подвижных и неподвижных точечных целей на подвижном (или неподвижном) звёздном фоне в последовательности изображений осложняется неинвариантностью формы объектов к сдвигу положения объекта относительно дискретного растра фотоприёмной матрицы, а малое отношение сигнал/шум требует совместной обработки нескольких кадров с целью достижения приемлемой достоверности обнаружения [4].

Рассмотрим задачу построения оптимальных структур данных изображения в виде неравномерных решёток с использованием рекурсивных алгоритмов разбиения исходного изображения на непересекающиеся области (полигоны), обеспечивающих предварительное сжатие изображения для сокращения вычислительных затрат при обнаружении и оценивания параметров цели. Результатом структуризации видеосигнала является иерархическая неравномерная решётка, логически связанная и построенная в результате разбиения изображения на полигоны. Целью данной работы является решение задачи обнаружения малоразмерных неподвижных целей с неизвестными координатами и яркостью на подвижном фоне звёзд в последовательности изображений на основе пространственно-рекурсивных решётчатых структур представления объектов.

Для достижения данной цели необходимо решить следующие задачи:

1. Разработать метод и алгоритмы построения регулярных решёток объектов (РРО) и оценки геометрических и амплитудных характеристик объектов изображения с заданной точностью при использовании рекурсивных процедур разбиения и структуризации объектов изображения для дальнейшей их параллельной реализации на программируемых схемах.

2. Повысить достоверность обнаружения при воздействии аддитивного шума датчика и фона путём управления локальными и глобальными порогами.

## Описание метода обнаружения объектов в звёздном небе

Рассмотрим метод поиска неподвижных точечных целей на подвижном звёздном фоне в последовательности изображений с использованием пространственно-рекурсивного разбиения на полигоны различной формы и площади и создания регулярных решёток объектов (рис. 1). Идея пространственно-рекурсивного метода (ПРМ) заключается в преобразовании каждого изображения видеопотока в такую структуру данных, которую сохраняла бы взаимосвязи объектов и имела минимальную длину описания с целью сокращения вычислительных затрат. Для проверки метода в работе были реализованы и протестированы различные алгоритмы построения структурированных иерархических неравномерных решёток.

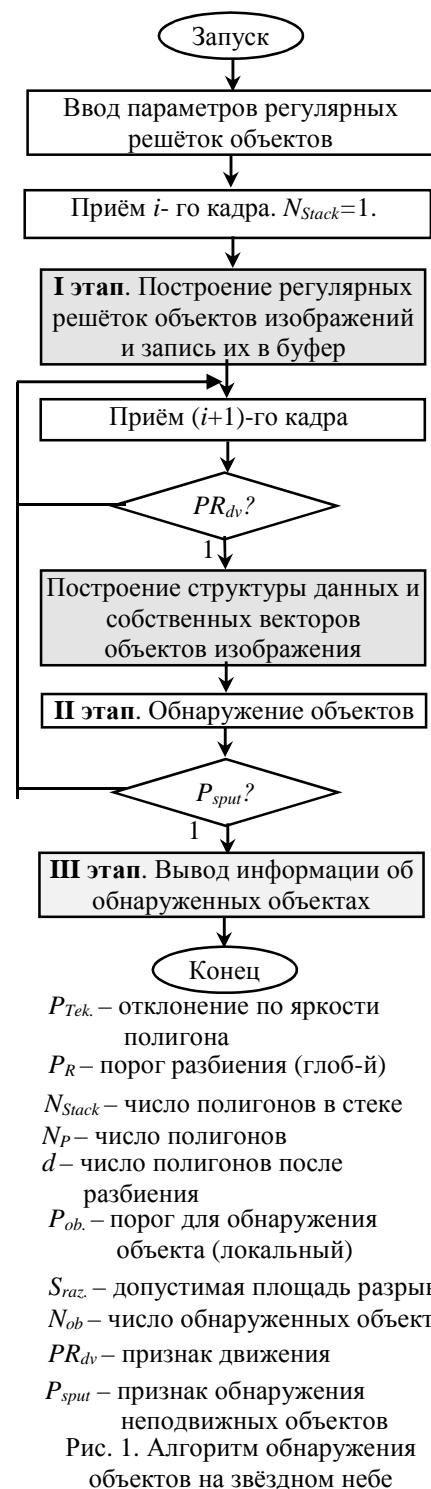
В режиме реального времени осуществляется подготовка и инициализация основных параметров: количество решёток, количество обнаруженных объектов, количество уровней пирамиды изображения, список решёток объектов, список объектов и т. д.) (рис. 1).

По сигналу «Запуск» алгоритм осуществляет формирование структуры данных изображения – РРО методом деления на полигоны в зависимости от заданного пользователем порога по яркости сигнала изображения ( $P_R$ ). Этот порог является глобальным, т. к. он предназначен для борьбы с фоновым шумом. Далее программа находится в состоянии ожидания до появления признака движения ( $PR_{dv}$ ). Следующим шагом является определение параметров объектов (номер уровня, число пикселов, координаты объекта, список пикселов контура объекта) и построение собственных векторов объектов изображения на основе созданных РРО.

Даная процедура необходимо для следующего основного этапа обнаружения объектов на изображении. В результате завершения процесса обнаружения выводится вся информация об обнаруженных объектах: координаты объекта, контур и площадь объекта и др.

Алгоритм включает в себя три основных этапа (рис. 1):

**Этап I. Построение регулярных решёток.** Регулярная решётка объектов изображений – это структура данных типа – «Планарный граф». С помощью такого графа осуществляется определение всех параметров объектов изображения с использованием, заранее созданных РРО.



На данном этапе выполняется рекурсивное разбиение исходного изображения на полигоны одинаковой формы и площади с целью формирования регулярной структуры с заданным глобальным порогом (рис. 2 а). Управление данным порогом позволяет удалить фоновый шум без влияния на качество геометрических величин и форм предполагаемых объектов. Затем процедура разбиения повторяется для каждого нового полигона до тех пор, пока условие ( $P_{Tek} < P_R$ ) выполняется в пределах полигона. Предельный случай разбиения – размер полигона равен размеру элемента исходного изображения. В результате данного этапа получаем неравномерную регулярную решётку объектов (рис. 2

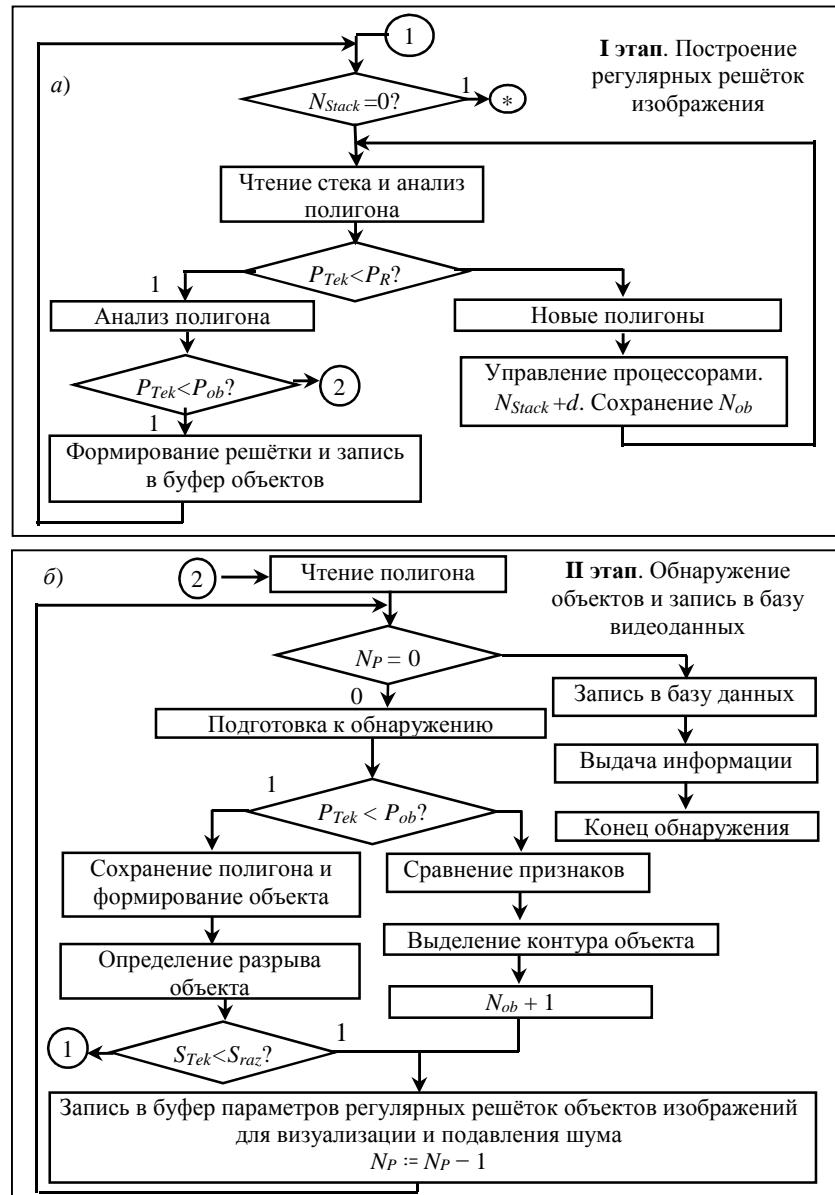


Рис. 2. Алгоритмы построения регулярных решёток – а;  
обнаружения объектов изображения – б

а), где каждому полигону в объектно-ориентированном описании приписываются: размер, номер уровня, номер объекта, признак разрыва объекта, координаты центра тяжести, яркость полигона для записи в буфер решёток объектов изображения [5], [6]. По завершению данного этапа формируется множество полигонов в виде регулярной неравномерной решётки.

Решётка является результатом структуризации исходного изображения с минимальной длиной описания (I) объектов и взаимосвязей между ними [11], [12]:

$$I = \log d (nM - (1/d^M) \sum_{l=1}^m \mathbf{1} \otimes \mathbf{1}) + \begin{cases} 3nE(E+1), & \text{при } d = 2; \\ n(2dE + d + 1)/(d - 1), & \text{при } d > 2, \end{cases}$$

где  $n$  – число полигонов, содержащих искомые объекты;  $M$  – максимальное число разбиений, приводящее к размеру полигона в один пиксель изображения;  $E = Dn/n$ , где  $Dn$  – полигоны, не содержащие полезной информации (т. е. объектов);  $\ell$  – текущий номер уровня пирамиды изображения;  $d$  – число полигонов, полученных при разбиении исходного полигона;  $N$  – общее число пикселов области объекта (звезды) на исходном изображении; очевидно, что  $N = d^M$ ;  $z_\ell$  – число полигонов, содержащих объекты уровня  $\ell$ .

Механизм анализа и формирования структуры изображения основан на использовании стандартной структуры данных типа «Стек» (*Stack*). Если значение отклонения по яркости в пределах анализируемого полигона превышает заданное, то формируются новые полигоны и записываются в стек для дальнейшего анализа. В противном случае проверяется условие ( $P_{Tek} < P_{ob}$ ) и вычисляются параметры полигона (размер, координаты и уровень), площадь, контур и координаты предполагаемого объекта. Далее вся информация записывается в буфер объектов для последующего этапа обнаружения.

**Этап II. Обнаружение объектов.** Данный этап (рис. 2б) является наиболее сложным с точки зрения организации параллельной обработки видеоинформации в видеосистеме на кристалле из-за необходимости синхронизации и согласования работ всех процессоров при анализе и выделении признаков объектов изображения.

На этом этапе процедура обнаружения объектов основана на фиксации неподвижных путём сравнения соответствующих фрагментов двух решёток  $i$ -го и  $(i+1)$ -го кадров.

*Выделение контура объекта.* Локальная обработка в области объектов заключается в выделении на изображении только пикселов, лежащих на контуре (границе). На практике это множество пикселов редко отображает контур достаточно точно по причине шумов, разрывов контуров из-за неоднородности освещения и т. п. Поэтому в работе авторами предложены дополнительные процедуры связывания (на основе, созданной на первом этапе пирамиды изображения и РРО), чтобы сформировать список контурных точек (рис. 3) [7], [8].

Подход к обнаружению и связыванию контуров на основе представления в виде регулярных решёток и поиска на этой решётке путей с наименьшим весовым коэффициентом, которые соответствуют значимым контурам, позволяет построить метод, хорошо работающий в присутствии шума. Такая процедура оказывается достаточно сложной и требует большого времени обработки работе с динамическими массивами [9].

Элемент контура – граница между двумя пикселями  $w$  и  $q$ , являющимися соседями. Элементы контура идентифицируются координатами точек  $w$  и  $q$ . Например, на рис. 3 элемент контура определяется парами  $(x_w, y_w)(x_q, y_q)$ . Контур – список или последовательность соединенных между собой элементов. Каждому элементу контура, заданному пикселями  $w$  и  $q$  (рис. 3ж), соответствует некий весовой коэффициент  $c(w, q) = Max - [f(w) - f(q)]$ , где  $Max$  – максимальный уровень яркости в изображении;  $f(w)$ ,  $f(q)$  – яркости пикселов  $w$  и  $q$  соответственно.

Задача отыскания на решётке пути минимального коэффициента является нетривиальной по вычислительной сложности, поэтому приходится жертвовать оптимальностью в пользу скорости вычислений [7].

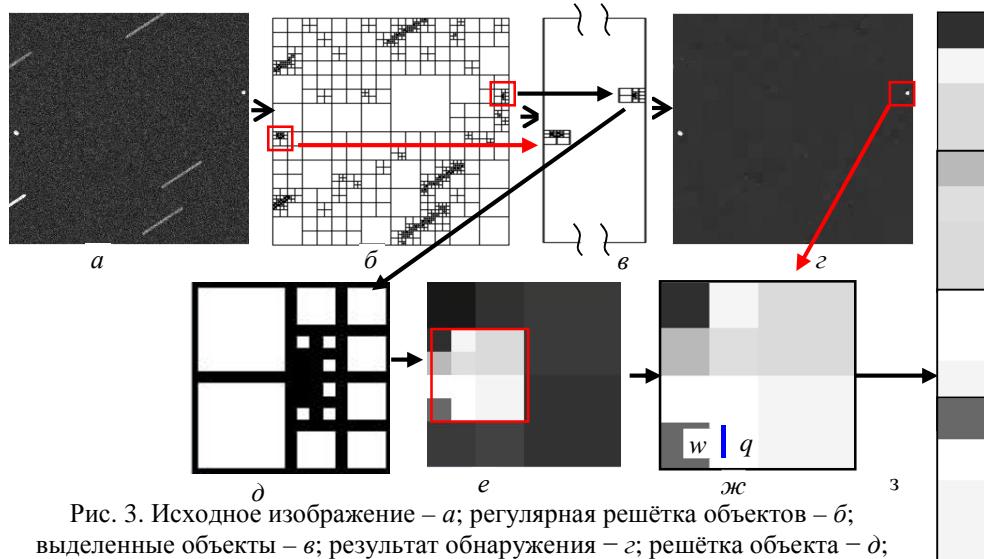


Рис. 3. Исходное изображение – *а*; регулярная решётка объектов – *б*; выделенные объекты – *в*; результат обнаружения – *г*; решётка объекта – *д*; область объекта – *е*; сам объект – *жс*; собственный вектор объекта – *з*

Учёт сложности реализации и большой ресурсоемкости на этапе проектировании видеосистемы обнаружения возможен только с использованием многопроцессорных программируемых вентильных матриц со встроенными процессорными компонентами [9].

Способы связывания точек контура объектов изображения, предложенные в [8] и [10], характеризуются анализом характеристик пикселов в небольшой окрестности каждой точки изображения, которая была отмечена как контурная. Все точки, являющиеся сходными в соответствии с некоторыми критериями, связываются и образуют контур, состоящий из отвечающих этим критериям пикселов. При этом, в указанных выше ссылках, используются два основных параметра для установления сходства пикселов контура: отклик оператора градиента, определяющий значение пикселя контура, и направление вектора градиента [12]–[14].

Предложенный авторами новый алгоритм обнаружения объекта на основе ООП осуществляет сравнение решёток *i*-го и (*i* + 1)-го кадров из видеопотока и перебор всех локальных областей (рис. 3б) для выявления и фиксации неподвижных объектов по геометрическим признакам (смещение объектов). Если смещение предполагаемого объекта равно нулю, то данная локальная область содержит неподвижный объект и в буфер обнаруженных объектов (база видеоданных) заносится соответствующая информация в двух видах: собственная матрица (рис. 3жс) и собственный вектор (рис. 3з) объекта. Далее выполняется вычисление параметров обнаруженных объектов, т. е. вычисляются функции, указанные в списке разработанного выше класса *class object\_sputnik*.

Например, на *i*-м кадре рассчитывался бинарный параметр правильного обнаружения  $I_{pr(i)}$  по следующему правилу:  $I_{pr(i)} = 1$ , если  $|T_{(x,y)i} - T_{(x,y)i+k}| < 1$  - отклонение значения координат центра объекта в течение последовательности из *k* – кадров, а  $I_i = 0$  в противном случае. Вычисление частоты правильного обнаружения по формуле  $p_{pr} = \sum I_{pr}/k$ , частоты ложных тревог  $I_{loj}$ , как усредненное отношение количества точек, неверно

отнесённых к объекту. При варьировании значений и глобальных ( $P_R$ ) и локальных ( $P_{ob}$ ) порогов можно найти зависимость  $I_{pr} = f(I_{loj})$  [12]–[15].

*Визуализация подавления шума.* Объединение результатов работы алгоритмов построения РРО с помощью операции «ИЛИ» мало исследовано в связи с возникающими трудностями из-за возрастающей вероятности повреждения изображения. В работе были исследованы различные дополнительные фильтры, позволяющие значительно снизить вероятность повреждения изображения в блоках подавления шума (в данной работе не рассматриваются). Например, к фильтрам подавления шума относится многошаговый медианный фильтр [15].

Для иллюстрации эффективности фильтрации изображений в работе [15] предложен способ визуализации «гистограмма связности». Каждый пиксель растрового изображения имеет вокруг себя некоторое количество других пикселей – множество, называемое окрестностью этого пикселя. Обозначим символом  $K_{ok}$  – количество анализируемых пикселей, окружающих пиксель с координатами  $[x, y]$ , т. е. количество проверяемых связей. Пиксель изображения считается «связанным» с соседним пикселием, принадлежащим окрестности, если их яркости одинаковы. Обозначим количество соседей пикселя символом  $K_{sos}[x, y]$ . Обозначим количество пикселей изображения, имеющих одинаковую связность, символом  $K_{os}$ , где  $K_{os} \in [0, \dots, K_{ok}]$ .

Некоторые понятия, характеризующие РРО:

1. Гистограмма связности изображения (области) – это гистограмма, в которой отражается статистика области  $St(K_{os}) = f(K_{os}/S_{Iz})$ , где  $S_{Iz}$  – площадь изображения. Под выделенными пикселями понимаются пиксели того цвета, который в данный момент представляет интерес; при обнаружении звёзд – белый.

2. Общее количество соседей у всех выделенных пикселей – это величина  $K_{ob} = SK'_{sos}[x, y]$ , где  $K'_{sos}$  – количество связей у выделенного пикселя.

3. Средняя связность изображения – это величина  $K_{sr} = K_{ob}/K_v$ , где  $K_v$  – общее количество выделенных пикселей [12]–[14].

**Этап III.** Выдача информации об обнаруженных объектах. Результатом данного этапа является выдача списка информации об обнаруженных объектах: длина, ширина, контур, координаты по отношению к началу координат, площадь, яркости пикселов и др.

### Устройства обнаружения объектов

Системы и видеосистемы на кристалле (СнК и ВСнК) в основном строятся на базе программируемых модулей (рис. 4) – сложно-функциональных блоков (СФ-блоков) и характеризуются высокой надежностью, низкой стоимостью, малыми габаритами, малым потреблением энергии и высокой агрегатной способностью.

Важнейшей особенностью предлагаемого метода обнаружения объектов изображений является возможность отображения рекурсивных вычислительных процедур разбиения и структуризации видеоданных на параллельную архитектуру устройств ВСнК. Построение параллельных алгоритмов функционирования ВСнК и оценка производительности системы в целом приведено в работе [16].

В данной статье приводится описание некоторых основных СФ-блоков, входящих в состав устройств СнК и ВСнК. Выбор наилучшей структурной схемы ВСнК для выполнения алгоритмов структуризации и обнаружения объектов на изображениях

видеопотока с использованием пирамидально-рекурсивного подхода необходимо выполнить с учётом следующих основных аппаратно-программных характеристик (рис. 4):

наличие элементарных процессоров в виде СФ-блоков (СФ-блок ПР<sub>ЭЛЕМ</sub>) для выполнения рекурсивных функций над большими объёмами видеоданных;

наличие управляющего процессора (СФ-блок ПР<sub>УПР</sub>), формирующего соответствующих сигналов управления и синхронизации всего процесса обнаружения объектов;

наличие процессора диспетчера (СФ-блок ПР<sub>ЭЛЕМ</sub>), осуществляющего организацию взаимодействия работы элементарного процессора и обслуживание очереди на обработку полигонов элементарными процессорами без задержек.

Все элементарные процессоры могут использовать как общую, так и собственную

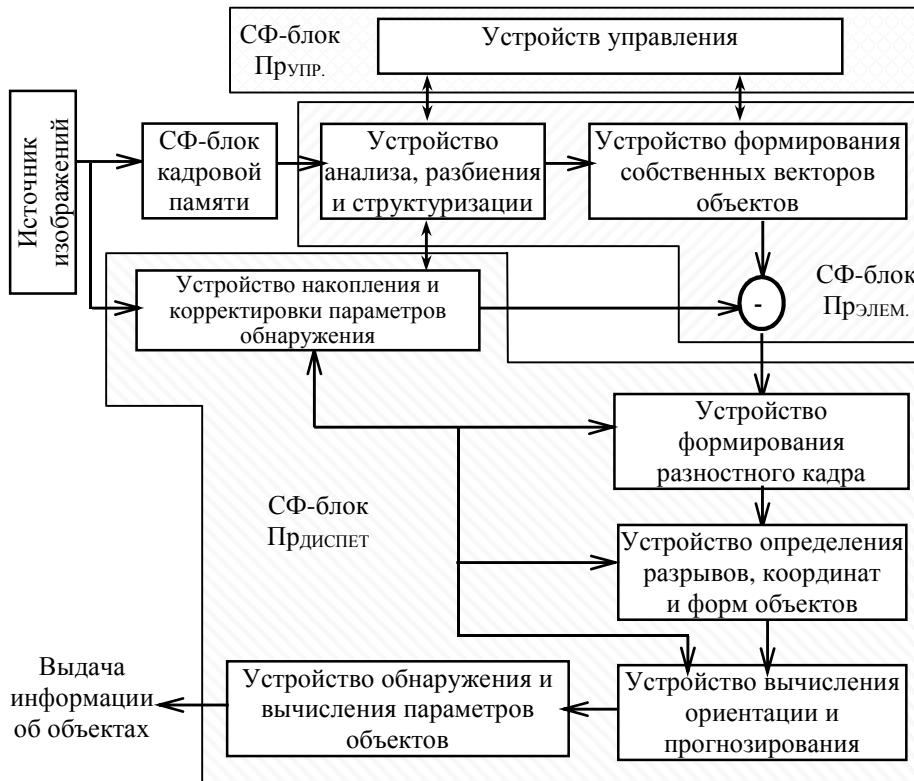


Рис. 4. Структурная схема ВСнК обнаружения объектов в звёздном небе

(локальную) память.

Аппаратно-программные средства могут функционировать как во взаимодействии друг с другом, так и параллельно.

Основной принцип, которым следует руководствоваться при построении специализированной ВСнК на основе ПРМ – это адекватность её структуры пространственно-временной структуре процесса вычислений, реализующего набор алгоритмов [16].

Приведённый универсальный алгоритм обнаружения объектов на изображениях, в основе которого лежит принцип рекурсивного деления на полигоны различной формы и размеров, позволяет реализовать различные устройства с параллельной обработкой, описание и реализации которых приведены в работе [16].

## Результаты моделирования алгоритмов

Моделирование алгоритмов структуризации и обнаружения объектов реализованы на языке C++ с использованием библиотеки OpenCV (рис. 5) и были получены результаты обработки видеопотока в двух режимах слежения: в процессе наблюдения за объектом при неподвижный объект на подвижном фоне (рис. 5 $a$ ) и при подвижном объекте на неподвижном фоне (рис. 5 $b$ ).

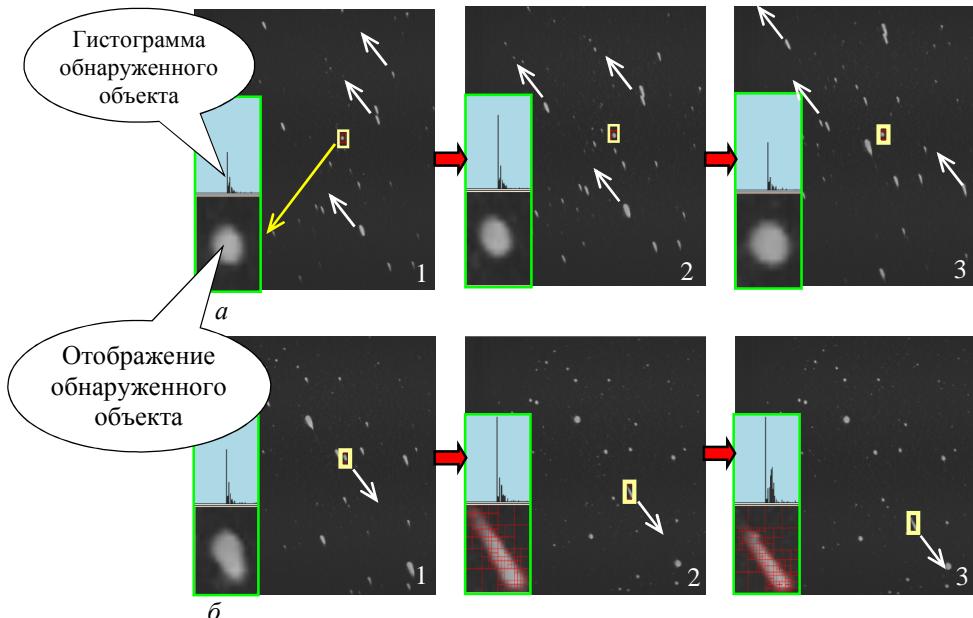


Рис. 5. Процесс наблюдения за объектом: неподвижный объект на подвижном фоне –  $a$ ; подвижный объект на неподвижном фоне –  $b$

В результате проведения экспериментов на натурных тестовых видео-последовательностях были получены количественные показатели качества обнаружения объектов, на основе которых можно сделать следующие выводы:

исследования показали, что ценой увеличения вычислительной сложности алгоритма обнаружения в 2...3 раза можно обеспечить улучшение точности и качества выделенных объектов (сигнал/шум) в ~2 раза;

использование новых рекурсивных алгоритмов построения неравномерных регулярных решёток объектов изображения позволяют создавать параллельные высокопроизводительные видеосистемы обнаружения на кристалле.

### Заключение

В процессе исследований задачи получены следующие результаты:

1. Показано, что построение неравномерных регулярных решёток при обнаружении и оценке координат объектов позволяет оценивать их параметры при минимальных вычислительных затратах и, благодаря реализованному управлению локальными и глобальными порогами, достичь достоверности обнаружения и оценки параметров, близкой к потенциальной в условиях ограничения преобразования непрерывного оптического сигнала в дискретный сигнал на выходе фотоприёмной матрицы.

2. Эксперименты подтвердили, что предложенные метод и алгоритмы обнаружения на основе неравномерных динамических решёток могут быть использованы для обнаружения объектов, как при нулевом, так и при произвольном векторе скорости цели относительно растра и фона звёзд, а также как на этапе послесеансной обработки

видеоинформации, так и на этапе сопровождения движущихся объектов в реальном масштабе времени.

3. Предложены алгоритмы обнаружения локальных неоднородностей на изображениях в условиях априорной неопределенности формы, размеров, координат и амплитуды локальной неоднородности, основанные на построении решёток и структуризации перепадов яркости с выделением связных границ неоднородностей.

4. Разработанные алгоритмы с использованием рекурсивных процедур разбиения и представления изображения могут быть реализованы аппаратно с использованием параллельных многопроцессорных видеосистем на кристалле, которые могут быть использованы в бортовых космических системах наблюдения.

### Библиографический список

1. L'Yvesque M. P., Leliuvre M., Improving satellite-streak detection by the use of false alarm rejection algorithms. DRDC Valcartier TR 2006-587. Defence R&D Canada – Valcartier. <http://pubs.drdc.gc.ca/PDFS/unc76/p530206.pdf>, accessed Apr. 2011.
2. Обнаружение объектов на звёздном фоне/ Левко Г. В., Бобровский А. И., Морозов А. В. и др. // Вопросы радиоэлектроники. Сер. Техника телевидения. 2016. Вып. 2. С. 29–38.
3. Pinkney F., Buteau S., L'Yvesque M. P., A Canadian Array of Ground-Based Small Optical Sensors for Deep Space Monitoring/ Wallace B., Rody J., Scott R. et all.// 2003 AMOS Technical Conference.
4. Сакович И. О., Белов Ю. С. Обзор основных методов контурного анализа для выделения контуров движущихся объектов // Инженерный журнал: наука и инновации. 2014. Вып. 12. URL: <http://engjournal.ru/catalog/it/hidden/1280.html>.
5. Александров В. В., Горский Н. Д. Представление и обработка изображений. Рекурсивный подход. Л.: Наука. 1985. 192 с.
6. Фахми Ш. С. Кодирование и декодирование видеоинформации // Вопросы радиоэлектроники. Сер. Техника телевидения. 2007. Вып. 2. С. 43–51.
7. Потапов А. С. Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания. СПб.: Политехника, 2007. 548 с.
8. Выделение и хранение опорных точек изображений на основе динамических рекурсивных структур / Бобровский А. И., Еид М. М., Альмахрук М. и др. // Вопросы радиоэлектроники. Сер. Техника телевидения. 2016. Вып. 3. С. 99–107.
9. <http://altera-plis.ru/production.html>
10. Шапиро Л., Стокман Дж. Компьютерное зрение. М.: Бином. Лаборатория знаний. 2009. 763 с.
11. Баян П. В., Фельдман А. Б. Обнаружение и оценка параметров воздушных объектов в видеопоследовательностях на основе кратномасштабной пространственной обработки // Цифровая обработка сигналов. 2010. № 4. С.7–11.
12. Ванияев А. А., Емельянов Г. М. Метод выделения быстродвижущихся объектов при использовании цифрового оптического локатора следящего типа // Компьютерная оптика. 2013. Т. 37. №4. С. 477–483.
13. Адаптивный метод и алгоритм обнаружения малоконтрастных объектов оптико-электронным средством/ Катулев А. Н., Колонсков А. А., Храмичев А. А. и др. // Оптический журнал. 2014. Т. 81. № 2. С. 29–39.
14. Кузьмин С. А. Сегментация последовательностей изображений с регулируемой точностью и визуализация эффективности // Технические науки – от теории к практике: сб. ст. по матер. XXV междунар. науч.-практ. конф. № 8(21). – Новосибирск: СибАК, 2013.
15. Чочия П. А. Пирамидальный алгоритм сегментации изображений // Информационные процессы. 2010. Т. 10. № 1. С. 23–35.
16. Фахми Ш.С., Бобровский А.И., Алексеенко Я.В. Реализация параллельной обработки видеоинформации на основе пространственно-рекурсивного метода // Вопросы радиоэлектроники. Сер. Техника телевидения. 2015. Вып. 4. С. 89–97.

### A NEW METHOD OF DETECTING OBJECTS IN THE STARRY BACKGROUND WITH THE USE OF REGULAR LATTICES

**Fahmi<sup>1,2,3</sup> Sh. S. , Mukalo<sup>4</sup> Y. I. , Gavrilov<sup>4</sup> I. A.**

<sup>1</sup>Saint-Petersburg state electrotechnical University "LETI"

<sup>2</sup>Scientific research Institute of television

<sup>3</sup>Institute of problems of transport after. N. With. Solomenko FANO

<sup>4</sup> State Research Institute of Applied Problems

*The algorithms of solving the problem of detecting objects in the starry background with the use of regular lattices, poluchennyh pyramidal-recursive partitioning images. The results of object detection in the two modes of motion relative to the starry background with the evaluation of the coordinates and brightness of the object.*

**Detection, partitioning, structuring, recursion, grid, polygon**

**И. С. Фомин, Д. А. Громошинский, А. В. Бахшиев**  
Центральный научно-исследовательский и опытно-конструкторский  
институт робототехники и технической кибернетики (ЦНИИ РТК)

## Обнаружение визуальных ориентиров в задачах стыковки с помощью глубокой нейронной сети

*Автоматическая и автоматизированная стыковка управляемых аппаратов – сложная и актуальная проблема современной техники и робототехники. Она может быть решена с использованием технического зрения, если на стыкующемся аппарате установлена камера, способная наблюдать стабильные визуальные ориентиры, относительное положение которых известно. Наиболее важным этапом в этом случае является задача обнаружения ориентиров. Задача будет рассмотрена на примере обнаружения визуальных ориентиров в процессе космической стыковки с использованием системы обнаружения объектов, основанной на глубокой нейронной сети.*

**Стыковка, управляемый аппарат, обнаружение объектов, обработка изображений, нейронные сети, глубокое обучение**

### Описание проблемы

Одной из сложных проблем в области работы автономных аппаратов, имеющих возможность автоматического или автоматизированного управления, является стыковка одного аппарата с другим либо специальным местом, предназначенном для стыковки, нахождения аппарата в неподвижном состоянии. При решении задачи стыковки важно иметь возможность определять относительное положение одного аппарата относительного другого либо относительно стыковочной позиции. Рассмотрим эту задачу на примере задачи стыковки космического аппарата (КА) и Международной Космической Станции (МКС). На данный момент стыковка осуществляется с использованием специальных радиоэлектронных и оптических систем, компоненты которых должны быть размещены как на МКС, так и на КА. Кроме того, все КА на протяжении более 40 лет оснащаются специализированной телевизионной системой, используемой в процессе сближения и стыковки в качестве средства визуального контроля.

Ранее в статьях [1], [2] описано применение телевизионной системы для определения параметров относительного положения КА относительно МКС. Для этого используется специальное программное обеспечение (ПО), которое может быть размещено на вычислителе МКС либо на стационарном ПК в Центре Управления Полетом. Данное ПО принимает сигнал с борта стыкующегося аппарата в процессе стыковки, производит совместное обнаружение и сопровождение хорошо видимых особенностей на МКС. Используя данные о положении этих особенностей, известную модель телевизионной камеры и 3D-модель МКС повышенной точности возможно с высокой точностью вычислить относительное положение КА и МКС путем решения PnP-проблемы.

### Постановка задачи

Одним из основных компонентов, который во многом определяет как быстродействие, так и качество работы системы, является модуль обнаружения и сопровождения особенностей. Тогда как расчетные методы, использующие модель камеры и модель МКС являются детерминированными и полностью математически

описанными, не изменяющимися отстыковки к стыковке, то система обнаружения объектов зависит от условий, в которых проходит стыковка, от качества телевизионного сигнала и получаемых изображений в процессе стыковки. От данных, получаемых системой обнаружения и сопровождения особенностей зависит результат работы прочих методов и конечные результаты определения относительного положения.

Существующая телевизионная система имеет ряд особенностей. В ней применяются аналоговые телекамеры и множество других аналоговых компонентов, участвующих в формировании сигнала, его обработке и передаче на МКС и в ЦУП. Эти компоненты вносят разнообразные искажения в передаваемый сигнал. Некоторые из подобных искажений, имевшие место в процессе одной из стыковок, приведены на рисунке 1. Кроме того, часто возникают проблемы, связанные с размытием изображения, засветкой камеры, быстрым изменением освещенности и, соответственно, яркости на изображении отдельных частей МКС.

Кроме подхода, в котором для компенсации изменений изображения используются алгоритмы разной степени сложности можно применить совершенно иной подход. Нейронные сети вообще и глубокие сверточные нейронные сети в частности обладают очень высокой способностью к обобщению информации. Не так давно была представлена работа [3], в которой комбинация нейронных сетей применяется для детектирования различных объектов на изображениях, при этом достигается точность обнаружение до 78.8%. Принято решение применить данный детектор для обнаружения объектов на видеозаписях стыковки, сравнить качество обнаружения с текущим, в дальнейшем, возможно, заменить текущий детектор на детектор на основе нейронной сети.

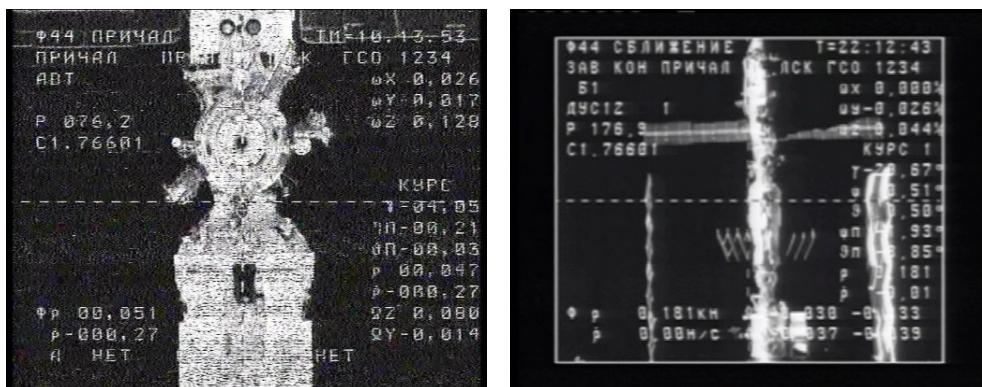


Рис. 1. Примеры шумов на изображениях

### Предыдущие работы

Система для обнаружения объектов Fast R-CNN [4] впервые показала достаточную для реального применения скорость обнаружения объектов. Данная модификация сверточной нейронной сети - улучшение оригинальной архитектуры R-CNN [5] разработанное для ускорения обнаружения объектов. Перед началом работы сети одним из специальных алгоритмов [6], [7] генерируются предположения регионов интереса (Region of Interest, RoI), далее все изображения отправляются на вход CNN. Сначала к обнаружениям применяется специальная регрессионная сеть, улучшающая положение каждого региона. Затем для каждого региона применяется подвыборка к финальному вектору признаков для получения вектора фиксированного размера, который подается на вход классификатора. Классификатор представляет собой полносвязную сеть с последующим выбором максимума по всем откликам классов.

При выполнении операции подвыборки выполняется проецирование найденного RoI на каждый из последних сверточных слоев с учетом изменения размера и положения региона после прохождения всех операций свертки. Далее, каждый RoI преобразуется в вектор стандартного размера.

Преимущество Fast R-CNN перед RCNN в том, что вычисление сверток производится не для каждого RoI, но один раз сразу для всего изображения, после чего полученные выходы сверток последовательно используются для каждого RoI. Таким образом, данная конфигурация сети обучается целиком за исключением метода обнаружения зон интереса, который работает отдельно.

Следующим логичным шагом после совместного использования результатов свертки стало применение нейронной сети не только для классификации найденных регионов, но и разработка специальной сети для формирования RoI, чтобы сделать метод полностью основанным на нейронных сетях и соответственно, полностью обучаемым. Также расчет нейронной сети ведется на GPU, а обнаружение RoI в Fast R-CNN выполняется на CPU, что сильно замедляет процесс.

Faster R-CNN [3] – это Fast R-CNN, объединенная с сетью предположения регионов (Region Proposal Network, RPN). Fast R-CNN и RPN разделяют между собой сверточные слои для ускорения работы.

Рассмотрим более подробно принцип работы RPN. Здесь используется скользящее окно, для генерации к прямоугольникам в каждой позиции на конечной карте признаков. Сначала выполняется свертка квадратным фильтром с картой признаков последнего сверточного слоя. Затем результат проецируется в пространство меньшей размерности, получая вектор фиксированного размера для каждого положения. Вектор передается в слой уточнения RoI и слои классификации RoI. В слое уточнения для каждого из прямоугольников генерируются 4 выхода, которые описывают уточненные положение и форму.

Слои классификации для каждого из прямоугольников генерируют пару выходов. Эти пары – значения вероятности того, является прямоугольник объектом или фоном. Для каждого из выходов сверточного слоя среди всех прямоугольников выбирают имеющий наибольшую вероятность принадлежности классу объекта. Затем из всех полученных прямоугольников оставляют имеющий максимальную принадлежность. Сеть RPN обучается по принципу, сходному с обучением основной сети Fast R-CNN. Предположение RoI помечается как хороший пример, если мало отличается от известной правильной рамки, или наиболее схоже с правильной среди всех, если рамок несколько. Предположение RoI отмечается как плохой пример, если сильно отличается от любой из правильных рамок.

Все хорошие примеры рамок передаются на вход детектора Fast R-CNN и классифицируются так, как описано ранее. Так как RPN и сеть детектора разделяют одни и те же сверточные слои, они могут быть обучены совместно. R-CNN сеть ожидает фиксированные предположения RoI. Положение прямоугольников предположений RoI меняется в процессе обучения RPN. Если RPN обучать одновременно с R-CNN, процесс обучения разойдется. Поэтому на практике сети в составе Faster R-CNN обучаются в несколько этапов. Сначала сама по себе обучается RPN. Затем сеть обнаружения (т.е. фактически Fast R-CNN) обучается с использованием полученных предположений RoI формируя отдельные сверточные слои. Затем вновь тренируется RPN, но уже используя сформированные на предыдущем этапе слои свертки. В конце слои свертки и RPN

фиксируются выполняется подстройка параметров полносвязных слоев в составе сети, отвечающих за классификацию объектов.

Для проведения исследований выбрана именно система Faster R-CNN, как имеющая наиболее высокое качество обнаружения объектов на изображениях в тесте PASCAL VOC на момент выполнения экспериментов. Точнее, была выбрана выложенная в свободном доступе модификация системы, выполненная с использованием языка python (py-faster-rcnn [8]), со стандартной конфигурацией сети VGG-CNN-M-1024, модифицированной под имеющееся число классов.

### **Описание обучающей выборки**

Для обучения выбранной нейронной сети была подготовлена выборка изображений, полученных в процессестыковки космических аппаратов с различными узлами МКС. Для каждой из стыковок была сделана разметка, отмечающая положение каждой из частей МКС, выбранной в качестве особенности, на каждом из изображений, если она там видна. Для каждого из 4 стыковочных узлов внешний вид МКС с камеры КА существенно отличается и поэтому набор особенностей для каждого из узлов свой. Данные о количестве особенностей, числе обучающих и тестовых изображений для каждого из узлов приведены в таблице 1.

Кроме того, следует отметить некоторые особенности выборки изображений, использовавшейся для обучения. Прежде всего, выборка, на которой была обучена нейронная сеть совпадает с той, которая была использована для обучения детектора объектов, который на данный момент используется в системе стыковки. Так как в имеющейся размеченной выборке изображений стыковки отсутствует вращение КА вокруг собственной оси, в обучающую выборку не добавлялись повернутые изображения для обеспечения инвариантности к повороту. Так как обучающие изображения представляют собой кадры видеозаписи стыковки, один и тот же объект присутствует на ней в существенно различных масштабах. В существующем детекторе данная проблема решена путем ввода так называемых дистанций, отрезков расстояния от КА до МКС в пределах каждого из которых обнаруживается только часть особенностей, которая хорошо видна в данный момент. Это связано с тем, что с ростом числа особенностей время работы существующего детектора увеличивается.

### **Описание экспериментов**

В процессе исследования возможности использования нейронной сети для обнаружения объектов в рамках космической стыковки было выполнено обучение нейронной сети для обнаружения особенностей для всех 4 узлов. Выполнена проверка на малой тестовой выборке, которая представляет собой 5% случайнным образом изъятых из обучающей выборки изображений. Проверка подтвердила, что обучение сети завершено успешно, так как большинство объектов на изображениях из тестовой выборки распознаны верно, подтверждение чему приведено в таблице 2.

Для сравнения результатов работы с существующим детектором была выбрана не применявшаяся при обучении детектора стыковочная последовательность изображений для узла №2. Данная последовательность была размечена, на ней были указаны верные положения для всех ориентиров, видимых большую часть времени стыковки, объекты, видимые только в начале или конце стыковки, не рассматривались.

Выбранная последовательность была обработана с использованием обученного детектора на основе нейронной сети, а также с использованием софта, включающего

текущий детектор. Результаты обнаружения объектов нейронной сетью и существующим детектором сравнивались с разметкой. Так как нейронная сеть ищет все возможные объекты на каждом кадре, сравнение производилось только по кадрам, соответствующим дистанциям и по объектам, которые обнаруживаются на этих дистанциях каскадный детектор.

Особенностью каскадного детектора является точное определение не ограничивающей объект рамки, но координат центра объекта, которые и являются выходом детектора. Поэтому вместо стандартной метрики IoU, используемой при определении правильности обнаружения объектов в случае известных правильных рамок, применяется мера пиксельной разности, когда объект отбрасывается при превышении этой разностью некоторого порога.

Параметры обучающей выборки

Таблица 1.

Параметры	Узел 1	Узел 2	Узел 3	Узел 4
Число объектов	20	33	38	51
Размер обучающей выборки	2307	3443	5070	5146
Размер тестовой выборки	122	182	267	271

Результаты обнаружения объектов

Таблица 2.

Параметры	Узел 1	Узел 2	Узел 3	Узел 4
Размер тестовой выборки	548	1890	2724	2806
Число верных обнаружений	452	1333	1643	1967
Верные обнаружения, %	82,5	70,5	60,3	70,1

Результаты работы для каждого из типов объектов и по всем сразу приведены в таблице 3.

Исходя из результатов, приведенных в таблице, можно сделать вывод о том, что классификатор на основе нейронной сети работает лучше. Полнота обоих алгоритмов приблизительно совпадает, то есть они обнаруживают один и тот же процент объектов. При этом точность классификатора на основе нейронной сети значительно выше. Кроме того, время работы каскадного детектора колеблется от 0.1 до 0.45 секунды на кадр, время работы нейронной сети (0.108 с) фиксировано и зависит только от разрешения изображения.

Результаты работы детекторов

Таблица 3.

Дистанции	Нейронная сеть		Каскадный детектор	
	Точность	Полнота	Точность	Полнота
110-75 м.	0.83	0.46	0.52	0.42
75-62 м.	0.84	0.54	0.65	0.54
62-52 м.	0.69	0.46	0.51	0.56
52-45 м.	0.75	0.48	0.42	0.46
45-30 м.	0.66	0.59	0.42	0.43
30-23 м.	0.61	0.56	0.46	0.58
23-19 м.	0.64	0.76	0.24	0.57

## Заключение

В работе рассмотрена возможность улучшения качества детектирования объектов в процессе процедуры космической стыковки с использованием принципиально нового детектора на основе нейронной сети.

Выполнено сравнение качества обнаружения объектов существующим детектором и детектором на основе нейронной сети. Исследована работа данного детектора на кадрах из видеозаписей стыковок по всем четырем узлам МКС, в результате чего получены результаты, отражающие качество обнаружения объектов нейронной сетью выбранной структуры.

Произведено сравнение детектора на основе нейронной сети с имеющимся детектором. Нейронная сеть обладает более высоким быстродействием, для нее отсутствует необходимость распределения по дистанциям, качество обнаружения объектов по некоторым показателям сопоставимо, по некоторым превышает качество работы каскадного детектора. Кроме того, нейронная сеть имеет резервы по качеству обнаружения и по быстродействию, такие как искусственное расширение выборки путем добавления поворотов и перспективных искажений изображений, обучение на одном изображении со смещенной рамкой, перекрестное обучение между узлами, учет геометрии станции и подача на вход детектора части изображения, содержащей нужную часть, искусственное понижение разрешения входных изображений и т.д.

### Библиографический список

1. D. Stepanov, A. Bakhshiev, D. Gromoshinskii, N. Krpan, and F. Gundelakh, “Determination of the relative position of space vehicles by detection and tracking of natural visual features with the existing TV-cameras,” Analysis of Images, Social networks and Texts, Four International Conference, AIST 2015, Yekaterinburg, Russia, April 9-11, 2015, Revised Selected papers. Communications in Computer and Information Science, vol. 542, pp. 431-442.
2. A.V. Bakhshiev, P.A. Korban, N.A. Kirpan, “Software package for determining the spatial orientation of objects by tv picture in the problem space docking,” Robotics and Technical Cybernetics, Saint-Petersburg, Russia, RTC, vol 1, pp. 71-75, 2013.
3. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” Neural Inormation Processing Systems (NIPS), 2015.
4. Ross Girshick, “Fast R-CNN,” Proceedings of the International Conference on Computer Vision (ICCV), 2015.
5. Ross Girshick, Jeff Donahue, Trevor Darrel, and Jitendra Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
6. J.R.R. Uijlings, K.E.A. van de Sande, T. Gevers, and A.W.M. Smeulders, “Selective search for object recognition,” International Journal of Computer Vision, 2013.
7. C. Lawrence Zitnick and Piotr Dollar, “Edge boxes: Locating object proposals from edges,” European Conference on Computer Vision (ECCV), September 2014.
8. Ross Girshick, “Faster R-CNN (Python implementation),” <https://github.com/rbgirshick/py-faster-rcnn>

## VISUAL LANDMARK DETECTION IN DOCKING TASKS USING DEEP NEURAL NETWORK

**I.S. Fomin, D.A. Gromoshinskii, A.V. Bakhshiev**

The Russian State Scientific Center for Robotics and Technical Cybernetics (RTC)

*Automatic and automated docking of controllable apparatus is very hard and actual problem of technics and robotics. This problem can be solved with technical vision if one or both of apparatus are equipped with video camera that able to observe visual landmarks with known relative positions. The most important step is detection of the landmarks in frames received from camera. In example, visual landmarks detection in images, received in the process of space docking using neural network-based object detection system will be discussed.*

**Docking, controllable apparatus, object detection, image processing, neural networks, deep learning**

**Л. И. Ивановский, В. В. Хрящев, Д. Е. Храбров**  
Ярославский государственный университет им. П.Г. Демидова

## Распознавание эмоций по изображению лица человека на основе сверточных нейронных сетей

*В данной работе рассмотрены алгоритмы детектирования улыбок и классификации эмоций. В их основе лежит реализация архитектуры свёрточной нейронной сети. В работе также представлены численные результаты качества работы алгоритмов.*

**Машинное обучение, сверточная нейронная сеть, детектирование эмоций, классификация эмоций.**

В настоящее время большой интерес исследователей в области машинного обучения связан с разработкой систем автоматического обнаружения людей. Современные алгоритмы в основном ориентированы на определение возраста или пола человека по изображению его лица, полученному с камеры наблюдения. Актуальной в этой сфере остается проблема распознавания эмоций человека.

На практике данная задача представляет наибольший интерес в области видеоаналитики для сбора статистики в масштабах торгового центра или части города, для анализа действия рекламы и для оценки работы персонала при общении с клиентом. Сложность задачи определяется разнообразием лиц, наличием оптических препятствий, а также различной мимикой человека.

Для решения поставленной задачи был разработан алгоритм, в основе которого лежит реализация свёрточной нейронной сети – специальной архитектуры, нацеленной на быстрое распознавание различных объектов на изображениях и их эффективную классификацию [1, 2].

Реализация архитектуры свёрточной нейронной сети осуществлялась с помощью фреймворка Caffe [3] на основе сети предложенной в [4]. Как показано на рис. 1, разработанная сеть состоит из 4 свёрточных слоёв, 4 слоёв, реализующих процесс локальной нормализации, а также из 3 слоёв, описывающих процесс дискретизации изображения с помощью операции max pooling.

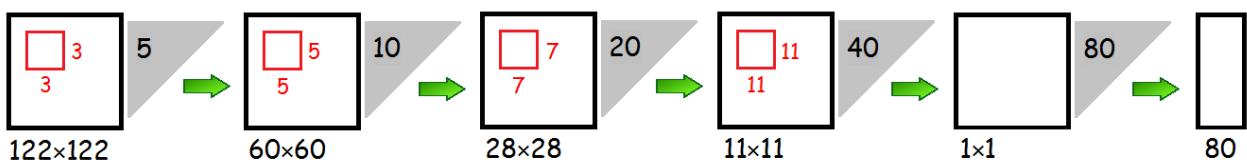


Рис. 1. Архитектура свёрточной нейронной сети

На вход нейронной сети подаются патчи, размером  $122 \times 122 \times 5$ , случайным образом вырезанные из полутоновых изображений разрешения  $128 \times 128$  пикселей. На первом свёрточном слое к патчам применяется 10 ядер размера  $3 \times 3 \times 5$  с шагом пикселя 1. После локальной нормализации и операции max pooling с фильтром  $2 \times 2$  и шагом пикселя 2, признаки будут иметь размерность  $11 \times 11 \times 10$ . Аналогичные операции повторяются на втором (20 ядер размера  $5 \times 5 \times 10$ ), третьем (40 ядер размера  $7 \times 7 \times 20$ ) и четвёртом (80 ядер размера  $11 \times 11 \times 40$ ) слоях нейронной сети. После всех проделанных операций, получается полно связанный слой из 80 нейронов. И в конце сеть соотносит исходную картинку с каким-нибудь классом в соответствии со значениями посчитанных признаков. Другими словами, нейронная сеть классифицирует входное изображение в соответствии с обнаруженной на

лице человека эмоцией. Для ускорения работы сети процессы обучения и тестирования свёрточной нейронной сети осуществлялись параллельно, на графическом процессоре видеокарты.

Эксперимент по оценке точности алгоритма проводился на стандартной базе изображений MultiPIE [5] (рис. 2).



Рис. 2. Примеры изображений из базы данных картинок MultiPIE

База изображений MultiPIE содержит более 750000 изображений 337 людей. Изображения представляют собой снимки, сделанные на фотокамеру, под углом обзора, не превышающим 90 градусов, и с разной степенью освещенности. В базе данных MultiPIE содержатся изображения людей с 6 различными типами эмоций: спокойствие, улыбку, удивление, заинтересованность (прищуривание), отвращение и усталость (зевоту) (рис 3).

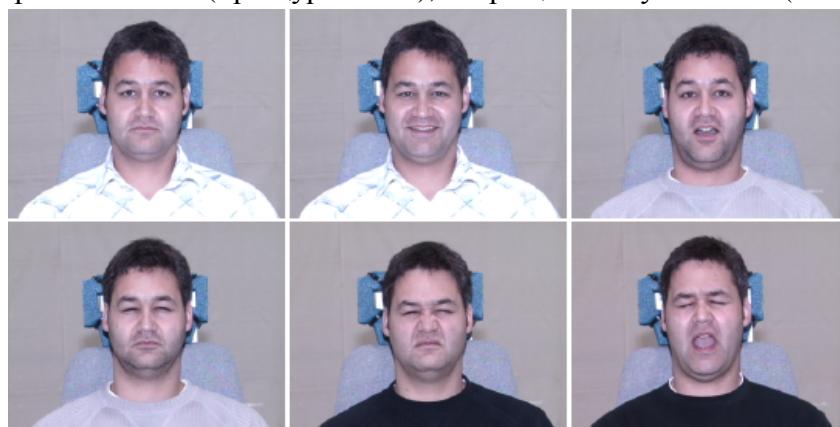


Рис. 3. Типы эмоций, представленными в базе данных MultiPIE

В ходе эксперимента бралась размеченная на 6 классов выборка из 20000 случайным образом отобранных снимков, сделанных в различных условиях освещенности, с углом обзора камеры не превышающим 30 градусов. Это множество разбивалось тренировочный и тестовый наборы данных в соотношении 80% и 20%. Тренировочный и тестовый набор не содержали одинаковых изображений. Результаты тестирования свёрточной нейронной сети приведены в табл. 1.

Итоговые результаты тестирования свёрточной нейронной сети

Таблица 1.

Задача	Доля правильных ответов
--------	-------------------------

Распознавание эмоций	90.23%
Детектирование улыбок	95.42%

Эта же сеть была протестирована и для задачи детектирования улыбки. Стоит отметить, что эксперимент был смоделирован точно так же, как и в задаче распознавания эмоций, за исключением того, что в исходной выборке картинки были размечены на 2 класса: улыбки и неулыбки. Полученные результаты тестирования также приведены в табл. 1.

В заключение необходимо отметить, что предложенный алгоритм классификации является простым в реализации, а также позволяет эффективно детектировать улыбки и распознавать эмоции на изображениях лиц людей.

#### Библиографический список

1. Goodfellow I., Bengio Y., Courville A. Deep Learning // MIT Press, 2016.
2. Mazakazu M., Katsuhiko M., Yusuke M., Yuji K. Subject independent facial expression recognition with robust face detection using a convolutional neural network // Neural Networks, 2003, №16 (5). P. 555–559.
3. Caffe Framework. <http://caffe.berkeleyvision.org>
4. Niu Z., Zhou M., Wang L., Gao X., Hua G. Ordinal Regression with Multiple Output CNN for Age Estimation // IEEE CVPR, 2016.
5. The CMU Multi-PIE Face Database. <http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/MultiPie/Home.html>.
6. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных // М.: ДМК Пресс, 2015.

### **EMOTIONS RECOGNITION ON FACIAL IMAGES BY MEANS OF CONVOLUTIONAL NEURAL NETWORK**

**L. Ivanovsky, V. Khryashchev, D.Khrabrov**  
P.G. Demidov Yaroslavl State University

*Algorithms for smile detection and emotions recognition is considered in this work. They are based on the architecture of convolutional neural network. Also in this article for these algorithms there are given numerical results of the quality of their work.*

**Machine learning, convolutional neural network, smile detection, emotions recognition.**

**А. А. Лебедев, В. В. Хрящев, В. А. Павлов**  
ЯрГУ им. П.Г. Демидова

## Разработка алгоритма биометрической идентификации по изображению лица на основе сверточных нейронных сетей

*Рассматривается алгоритм биометрической идентификации по изображению лица с использованием глубокой сверточной нейронной сети. Представлены численные результаты исследования алгоритма на унифицированной тестовой базе LFW.*

**Машинное обучение, сверточная нейронная сеть, распознавание лиц, биометрическая идентификация**

На современном этапе развития науки и техники, наибольший интерес в областях машинного обучения и компьютерного зрения представляют интеллектуальные системы анализа аудитории, позволяющие идентифицировать личность человека по изображению лица. Областью применения таких систем могут являться:

- системы обеспечения безопасности в местах массового скопления людей: вокзалы, стадионы;
- системы анализа посетителей в магазинах и торговых центрах (собственные базы постоянных клиентов);
- системы верификации по предъявленному фото в паспорте (например, при получении быстрого кредита в терминале самообслуживания банка);
- вендинг и встроенные системы;
- системы «умный дом»;
- системы взаимодействия людей и роботов;
- системы аутентификации в мобильных устройствах.

В связи с многообразием практических приложений алгоритмов биометрической идентификации, следует учитывать факторы негативно влияющие на их точность работы: условия естественного неравномерного освещения (вокзалы, аэропорты, промышленные объекты, торговые и деловые центры), деградация уровня распознавания по прошествии интервала времени от полугода до года между моментом помещения изображения лица в базу и моментом идентификации, дополнительная вариативность, вносимая изменением угла обзора лица в кадре.

Таким образом, задача разработки устойчивого к освещению и возрастным изменениям алгоритма распознавания лиц на видеоизображениях является актуальной для систем биометрической идентификации по видеоданным.

Для решения данной проблемы был использован алгоритм, основанный на использовании глубокой сверточной нейронной сети, показывающий высокую степень рабочести к выше перечисленным факторам [1, 2].

На первом этапе алгоритма входное изображение приводится к нормальному виду – масштабируется до размера 250×250 пикселей и центрируется. Затем из полученного изображения вырезаются 5 фрагментов размером 224×224 пикселей – с каждого из четырех углов и из центра изображения, как показано на рис. 1. После этого создается зеркальная копия входного изображения, для которого описанная ранее процедура

получения фрагментов изображения повторяется. На выходе последнего сверточного слоя вектора признаков для каждого фрагмента изображения усредняются, образуя дескриптор изображения. Полученный дескриптор используется для классификации входного изображения. Сверточная нейронная сеть имеет 13 сверточных слоев с функцией активации ReLU [3], два полно связных слоя так же с функцией активации ReLU и выходной полно связный слой с сигмоидальной функцией активации. Для оценки точности работы предложенного алгоритма использовалась тестовая база изображений LFW [4](рис. 2).



Рис. 1. Формирование фрагментов изображения



Рис. 2. Примеры изображений из тестовой базы изображений LFW

Тестовая база изображений LFW содержит 13233 изображения 5749 различных людей. 1680 людей имеют два или более изображения, оставшиеся 4096 людей представлены единственными изображениями. Изображения имеют разрешение  $250 \times 250$  пикселей, большинство из них цветные, несколько изображений полутоновые. Данная база является популярной для тестирования алгоритмов распознавания лиц, благодаря следующим особенностям: все изображения получены путем применения свободной реализации алгоритма детектирования лиц Виолы-Джонса [5] из открытой библиотеки OpenCV, результаты работы которого проходили ручную фильтрацию ложноположительных срабатываний детектора, а также и приводились к размеру  $250 \times 250$ , данная база изображений имеет большую вариативность по углам поворота лиц в кадрах,

освещению, а также по расовым, гендерным и возрастным признакам людей в базе данных, все снимки сделаны в естественных условиях.

Для проведения теста верификации использовался стандартный протокол тестирования, определенный авторами базы LFW в [4] и дополненный в [6]. Кроме того, проведен более сложный тест – идентификация с открытым набором, по протоколу, предложенному в [7]. В обоих тестах использовались 10 наборов, сформированных с помощью алгоритма перекрестной проверки из базы LFW. Алгоритм выравнивания лиц не рассматривается в данной работе, для оценки влияния применения алгоритма выравнивания на конечное качество распознавания использовалась модификация базы LFW – LFWA, полученная путем применения коммерческого алгоритма выравнивания лиц на оригинальной базе LFW. В качестве метрик оценки качества алгоритма распознавания используются ROC-кривая и площадь под ROC-кривой (AUC) [8]. Так как тесты проводились с применением перекрестной проверки, в качестве конечного результата приводится кумулятивная ROC-кривая, получаемая при помощи усреднения выходных результатов с каждого набора, сформированного перекрестной проверкой (рис. 3) [9]. Результаты тестирования рассматриваемого алгоритма приведены в табл. 1.

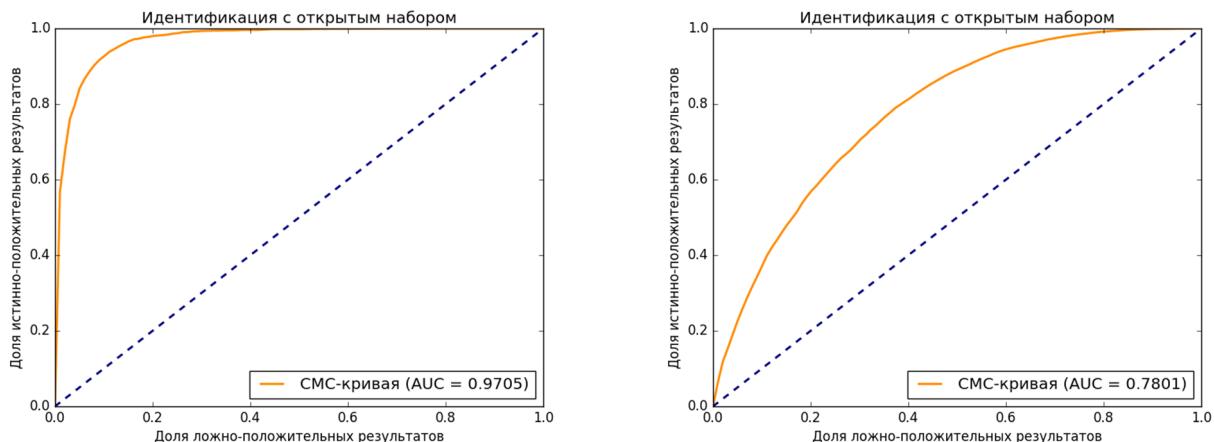


Рис. 3. Кумулятивные ROC-кривые тестируемого алгоритма

Итоговые результаты тестирования свёрточной нейронной сети

Таблица 1.

Тест	Без дополнительной обработки	Дополнительная обработка - выравнивание	Дополнительная обработка - патчи	Дополнительная обработка - выравнивание и патчи
Верификация	0,6325	0,7023	0,7548	0,7801
Идентификация с открытым набором	0,9169	0,9462	0,9647	0,9705

Применение глубоких свёрточных сетей позволяет получать робастные, точные алгоритмы биометрической идентификации по изображению лица, однако данный подход является довольно сложным в реализации и имеет высокую вычислительную сложность, для обучения данного алгоритма требуется применение высокопроизводительных графических процессоров и/или вычислительных кластеров. Работа выполнена при поддержке гранта РФФИ № 15-07-08674.

### Библиографический список

1. Datta A.K., Datta M., and Banerjee P.K. Face Detection and Recognition Theory and Practice // CRC Press, 2016.
2. Лукьяница А.А. Цифровая обработка изображений / Лукьяница А.А., Шишкун А.Г. – М.: Ай-Эс-Эс Пресс, 2009. - 518 с.
3. Nair V., and Hinton G. Rectified linear units improve restricted Boltzmann machines // In Proc. ICML, 2010.
4. Huang G.B., Ramesh M., Berg T., and Learned-Miller E. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments // University of Massachusetts, Amherst, Technical Report 07-49, October, 2007.
5. Jones M., and Viola P. Rapid object detection using a boosted cascade of simple features // Proc. International Conference on Computer Vision and Pattern Recognition, № 1, P. 511-518. 2001.
6. Huang G.B., and Learned-Miller E. Labeled Faces in the Wild: Updates and New Reporting Procedures // University of Massachusetts, Amherst, Technical Report UM-CS-2014-003, May, 2014.
7. Liao S., Lei Z., Yi D., and Li S.Z., A Benchmark Study of Large-scale Unconstrained Face Recognition // In IAPR/IEEE International Joint Conference on Biometrics, Clearwater, Florida, USA, 2014.
8. Fawcett T. An introduction to ROC analysis // Pattern Recognition Letters, V. 27, № 8, P. 861–874. 2006.
9. Phillips P.J., Grother P., and Micheals R. Evaluation methods in face recognition // In S. Z. Li and A. K. Jain, editors, Handbook of Face Recognition, chapter 21, P. 551–574. Springer, 2nd edition, Sep. 2011.

### **DEVELOPMENT OF THE BIOMETRIC IDENTIFICATION ALGORITHM FOR THE FACE IMAGE BASED ON CONVOLUTIONAL NEURAL NETWORKS**

**A. Lebedev, V. Khryashchev, V. Pavlov**  
P.G. Demidov Yaroslavl State University

*In this paper, we consider the algorithm for biometric identification by a face image using a deep convolutional neural network. Numerical results of testing our approach on LFW image database are presented.*

**Machine learning, convolutional neural network, face recognition, biometric identification.**

**М. В. Кипке<sup>1</sup>, Н. Н. Гаврютин<sup>1</sup>, Т. М. Сухов<sup>1</sup>, А. А. Сорокин<sup>1</sup>,  
Ю. А. Балошин<sup>2</sup>, М. А. Сухова<sup>3</sup>, А. В. Арсеньев<sup>3</sup>, С. В. Василевич<sup>3</sup>**

<sup>1</sup>БГТУ «ВОЕНМЕХ» им. Д.Ф. Устинова

<sup>2</sup>Университет «ИТМО»

<sup>3</sup>ООО «Смарт-Орто»

## Выделение контура тела человека по цвету кожи на бытовом фоне в фотоизображении

*Работа посвящена рассмотрению методов обработки изображений с целью выделения контуров тела человека на бытовом фоне. Задача актуальна для получения информации об антропометрических характеристиках человека с последующим анализом. Практическое применение в качестве диагностики возможно в медицинской практике (ортопедия, спортивная медицина, мануальная терапия, кинезотерапия) и спорте. В работе рассмотрены методы обработки изображений на основе цвета кожных покровов человека. Предложена метрика, на основе которой определяется местоположение человека на изображении.*

### Обработка изображений, цвет кожи, метод Отцу, гистограмма изображения

Детектирование на основе цвета кожи человека широко используется для распознавания лиц и человеко-машинного взаимодействия, являясь одним из наиболее эффективных подходов для поиска и выделения лица человека на изображениях. Известно, что для представителей различных рас составляющие цветового тона кожных покровов изменяются незначительно [1]. Выделение контура тела человека по цвету кожи – распространенная задача в медицине [2, 3].

Характерным признаком того, что область на изображении соответствует отображению человеческой кожи, является доминирование интенсивности красной составляющей аддитивного смешения цветов модели цветового пространства RGB. На рис. 1 приведены интенсивности красной (а), зеленой (б) и синей (в) составляющей изображения в целом.

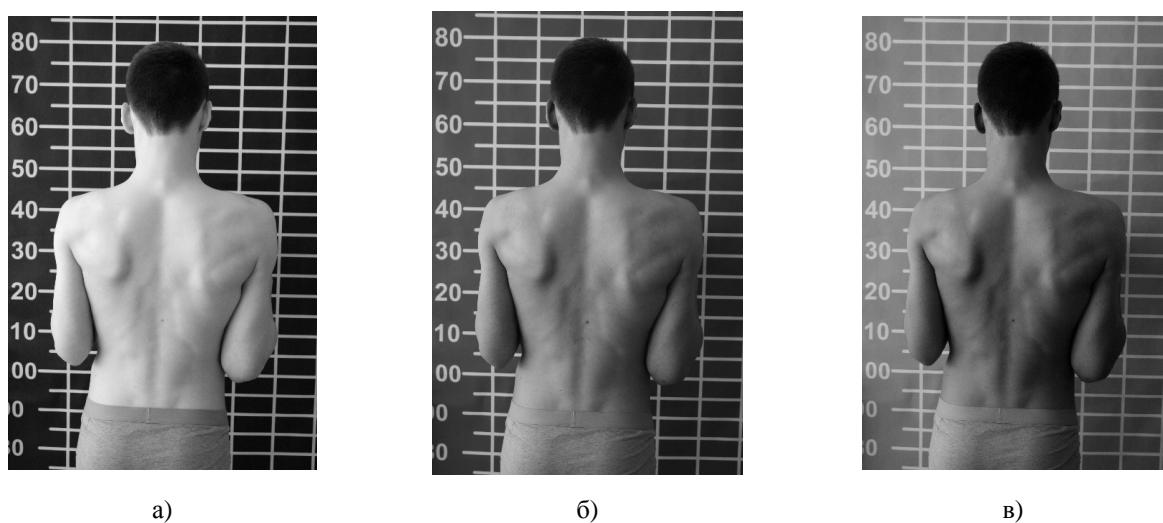


Рис. 1. RGB составляющие изображения

На рис. 2 приведены распределения интенсивностей красной, зеленой и синей составляющей в одной из строк изображения. Цвет кривых на рис. 2 соответствует составляющей RGB.

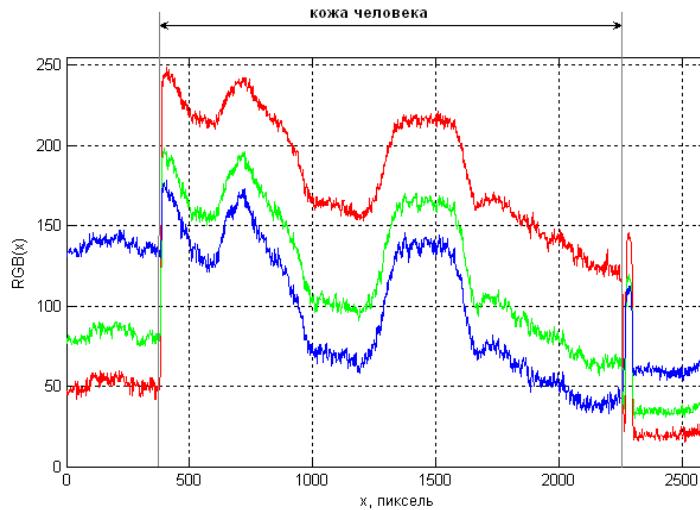


Рис. 2. Интенсивности красной, зеленой и синей составляющей в строке изображения

Используя все три цветовых составляющих, получим нормированную характеристику учитывающую доминирование красной составляющей, как информативного признака в виде:

$$I(x, y) = \frac{(R(x, y) - G(x, y) + 2^n - 1)(R(x, y) - B(x, y) + 2^n - 1)}{(2^{n+1} - 2)^2} \quad (1)$$

где  $n$  – количество бит для кодирования каждой из цветовых составляющих палитры (учитывает исключительно случай, когда для кодирования  $RGB$  составляющих используется одинаковое количество бит).

Для палитры с глубиной цвета 24 бита (8 бит на составляющую  $RGB$ ), формула (1) примет вид

$$I(x, y) = \frac{(R(x, y) - G(x, y) + 255)(R(x, y) - B(x, y) + 255)}{260100}.$$

Значения характеристики  $I(x, y)$ , рассчитанные для основных цветов, приведены в таблице 1. Максимум нормированной характеристики соответствует окрашиванию пикселя в чистый красный цвет ( $RGB(x_0, y_0) = (255; 0; 0)$ ).

Значения характеристики  $I(x)$  для основных цветов

Таблица 1.

Цвет	$R(x_0, y_0)$	$G(x_0, y_0)$	$B(x_0, y_0)$	$I(x_0, y_0)$
Черный	0	0	0	0,25
Красный	255	0	0	1
Зеленый	0	255	0	0
Синий	0	0	255	0
Желтый	255	255	0	0,5
Фиолетовый	255	0	255	0,5
Голубой	0	255	255	0
Белый	255	255	255	0,25

На рис. 3, а приведена характеристика (1) рассчитанная по всем пикселям исходного изображения. Учитывая то, что появление черного пикселя ( $I(x_0, y_0) = 0,25$ ) означает

отсутствие в данном месте объекта поиска, либо дефект изображения (что в данном случае равнозначно), отмасштабируем динамический диапазон характеристики (рис. 3, б)

$$I_{norm}(x, y) = \frac{I(x, y) - \min(I(x, y))}{\max(I(x, y)) - \min(I(x, y))}, \quad \min(I(x, y)) > 0.25 \quad (2)$$

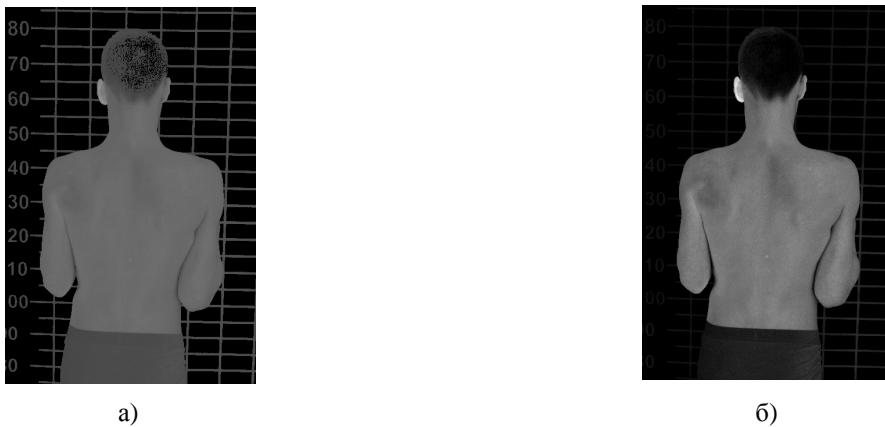


Рис 3. Результат обработки исходного изображения: а –  $I(x, y)$ ; б –  $I_{norm}(x, y)$

В качестве метода выделения предметной области возможно применение метода бинаризации с глобальным аддитивным порогом - метода Оцу [4]. Алгоритм вычисляет порог бинаризации для полутонового изображения (которым, по сути является  $I_{norm}(x, y)$ ), и позволяет разделить пиксели двух классов («полезные» и «фоновые»), рассчитывая такой порог, так, чтобы внутриклассовая дисперсия была минимальной. Алгоритм использует для расчетов гистограмму интенсивностей пикселей. На рис. 4 представлена гистограмма интенсивностей пикселей изображения

$$h(i) = \sum_{x=1}^X \sum_{y=1}^Y \mathbf{1}_{\{I_{norm}(x, y) \in (a_{i-1}, a_i]\}}, \quad h(i) = \sum_{x=1}^X \sum_{y=1}^Y \mathbf{1}_{\{I_{norm}(x, y) \in (a_{i-1}, a_i]\}}, \quad i = 1,..256.$$

Как видно из приведенного графика, весомый вклад принадлежит нулевым пикселям (черного цвета) – тем, которые на предыдущем этапе обработки были отнесены к фону. В случае, если искомый объект занимает значительную площадь изображения, бинаризация методом Оцу дает удовлетворительный результат (рис. 5, а), однако в случае, когда размер объекта по площади меньше фона, это может приводить к неудовлетворительным результатам (рис. 5, б).

Для коррекции данного нюанса, необходимо исключить из расчета пиксели заведомо принадлежащие фону, т.е. интенсивность которых равна нулю. Для этого, необходимо обнулить в гистограмме количество фоновых пикселей и домножить каждый элемент гистограммы на весовой коэффициент, таким образом, чтобы ее сумма осталась равной общему количеству пикселей изображения (рис. 6)

$$h_{norm}(i) = \frac{h(i) \cdot X \cdot Y}{\sum_{i=2}^{256} h(i)}.$$

Результаты бинаризации изображений методом Оцу с использованием нормированной гистограммы, представлены на рис. 7.

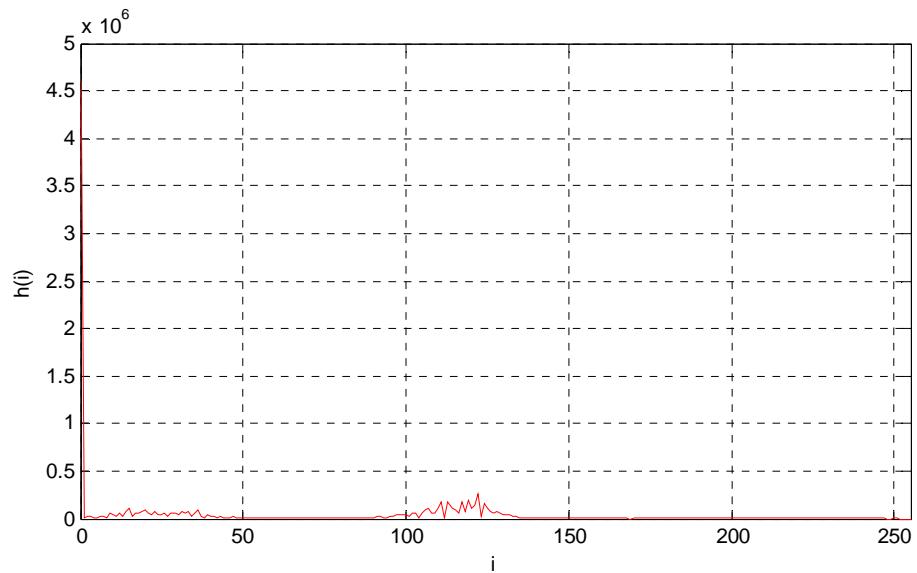


Рис. 4. Гистограмма изображения  $I_{norm}(x, y)$

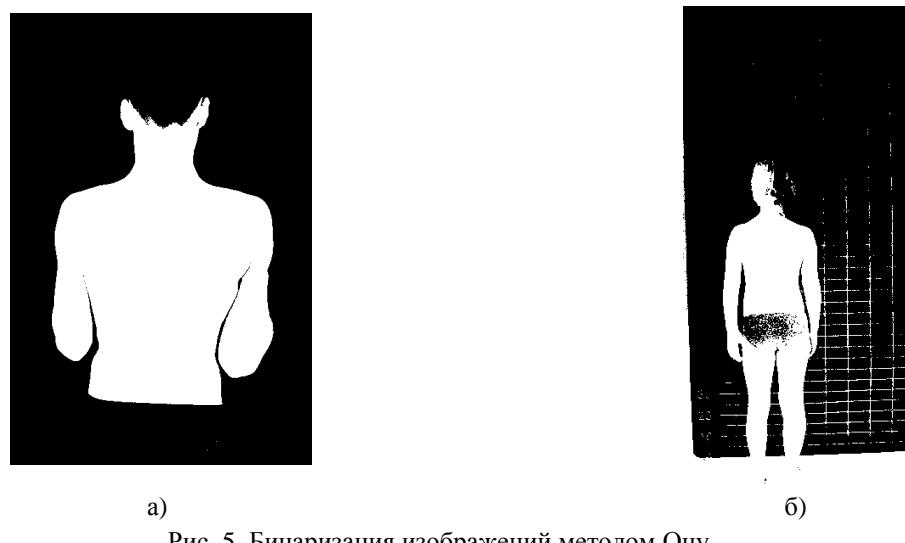


Рис. 5. Бинаризация изображений методом Оцу

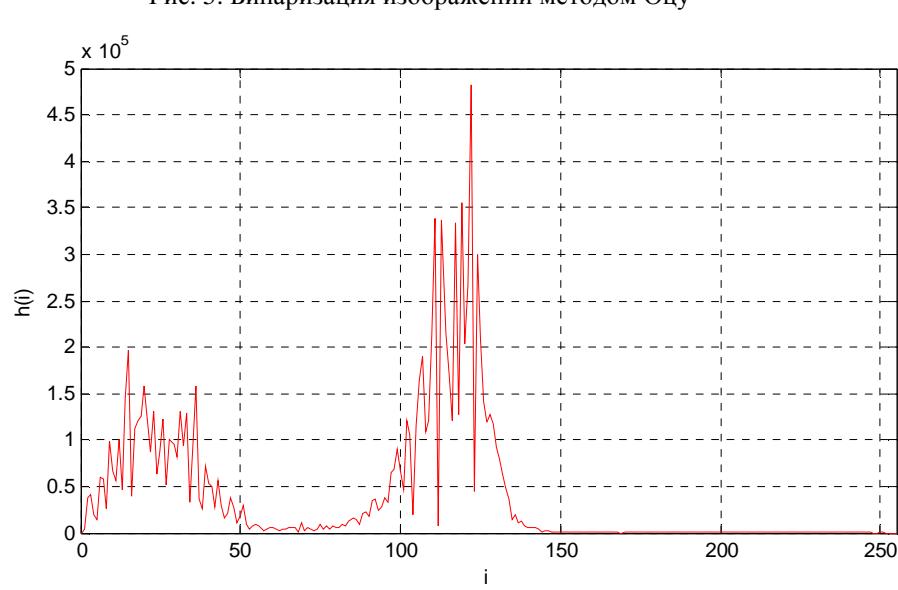


Рис. 6. Нормированная гистограмма



a)

б)

Рис. 7. Бинаризация изображений методом Оцу с использованием нормированной гистограммы

Предложенные методы предварительной обработки изображений позволяют с большой точностью выделять контуры человеческого тела на фоне повседневной (бытовой) обстановки и являются основой для разработки доступных методов диагностики с возможностью широкого применения.

#### Библиографический список

1. Хомяков М.Ю. Классификация цвета кожи человека на цветных изображениях / Компьютерная оптика. – том 35, №3 – 2011. – С. 373–379.
2. Hardware-software method for postural abnormalities and spine deformities diagnosis based on digital photo analysis. Arsenev A., Baloshin Y., Vasilevich S., Dudin M., Kipke M., Sorokin A., Sukhov T., Sukhova M. J.LOCOMOTOR SYSTEM: Advances in Research, Diagnostics and Therapy. ISSN 2336-4777, ročník 23, 2016, číslo 1+2, P.73–76.
3. Арсеньев А.В., Балошин Ю.А., Васильевич С.В., Дудин М.Г., Кипке М.В., Сорокин А.А., Сухов Т.М., Сухова М.А. Новые возможности в диагностике пациентов с нарушением осанки и деформациями позвоночного столба // Сборник трудов VII научно-практической конференции с международным участием «Актуальные вопросы реабилитации», часть II. СПб.: СПбГПИМУ, 2016. С.50–52.
4. N. Otsu. A threshold selection method from gray-level histograms // IEEE transactions on systems, man, and cybernetics, vol. SMC-9, NO. 1, January 1979. – pp. 62–66.

#### ISOLATION OF THE CONTOUR OF THE HUMAN BODY BY THE COLOR OF THE SKIN ON A HOUSEHOLD BACKGROUND IN A PHOTO IMAGE

<sup>1</sup>**Kipke M.V., <sup>1</sup>Gavryutin N.N., <sup>1</sup>Sukhov T.M., <sup>1</sup>Sorokin A.A., <sup>2</sup>Baloshin Yu.A., <sup>3</sup>Sukhova M.A.,  
<sup>3</sup>Arsenev A.V., <sup>3</sup>Vasilevich S.V.**

<sup>1</sup>Baltic State Technical University «VOENMEH» named after D.F. Ustinov

<sup>2</sup>ITMO University

<sup>3</sup>LLC “Smart-Ortho”

*The article is devoted to the consideration of image processing methods with the purpose of isolating the contours of the human body on a household background. The task is relevant for obtaining information about human anthropometric characteristics with subsequent analysis. Practical application as a diagnostic is possible in medical practice (orthopedics, sports medicine, manual therapy, kinesitherapy) and sports. In the article methods of image processing on the basis of color of human integuments are considered. A metric is proposed, on the basis of which the position of the person on the image is determined.*

**Image processing, skin color, Otsu method, image histogram**

**Е. П. Петров, Н. Л. Харина, Е. Д. Ржаникова**  
Вятский государственный университет, г. Киров

## Метод сокращения времени передачи видео с борта микро беспилотного летательного аппарата

*В работе предлагается метод сокращения времени передачи динамических многоразрядных цифровых изображений (МЦИ) с борта микро БПЛА (МБПЛА). Метод включает две части. Первая часть - преобразование динамических МЦИ без искажений и вычислительных операций на борту МБПЛА для передачи по радиоканалу многофазными (МФМ-4,8,16) сигналами, сокращающими время передачи преобразованных динамических МЦИ. Вторая часть – реализация статистической избыточности оригиналов динамических МЦИ многомерной нелинейной фильтрацией МФМ сигналов для компенсации потерь помехоустойчивости, вызванных их применением*

**Беспилотный летательный аппарат, динамические цифровые изображения, статистическая избыточность, помехоустойчивость приема видео, многомерная цепь Маркова**

Авиационные беспилотные летательные аппараты (БПЛА) в настоящее время получили широкое распространение во многих приложениях гражданского и военного назначения.

В связи с успехами нанотехнологий и миниатюризации сенсоров классификация БПЛА расширилась в область разработки малоразмерных БПЛА, таких как микро и нано БПЛА. В системах дистанционного видеонаблюдения большое внимание уделяется методам передачи по радиоканалу динамических многоразрядных (более 8 разрядов на цветовую компоненту) цифровых изображений (МЦИ) высокого пространственного разрешения с микро БПЛА (МБПЛА) за минимально короткое время без вычислительных операций и искажений оригиналов МЦИ на борту МБПЛА, имеющих ограниченные энергетические ресурсы.

Поиск методов сокращения времени передачи динамических МЦИ высокого разрешения привел к разработке методов сжатия динамических МЦИ без вычислительных операций, не критичных к разрядности динамических МЦИ, превосходящих известные методы сжатия видео, такие как MPEG2, MPEG4. Перечисленные кодеки требуют для реализации больших временных и энергетических ресурсов, поэтому не могут быть реализованы на борту МБПЛА. Отсюда следует, что разработка новых и совершенствование известных методов сокращения времени передачи оригиналов динамических МЦИ без искажений и вычислительных операций на борту МБПЛА является актуальной.

В работе предлагается один из возможных вариантов решения задачи сокращения времени передачи динамических МЦИ с борта МБПЛА, состоящий из двух частей. Первая часть включает преобразование оригиналов динамических МЦИ, состоящих из разрядных двоичных изображений (РДИ) (рис. 1) в МЦИ, состоящие из групп соседних РДИ (ГРДИ), включающих два и более РДИ (рис. 2), которые могут передаваться согласованными с ГРДИ сигналами, например, многофазными (МФМ-4,8,16) или многочастотными (МЧМ). Число ГРДИ меньше, чем РДИ, поэтому время передачи динамических МЦИ сокращается в число раз, равное числу РДИ, объединенных в ГРДИ. Переход в МЦИ от РДИ к ГРДИ осуществляется без искажений и вычислительных операций на борту МБПЛА. Вторая часть, более сложная, заключается в разработке метода реализации статистической

избыточности оригиналов динамических МЦИ многомерной нелинейной фильтрацией ГРДИ для повышения помехоустойчивости приема динамических МЦИ, передаваемых МФМ сигналами.

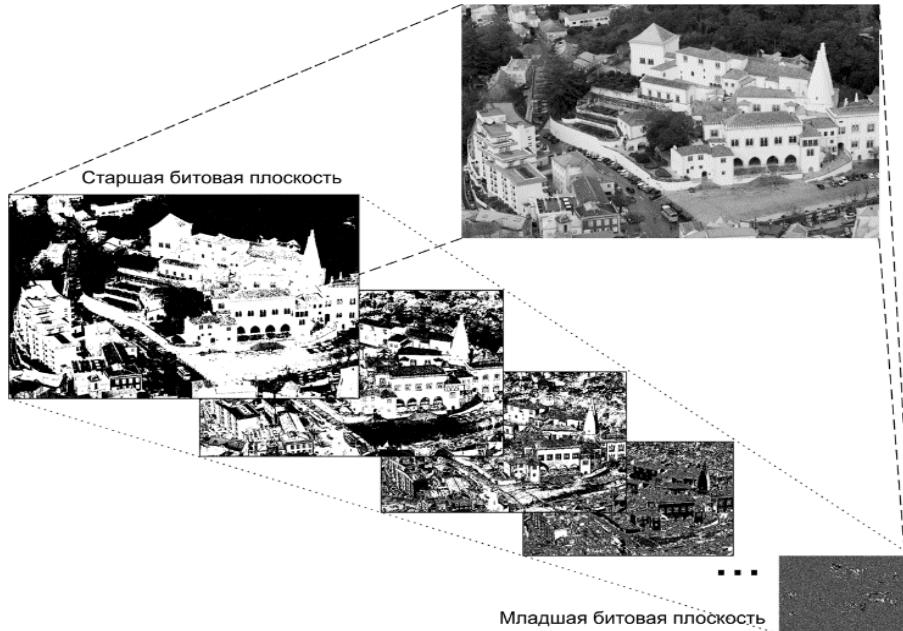


Рис. 1. Пример разделения кадра динамического МЦИ на РДИ

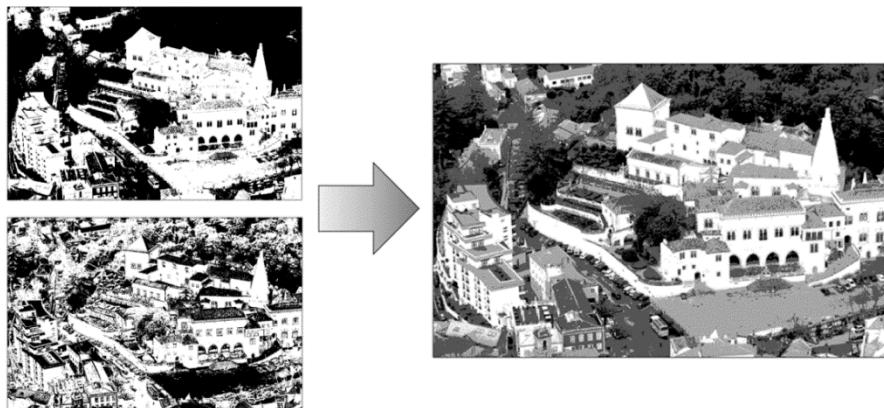


Рис. 2. Объединение двух РДИ в ГРДИ

Будем полагать, что в каждом кадре в ГРДИ объединяется  $g$  РДИ и динамическое МЦИ аппроксимируется трехмерной цепью Маркова с несколькими состояниями  $N = 2^g$

с вектором вероятностей начальных состояний  $P = \{p_1, p_2, \dots, p_N\}^T$  и матрицами вероятностей переходов (МВП)  ${}^1P$ ,  ${}^2P$  и  ${}^4P$  из состояния  $M_i$  в соседнее состояние  $M_j$  ( $i, j \in N$ ) по горизонтали, вертикали и между кадрами, соответственно [1,2]:

$${}^1\Pi = \begin{vmatrix} {}^1p_{11} & {}^1p_{12} & \dots & {}^1p_{1N} \\ {}^1p_{21} & {}^1p_{22} & \dots & {}^1p_{2N} \\ \dots & \dots & \dots & \dots \\ {}^1p_{N1} & {}^1p_{N2} & \dots & {}^1p_{NN} \end{vmatrix}, \quad {}^2\Pi = \begin{vmatrix} {}^2p_{11} & {}^2p_{12} & \dots & {}^2p_{1N} \\ {}^2p_{21} & {}^2p_{22} & \dots & {}^2p_{2N} \\ \dots & \dots & \dots & \dots \\ {}^2p_{N1} & {}^2p_{N2} & \dots & {}^2p_{NN} \end{vmatrix},$$

$${}^4\Pi = \begin{vmatrix} {}^4p_{11} & {}^4p_{12} & \dots & {}^4p_{1N} \\ {}^4p_{21} & {}^4p_{22} & \dots & {}^4p_{2N} \\ \dots & \dots & \dots & \dots \\ {}^4p_{N1} & {}^4p_{N2} & \dots & {}^4p_{NN} \end{vmatrix}. \quad (1)$$

Математическая модель динамического ГРДИ представлена на рис.3.

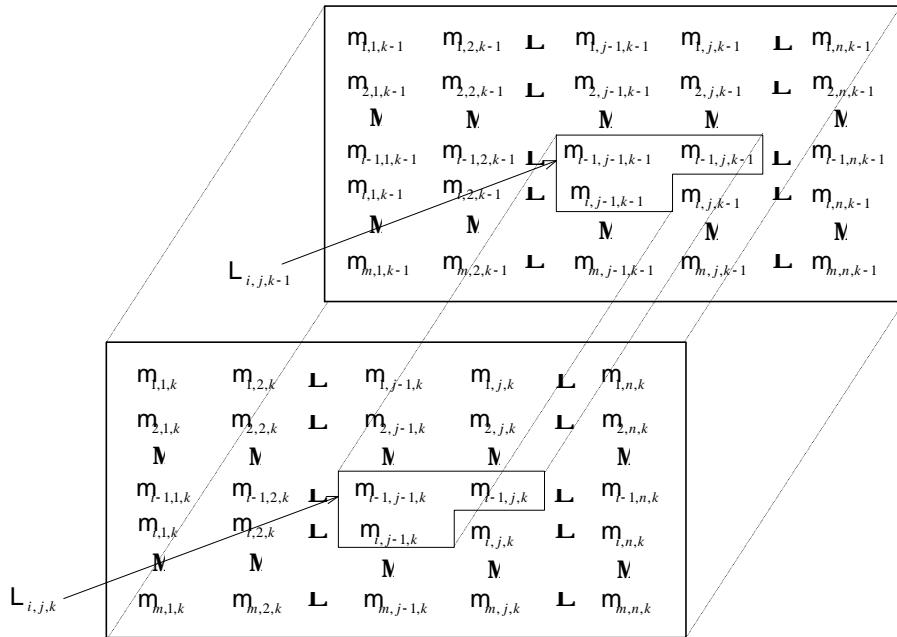


Рис. 3. ММ динамического ГРДИ

Состояние элемента  $v_4$  зависит только от известных элементов некоторого подмножества  $L_{i,j,k}$ , называемого окрестностью элемента. Лучшим образом удовлетворяющее условию каузальности является конфигурация окрестности (рис. 4)

$$L_{i,j,k} = \{n_1; n_2; n_3; n_4; n_{\phi}; n_{\phi}; n_{\phi}\} \quad (2)$$

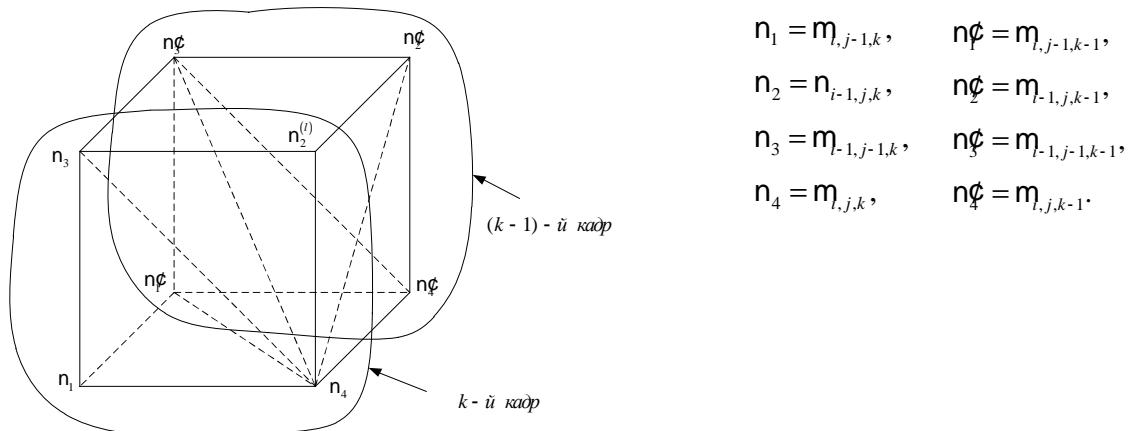


Рис.4. Окрестность пикселя  $N_4$  динамического МЦИ, состоящего из ГРДИ с РДИ

Вероятности перехода от состояний элементов  $L_{i,j,k}$  к состоянию  $N_4$  образуют МВП вида [1,2]:

$$P = \begin{matrix} \hat{e} p_{iiii} & p_{iiji} & L & p_{iil} & L & p_{iiIN} & \dot{u} \\ \hat{e} p_{ijji} & p_{ijji} & L & p_{ijkl} & L & p_{ijjN} & \dot{u} \\ \hat{e} M & M & L & M & L & M & \dot{u} \\ \hat{e} p_{ijkl} & p_{ijkl} & L & p_{ijkl} & L & p_{ijkN} & \dot{u} \\ \hat{e} M & M & L & M & L & M & \dot{u} \\ \hat{e} p_{NNNi} & p_{NNNj} & L & p_{NNNI} & L & p_{NNNN} & \dot{u}_{N^3 \cdot N} \end{matrix} \quad i, j, k, l = 1, N \quad (3)$$

При известных матрицах  ${}^1 P$ ,  ${}^2 P$ ,  ${}^4 P$  для вычисления элементов матрицы  $P$  (3) необходимо предварительно вычислить матрицы

$$\begin{aligned} {}^3 P &= {}^1 P \times {}^2 P; \quad {}^5 P = {}^1 P \times {}^4 P; \quad {}^6 P = {}^2 P \times {}^4 P; \\ {}^7 P &= {}^3 P \times {}^4 P = {}^1 P \times {}^2 P \times {}^4 P. \end{aligned} \quad (4)$$

определяющие статистическую связь элементов окрестности  $L_{i,j,k}$  с элементом  $n_4$ .

Значения элементов матрицы  $P$  (4) могут быть вычислены следующим образом:

$$\begin{aligned} p_{iiii} &= \frac{{}^1 p_{ii} \times {}^2 p_{ii} \times {}^4 p_{ii} \times {}^7 p_{ii}}{{}^3 p_{ii} \times {}^5 p_{ii} \times {}^6 p_{ii}}, \quad p_{iiji} = \frac{{}^1 p_{ij} \times {}^2 p_{ij} \times {}^4 p_{ij} \times {}^7 p_{ij}}{{}^3 p_{ii} \times {}^5 p_{ii} \times {}^6 p_{ii}}, \quad L, \\ p_{iil} &= \frac{{}^1 p_{il} \times {}^2 p_{il} \times {}^4 p_{il} \times {}^7 p_{il}}{{}^3 p_{ii} \times {}^5 p_{ii} \times {}^6 p_{ii}}, \quad L, \quad p_{iiIN} = \frac{{}^1 p_{iN} \times {}^2 p_{iN} \times {}^4 p_{iN} \times {}^7 p_{iN}}{{}^3 p_{ii} \times {}^5 p_{ii} \times {}^6 p_{ii}}, \end{aligned} \quad (5)$$

Приведены выражения для вычисления элементов первой строки матрицы  $P$ . Вычисление остальных элементов матрицы при различных сочетаниях значений пикселей окрестности  $L_{i,j,k}$  осуществляется аналогично.

### Алгоритм многомерной нелинейной фильтрации ГРЦИ из $g$ РДИ

Систему уравнений многомерной нелинейной фильтрации динамического МЦИ, состоящего из ГРДИ с  $g$  РДИ, передаваемых МФМ сигналами, представим в виде [3]:

$$\begin{aligned} u_{1(k+1)}(v_4) &= [f(M_1(v_4)) - f(M_N(v_4))] + u_{1(k)}(v_1) + z_1(u(v_1), {}^1\pi_{ij}) + \\ &+ u_{1(k)}(v_2) + z_1(u(v_2), {}^2\pi_{ij}) + u_{1(k)}(v_3) + z_1(u(v_3), {}^4\pi_{ij}) + u_{1(k)}(v_4) + z_1(u(v_4), {}^7\pi_{ij}) \\ &- u_{1(k)}(v_5) - z_1(u(v_5), {}^3\pi_{ij}) - u_{1(k)}(v_6) - z_1(u(v_6), {}^5\pi_{ij}) - u_{1(k)}(v_7) - z_1(u(v_7), {}^6\pi_{ij}) \\ \\ u_{2(k+1)}(v_4) &= [f(M_2(v_4)) - f(M_N(v_4))] + u_{2(k)}(v_1) + z_2(u(v_1), {}^1\pi_{ij}) + \\ &+ u_{2(k)}(v_2) + z_2(u(v_2), {}^2\pi_{ij}) + u_{2(k)}(v_3) + z_2(u(v_3), {}^4\pi_{ij}) + u_{2(k)}(v_4) + z_2(u(v_4), {}^7\pi_{ij}) \\ &- u_{2(k)}(v_5) - z_2(u(v_5), {}^3\pi_{ij}) - u_{2(k)}(v_6) - z_2(u(v_6), {}^5\pi_{ij}) - u_{2(k)}(v_7) - z_2(u(v_7), {}^6\pi_{ij}) \\ \\ M & \\ u_{(N-1)(k+1)}(v_4) &= [f(M_{(N-1)}(v_4)) - f(M_N(v_4))] + u_{(N-1)(k)}(v_1) + z_{(N-1)}(u(v_1), {}^1\pi_{ij}) + \\ &+ u_{(N-1)(k)}(v_2) + z_{(N-1)}(u(v_2), {}^2\pi_{ij}) + u_{(N-1)(k)}(v_3) + z_{(N-1)}(u(v_3), {}^4\pi_{ij}) + \\ &+ u_{(N-1)(k)}(v_4) + z_{(N-1)}(u(v_4), {}^7\pi_{ij}) - u_{(N-1)(k)}(v_5) - z_{(N-1)}(u(v_5), {}^3\pi_{ij}) - \\ &- u_{(N-1)(k)}(v_6) - z_{(N-1)}(u(v_6), {}^5\pi_{ij}) - u_{(N-1)(k)}(v_7) - z_{(N-1)}(u(v_7), {}^6\pi_{ij}) \end{aligned} \quad (6)$$

где  $u_j(v_4) = \ln \frac{\hat{e} p_j(v_4)}{\hat{e} p_N(v_4)}$  ( $j = \overline{1, N-1}$ );  $p_j(v_4)$  ( $j = \overline{1, N}$ ) – апостериорная вероятность

состояния дискретного параметра МФМ сигнала в пикселе  $n_4$  ММ ГРДИ (рис. 4),

$f(M_j(v_4)) - f(M_N(v_4)) \ (j = \overline{1, N-1})$  – разность логарифмов функции правдоподобия состояний дискретного параметра МФМ сигнала в пикселе  $n_4$  ГРДИ;  $z_j(\lambda)$  – нелинейная функция вида:

$$z_q(u(v_l), l \pi_{ij}) = \ln \frac{\sum_{i=1, i \neq j}^{N-1} \left\{ \exp(u_i(v_l) - u_j(v_l))^l \pi_{ij} \right\} + \exp(-u_j(v_l))^l \pi_{Nj} + \pi_{jj}}{\sum_{i=1}^{N-1} \left\{ \exp(u_i(v_l))^l \pi_{iN} \right\} + \pi_{NN}}, \quad (7)$$

$$(q = \overline{1, N-1}, l = \overline{1, 7}).$$

В качестве критерия различения состояний элементов ЦПИ примем критерий максимума логарифма отношения апостериорных вероятностей  $u_j(n_4)$ , в соответствии с которым, если  $u_j(v_4) > u_i(v_4), i, j = \overline{1, N-1}; i \neq j$ , то принимается решение о состоянии элемента изображения  $n_4 = M_j$ ; если все значения  $u_j(v_4) \leq 0 \ (j = \overline{1, N-1})$ , то принимается решение о состоянии элемента изображения  $v_4 = M_N$ .

### Анализ результатов исследования метода сокращения времени передачи динамических МЦИ

Для демонстрации предложенного метода проведено моделирование процесса передачи и восстановления динамического МЦИ, состоящего из ГРЦИ с  $g=4$  РДИ при отношении сигнал/шум по мощности сигнала на входе нелинейного фильтра 0 дБ. На рис. 5 показан кадр динамического МЦИ, представленный набором из двух ГРЦИ, каждое из которых включает по четыре соседних РДИ (рис. 5а), искаженный (рис. 5б) и восстановленный кадр МЦИ (рис. 5в). На рис. 6а, б приведены искаженный и восстановленный кадр динамического МЦИ при отношении сигнал/шум на входе нелинейного фильтра 3 дБ. Для оценки качества нелинейной фильтрации динамических МЦИ взят критерий среднеквадратической ошибки (СКО).

Анализ результатов показывает, что, несмотря на снижение помехоустойчивости, возникающее вследствие применения МФМ сигналов, потери в помехоустойчивости полностью скомпенсированы за счет реализации статистической избыточности МЦИ. Время передачи сократилось в четыре раза по сравнению с прямой передачей МЦИ бинарными ФМ сигналами. Потери в помехоустойчивости приема МЦИ, состоящих из ГРЦИ, полностью скомпенсированы реализацией статистической избыточности МЦИ.



Рис. 5а. Кадр динамического МЦИ на борту МБПЛА

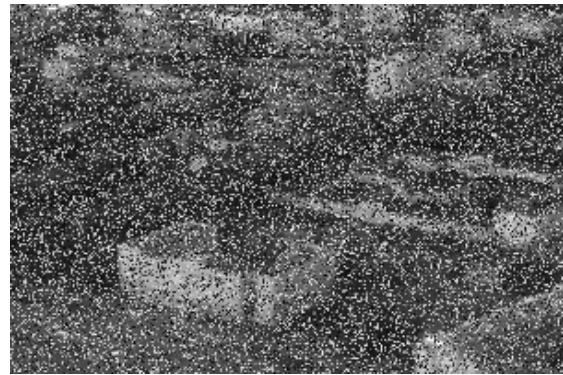


Рис. 5б. Кадр динамического МЦИ на входе приемного устройства, СКО =  $37 \times 10^{-2}$



Рис. 5в. Восстановленный кадр динамического МЦИ, СКО =  $2,75 \times 10^2$



Рис. 6а. Кадр динамического МЦИ на входе приемного устройства, СКО =  $15,9 \times 10^2$



Рис. 6б. Восстановленный кадр динамического МЦИ, СКО =  $0,73 \times 10^2$

#### Библиографический список

1. Петров Е.П., Трубин И.С., Харина Н.Л. Математическая модель двумерного цифрового полутонового изображения марковского типа // Вестник ВНЦ Верхне-Волжского отделения АТН РФ, серия «Проблемы обработки информации», Нижний Новгород. 2005. № 1(6). С. 41–46.
2. Петров Е.П., Харина Н.Л., Ржаникова Е.Д. Математическая модель цифровых полутоновых изображений на основе цепей Маркова с несколькими состояниями // Нелинейный мир. 2013. Т. 11. № 7. С. 487–492.
3. Петров Е.П., Харина Н.Л., Ржаникова Е.Д. Нелинейная обработка многоразрядных цифровых полутоновых изображений в системах дистанционного зондирования Земли // Нелинейный мир. 2015. Т. 13. № 6. С. 18–23.

#### METHOD FOR THE REDUCTION OF VIDEO TRANSMISSION TIME FROM A MINI UNMANNED AERIAL VEHICLE

**E.P.Petrov, N.L.Kharina, E.D.Rzhanikova**  
Vyatka State University, Kirov

*In this paper a method of reducing transmission time of dynamic multi-bit digital images (MDIs) from micro unmanned aerial vehicles (MUAVs). The method consist of two parts. The first part is distortionless transform of MDIs without computing operations on board the MUAV. The transmission of MDIs by MPSK (QPSK,8-PSK,16-PSK) signals is proposed. The second part is the compensation of noise immunity losses of MPSK signals receiving by multidimensional nonlinear filtering of MPSK signals. It is based on the effective utilizing of statistical redundancy of dynamic MDIs, which is caused by intra and inter-digit correlation of pixels in the MDIs.*

**Unmanned aerial vehicle, dynamic digital images, statistical redundancy, noise stability of video reception, multidimensional Markov chain**

**В. П. Кирнос, А. Л. Приоров, В. А. Антипов, В. А. Ковкина**  
Ярославский государственный университет им. П.Г. Демидова

## Алгоритм генерации базы изображений подстилающей поверхности

*Предлагается метод того, как, имея несколько изображений подстилающей поверхности и небольшой размер базы данных изображений, можно ее увеличить в 10-100 раз при помощи карты теней. Рассмотрены два алгоритма: синтез фотопроявленного изображения и добавление изображению теней на основе карты теней.*

### Шум Перлина, карта теней, подстилающая поверхность, нейронная сеть

При разработке алгоритмов детектирования изображений различного вида часто приходится использовать сверточные нейронные сети. В этом случае необходима база данных значительного размера для обучения сети и последующего ее тестирования. С этим связана другая проблема – ограниченность базы данных для конкретных приложений технического зрения [1–4]. Например, в настоящее время достаточно сложно найти соответствующую базу данных для решения задачи детектирования подстилающей поверхности в рамках решения задач мобильной робототехники.

Рассмотрим пример того, как, имея изображение подстилающей поверхности, получить дополнительный набор соответствующих изображений.

### Генерация фотопроявленного изображения

Для генерации фотопроявленных изображений подстилающей поверхности необходимо иметь заранее подготовленный шаблон. Для простоты реализации использовался шаблон типа «линолеум», представленный на рис. 1. В данном случае, для отладки алгоритма синтеза используются шаблоны с не очень высокой контрастностью. Само тестовое изображение из представленных шаблонов получается методом создания сетки, в ячейки которой вставляется указанный шаблон. После чего на полученное изображение накладывается структурный шум.



Рис. 1. Шаблоны типа «линолеум»

При имитации структурного шума здесь используется шум Перлина из-за его визуального подобия соответствующим помехам.

Шум Перлина – математический алгоритм по генерированию процедурной текстуры псевдослучайным методом. Это градиентный шум, состоящий из набора псевдослучайных единичных векторов (направления градиента), расположенных в определенных точках пространства и интерполированных функцией сглаживания между этими точками. Данный шум был создан Кеном Перлином в 1983 году и впоследствии был назван в честь своего создателя. Отличительной особенностью такого шума от других алгоритмов по

генерированию процедурных текстур является одинаковый размер визуальных деталей текстуры. Это свойство делает шум Перлина легко управляемым: множество масштабированных копий такого шума могут быть вставлены в математические выражения для создания самых разнообразных процедурных текстур. Шум Перлина широко используется в двумерной и трехмерной компьютерной графике для создания таких визуальных эффектов, как дым, облака, туман, огонь и пыль.

Выбранный масштаб сетки влияет на сложность шума. Большее число квадратов на сетке изображения создаст более «плотно упакованный» шум, подобный белому шуму. Меньшее число квадратов – порождает «клубящийся» шум, внешне похожий на облака. В каждой точке на сетке строится случайный вектор нормали. Это обычный двумерный вектор единичной длины, который указывает в случайном направлении в пределах каждого из квадратов. Традиционный способ создания таких векторов – организация справочной таблицы из векторов, которые охватывают полный круг, и последующий случайный выбор одного из них для каждой точки на сетке. Это гарантирует случайное распределение векторов, которые могут с равной вероятностью указывать в любом направлении [5–8].

Для каждого пикселя находят значение, основанное исключительно на данных ячейки, в которой он находится. Затем создаются четыре диагональных вектора, соединяющих углы ячейки с текущим пикселям (рис. 2).

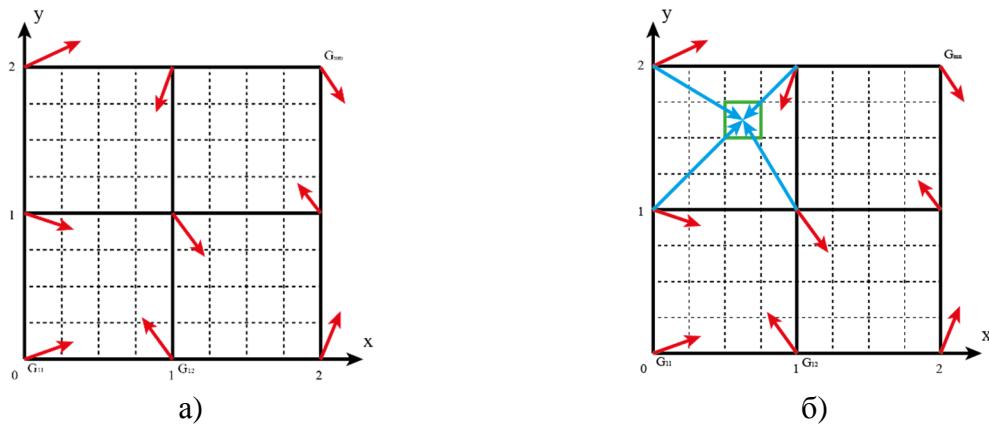


Рис. 2. Этапы формирования шума Перлина: а) формирование сетки и векторов нормалей; б) вычисление векторов от углов ячейки решетки к направлению рассматриваемого пикселя

Каждый угол ячейки теперь является базой для двух векторов – случайного единичного вектора и вектора в направлении рассматриваемого пикселя. Для каждой пары таких векторов находят скалярное произведение. Далее эти четыре значения объединяются. Объединять их можно по-разному, получая различные результаты, однако, чаще всего, используют взвешенную интерполяцию четырех значений с учетом близости текущей позиции к каждому углу ячейки сетки:

$$I(i,j) = g_a(w(x,y)) + g_b(1 - w(x,y)), \\ w = (6x^5 - 15x^4 + 10x^3)(6y^5 - 15y^4 + 10y^3),$$

где  $g_a$  и  $g_b$  – значения рассматриваемых углов ячейки;  $w(x,y)$  – вес, учитывающий близость пикселя к углу ячейки;  $x$  и  $y$  – координаты пикселя относительно рассматриваемого угла ячейки.

Сначала объединяют результаты произведения векторов верхней пары углов ячейки. Та же процедура затем повторяется и для двух значений произведения нижних углов. Наконец, результаты двух объединений также объединяются [9–11].

Примеры результатов расчета шума Перлина показаны на рис. 3.

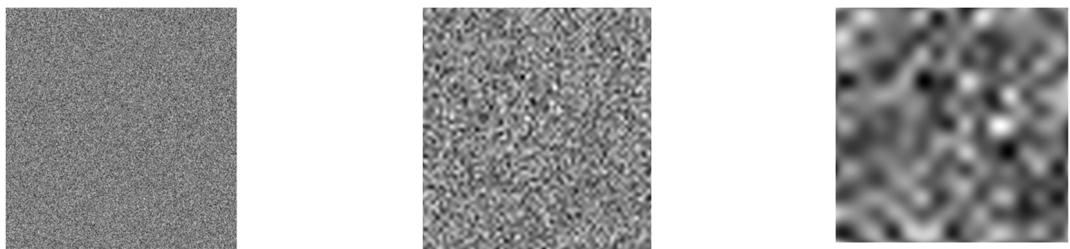


Рис. 3. Примеры карт шума Перлина (октав)

Отдельные карты шума носят название октав, так как одна из них имеет вдвое больший масштаб, чем другая. Объединение октав путем их сложения, перемножения и т. д. ведет к образованию новых шаблонов шума (рис. 4).

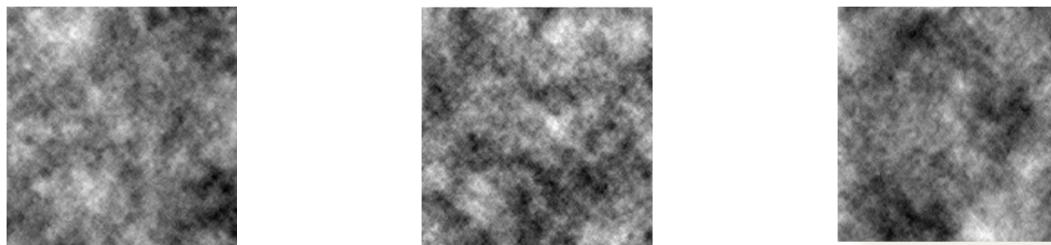


Рис. 4. Примеры изображений шума Перлина

Полученная маска добавляется к изображению, при этом самые контрастные участки считаются объектами теней. После получения изображения шума Перлина необходима дополнительная обработка для получения эффекта объекта тени.

#### Алгоритм добавления теней на изображение

Рассматривается случай, когда есть уже заранее подготовленное изображение с маской теней. Синтез теней производился при помощи библиотеки [2]. Задачей было получить результат не хуже, чем алгоритм, описанный в [3]. В качестве опоры используется карты google maps, в качестве источника света – направленный пучок, имитирующий поведения солнца. Пример получаемых изображений представлен на рис. 5.



Рис. 5. Примеры изображений теней на карте

Важно отметить, что на карте уже есть подготовленные отметки для создания теней – это высотные здания и мосты.

Объединение двух алгоритмов позволяет создать пример подстилающей поверхности, которую можно наблюдать в помещении, где в качестве источника света используется окно, так как представленный пример теней больше похож на структуру деревьев, нежели на структуру бытовой обстановки помещения.

Причиной такой особенности явился шум Перлина. Так как по своей структуре он напоминает облака или клубы дыма, он больше напоминает вид сверху дерева, объекты,

представленные внутри помещений, имеют правильные формы и не могут быть представлены в виде шумовых моделей.

При замене изображения типа «линолеум» на изображение подстилающей поверхности типа «дорога» или «тропа» получается алгоритм анализа поверхности вне помещения. Если использовать поверхность типа «дорога» и многопозиционное освещение, то получится алгоритм анализа поверхности вне помещения в вечернее или ночное время.

### Библиографический список

1. Трапезников И.Н. Разработка и анализ системы распознавания автомобильных регистрационных знаков. Автореферат диссертации. Владимир. 2014.
2. Алгоритм синтеза <https://github.com/perliedman/shadow-mapper>
3. Дебелов В.А., Новиков И.Е. Генерация мягких теней при использовании алгоритма трассировки лучей // Вестник Новосибирского государственного университета. Серия: Информационные технологии. Вып. № 2. Т. 7, 2009. С. 50–66.
4. Дебелов В.А., Васильева Л.Ф., Смирнова Г.Г. Метод световых сеток для случая полупрозрачных поверхностей // Тр. 16-й междунар. конф. по компьютерной графике и машинному зрению Графикон'2006. – Новосибирск, 2006. С. 298–302.
5. Agrawala M., Ramamoorthi R., Heirich A., Moll L. Efficient image-based methods for rendering soft shadows. In SIGGRAPH 2000, pp. 375–384, July 2000.
6. Atty L., Holzschuch N., Lapierre M., Hasenfratz J.-M., Hansen C., Sillion F. Soft shadow maps: Efficient sampling of light source visibility. Computer Graphics Forum 25, 4 (2006).
7. Brabec S., Seidel H.-P. Single Sample Soft Shadows Using Depth Maps // Proc. GraphicsInterface, pp. 219–228, Calgary, Alberta, 2002.
8. Eisemann E., Décoret X. Plausible image based soft shadows using occlusion textures. In Proc. of the Brazilian Symposium on Computer Graphics and Image Processing (2006), IEEE.
9. Soler C., Sillion F.X. Fast calculation of soft shadow textures using convolution. In Computer Graphics (SIGGRAPH 1998), Annual Conference Series, pp. 321–332. ACM SIGGRAPH, 1998.
10. St-Amour J.-F., Paquette E., Poulin P. Soft Shadows from Extended Light Sources with Penumbra Deep Shadow Maps. Graphics Interface 2005, pp. 105–112. 2005.
11. Williams L. Casting curved shadows on curved surfaces. Computer Graphics. 1978. Vol. 10, № 2. pp. 270–274.

### THE SPREADING SURFACE IMAGES BASE GENERATION ALGORITHM

**Kirnos Vasiliy Pavlovich, Priorov Andrey Leonidovich, Antipov Vladimir Alekseevich, Kokovkina**

**Vera Andreevna**

P.G. Demidov Yaroslavl State University

*The method of , having several images of the spreading surface and the small size of the database of images is suggested, it is possible to increase it by 10-100 times by means of a chart of shadows is offered. Two algorithms are considered: synthesis of the photorealistic image and addition to the image of shadows on the basis of the shadows chart.*

**Perlin noise, shadow mapping, the spreading surface, neural network.**

**В. П. Кирнос, А. Л. Приоров, В. А. Антипов, В. А. Ковкина**  
Ярославский государственный университет им. П.Г. Демидова

## Проблемы картирования пространства в задаче SLAM

Производится анализ границ применимости базовых алгоритмов картирования пространства в задаче SLAM, полученных на основе объемных телевизионных изображений с RGB-D сенсоров. Рассматриваются проблемы, которые влияют на составление карты пространства.

### Алгоритм SLAM, сенсор RGB-D, фильтр Калмана, карта глубины

В реальных условиях работа алгоритма картирования пространства мобильным роботом подразумевает ряд аспектов, которые могут привести к ухудшению точности локализации и построения карты пространства. Так, к примеру, карта глубины, получаемая на выходе RGB-D камеры подвержена различным дефектам, вызываемым способом получения дальномерических данных, недостаточным уровнем освещённости, характером материалов объектов сцены. Для задач технического зрения в большинстве случаев требуется предобработка карты глубины для минимизации влияния подобных дефектов [1].

### Проблемы основных алгоритмов картирования

Для реализации SLAM очень часто используются алгоритмы FastSLAM, EKF SLAM [2]. Сравним их реализацию на базе стандартной алгоритмов в среде Matlab. Для симуляции движения используется предопределенная траектория движения и набор ориентиров, ошибка моделируется при помощи АБГШ.

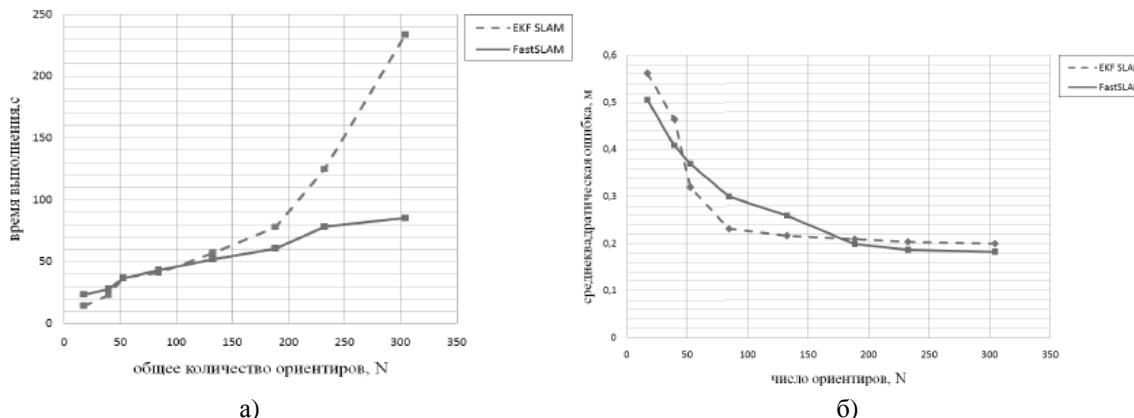


Рис. 1. Сравнение алгоритмов FastSLAM и EKF SLAM а) время выполнения алгоритма, б) СКО

При числе точек меньших 50 алгоритм EKF SLAM показывает более быстрое выполнение по сравнению с алгоритмом FastSLAM, но при этом проигрывает по СКО. Совершенно иная картина наблюдается на участке больше 200 точек: алгоритм FastSLAM начинает выигрывать и в СКО и в времени выполнение.

Как показывает практика, ошибки вследствие недостаточного числа особых точек имеют серьезные последствия при вычислении траектории движения камеры и построении карты. Достаточно тяжело порой добиться числа точек порядка 180-200 при условии разрешения камеры 640 на 480 пикселей.

## Причины предобработки карты глубины

Карта глубины, получаемая на выходе RGB-D сенсора при помощи искусственного светового текстурирования сцены, обычно содержит множество различных искажений и дефектов. Среди ряда причин, приводящих к подобным искажениям можно выделить следующие явления [3]:

1. Поглощающие и отражающие поверхности приводят к появлению на карте глубины дефектов, что вызвано невозможностью определения шаблона в данной точке сцены.
2. Ненулевая стереобаза системы проектора приводят к взаимным перекрытиям объектов на изображениях, поэтому часть точек проецируемого шаблона оказывается вне поля зрения камеры.
3. Шумы проектора и камеры приводят к ошибкам детектирования точек шаблона и, как следствие, к ошибкам карты глубины.
4. Триангуляционный метод расчета расстояния характеризуется тем, что погрешность измерения глубины растет с увеличением расстояния до объектов сцены.
5. Сильные искажения проецируемого шаблона на поверхностях объектов, ориентированных вдоль оптической оси камеры или вблизи этого положения.

В зависимости от причин возникновения и характера искажений, дефекты можно разделить на три группы [4]:

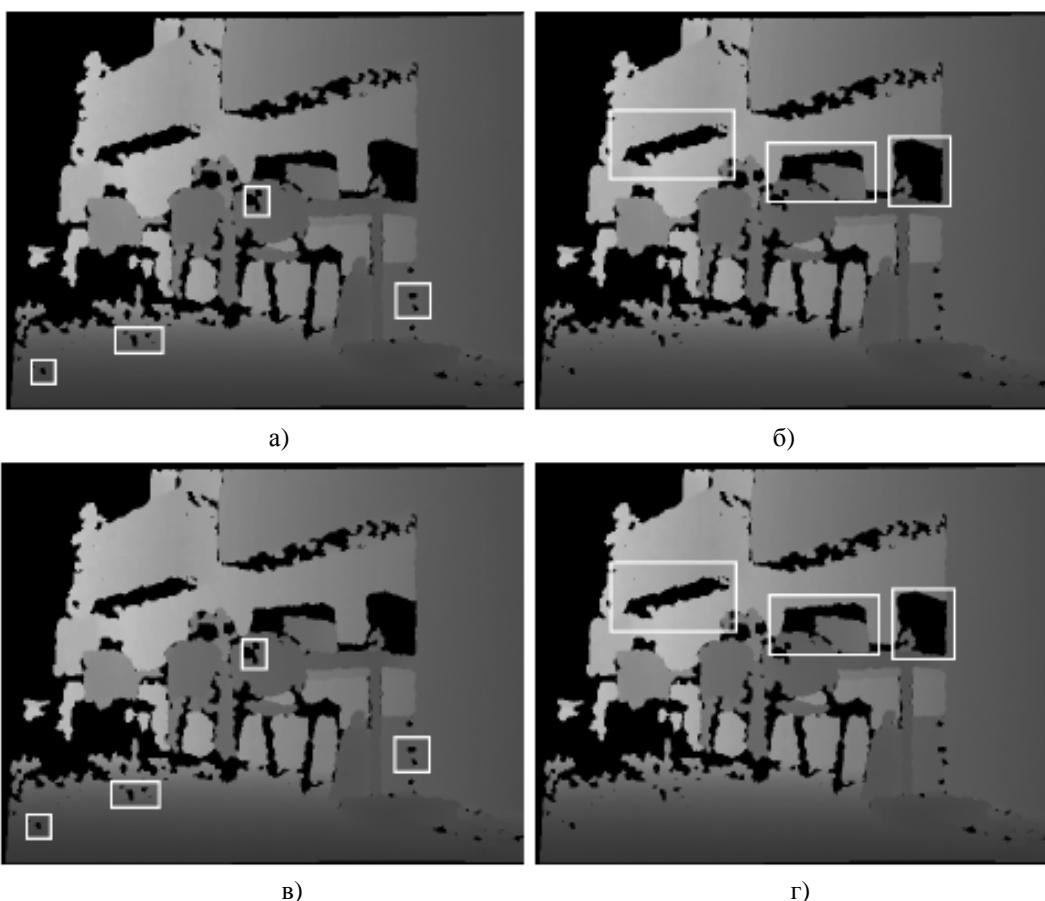


Рис. 2. Различные виды дефектов карты глубины, характерные для метода сканирования сцен с помощью инфракрасной локации

1. Случайно расположенные области на поверхностях объектов, возникающие вследствие ошибок распознавания светового шаблона (рисунок 2а), в том числе из-за поглощающего/отражающего характера поверхности объектов (рисунок 2б).

2. Области на границах объектов, появляющиеся вследствие перекрытия проецируемого шаблона объектами сцены, (рисунок 2в).
3. Ошибками детектирования светового шаблона на большом расстоянии от камеры за границей разрешающей способности сенсора камеры (рисунок 2г).

### Проблемы с сенсорами

Большинство существующих систем SLAM используют RGB-D камеру для получения визуальной информации, т.е. воспринимают одновременно и глубину, и цветовые компоненты получаемого изображения. Поскольку эти камеры имеют ограниченную частоту кадров, ограниченную пропускную способность для передачи данных и обычно прокатный затвор, производительность таких систем ограничивается аппаратными характеристиками [5].

#### 1. Пропускная способность

Системы типа Kinect и другие, как правило, имеют ограниченную пропускную способность для передачи и приема кадров камеры. Можно, конечно, использовать отправку кадров низкого разрешения, либо сжимать кадры, но оба эти решения могут привести нас к ухудшению производительности SLAM системы в целом. В частности, изображения с более низким разрешением будет не в состоянии найти мелкие детали на получаемой сцене. В режиме реального времени кодирование и декодирование требует значительных вычислительных мощностей, при этом могут появляться артефакты сжатия.

#### 2. Искажения, вызванные затвором камеры

Поскольку камеры, используемые в системах SLAM, не могут моментально записать кадр, в системе может появляться смазывание кадра (motion blur). Большинство камер для систем SLAM реального времени работают на частоте 30 кадров в секунду. Порой этого достаточно для движения некоторых объектов, но не для быстро движущихся объектов (лопасти вертолета, колеса автомобиля). Эта проблема решается увеличением частоты кадров при записи видео, но данное решение к проблеме полосы пропускания и к увеличению вычислительной сложности.

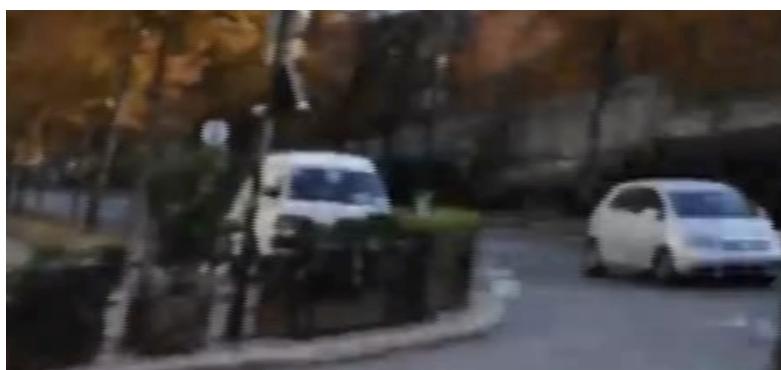


Рис. 3. Рулонный эффект

Еще одно ограничение вызвано тем, что используется прокатный затвор. Большинство бюджетных камер используют механический прокатный затвор, который приводит к появлению рулонного эффекта.

Более качественные камеры имеют глобальный затвор, который захватывает целиком сенсор, при этом рулонный эффект отсутствует.

Многие системы полагают, что построенная карта является априорной и не требует дополнительных уточнений. Но это далеко не так. Нельзя быть уверенным, что на

представленной карте по итогам SLAM нет движущихся объектов, для этого используется вероятностный подход, который позволяет учесть все подобные изменения [5,6].

Стерео система SLAM имеет две камеры, значит, может оценить глубину от вычисления разности между двумя слегка смещенными изображениями камеры. Тем не менее, хорошая оценка глубины нуждается в приличном расстоянии между главными оптическими осями объективов. Как правило, чем шире это расстояние, тем лучше оценка глубины. Компании разрабатывают мобильные устройства с бинокулярными камерами для использования их в приложениях дополненной реальности (AR), но рекомендуемое расстояние базовой линии для стереопары в районе 20 см [7]

Работа выполнена при финансовой поддержке РФФИ (грант № 15-07- 08674).

#### Библиографический список

1. Прозоров А.В.. Повышение точности позиционирования камеры в задаче одновременной локализации и картирования для систем прикладного объемного телевидения. Автореферат диссертации. Москва. 2016.
2. Bailey T. SLAM Package of Tim Bailey // <https://svn.openslam.org/data/svn/bailey-slam>.
3. Шапиро Л., Стокман Д. Компьютерное зрение // М.: БИНОМ, 2006. – С. 752.
4. Фисунов А.В., Воронин В.В., Марчук В.И., Свирин И.С., Петров С.П. Восстановление карты глубины на основе комбинированной обработки многоканального изображения // Современные проблемы науки и образования. – 2014. – № 6, <https://www.science-education.ru/ru/article/view?id=16360>
5. Newcombe, R. A., Fox, D. & Seitz, S. M. DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time, 2015.
6. Thrun S., Burgard W., Fox D. Probabilistic Robotics, The MIT Press, 2006, p. 668
7. Boonsuk W. Investigating Effects of Stereo Baseline Distance on Accuracy of 3D Projection for Industrial Robotic Applications, 5th IAJC/ISAM Conference Proceedings, 2016

#### SPACE MAPPING PROBLEMS IN THE SLAM TASK

**Kirnos V. P., Priorov A. L., Antipov V. A., Kokovkina V. A.**

P.G. Demidov Yaroslavl State University

*The analysis of limits of applicability of basic algorithms of mapping of space in a task of SLAM received on the basis of volume television images with RGB-D of sensors is made. Problems which influence drawing up the card of space are considered*

**SLAM, RGB-D, EKF, depth map**

**А. А. Чиркунова**  
СПБГЭТУ каф ТВ

# Оптимизация спектральной характеристики телекамеры при обнаружении малоконтрастных объектов

*Рассмотрены особенности выбора спектральных диапазонов при обнаружении малоконтрастных объектов.*

**Спектральная селекция, спектрозональная ТВ система, спектрально-энергетические признаки, малоконтрастный объект, контрастная чувствительность**

Одним из важных, но редко учитываемых параметров телекамеры является контрастная чувствительность, которая определяет возможность телекамеры обнаруживать и различать сигналы от малоконтрастных объектов.

Контрастная чувствительность телекамеры определяется через формулу А. Роуза [1] и задает минимальный контраст  $K$ , формирующий перепад видеосигнала с заданным отношением сигнал/шум  $\Psi$ , необходимым для различения изображений объектов [2]. Ее повышение особенно актуально при решении задач обнаружения малоконтрастных объектов и может быть достигнуто за счет оптимизации параметров и характеристик телекамеры [3 - 7].

В общем случае оптимизация должна проводиться по всем направлениям, т. е. по длине волны регистрируемого излучения (спектральная характеристика телекамеры), пространственным координатам (площадь пикселя фотоприемника (ФП), времени (время накопления) и на этапе аналого-цифрового преобразования – применением нелинейного квантования, количество и значение уровней при котором согласовано с зависимостью уровней фотонного шума и шума считывания от освещенности.

Однако, как правило, при построении радиотехнических систем наиболее жесткие требования предъявляются к входному звену, поскольку при последовательном прохождении сигналов последующие звенья только усиливают ошибки, возникающие при прохождении предыдущих. В роли такого входного звена в телекамере выступает оптическая система.

Таким образом, при обнаружении малоконтрастных объектов важным является корректный выбор спектральной характеристики оптической системы.

Спектральная характеристика телекамеры  $Q(\lambda)$  определяется совокупностью спектрального коэффициента пропускания светофильтров  $\tau_{\text{сф}}(\lambda)$ , спектрального коэффициента пропускания объектива  $\tau_{\text{об}}(\lambda)$  и зависимостью квантовой эффективности ФП от длины волны  $QE(\lambda)$

$$Q(\lambda) = \tau_{\text{сф}}(\lambda) \tau_{\text{об}}(\lambda) QE(\lambda).$$

Квантовая эффективность и коэффициент пропускания объектива являются паспортными характеристиками ФП и объектива и определяются материалами и технологией изготовления. Ввиду этого оптимизация спектральной характеристики телекамеры должна проводиться за счет определения полос пропускания спектральных светофильтров.

В роли критерия качества может быть выбран, например,  $\max(\Psi \times K)$ . Данный

критерий отражает компромисс между максимумом отношения сигнал/шум и максимумом контраста и подчеркивает, что данные максимумы достигаются на разных длинах волн. Как правило, максимум отношения сигнал/шум достигается на более широком диапазоне, максимум контраста, наоборот, в более узком.

Согласно выбранному критерию оптимизации подвергаются два параметра: нижняя граница пропускания спектрального светофильтра и ширина полосы пропускания (верхняя граница полосы пропускания).

Алгоритм оптимизации представлен на рисунке 1.

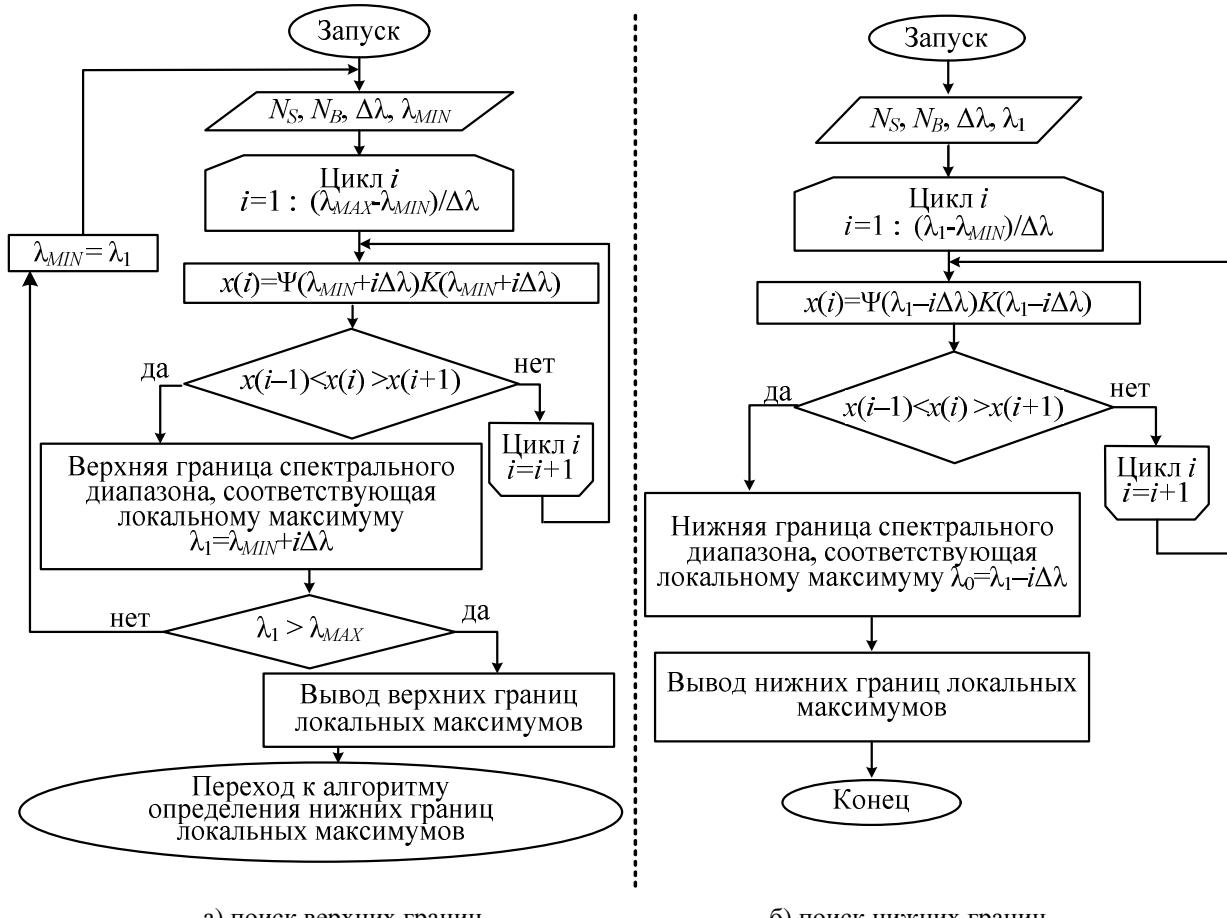


Рис. 1. Алгоритм оптимизации спектральной характеристики

Рассмотрим пример оптимизации спектральной характеристики телекамеры, работающей в режиме обнаружения объекта со спектрально-энергетической характеристикой  $I_1$  на фоне, имеющем спектрально-энергетическую характеристику  $I_2$ .

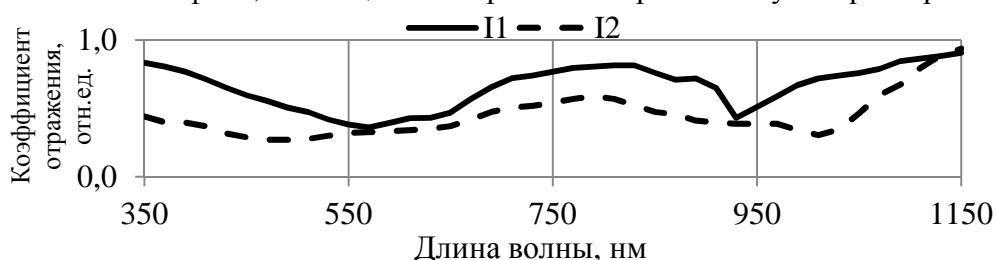


Рис. 2. Спектрально-энергетические характеристики объекта и фона

При работе в спектральном диапазоне чувствительности кремниевых ФП (350...1100 нм) телекамерой регистрируется отраженное объектами излучение. Вид зарегистрированных спектрально-энергетических характеристик обнаруживаемых объектов и фонов зависит от спектрального состава потока излучения источника,

освещдающего наблюдаемую сцену, спектральных характеристик коэффициента отражения, среды распространения излучения, оптической системы и зависимости квантовой эффективности от длины волны. С учетом вышеперечисленных факторов результирующие спектрально-энергетические характеристики представлены на рисунке 3.

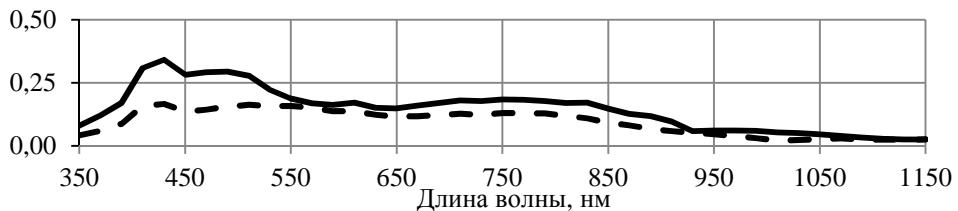


Рис. 3. Вид спектрально-энергетических характеристик объекта и фона

Для выявления отличия в сигналах от объекта и фона на рисунке 4 представлен график разности спектрально-энергетических характеристик объекта и фона.

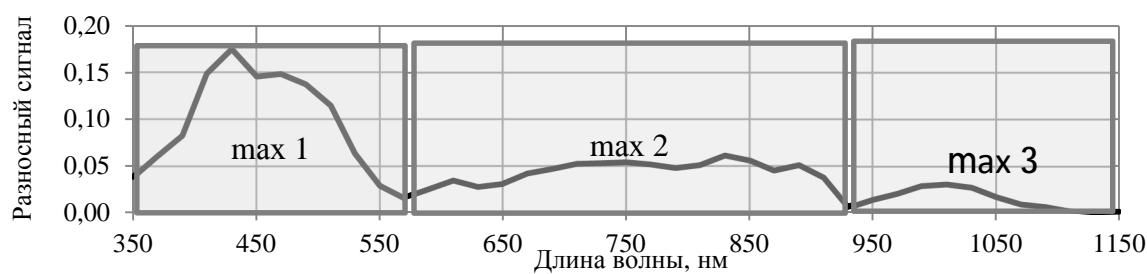


Рис. 4. Разностный сигнал

Из рисунка 4 видно, что на графике четко прослеживаются 3 максимума, при этом определение границ спектральных диапазонов обнаружения должно производиться по алгоритму, представленному на рисунке 1.

В результате оптимизации были получены зависимости произведения отношения сигнал/шум на контраст от нижней и верхней границ пропускания спектрального светофильтра для каждого из трех максимумов (рисунок 5).

Шаг дискретизации при работе алгоритма оптимизации составлял 20 нм, были получены границы трех спектральных диапазонов, которые могут быть использованы для обнаружения: (350...510 нм), (670...910 нм) и (970...1070 нм).

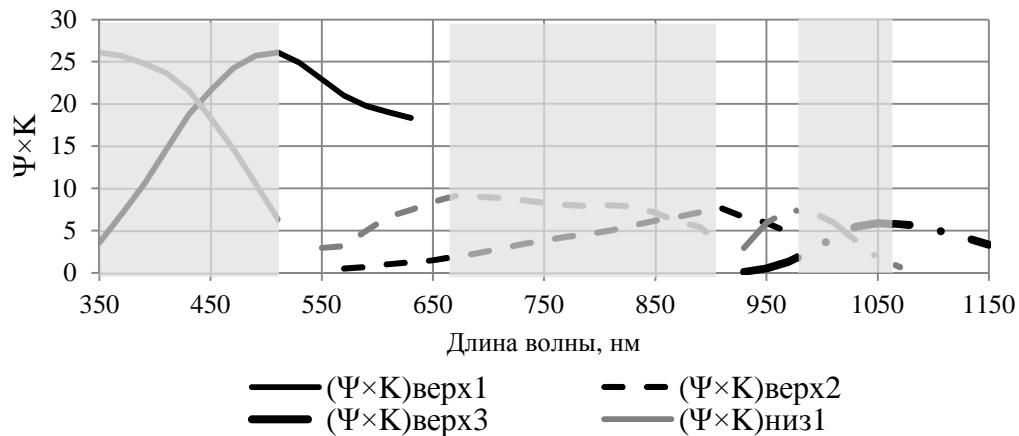


Рис. 5. Определение нижних и верхних границ спектрального диапазона

Уточненные границы, полученные в результате оптимизации, показаны на рисунке 6.

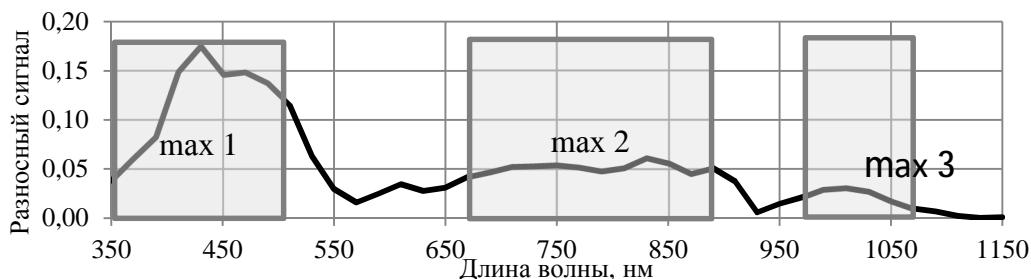


Рис. 6. Результат оптимизации границ спектральных диапазонов

Ширина и границы найденных спектральных диапазонов считаются оптимальным согласно критерию  $\max(\Psi \times K)$  при заданных внешних условиях обнаружения. Смена условий обнаружения, а также приемного устройства влечет за собой необходимость коррекции полученных значений.

Рассмотрим вопрос выбора спектрального диапазона обнаружения среди найденных по алгоритму оптимизации.

Байесовский критерий обнаружения гласит, что с целью повышения вероятности правильного обнаружения и качества передачи изображений нельзя терять никакой информации. Применительно к обозначенной задаче оптимизации это означает, что сигналы найденных трех диапазонов должны суммироваться.

Проверим данную гипотезу на примере суммирования сигналов найденных спектральных диапазонов.

Результаты вычисления отношения сигнал/шум, контраста и произведения отношения сигнал/шум на контраст представлены в таблице 1.

Расчет  $\Psi$ ,  $K$ ,  $\Psi \times K$ 

Таблица 1.

Диапазон №	Нормированные сигналы		$\Psi$	$K$	$\Psi \times K$
	Объект	Фон			
max 1	0,48	0,25	81,8	0,32	26,13
max 2	0,45	0,31	49,21	0,19	9,16
max 3	0,07	0,04	26,27	0,27	6,97
max 1+max 2+ max 3	1	0,60	95,98	0,25	24,16
max 1+max 2	0,93	0,56	92,33	0,25	23,15
max 1+ max 3	0,52	0,35	55,29	0,20	10,84
max 2+ max 3	0,55	0,29	85,83	0,31	26,79

Из полученных результатов видно, что при комбинировании спектральных диапазонов положительные результаты по сравнению с использованием одного спектрального канала наблюдаются только в том случае, когда величины сигналов спектральных каналов соизмеримы, в противном случае суммирование сигналов спектральных каналов приводит к снижению  $\Psi \times K$ .

В остальных же случаях  $\max(\Psi \times K)$  наблюдается при использовании информации из одного спектрального диапазона.

В рассмотренном примере для обнаружения объекта со спектрально-энергетической характеристикой  $I_1$  на фоне объекта со спектрально-энергетической характеристикой  $I_2$  может быть использован либо спектральный канал 1, либо комбинированный спектральный диапазон 2+3.

## Выводы

1. Алгоритм оптимизации спектральной характеристики разработан с учетом требований к контрастной чувствительности, позволяет избежать ее потерь на входе телевизионной системы.
2. Предложенный алгоритм позволяет определить полосы пропускания спектральных светофильтров, которые должны быть использованы при обнаружении объектов на фоне, при этом обеспечивая минимум ошибки на входе.
3. Выбор границ спектральных светофильтров должен производиться на основе критерия  $\max(\Psi \times K)$ , который подчеркивает важность достижения не только  $\max(\Psi)$ , но и  $\max(K)$ .

## Библиографический список

- 1 Роуз, А. Зрение человека и электронное зрение [Текст]: монография / А. Роуз. – М.: Мир, 1977. – 216с.
- 2 Цыбулин, А. К. Твердотельные телекамеры: накопление качества информации [Текст]: монография / А. К. Цыбулин, Д. Ю. Адамов, А. А. Манцевов, И. А. Зубакин. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2014. – 272 с.
- 3 Сагдуллаев, Ю. С. Восприятие и анализ разноспектральных изображений [Текст]: монография / Ю. С. Сагдуллаев, С. Д. Ковин. – М.: Изд-во «Спутник+», 2016. – 251 с.
- 4 Сагдуллаев, Ю. С. Информационно-измерительные системы телевидения [Текст]: монография / Ю. С. Сагдуллаев, С. Д. Ковин, Т. Ю. Сагдуллаев, А. И. Смирнов. – М.: Изд-во «Спутник+», 2013. – 199 с.
- 5 Сагдуллаев, Т. Ю. К вопросу выбора зон регистрации в спектрональном телевидении [Текст] / Т. Ю. Сагдуллаев, Ю. С. Сагдуллаев // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2011. – Вып. 2. С. 3-25
- 6 Chirkunova, A. A Method of Increasing the Contrast of Low-Contrast Objects by Combining Spectral Bands (Метод повышения контраста малоконтрастных изображений за счет комбинирования спектральных диапазонов) / N. Lysenko, A. Chirkunova // Proceedings of the 2016 IEEE ElConRusNW, St. Petersburg, 22 june – 3 july 2015.– Р. 283-287.
- 7 Чиркунова А. А. Методы и аппаратно-программные средства оптимизации и адаптации телевизионных систем обнаружения малоконтрастных объектов: дис. канд. тех. наук: 05.12.04: защищена: 19.04.17 – СПб., -2017 – 178 с.

## OPTIMIZATION OF TV CAMERA'S SPECTRAL RANGE IN THE DETECTION OF LOW-CONTRAST OBJECTS

**Chirkunova A. A.**

Saint Petersburg Electrotechnical University "LETI", Television subdepartment, AO "NIIT"

*The features of the choice of spectral ranges for the detection of low-contrast objects are considered.*

**Spectral selection, multispectral TV system, spectral-energy features, low-contrast object, contrast sensitivity**

**М. А. Занина, В. А. Павлов**  
ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра  
Великого»

## Совместное применение алгоритма TLD и фильтра частиц для задачи сопровождения объектов

*Работа посвящена совместному применению алгоритма сопровождения TLD (Tracking-Learning-Detection, Сопровождение-Обучение-Обнаружение) и фильтра частиц для сопровождения объектов. Описана работа алгоритма TLD и фильтра частиц.*

### **TLD, фильтр частиц, сопровождение объекта**

Задачи сопровождения объектов являются неотъемлемой частью различных областей, таких как наблюдение, навигация, мониторинг, контроль качества продукции, обеспечение безопасности и т.д.

Большинство известных методов обработки изображений (кадров видео) либо получены для простейших моделей, либо имеют эвристическое происхождение и не обеспечивают точного решения задач при наличии осложняющих факторов (различные помехи, неоднородность фона, слабую контрастность и т.д.). Разработка эффективных алгоритмов обработки изображений для работы в реальных условиях сложной и изменяющейся во времени фоново-целевой обстановки, при наличии априорной неопределенности, является актуальной задачей. Также данные алгоритмы должны адаптироваться к изменениям фоново-целевой обстановки.

В этой статье рассмотрена совместная работа алгоритма TLD и фильтра частиц, с помощью которой можно осуществлять сопровождение объекта в реальном времени по известному начальному положению.

Алгоритм TLD (Tracking-Learning-Detection, Сопровождение-Обучение-Обнаружение) состоит из трёх компонентов: модуль краткосрочного сопровождения, модуль обнаружения и модуль обучения.

На этапе обнаружения детектор сканирует входное изображение и принимает решение о наличии или отсутствии объекта на каждом участке изображения. Генерируются все возможные изменения масштабов и сдвигов исходной ограничивающей рамки, в результате получается множество ограничивающих прямоугольников. В дальнейшем, каждый отдельный участок классифицируется (используется классификатор по ближайшему соседу).

Модуль обучения оценивает действия краткосрочного алгоритма сопровождения, определяет его ошибки и обновляет его. Основная идея *P-N* обучения заключается в том, что ошибки детектора можно разделить на два типа. *P*-эксперт идентифицирует только ложно-отрицательные примеры (ответы, которые ожидались на выходе, но алгоритм их не получил), а *N*-эксперт – ложно-положительные (ответы, которых не должно было быть на выходе, но алгоритм ошибочно получил их). *P-N* обучение состоит из четырех блоков: обучающий классификатор, обучающий набор (помеченные тренировочные примеры), контролируемое обучение (метод, который обучает классификатор на тренировочном наборе) и *P-N* эксперты (функции, генерирующие положительные и отрицательные тренировочные). Схема работы обучающего модуля изображена на рис. 1.

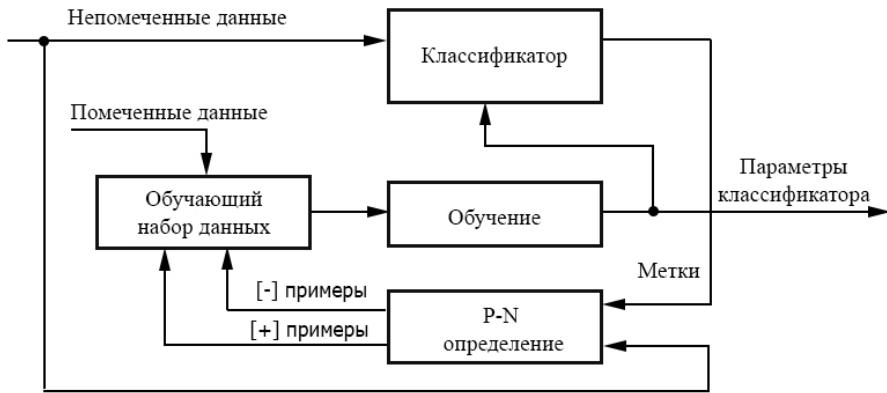


Рис. 1. Модуль *P-N* обучения

Компонент сопровождения в TLD представлен алгоритмом на основе вычисления оптического потока (Median-Flow tracker, MFT [1]). Алгоритм MFT представляет объект ограничивающей рамкой и покадрово оценивает его движение. Затем он оценивает перемещение набора точек внутри ограничивающей рамки. Выбирая 50% самых надежных перемещений, данный метод представляет движение ограничивающей рамки, используя медианные значения. Принцип работы MFT предполагает видимость объекта, поэтому алгоритм неустойчив к исчезновению объекта из кадра или его перекрытию.

Вместо алгоритма MFT в данной работе предлагается использовать алгоритм на основе фильтра частиц. Фильтр частиц [2] применяется для отслеживания объектов при наличии шума, в котором апостериорная плотность и плотность наблюдений имеют распределение, отличное от нормального. Характеристики отслеживаемого объекта описываются вектором состояний  $X_t$ , а вектор  $Z_t$  обозначает все наблюдения  $\{z_1, z_2, \mathbf{K}, z_t\}$  за промежуток времени, равный  $t$ .

Ключевая идея фильтрации с помощью фильтра частиц – это аппроксимация распределения вероятности с помощью взвешенного набора отсчетов (частиц). Каждая частица состоит из элемента  $s$ , который представляет гипотетическое состояние объекта, и соответствующей ему вес  $w$ , удовлетворяющей условию нормировки [3]. Чаще всего алгоритмы фильтра частиц строятся на основе метода воспроизведения условной плотности вероятности Condensation (Conditional Density Propagation) [4].

Пусть система может находиться в состояниях  $X_t = \{x_1, x_2, \mathbf{K}, x_t\}$   $X_t = \{x_1, x_2, \dots, x_t\}$  и существует последовательность наблюдений  $Z_t = \{z_1, z_2, \mathbf{K}, z_t\}$ . Предполагая, что текущее состояние системы зависит только от предыдущего, получим систему с независимым набором наблюдений. Данное множество

$$S_t = \left\{ \left( s_i^t, w_i^t \right), i = \overline{1, N}, \sum_{i=1}^N w_i^t = 1 \right\} \quad (1)$$

определяет плотность вероятности для состояния системы  $x_t$  при имеющемся наборе наблюдений  $Z_t$ . Множество  $S_t$  необходимо построить на основании  $S_{t-1}$  [2, 5]. Для этого необходимо выполнить следующие шаги:

1. Имеется взвешенный набор

$$S_{t-1} = \left\{ \left( s_i^{t-1}, w_i^{t-1} \right), i = \overline{1, N}, \sum_{i=1}^N w_i^{t-1} = 1 \right\} \quad (2)$$

Вычисляются интегральные весовые коэффициенты:

$$c_i = c_{i-1} + w_i^{t-1}, i = \overline{1, N}, c_0 = 0 \quad (3)$$

2. Определяется  $n$ -ый экземпляр выборки  $S_t$ . Вычисляется

$$j = \arg \min_i \{c_i > r\}, \quad (4)$$

где  $r$  – случайное число из интервала  $[0, 1]$ . В итоге частица  $s_j^{t-1}$  попадает в множество  $S_t$  с вероятностью, равной ее весовому коэффициенту.

3. Выполняется предсказание:

$$s_n^t = F_{t-1} s_j^{t-1} + X_{t-1} s_n^t, \quad (5)$$

4. Полученное значение корректируется (для  $s_n^t$  устанавливается весовой коэффициент):

$$w_n^t = p(z_t | x_t = s_n^t) \quad (6)$$

5. Шаги 2-4 повторяются  $N$  раз, результатом является множество частиц  $S_t$ .

Последовательность весовых коэффициентов нормализуется.

6. Вычисляется оценка состояния  $x_t$ , в данном случае как среднее значение:

$$x_t = \sum_{i=1}^N w_i^t s_i^t \quad (7)$$

### Заключение

Фильтр частиц устойчив к помехам и шумам, а также быстро работает при правильном наборе параметров. Также для него существует возможность задавать разные метрики для сравнения моделей. Данный метод позволяет рассматривать несколько гипотез состояния объекта одновременно. Поскольку менее вероятные состояния объектов имеют шанс храниться какое-то время в процессе отслеживания, фильтр частиц можетправляться с кратковременными окклюзиями [3], что важно в реальных условиях сложной и изменяющейся во времени фоново-целевой обстановки. В дальнейшем предполагается экспериментальное исследование предложенной модификации и сравнение с оригинальным алгоритмом TLD.

### Библиографический список

1. Kalal Z., Mikolajczyk K., Matas J. Forwardbackward error: automatic detection of tracking failures // Proceedings of the ICPR'10. 2010. p. 2756–2759.
2. Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2010. 979 p.
3. K. Nummiaro, E. Koller-Meier, L. Van Gool, A Color-Based Particle Filter // First International Workshop on Generative-Model-Based Vision, in conjunction with ECCV'02 (2002) p. 53– 60.
4. Yilmaz A., Javed O., Shah M. Object tracking: A survey // ACM Computing Surveys. 2006. V. 38. No 4, Article 13.
5. Isard M., Blake A. Condensation - conditional density propagation for visual tracking. Int. J. Comput. Vision. 1998. V.29. No 1. P. 5 – 28

**THE JOINT APPLICATION OF THE TLD ALGORITHM AND THE PARTICLE FILTER FOR THE OBJECT-TRACKING TASK**

**Zanina M. A., Pavlov V. A.**

Peter the Great Saint-Petersburg Polytechnic University

*The paper is devoted to the joint application of the TLD (Tracking-Learning-Detection) tracking algorithm and the particle filter algorithm for the object tracking task. The work of the TLD method and the particle filter is described.*

**TLD, particle filter, object tracking**

## Секция 3. Цифровая обработка и кодирование видеинформации

Руководитель секции – д.т.н., проф. Б. С. Тимофеев(СПбГУАП)

Зам. руководителя – к.т.н., доц. П. С. Баранов (СПбГЭТУ «ЛЭТИ»)

Секретарь секции – асс. А. В. Морозов (СПбГЭТУ «ЛЭТИ»)

---

**Д. С. Андреев**  
СПбГЭТУ «ЛЭТИ»

### Метод обработки разреженной карты глубины

*Рассмотрен метод обработки разреженной карты глубины, состоящий в фильтрации с быстрым глобальным сглаживанием и основанный на взвешенном методе наименьших квадратов.*

**Карта глубины, взвешенный метод наименьших квадратов, фильтрация, трехмерная реконструкция, трехмерная модель**

Восстановление трехмерной модели сцены по набору изображений – классическая задача компьютерного зрения. Развитие виртуальной реальности, систем управления транспортными средствами, визуализации на основе анализа изображений, медицинской промышленности, а также других областей, требующих построения трехмерных моделей, привело к увеличению внимания к этой области. В том числе, задача трехмерной реконструкции нашла свое место и в телевизионных системах для космического применения [1,2]. Так, для ряда задач возникает необходимость в построении объемной модели видимой части объекта в космическом пространстве [3,4].

Под трехмерной реконструкцией понимают процесс получения цифрового представления трехмерной сцены реального мира [5]. Одним из ключевых этапов трехмерной реконструкции является построение так называемой карты глубины. Известно несколько методов для ее формирования, среди которых весьма популярным подходом является стереозрение [5,6,7]. Для его реализации требуется две или более камер, направленных на сцену, предварительно откалиброванных. Т.к. области зрения камеры в стереопаре пересекаются - могут быть найдены соответствия одним и тем же частям сцены на изображениях стереопары [5]. На основе полученных снимков строится карта глубины, отражающая дальность каждой точки наблюдаемой сцены от плоскости камеры. Таким образом, формируется трехмерная модель объекта [3-6].

Последние десятилетия системы компьютерного зрения активно развиваются и к настоящему моменту решены многие частные задачи. Кроме того, накоплен обширный практический и теоретический материал, облегчающий разработку новых систем [5].

Тем не менее, известные решения в ряде случаев оказываются малоэффективны. Применение алгоритмов стерео-зрения, особенно высокооптимизированных и предназначенных для обработки в реальном времени на встраиваемых процессорах, влечет за собой возникновение ошибок при работе со сложными сценами [8]. Так, актуальной остается проблема построения карты глубины для малодетальных объектов, когда крупные области на изображении объекта имеют приблизительно равное значение яркости для всех входящих в них пикселей, областей с почти бесконечной дальностью и областей, содержащих блики [9].

В таком случае, методы минимизации ошибки, основанные на яркости, могут учитывать крайне малые отличия и определять произвольные значения дальностей для таких областей [9]. Карта глубины получается в значительной мере зашумленной и разреженной. Так происходит при обработке сцен, на которых присутствует, к примеру, ровная стена, тень от объекта, предмет, имеющий равномерную окраску и малое количество характерных деталей.

Частично подобные проблемы можно решить подбором оптимального размера окна. При увеличении размера окна в определенных пределах наблюдается снижение шумовой составляющей и числа ошибок обнаружения соответствующих областей, однако карта глубины становится сильно размытой, падает детальность. Подобное решение подходит лишь для задач с невысокими требованиями к точности реконструкции объема и детальности формируемой карты глубины.

При более строгих требованиях к точности для снижения числа ошибок используют различные алгоритмы детектирования точек с потенциально неточными значениями дальности. Подобные алгоритмы присутствуют в большинстве реализаций построения карты глубины, например, в методах StereoBM и StereoSGBM из библиотеки OpenCV [8]. Значения глубины с низкой достоверностью помечаются невалидными и не включаются в карту глубины, в результате чего она становится разреженной (см. рисунок 1).



Рис. 1. Изображение с левой камеры и разреженная карта глубины

Для подобных разреженных и зашумленных карт глубины свою эффективность показало применение процедур фильтрации. Такая фильтрация осуществляет сопоставление границ карты глубины с границами исходного изображения и распространение значения дальностей с участков с высокой достоверностью на участки с низкой достоверностью. Последние достижения в методах технического зрения обеспечивают осуществление подобной пост-фильтрации в условиях реального времени [10].

Одним из методов фильтрации является алгоритм быстрого глобального сглаживания, основанный на взвешенном методе наименьших квадратов. Этот метод относится к группе сглаживающих фильтров, обеспечивающих сохранение границ и был предложен Донбо Мином и др. [10] в качестве альтернативы имеющимся подходам. Такие алгоритмы подавляют шум, но сохраняют четкость изображения.

По принципу действия фильтры с сохранением границ могут быть разделены на две группы. В первую группу входят фильтры, которые в явном виде вычисляют результирующее изображение как взвешенное среднее, иногда итеративным путем. Ранние работы в этом направлении включают в себя анизотропную диффузию и билатеральный фильтр [10]. Несмотря на варьирующееся качество обработки и

ограниченные возможности применения такие фильтры обычно эффективны и как правило обеспечивают линейную сложность, зависящую только от числа пикселов в изображении. Однако подобные локальные фильтры не в состоянии полностью разрешить неоднозначность в отношении того, сглаживать определенную границу или нет и напрямую неприменимы для ряда задач обработки изображений.

Вторая группа методов основывается на глобальной оптимизации [10]. Они ищут глобально-оптимальное решение для целевой функции, обычно учитывая члены ограничения и сглаживания. Благодаря глобальному решению таких целевых функций данные методы часто достигают высочайшего качества обработки во множестве задач обработки изображений и компьютерной графики и превосходят многие ограничения фильтров первой группы, в том числе и эффекта «гало». Однако такое качество достигается за счет высокой вычислительной сложности, существенно возрастающей при решении огромных линейных систем.

Метод быстрого глобального сглаживания, основанный на взвешенном методе наименьших квадратов, является быстрым алгоритмом, выполняющим неоднородное сглаживание с сохранением границ. Подход, как и в глобальных алгоритмах, состоит в оптимизации глобальной целевой функции, заданной с ограничивающим и сглаживающим членами, но предлагается альтернативный метод решения систем линейных уравнений [10]. Решение исходной системы линейных уравнений аппроксимируется неоднородной матрицей Лапласиана путем решения последовательности одномерных линейных подсистем, основанных на глобальной оптимизации. Подобный метод разложения позволяет итеративным путем применять высокоэффективный трехдиагональный матричный алгоритм [10] в каскадной форме. Как итог, алгоритм имеет линейную вычислительную сложность, зависящую только от числа пикселов в изображении [10].

Алгоритм основывается на глобальных методах и выполняет минимизацию функционала взвешенного метода наименьших квадратов. В общем случае, для задач сглаживания с сохранением границ функционал имеет вид:

$$J(u) = \sum_p \left( (u_p - f_p)^2 + \lambda \sum_{q \in N(p)} w_{p,q}(g) (u_p - u_q)^2 \right),$$

где  $f$  – входное изображение,  $g$  – управляющее изображение,  $u$  – ожидаемый результат,  $f_p$  и  $g_p$  – значение яркости пикселя  $p(x,y)$ . Управляющее изображение  $g$  может быть задано входным изображением  $f$  (для задач фильтрации изображений) или другим управляющим сигналом, согласованным с  $f$ , в зависимости от применения.  $S=W \times H$  – размер изображения,  $N(p)$  – окрестность пикселя  $p$  – набор соседних пикселов, как правило 4.  $\lambda$  – коэффициент, отвечающий за баланс между двумя слагаемыми. Увеличение  $\lambda$  приводит к большему сглаживанию результата  $u$ .

Ограничение на сглаживание усиливается весовой функцией  $w_{p,q}(g)$ , представляющей схожесть между двумя пикселями  $p$  и  $q$  и может быть задана как:

$$w_{p,q}(g) = \exp\left(-\frac{\|g_p - g_q\|}{\sigma_c}\right).$$

Приравнивая градиент функционала к нулю, результат  $u$  получают решением системы линейных уравнений, основанных на разреженной матрице:  $(I + \lambda A)u = f$ , где  $u$  и  $f$  обозначают вектор-столбец размерностью  $S$ , содержащий значения яркостей  $u$  и  $f$

соответственно,  $\mathbf{I}$  – единичная матрица. Матрица Лапласиана  $\mathbf{A}$  размерностью  $S \times S$  задается следующим образом:

$$A(m, n) = \begin{cases} \sum_{l \in N(m)} w_{m,l}(g) & n = m \\ -w_{m,l}(g) & n \in N(m) \\ 0 & \text{otherwise} \end{cases}$$

где  $m$  и  $n$  обозначают индекс, соответствующий пиксели  $p$ . Значение пикселя  $m$  результирующего изображения теперь определяется следующим образом:

$$\mathbf{u}(m) = ((\mathbf{I} + \lambda \mathbf{A})^{-1} \mathbf{f})(m)$$

Первым отличием метода является разложение глобального функционала на ряд одномерных:

$$\sum_x \left( (u_x^h - f_x^h)^2 + \lambda_t \sum_{i \in N_h(x)} w_{x,i}(g^h) (u_x^h - u_i^h)^2 \right),$$

где  $f^h$ ,  $u^h$  и  $g^h$  – одномерные сигналы изображений вдоль измерения  $x$ .

Тогда одномерное решение  $\mathbf{u}$ , минимизирующее представленное выражение записывается в виде линейной системы:

$$(\mathbf{I}_h + \lambda_t \mathbf{A}_h) \mathbf{u}_h = \mathbf{f}_h$$

Та же система:

$$\begin{bmatrix} b_0 & c_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & a_x & b_x & c_x & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{W-1} & b_{W-1} \end{bmatrix} \begin{bmatrix} u_0^h \\ \vdots \\ u_x^h \\ \vdots \\ u_{W-1}^h \end{bmatrix} = \begin{bmatrix} f_0^h \\ \vdots \\ f_x^h \\ \vdots \\ f_{W-1}^h \end{bmatrix}$$

с граничными значениями  $a_0 = 0$  и  $c_{W-1} = 0$ . Здесь  $u_x^h$  и  $f_x^h$  –  $x$ -е элементы  $u^h$  и  $f^h$ ,  $a_x$ ,  $b_x$  и  $c_x$  представляют три ненулевых элемента  $x$ -той строки  $(\mathbf{I}_h + \lambda_t \mathbf{A}_h)$ , которая может быть записана как:

$$\begin{aligned} a_x &= \lambda_t \mathbf{A}_h(x, x-1) = -\lambda_t w_{x,x-1}, \\ b_x &= 1 + \lambda_t \mathbf{A}_h(x, x) = 1 + \lambda_t (w_{x,x-1} + w_{x,x+1}), \\ c_x &= \lambda_t \mathbf{A}_h(x, x+1) = -\lambda_t w_{x,x+1}. \end{aligned}$$

Такая система имеет точное решение методом Гаусса, вычислительная сложность которого  $O(N)$ .

Распространенный подход к раздельному сглаживанию состоит в последовательном применении одномерных решений вдоль каждого измерения сигнала [10]. Для заданного двухмерного изображения одномерное решение применяется итеративно вдоль строк и столбцов изображения. Однако этого недостаточно для распространения информации вдоль границ, и в результате возникают «полосовые артефакты», свойственные раздельным алгоритмам [10].

Чтобы избежать этого, двухмерное сглаживание осуществляется путем многократного применения одномерного сглаживания с изменением коэффициента сглаживания на каждой итерации. Суммарный коэффициент определяется по формуле:

$$\lambda = \sum_{t=1}^T 2\lambda_t = \sum_{t=1}^T \left(\frac{1}{2}\right)^{2t} 2\lambda_t,$$

где  $\lambda_t$  – коэффициент сглаживания на  $t$ -й итерации. Тогда, задавшись необходимым итоговым значением коэффициента, можно определить значение для каждого прохода фильтра:

$$\lambda_t = \frac{3}{2} \frac{4^{T-t}}{4^T - 1} \lambda.$$

Минимальное число итераций, дающих результаты обработки, не уступающие по качеству глобальным методам, было установлено экспериментально и равняется 3.

Для задач обработки карты глубины полезным оказалось расширение алгоритма для задач интерполяции разреженных данных. Имея управляющее изображение, например, исходное изображение левой камеры, и разреженное входное – карту глубины, функционал ошибки можно выразить следующим образом:

$$\sum_p \left( h_p (u_p - f_p)^2 + \lambda \sum_{q \in N(p)} w_{p,q}(g) (u_p - u_q)^2 \right),$$

где  $h_p$  – индексирующая функция, принимающая значение 1 для валидных пикселов и 0 для невалидных.

Результат можно получить решением данной системы:

$$(\mathbf{H} + \lambda \mathbf{A})\mathbf{u} = \mathbf{H}\mathbf{f},$$

где  $\mathbf{H}$  – диагональная индексирующая матрица. При решении системы разреженные данные распространяются с учетом структуры управляющего изображения  $g$ . Однако разложение на  $x$  и  $y$  составляющие, аналогичное изначальному подходу, приводит к нестабильным результатам ввиду разреженности входных данных.

Вместо этого сглаживается по-отдельности и входное изображение, и индексирующий вектор. Пусть  $\tilde{\mathbf{u}}$  и  $\tilde{\mathbf{u}}_h$  обозначают сглаженные сигналы входного изображения и индексирующего вектора, тогда результатом обработки будет  $\tilde{\mathbf{u}} ./ \tilde{\mathbf{u}}_h$ , где  $./$  – поэлементное деление матриц.

$$\mathbf{u}(m) = \frac{((\mathbf{I} + \lambda \mathbf{A})^{-1}\mathbf{f})(m)}{((\mathbf{I} + \lambda \mathbf{A})^{-1}\mathbf{h})(m)}.$$

Результаты применения алгоритма представлены на рисунке 2.

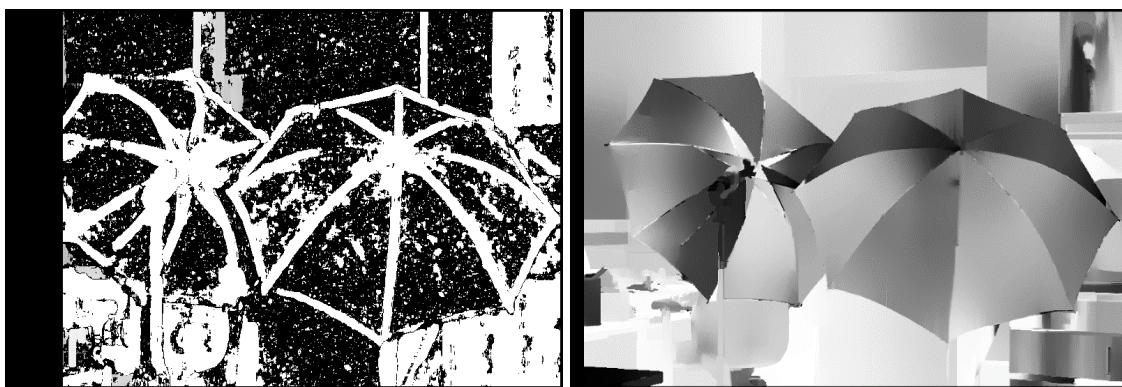


Рис. 2. Карта глубины до и после применения алгоритма

Как видно, метод быстрого глобального сглаживания, основанный на взвешенном методе наименьших квадратов является эффективным инструментом обработки разреженных карт глубины. Он позволяет повысить детальность построения карты глубины для сцен, содержащих малодетальные объекты, обеспечивая при этом высокое качество обработки. В то же время, алгоритм имеет линейную вычислительную сложность, что немаловажно для систем, требовательных к быстродействию.

### Библиографический список

1. Степанов Д. Н., Бахшиев А. В., Смирнова Е. Ю., Половко С. А. Определение положения космических аппаратов путём обнаружения и сопровождения естественных конструктивных особенностей // Вопросы радиоэлектроники, Серия «Техника телевидения». 2016. Вып. 1. С. 61-77
2. Ази С.Н., Бачевский С.В. Космическая стереофотометрическая телевизионная система в задачах сближения истыковки космических аппаратов // Вопросы радиоэлектроники, Серия «Техника телевидения». 2009. Вып. 1. С. 65-84
3. Андреев Д.С. Определение ориентации объекта в пространстве с помощью телевизионной системы // Сборник докладов научно-технической школы-семинара «Инфокоммуникационные технологии в цифровом мире». 2016.
4. Denis S. Andreev. Object's Spatial Orientation Detection in Television System. 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering
5. Gagolowicz A., Philips W. Computer Vision. Computer Graphics Collaboration Techniques. 4th International Conference, MIRAGE, Springer, France, 2009
6. Wöhler C. 3D Computer Vision. Efficient Methods and Applications. 2009. P. 389.
7. Бабенко В.С. Оптика телевизионных устройств. М.: Радио и связь, 1982. 256 с.
8. [http://docs.opencv.org/trunk/d3/d14/tutorial\\_ximgproc\\_disparity\\_filtering.html](http://docs.opencv.org/trunk/d3/d14/tutorial_ximgproc_disparity_filtering.html)
9. <http://ai.stanford.edu/~mitul/cs223b/dp.html>
10. Dongbo Min, Sunghwan Choi, Jiangbo Lu, Bumsub Ham, Kwanghoon Sohn, and Minh N Do. Fast global image smoothing based on weighted least squares. Image Processing, IEEE Transactions on

### SPARSE DEPTH MAP PROCESSING METHOD

**Andreev D.S.**

Saint Petersburg Electrotechnical University "LETI"

*Paper covers filtering procedure to align the disparity map edges with those of the source image and to propagate the disparity values from high- to low-confidence regions.*

**Depth map, weighted least squares, filtering, 3D reconstruction, 3D model**

**И. В. Жигулина**  
ВУНЦ ВВС «ВВА»

## Выбор апертуры фильтра при двумерной дискретной квазиоптимальной фильтрации

*В работе предложен алгоритм нахождения апертуры фильтра при согласованной со свойствами изображения двумерной дискретной фильтрации. Он основан на использовании функционала энергии изображения. Определено соответствие между апертурой фильтра и количеством нормированных амплитуд гармоник энергетического спектра изображения, которые используются при нахождении отсчетов импульсной характеристики фильтра.*

### **Дискретный фильтр, апертура, изображение, эталон, квазиоптимальная фильтрация**

При синтезе квазиоптимальных двумерных фильтров и адаптивных фильтров верхних пространственных частот существует ряд проблем, без решения которых невозможно говорить о том, что процедура синтеза имеет законченный вид. Одной из них является проблема выбора апертуры дискретного фильтра (ДФ). Такая проблема возникает, в частности, когда дискретная фильтрация осуществляется с целью выделить на изображении некоторый объект. Для этого используют заранее сформированное изображение (эталон) с присущими искомому объекту основными свойствами (размер, ориентация, характеристики формы).

Традиционно процедуры анализа или синтеза ДФ основаны том, что апертура фильтра заранее известна [1, 2]. Очевидным является то, что существует оптимальная апертура фильтра, поскольку в предельных случаях - очень малой или очень большой апертуры - показатели качества обработки могут быть минимальны. Поэтому сначала следует определять именно апертуру фильтра, а затем - отсчеты его импульсной характеристики (ИХ). Процедура нахождения апертуры должна отражать потенциально достижимые характеристики обработки и основываться только на свойствах входного изображения и эталона. Такой подход возможен на базе функционала энергии [3].

### **Критерии выбора оптимальной апертуры фильтра**

Оптимальная фильтрация подразумевает априорное знание о положении искомого объекта на поле изображения. Если же необходимо выделить объекты, не имея информации об их наличии и местоположении, то целесообразнее будет квазиоптимальная фильтрация, при которой пренебрегают взаимной корреляционной функцией объекта и фона.

Отсчеты ИХ ДФ определяются так, чтобы минимизировалась некоторая целевая функция (действительная и неотрицательная), являющаяся функцией входного, выходного и эталонного изображений. Минимум целевой функции является критерием синтеза фильтра [4].

При построении целевой функции удобно использовать *квадратичную форму функционала выходной энергии изображения* [3]:  $e(\mathbf{H}) = \mathbf{H}^T \mathbf{S} \mathbf{H}$ , где  $\mathbf{H}$  - вектор, компонентами которого являются отсчеты ИХ ДФ  $h_{l,p}$ , расположенные в определенном

порядке;  $\mathbf{S} = \|s_{k,r}\|$  - матрица входного изображения;  $s_{k,r} = \frac{1}{E_{\text{вх}}} \sum_m \sum_n f_{m,n} f_{m+k,n+r}$  -

нормированные амплитуды гармоник энергетического спектра входного изображения;  $m+k, m \hat{=} [-M; M]$ ,  $n+r, n \hat{=} [-N; N]$ ;  $f_{m,n}$  - отсчеты видеосигнала входного изображения;  $l \hat{=} [-a; b]$ ,  $p \hat{=} [-c; d]$ . «т» - транспонирование.

Минимизируя разность выходной энергии изображения и величины, которая определяется энергетическими свойствами эталона, получим

$$r = E_{\text{вых}} - \mathbf{H}^T \times \mathbf{B} = E_{\text{вх}} \times e - E_{\text{вх}} \times \mathbf{H}^T \times \mathbf{S}^0 = E_{\text{вх}} (e - \mathbf{H}^T \times \mathbf{S}^0), r \hat{=} 0, \quad (1)$$

где  $\mathbf{S}^0 = \{s_{l,p}^0\}$  - вектор, размерность и развертка которого определяется вектором  $\mathbf{H}^T$ ;

$s_{l,p}^0 = \frac{b_{l,p}}{E_{\text{вх}}}$  - коэффициенты разложения энергетического спектра изображения эталонного объекта в двумерный ряд Фурье по косинусам.

Минимум  $r = 0$  достигается при условии, что энергия выходного изображения определяется только эталоном, т.е.  $e(\mathbf{H}) = \mathbf{H}^T \times \mathbf{S}^0$ . Используя две разных формы записи функционала  $e(\mathbf{H})$ , можно получить соотношение для нахождения отсчетов ИХ ДФ

$$\mathbf{H} = \mathbf{S}^{-1} \times \mathbf{S}^0. \quad (2)$$

При этом энергия выходного изображения в наименьшей возможной степени будет отличаться от энергии обработанного подобным фильтром изображения эталона. Таким образом, (2) описывает потенциально лучший из возможных фильтров.

Для сравнения результатов обработки изображения различными фильтрами необходим некоторый показатель качества обработки, в качестве которого удобно использовать величину  $h = \frac{1}{e}$ .

При квазиоптимальной фильтрации  $e = (\mathbf{S}^0)^T \times \mathbf{S}^{-1} \times \mathbf{S}^0$ . Это соотношение позволяет найти величину функционала энергии только по свойствам входного изображения (матрица  $\mathbf{S}$ ) и эталона (матрица-столбец  $\mathbf{S}^0$ ). С его помощью можно достаточно просто, меняя порядок  $\mathbf{S}$ -матрицы и размерность  $\mathbf{S}^0$ , найти минимальное значение  $e$ , и, следовательно, максимальное значение показателя  $h$ . Размерность  $\mathbf{S}^0$  определит размерность вектора  $\mathbf{H}$ .

Подобный подход применим и для оценки характеристик оптимальной фильтрации, в этом случае вместо  $\mathbf{S}^0$  необходимо использовать  $R_b(l, p) + R_{a,b}(l, p)$ , где  $R_b(l, p) = \frac{K_b(l, p)}{E_{\text{вх}}}$ ;  $R_{a,b}(l, p) = \frac{K_{a,b}(l, p)}{E_{\text{вх}}}$ ;  $K_b(l, p)$  - отсчеты функции автокорреляции изображения объекта;  $K_{a,b}(l, p)$  - отсчеты функции взаимной корреляции изображений объекта и фона.

В действительности для определения апертуры фильтра необходимо вводить еще один критерий, т.к. с ростом апертуры характеристики обработки обычно улучшаются, а вычислительная сложность растет. И если сложность возрастает линейно, то качество обработки изменяется нелинейно. Поэтому, определить апертуру оптимального фильтра

можно по критерию «минимум произведения функционала энергии на вычислительную сложность», считая, что вычислительная сложность определяется количеством отсчетов ИХ:

$$k(w) = e \times w, \quad (3)$$

где  $w$  - количество отсчетов ИХ ДФ.

Для полного выделения объекта размерами  $1' k$  с наибольшим возможным показателем  $h$  необходимо, чтобы эталон имел размеры  $1' \frac{ek+1}{2}$ , а фильтр - апертуру  $3' (k+1)$ , где  $[ \cdot ]$  означает целую часть числа.

### Алгоритм нахождения оптимальной апертуры фильтра

Назовем квадратные апертуры фильтра *базовыми*. Для них  $a+b+1=c+d+1$ , т. е. количество отсчетов ИХ равно  $(a+b+1)^2 = 1; 4; 9; 16; 25; \mathbf{K}$

При спиральной развертке каждая следующая базовая апертура может быть получена из предыдущей путем последовательного добавления отсчетов ИХ (рис.1).

Количество отсчетов ИХ (вычислительная сложность) получаемых при этом фильтров различной конфигурации может быть определено по формуле

$$w = (a+b+1)^2 + y, \quad (4)$$

где  $(a+b+1)^2$  - соответствует предыдущей базовой апертуре;  $y=1, \mathbf{K}, (2(a+b+1)+1)$ .

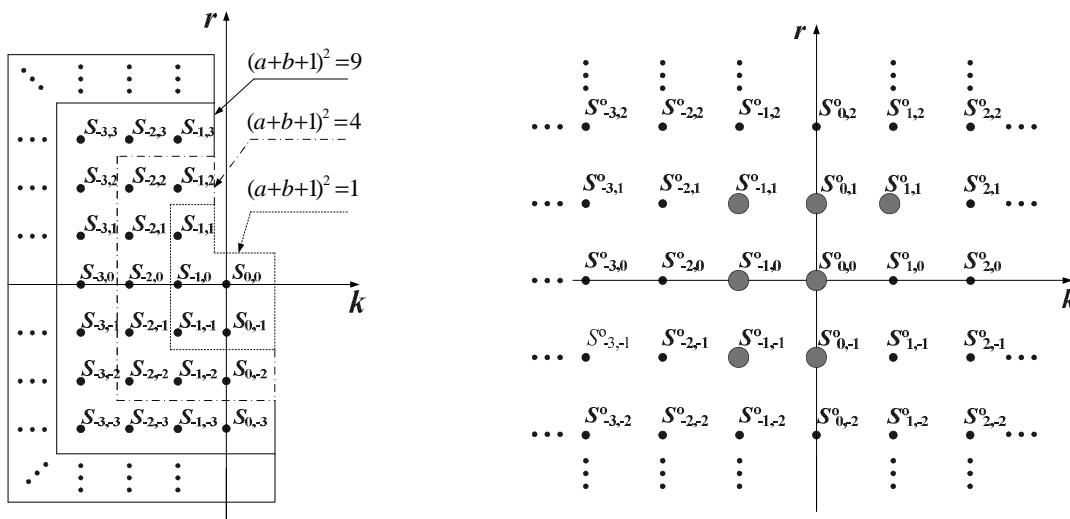


Рис. 1. Группы значений  $s_{k,r} \in E$ , соответствующие базовым апертурам.

Рис. 2. Компоненты  $s_{k,r}^0$  эталона, если количество отсчетов ИХ ДФ  $w=7$ .

В таблице 1 приведен пример соответствия ИХ фильтров различной конфигурации и тех элементов  $s_{k,r}$ , которые необходимо взять для оценки качества обработки (жирные точки).

Таблица 1

Количество отсчетов предыдущей базовой апертуры $(a+b+1)^2 = 4$
---

Величина $y$	1	2	3	4	5
ИХ ДФ (спиральная развертка)					
Элементы $s_{k,r}$					

Обобщая все вышеизложенное, можно предложить следующий алгоритм определения оптимально апертуры дискретного фильтра.

**Шаг 1.** Найти элементы  $s_{k,r}$  входного изображения, каждый из которых вычисляется по всему полю входного изображения, т.е. его величина не зависит от апертуры фильтра. Какие из этих отсчетов будут использованы в  $\mathbf{S}$ -матрице определяется апертурой фильтра.

**Шаг 2.** Вычислить  $w$  значений амплитуд гармоник энергетического спектра изображения эталона по формуле  $s_{k,r}^o = \frac{1}{E_{\text{вх}}} \sum_i \sum_j f_{k+i, r+j}^o f_{i,j}^o$ , где  $(k+i, r+j) \in W$ ,  $(r+j, j) \in W$ ,  $W$  - область эталона. Так как для выделения даже одномерных объектов необходимо использовать двумерный фильтр, то количество отсчетов ИХ следует выбирать, начиная с  $w=3$ . Для элементов множества  $E^o = \{s_{k,r}^o\}$  необходимо также использовать спиральную развертку (рис.1б).

**Шаг 3.** Найти значение  $e$  по формуле  $e = (\mathbf{S}^o)^T \times \mathbf{S}^{-1} \times \mathbf{S}^o$ .

**Шаг 4.** Вычислить показатель  $h$  и величину критерия  $K(w)$ .

**Шаг 5.** Увеличить значение  $w$  на единицу. Повторить шаги 1 - 4. Сравнить полученное значение  $K$  с предыдущим. Как только  $K(w+1) < K(w)$ , то  $w$  - количество отсчетов в оптимальной апертуре.

#### Библиографический список

- Гонсалес, Р Цифровая обработка изображений/ Р. Гонсалес, Р. Вудс. М.: Техносфера. 2005. 1072с.
- Визильтер, Ю.В. Обработка и анализ изображений с примерами на LabVIEW IMAQ Vision/ Ю.В. Визильтер [и др.]. М.: ДМК Пресс. 2007. 464 с.

3. Bogoslovsky A.V., Zhigulina I.V. A Way of Energy Analysis for Image and Video Sequence Processing/ Intelligent Systems Reference Library. Volume 73. Computer Vision in Control Systems-1. Mathematical Theory. - Springer. – 2015. – P.183-210. – DOI 10.1007/978-3-319-10653-3
4. Джиган, В.И. Адаптивные фильтры и их приложения в радиотехнике и связи/ В.И. Джиган//Современная радиоэлектроника, №9. 2009. С.56-63
5. Обработка многомерных сигналов. В 2-х книгах. Кн.1. Линейная многомерная дискретная обработка сигналов. Методы анализа и синтеза/ под ред. А.В. Богословского. М.: Радиотехника. 2013. 168с.

## **SELECTING A FILTER APERTURE FOR TWO-DIMENSIONAL DISCRETE QUASI-OPTIMAL FILTERING**

**Zhigulina I.V.**

Air Force Education and Research Center «The Zhukovsky and Gagarin Air Force Academy»  
(Voronezh)

*An algorithm for finding the aperture of a two-dimensional discrete filter consistent with image properties is proposed. It is based on the use of the image energy functional. A correspondence is established between the filter aperture and the number of normalized amplitudes of the image energy spectrum harmonics, which are used for finding the impulse response of the filter.*

**Discrete filter, aperture, image, standard, quasi-optimal filtering.**

**Н. А. Обухова, А. А. Поздеев**  
Санкт-Петербургский государственный электротехнический  
университет «ЛЭТИ им. Ульянова (Ленина)»

## Методы контрастирования медицинских ТВ изображений

*В статье приведен результат анализа эффективности применения различных методов коррекции яркости и контраста, а также их адаптация для обработки медицинских (эндоскопических) изображений. Описаны условия применимости этих методов к изображениям, полученным в условиях неравномерной освещенности*

**Локальные методы контрастирования, глобальные методы контрастирования, медицинские изображения, неравномерная освещенность**

Актуальность создания новых методов обработки и визуализации медицинских изображений, обеспечивающих существенное повышение их качества и высокую эффективность визуального анализа, не вызывает сомнений. Анализ изображения исследуемого органа является важнейшей составляющей частью при обследовании пациента. Соответственно, качество предъявляемых врачу изображений существенно влияет на корректность и точность формулировки диагноза.

Эндоскопические изображения – это один из наиболее значительных как по объему, так и по практической значимости класс медицинских изображений. Это изображения, получаемые при обследовании различных органов и полостей организма с помощью специального опто-электронного устройства (эндоскопа). Видами эндоскопии являются гистероскопия – осмотр полости матки, кольпоскопия – осмотр шейки матки, гастроскопия - осмотр желудка, лапароскопия осмотр брюшной полости и др. Насчитывается более 30 видов. В ходе эндоскопии врач выполняет визуальный осмотр поверхности органа, принимает решения о необходимости взятия биопсии, может проводить хирургические вмешательства (эндохирургия).

К наиболее значимому артефакту эндоскопических изображений, обусловленному сложными условиями их получения и спецификой объектов интереса, следует отнести выраженную неравномерность яркости и контраста: одновременное присутствие значительных по площади светлых и темных областей. Дополнительно, при обработке медицинских изображений должна быть учтена необходимость сохранения всей присутствующей в них изначально информации.

Целью исследования являлся анализ эффективности применения методов контрастирования различного типа, а также их адаптация для обработки эндоскопических изображений с учетом сформулированных выше их особенностей и ограничений.

Методы коррекции контраста можно разделить на *глобальные* и *локальные*. Глобальные методы используют одну функцию трансформации яркости для всех элементов изображения. Локальные методы применяют для каждого элемента изображения свою функцию преобразования яркости. Ее вид определяют на основе анализа окрестности, сформированной вокруг корректируемого элемента изображения.

К наиболее известным и часто употребляемым *глобальным методам* относятся метод линейной «растяжки» яркостной гистограммы (линейное контрастирование), гамма–коррекция, эквалайзация (выравнивание) яркостной гистограммы.

*Линейная «растяжка» яркостной гистограммы* – это наиболее простой метод глобального контрастирования. Его целесообразно применять, если гистограмма имеет вид близкий к форме кривой Гаусса с узким динамическим диапазоном. Растигивание диапазона на всю допустимую шкалу увеличивает разницу между соседними уровнями яркости, присутствующими на изображении, что приводит к увеличению контраста.

Метод *гамма-коррекции* реализует преобразование яркостной характеристики с помощью степенной функции. Варьируя показатель степени в функции преобразования яркости можно менять ее форму в достаточно широких пределах.

Метод *эквализации гистограммы* реализует преобразование уровней яркости таким образом, чтобы они следовали с примерно одинаковой частотой, а полученная яркостная гистограмма соответствовала бы равномерному закону распределения.

Как было указано выше, существенной особенностью эндоскопических изображений является одновременное присутствие светлых и темных областей. Контрастирование этих изображений путем применения глобальных методов малоэффективно. Например, при использовании гамма-коррекции для контрастирования темных участков изображения необходимо, чтобы крутизна трансформирующей функции убывала с увеличением яркости. Это обеспечивает сдвиг изображения в область светлого, в результате чего повышается контраст темных фрагментов изображения. Одновременно результатом такого сдвига будет потеря на изображении малоконтрастных деталей в светлой области, что является серьезным недостатком применительно к изображениям с неравномерными яркостной характеристикой и контрастом.

Применение метода эквализации гистограммы основанное на перераспределении уровней яркости в рамках сформулированной задачи имеет аналогичный недостаток, что и гамма-коррекция. Улучшение контраста в области темного приводит к его ухудшению в области светлого и наоборот.

Рисунок 1(а-г) демонстрирует результаты применения перечисленных глобальных методов контрастирования к реальным эндоскопическим изображениям, подтверждая приведенные выше положения.

Методами локального контрастирования являются метод адаптивной эквализации гистограммы с ограничением (CLAHE) [1], метод многомасштабного контрастирования, метод адаптивного нелинейного контрастирования (AINDANE)[3], адаптивный метод повышения качества цветных изображений (LTSNE) [4], метод с адаптивной глубиной коррекции[5].

*Метод адаптивной эквализации гистограммы с ограничением (CLAHE)* предполагает перераспределение уровней яркости в отдельных фрагментах изображения путем использования сканирующего окна. Этот способ обработки часто приводит к значительному усилению шумов, поэтому дополнительно вводят ограничение по контрасту.

*Метод многомасштабного контрастирования* основан на следующих нелинейных функциях трансформации :

$$G(x)=\frac{(B+A)x}{A+x} \quad "x \in [0, B], \quad A, B \in R$$
$$H(x)=\frac{A \times x}{A+B-x} \quad "x \in [0, B], \quad A, B \in R$$

где  $x$  - входные данные,  $B$  - максимальное значение  $x$ ,  $A$  – коэффициент, который регулирует вид функции трансформации.  $G(x)$  или  $H(x)$  - это функции, реализующие противоположенное изменение входных данных.

Идея преобразования заключается в следующем: для каждого пикселя  $(i, j)$  исходного изображения разность между значением его яркости  $Y_{ij}$  и средней яркостью его окрестности  $S_{ij}$  должна быть увеличена. Если верно условие  $Y_{ij} > S_{ij}$ , то используют функцию  $G(x)$  для того, чтобы вычислить новое увеличенное значение яркости анализируемого пикселя и, таким образом, сделать разность между яркостью пикселя и средней яркостью его окрестности  $S_{ij}$  больше. Если верно условие  $Y_{ij} < S_{ij}$ , то используется функция  $H(x)$  для того, чтобы вычислить новое уменьшенное значение яркости пикселя и, таким образом, увеличить разность между яркостью обрабатываемого пикселя и средней яркостью его окрестности  $S_{ij}$ . Используют размер окрестности трех разных масштабов, что позволяет использовать для анализа как низкочастотные, так и высокочастотные составляющие изображения.

Ключевой особенностью методов адаптивного нелинейного контрастирования (AINDANE) и адаптивного метода повышения качества цветных изображений (LTSNE) является одновременная коррекция яркостной характеристики и повышение контрастности. Методы близки между собой и имеют различие только в шаге коррекции яркостной характеристики. В методе AINDANE этот шаг основан на использовании глобальной информации об изображении, в LTSNE на локальной, что позволяет ему реализовать адаптивную коррекцию яркости. На этапе контрастирования оба метода используют локальную информацию и дают возможность применять различные функции трансформации к каждому элементу изображения.

$$S(x, y) = 255 \cdot I_{E_n}(x, y)^{E(x, y)},$$

$$E(x, y) = r(x, y)^p = \left( \frac{I_M(x, y)}{I(x, y)} \right)^p.$$

где  $S(x, y)$  – яркость пикселя после контрастирования,  $I_{E_n}(x, y)$  – нормированная яркость пикселя в изображении (после коррекции яркостной характеристики),  $I(x, y)$  – яркость пикселя в исходном изображении,  $p$  - параметр, регулирующий глубину коррекции контраста,  $I_M(x, y)$  – яркость фрагмента изображения после применения фильтра Гаусса.

$$I_M(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) G_i(m + x, n + y),$$

где  $G_i$  - коэффициенты фильтра Гаусса для структурных элементов разного размера ( $i = 3$ ).

Применение метода LTSNE на этапе коррекции контраста наряду с подчеркиванием мелких деталей усиливает шумовую составляющую. Это может приводить к потере эргономичности предъявляемого врачу изображения, которую нельзя скорректировать фильтрами подавляющими шумы, вследствие требования сохранить в изображении всю исходную информацию.

Этого недостатка в значительной мере лишен метод с адаптивной глубиной коррекции [5]. В нем коэффициент, определяющий глубину коррекции контраста (параметр  $p$ ), является адаптивным и определяется уровнем текстурированности фрагмента изображения.

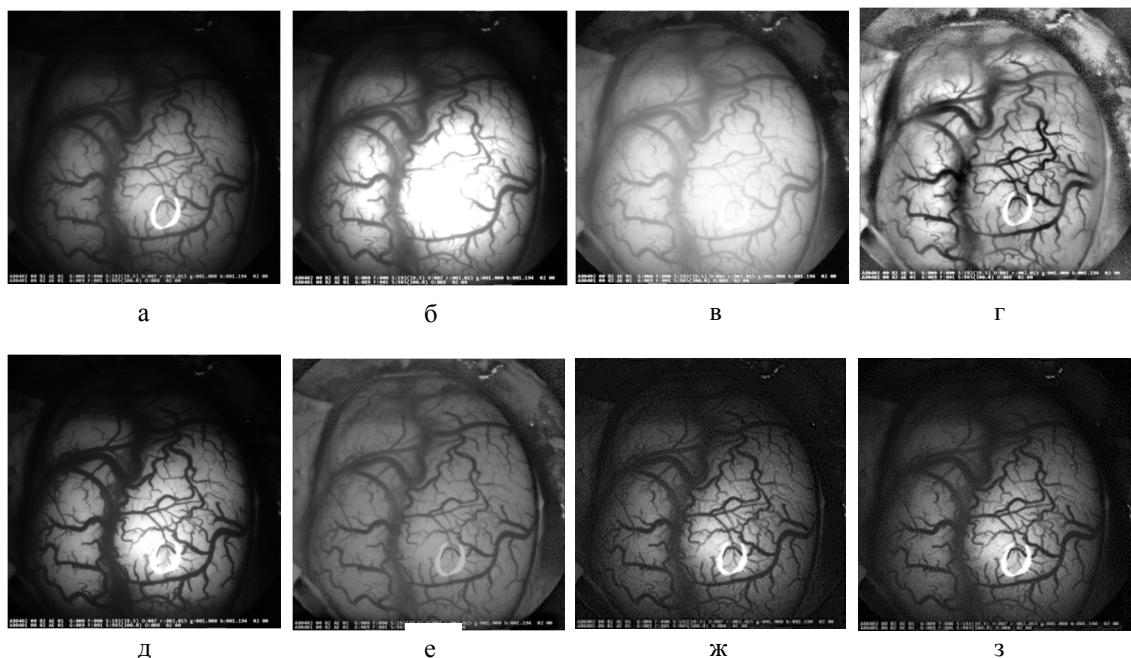
$$p = \log\left(\frac{k\sigma^2}{\sigma_{min}^2}\right),$$

где  $\sigma$  – среднеквадратическое отклонение яркости во фрагменте изображения;  $\sigma_{min}$  – минимальное среднеквадратическое отклонение яркости среди всех фрагментов изображения;  $k$  – параметр модели.

В рамках проводимых авторами НИР все перечисленные локальные методы были использованы в экспериментах для повышения качества эндоскопических изображений (рис. 1). Сравнение результатов, полученных при обработке реальных эндоскопических изображений, позволяет сделать следующие выводы.

Для изображений с умеренными артефактами метод многомасштабного контрастирования обеспечивает достаточно высокое качество коррекции. Получаемые изображения эргономичны (рис. 1е). Для выражено сложных случаев, например, для изображений с очень темными участками и очень светлыми участками одновременно, метод не является эффективным. В этих случаях приемлемый результат обеспечивает применение методов AINDANE и LTSNE, а также метода с адаптивной глубиной коррекции.

Результаты методов AINDANE и LTSNE близки между собой, но в большей части случаев визуальное качество изображения, согласно экспертным оценкам, после LTSNE выше, чем после AINDANE. Существенным артефактом при применении этих методов является подчеркивание шумов, особенно на слабо текстурированных участках изображения. Избежать этого позволяет метод с адаптивной глубиной контрастирования, который можно рассматривать как наиболее эффективный для коррекции изображений, имеющих одновременно очень светлые и очень темные области.



(а) Исходное изображение, (б) Растворка гистограммы, (в) Эквалайзация гистограммы, (г) CLAHE,  
 (д) Многомасштабное контрастирование, (е) AINDANE, (ж) LTSNE, (з) метод с адаптивной  
 глубиной коррекции

Рис. 1. Коррекция контраста

В качестве иллюстрации результативности последнего на рисунке 2 приведены контурные препараты исходного и обработанных изображений, бинаризованные по одному и тому же пороговому значению. После коррекции яркости и контраста стали видны детали, неразличимые на исходном изображении. В то же время метод с

адаптивной глубиной коррекции обеспечивает значительно более низкий уровень шумов по сравнению с алгоритмом LTSNE.

Таким образом, на основании проведенных исследований могут быть сформулированы следующие общие рекомендации по коррекции яркостной характеристики и контраста эндоскопических изображений.

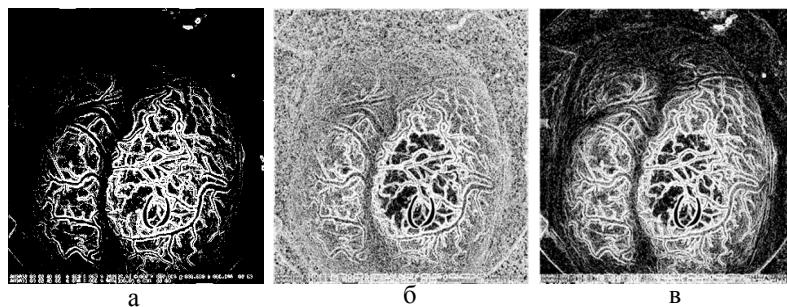
- Нецелесообразно использование глобальных методов: применение одной функции трансформации ко всем пикселям влечет за собой потерю эргономичности отдельных фрагментов изображения.

- Повысить качество одних областей интереса без ухудшения качества в других областях позволяют локальные методы, учитывающие особенности фрагментов изображения. Выбор конкретного метода из группы локальных методов контрастирования следует осуществлять исходя из особенностей изображений, которые в большой степени обусловлены условиями их формирования.

- Многомасштабное контрастирование применимо, когда изображение имеет умеренные деградации.

- При обработке изображений, полученных в условиях существенно неравномерной освещенности, наиболее целесообразно применение метода с адаптивной глубиной коррекции, включающего в себя одновременно коррекцию яркости и контраста, а также позволяющего избежать увеличения шумовой составляющей на слаботекстурированных участках.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 17-07-00045.



(а) Исходное изображение, (б) LTSNE, (в) Усиление контраста с использованием адаптивной глубины коррекции.

Рис. 2. Контурные препараты

### Библиографический список

1. S. M. Pizer, E. P. Amburn, J. D. Austin, et al.: Adaptive Histogram Equalization and Its Variations. Computer Vision, Graphics, and Image Processing 39 (1987), 355-368.
2. V. Vonikakis, I. Andreadis: Multi-Scale Image Contrast Enhancement. 10th Intl. Conf. on Control, Automation, Robotics and Vision, (2008) 856-861.
3. L. Tao and K.V. Asari, “An adaptive and integrated neighborhood dependent approach for nonlinear enhancement of color images,” SPIE Journal of Electronic Imaging, Vol. 14, No. 4, 2005, pp. 1.1-1.14.
4. Arigela, S., Asari, V.K.: A Locally Tuned Nonlinear Technique for Color Image Enhancement. WSEAS Trans. Signal Process. 4(8), (2008) 514-519.
5. Н.А. Обухова, А.А. Поздеев: «Метод нелинейного контрастирования медицинских изображений»/19 Международная конференция «Цифровая обработка сигналов и ее применение DSPA -2017», Москва, март 2017.

## TECHNIQUES FOR THE MEDICAL TV IMAGES ENHANCEMENT

**Natalia A. Obukhova, Alexandr A. Pozdeev**

Saint-Petersburg State Electrotechnical University "LETI", student

*The effectiveness of different enhancement methods, as well as their adaptation for the medical (endoscopic) images processing are considered in this paper. The conditions for the applicability of these methods to images captured under extremely non-uniform lighting conditions are described.*

**Local contrast enhancement, global contrast enhancement, medical images, non-uniform lighting conditions.**

**А. Г. Шоберг, К. А. Шоберг**  
Тихоокеанский государственный университет

## Сравнение изображений в частотной области при блочной форме выполнения преобразования

*Рассмотрены вопросы сравнения изображений на основе блочного частотного преобразования, выполненного различными способами. Дано математическое описание предлагаемого подхода. Представлены результаты сравнения традиционного и разработанного подходов к рассматриваемому преобразованию.*

### **Интегральное преобразование, модифицированная схема, сравнение изображений, блочная форма**

Многомерный сигнал, математически описываемый как  $f(n_1, n_2, \mathbf{K}, n_m)$ , может быть представлен в другом пространстве (домене) в виде  $F(w_1, w_2, \mathbf{K}, w_m)$  с помощью интегрального преобразования [1,2,3]. Такая операция сохраняет часть свойств исходного сигнала, что позволяет выполнять часть операций проще или быстрее. Но результаты операции поворота многомерного сигнала в пространственной области отличаются от результатов этой же операции в частотной области. [4,5].

Рассмотрим двумерный случай. Интегральное преобразование целого набора данных (для использования частотных характеристик) делится на набор стандартных блоков. Это выполняется для различных целей, например, для эффективного хранения [6,7]. Такие блоки  $b_i$  имеют  $n' n$  отсчетов. Эта процедура выполняется для  $b_i = b(0\mathbf{K}N - 1, 0\mathbf{K}N - 1)$ .

Дискретное преобразование в блочной форме может быть записано

$$F(X(m, n, u, v)) = C(u)C(v) \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(x(m, n, j, k)) F_B(u, v, j, k), \quad (1)$$

Где  $m = 1, \mathbf{K}, M$  – номер блока в столбце;

$n = 1, \mathbf{K}, L$  – номер блока в строке;

$j, k$  – индексы внутри блока;

$u, v$  – индексы в частотном домене;

$f(x(m, n, j, k))$  – исходный двумерный набор данных;

$x(m, n, j, k)$  – функция, определяющая выбор блока для чтения;

$F(X(m, n, u, v))$  – двумерный выходной набор данных;

$X(m, n, u, v)$  – функция, определяющая выбор блока для записи;

$F_B(u, v, j, k)$  – набор базисных функций;

$C(u), C(v)$  – коэффициенты, зависящие от набора базисных функций.

Выборка блока данных при традиционном подходе будет следующая

$$x(m, n, j, k) = mN + j, nN + k. \quad (2)$$

Выборка блока данных для записи при традиционном подходе

$$X(m, n, u, v) = mN + u, nN + v. \quad (3)$$

Обратное преобразование записывается следующим образом

$$f(x(m, n, u, v)) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} C(u)C(v) F(X(m, n, j, k)) F_B(u, v, j, k). \quad (4)$$

Выражения (2) и (3) используются в обратном преобразовании в левой и правой части соответственно. Это позволяет производить обратное преобразование в любом случае, если  $m$ ,  $n$  фиксированы.

Базовый набор функций для дискретного косинусного преобразования будет следующий

$$F_B(u, v, j, k) = \cos \frac{\pi u(2j+1)}{2N} \phi \cos \frac{\pi v(2k+1)}{2N} \phi.$$

Наш подход подразумевает изменение выборки блоков и направление их обработки в следующем виде [8,9]. Для входного набора

$$x_i(m, n, j, k) = \begin{cases} x_1((2m+1)N - (j+1), (2n+1)N - (k+1)) \\ x_2((2m+1)N - (j+1), (2n+1)N + k) \\ x_3((2m+1)N + j, (2n+1)N - (k+1)) \\ x_4((2m+1)N + j, (2n+1)N + k) \end{cases} \quad (8)$$

для выходного набора

$$X_i(m, n, u, v) = \begin{cases} X_1((2m+1)N - (u+1), (2n+1)N - (v+1)) \\ X_2((2m+1)N - (u+1), (2n+1)N + v) \\ X_3((2m+1)N + u, (2n+1)N - (v+1)) \\ X_4((2m+1)N + u, (2n+1)N + v) \end{cases}. \quad (9)$$

Производилось моделирование стандартных наборов тестовых изображений с вращением фигуры на 5 градусов в диапазоне 360 градусов. Примеры тестовых изображений представлены на рис. 1.

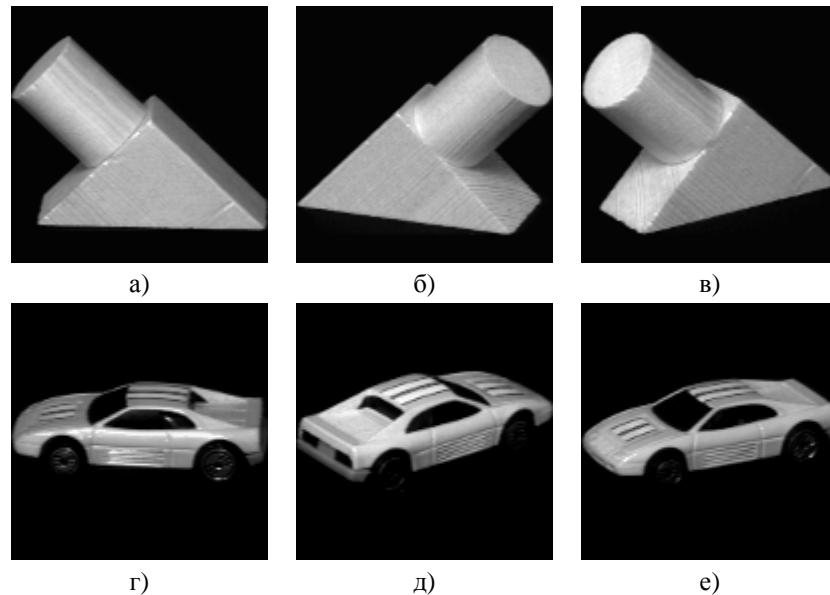


Рис. 1. Пример тестовых изображений

Рассчитывался коэффициент корреляции для частотного диапазона от  $1 \times 1$  до  $N/2 \times N/2$ . Производились сравнения исходного тестового изображения (поворот 0 градусов) с изображениями с углами поворота от 0 до 360 градусов. Коэффициент корреляции для традиционного и предлагаемых подходов представлены на рис. 2.

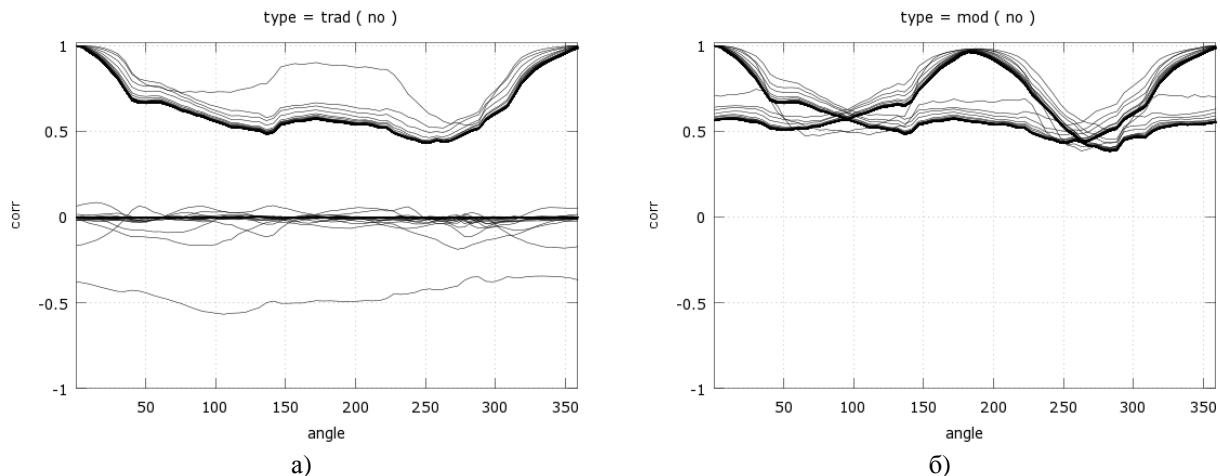


Рис. 2. Коэффициенты корреляции двух изображений для традиционного (а) и предлагаемого(б) подхода

При традиционном подходе коэффициент корреляции близок к единице (более 0,8) при углах от 0 до 10 и от 350 до 360 градусов. Предлагаемый подход дополнительно имеет коэффициент корреляции близкий к единице от 170 до 190 градусов.

Данный эффект позволяет расширить диапазон сравнения изображений тестовой фигуры в области поворота, что снижает использование вычислительных ресурсов или времени поиска.

#### Библиографический список

1. R. Bracewell, The Fourier Transform and Its Applications, Boston: McGraw-Hill, 2000.
2. R. Gonzalez, R. Woods, Digital image processing. Addison-Wesley, 2002.
3. E. Ifeachor, B. Jervis, Digital signal processing. A practical approach. Addison-Wesley Publishers Ltd, 1993.
4. A. Oppenheim, R. Schafer, J. Buck, Discrete-time signal processing. Upper Saddle River, N.J.: Prentice Hall, 1999.
5. B. Shen, I. K. Sethi, and V. Bhaskaran, “DCT convolution and its application in compressed domain,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, 1998, pp. 947-952.
6. C. Yim, “An efficient method for DCT-domain separable symmetric 2-D linear filtering,” IEEE Trans. Circuits Syst. Video Technol., vol. 14, pp. 517-521, Apr. 2004.
7. R. Kresch, N. Merhav, “Fast DCT domain filtering using the DCT and the DST”, IEEE transactions on Image Processing, vol.8(6), 1999, pp.821-833.
8. Шоберг А.Г. Повышение устойчивости к повороту при разделимом частотном преобразовании изображений // Вопросы радиоэлектроники, сер. Техника телевидения, 2015, № 6(23), С. 75-80
9. Shoberg A, Sai S, Shoberg K. Frequency transform basis with even symmetry elements. International journal of applied mathematics and informatics, Volume 10, 2016 pp 101-104

#### AN IMAGES COMPARISON IN FREQUENCY DOMAIN WITH BLOCK FORM OF TRANSFORM

**Shoberg A.G, Shoberg K.A.**  
Pacific National University

*It's considering problems of image comparison based on block frequency transform, performed with different ways. We give mathematical definition of our approach. We show results of the comparison of traditional and proposed approach to integral transform.*

**Integral - transform, modification scheme, image comparison, block form**

**Н. А. Андриянов**  
Ульяновский государственный технический университет

## Сжатие изображений на основе моделей случайных полей с изменяющимися параметрами

*В статье предложен метод сжатия изображений, базирующийся на разбиении исходного изображения на участки и идентификации параметров моделей по яркостным характеристикам этих участков. При этом восстановление предложено выполнять путем реализации модели с полученными параметрами.*

**Дважды стохастические изображения, идентификация параметров, сжатие изображений, восстановление изображений**

В настоящее время происходит постоянное увеличение объема хранимых и передаваемых изображений и их последовательностей. Это связано с широким распространением различных методов регистрации изображений, в том числе гиперспектральной и мультиспектральной съемки, а также с бурным развитием Интернета вещей [1,2]. Например, при обработке спутниковых изображений [3] объемы многозонального изображения могут достигать десятков гигабайт. Кроме того, развитие технологий в области телевидения позволяет сегодня передавать высококачественную картинку без существенных задержек. Понятно, что для этого требуется применение эффективных методик сжатия изображений, так как разрешение передаваемых картинок также достаточно велико.

При этом известно множество алгоритмов сжатия изображений, которые применяются в современных системах. К ним относятся, например, алгоритмы, основанные на вейвлет-преобразовании [4] или фрактальные методы сжатия [5]. При этом основная идея всех алгоритмов заключается в поиске некой избыточной информации, которая может быть безболезненно отброшена.

Однако развитие в последнее время математических моделей изображений с изменяющимися параметрами [6-8] совместно с идеей о наличии на изображении однородных областей позволяет сделать вывод о том, что сжатие может быть выполнено путем оценивания параметров выбранной модели или моделей в этих областях. Так, в работе [9] было предложено производить восстановление изображений на основе дважды стохастической модели случайных полей.

Рассмотрим следующую идею сжатия. Пусть имеется изображение размером  $M_1 \times M_2$ . Разобьем это изображение на области размером  $N_1 \times N_2$ . При этом необходимо выполнить условие о том, что все области не пересекаются друг с другом, т.е. выбранные размеры областей должны быть соответствующей кратности. В противном случае можно использовать параметры  $N_3 \times N_4$  для краевых ситуаций.

На следующем шаге выполним идентификацию параметров модели в каждой из полученных областей. В простейшем случае может быть использована двумерная авторегрессия первого порядка, которая записывается в следующем виде

$$x_{ij} = r_x x_{i-1,j} + r_y x_{i,j-1} - r_x r_y x_{i-1,j-1} + \sqrt{s_x^2 (1 - r_x^2) (1 - r_y^2)} \epsilon_{ij}, \quad (1)$$

где  $r_x$  и  $r_y$  - коэффициенты корреляции по строке и по столбцу соответственно;  $S_x^2$  - дисперсия случайного поля  $\{X\}$ ;  $X_{ij}$  - гауссова случайная добавка с нулевым математическим ожиданием и единичной дисперсией.

Отметим, что модель (1) имеет нулевое среднее значение, поэтому в ее каждой точке, вообще говоря, необходимо преобразование  $\tilde{x}_{ij} = x_{ij} + m_x$ , где параметр  $m_x$  характеризует среднее значение случайного поля.

Таким образом, необходимо определить 4 параметра модели для каждой области. Рассмотрим это на примере изображения с градациями серого. Тогда выполним идентификацию для:

$$1) \text{математического ожидания области } m_x = \frac{1}{N1 \cdot N2} \sum_{k=1}^{N1} \sum_{l=1}^{N2} I_{kl}, \text{ где } I - \text{представляет}$$

собой совокупность значений яркости изображения, которое требуется сжать;

$$2) \text{дисперсии области } S_x^2 = \frac{1}{(N1 - 1)(N2 - 1)} \sum_{k=1}^{N1} \sum_{l=1}^{N2} (I_{kl} - m_x)^2;$$

$$3) \text{коэффициента корреляции по строке } r_x = \frac{1}{S_x^2 (N1 - 1) N2} \sum_{k=2}^{N1} \sum_{l=1}^{N2} (I_{kl} - m_x)(I_{k-1,l} - m_x);$$

$$4) \text{коэффициента корреляции по столбцу } r_y = \frac{1}{S_x^2 (N1 - 1) (N2 - 1)} \sum_{k=1}^{N1} \sum_{l=2}^{N2} (I_{kl} - m_x)(I_{k,l-1} - m_x).$$

Помимо полученный выражений, для последующего восстановления необходимо значения первого параметра исходного изображения  $I_{11}$ . Сразу стоит отметить, что повышение порядка авторегрессии, основанное на модели с кратными корнями характеристических уравнений [10], не приведет к увеличению числа используемых параметров. Тогда, принимая одинаковым количество бит для каждого из параметров, получаем следующее отношение сжатия:

$$\frac{\text{Compressed}}{\text{Full}} = \frac{N_{\text{nap}}}{N1 \cdot N2} \cdot 100\%, \quad (2)$$

где  $N_{\text{nap}}$  - число параметров (в рассмотренном случае  $N_{\text{nap}} = 5$ ).

Из выражения (2) очевидно, что увеличивая размеры областей можно значительно сжать изображение. Однако для наилучшего восстановления необходимо соблюдать требования по однородности этих областей. На рисунке 1 представлена зависимость отношения сжатия от размеров квадратной области для рассмотренной модели изображений.

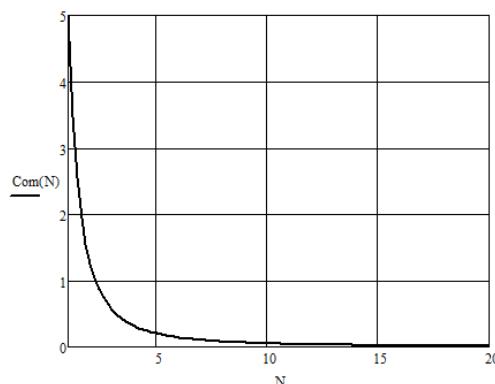


Рис. 1. Зависимость отношения сжатия от размеров областей деления исходного изображения

Таким образом, оптимальными в плане сжатия и вычислительных затрат будут размеры  $5 < N < 15$ , так как дальнейшее увеличение размеров  $N$  не приводит к значительному повышению эффективности сжатия и увеличивает ошибки при восстановлении.

На рисунке 2 показаны примеры сжатия некоторых изображений с их последующим восстановлением при разных параметрах сжатия. При этом исходное изображение предварительно было сжато и сохранено в формате jpeg. На рисунке 2а показана обработка изображения природных объектов, на рисунке 2б - обработка изображения человека. Слева направо: исходное изображение, сжатие при  $N = 3$ , сжатие при  $N = 5$ , сжатие при  $N = 10$ .

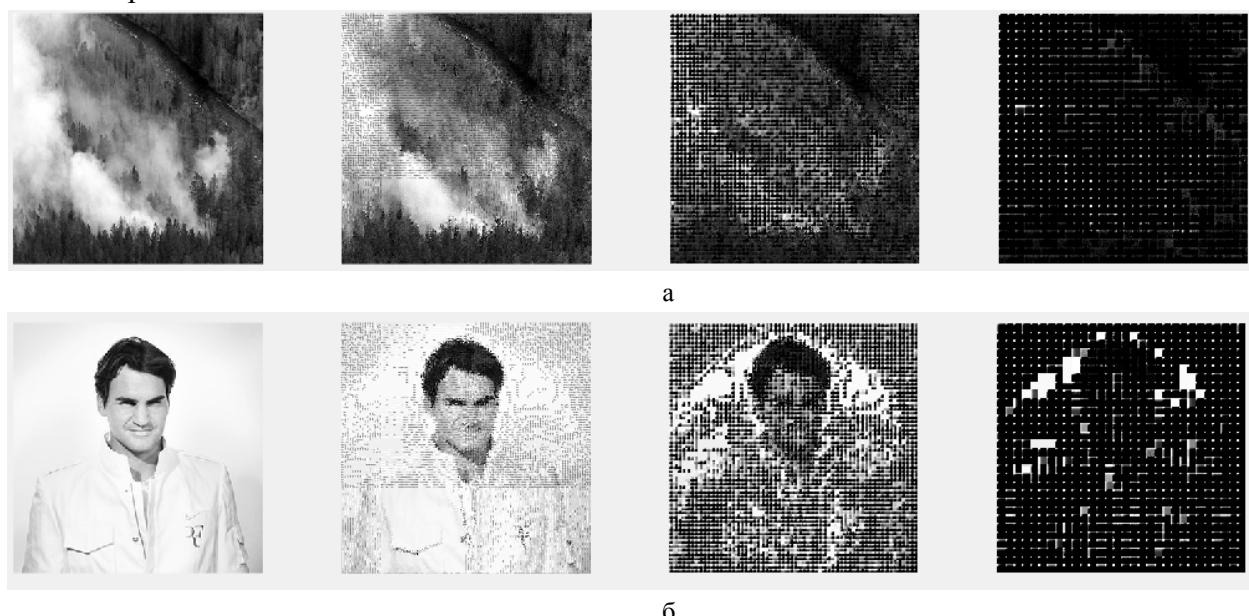


Рис. 2. Сжатие и восстановление на базе моделей изображений с изменяющимися параметрами

Анализ результатов восстановления (рисунок 2) показывает, что при восстановлении дисперсии ошибок составили следующие величины при разных  $N$ :

а: 3 - 0.6275, 5 - 1.1863, 10 - 1.9065;

б: 3 - 1.4832, 5 - 1.6722, 10 - 1.7253.

Размер исходного изображения 300x300.

Полученные результаты, как показывает исследование, значительно зависят от типа исходного изображения, а также от коэффициента сжатия. При этом наилучшую точность восстановления удалось получить для случая небольших  $N$ , что объясняется большим количеством опорных точек при таком восстановлении. В целом, предложенный подход обеспечил удовлетворительный результат только для изображения рисунок 2а. Однако анализируя изображения на рисунке 2б, можно сделать вывод о том, что предложенная методика может быть использована для определения границ объектов на изображении. Такая методика в этом случае может повысить качество сегментации изображений в случае совместного использования с сегментацией по цветности.

Таким образом, был предложен алгоритм сжатия и восстановления изображения, базирующийся на реализации моделей изображений с идентифицированными параметрами.

### Библиографический список

1. Dave Evans The Internet of Things. How the Next Evolution of the Internet Is Changing Everything // Cisco White Paper. Cisco Systems, 11 April 2011, 11 p.
2. Андриянов Н.А., Дементьев В.Е. Низкое энергопотребление в беспроводных сетях: BluetoothLowEnergy или ANT? // Радиоэлектронная техника. 2015. №2 (8) С. 254-257
3. Васильев К.К., Дементьев В.Е., Андриянов Н.А. Восстановление спутниковых изображений и оценка параметров с помощью моделей со сложной структурой // Компьютерный анализ изображений: Интеллектуальные решения в промышленных сетях (CAI-2016) сборник научных трудов по материалам I Международной конференции. Уральский федеральный университет имени первого Президента России Б.Н. Ельцина; Под общей редакцией А. Г. Тягунова. 2016. С. 155-160.
4. Штарт Г.-Г. Применение вейвлетов для ЦОС. — М.: Техносфера, 2007 - 181 с.
5. Пат. Michael F. Barnsley, Alan D. Sloan, John H. Elton, Charles S. Moreman, Guy A. Primiano Fractal transform compression board US 5430812 A
6. Vasil'ev K.K., Dement'ev V.E., Andriyanov N.A. Doubly stochastic models of images // Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications). 2015. Т. 25. № 1. Р. 105-110.
7. Андриянов Н.А., Дементьев В.Е. Модели сигналов со случайными параметрами // Актуальные вопросы технических наук в современных условиях. Сборник научных трудов по итогам международной научно-практической конференции. 2015. С. 37-41
8. Андриянов Н.А. Дискретные дважды стохастические авторегрессионные модели случайных полей // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2014. № 1 (9). С. 69-72.
9. Vasil'ev K.K., Dement'ev V.E., Andriyanov N.A. Application of mixed models for solving the problem on restoring and estimating image parameters // Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications). 2016. V. 26. № 1. P. 240-247.
10. Васильев К.К., Гаврилина Ю.Н., Андриянов Н.А. Эффективность фильтрации авторегрессионной модели с кратными корнями характеристических уравнений // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2016. № 1 (10). С. 130-133.

### COMPRESSION OF IMAGES BASED ON MODELS OF RANDOM FIELDS WITH VARYING PARAMETERS

**N.A. Andriyanov**

*Ulyanovsk State Technical University (Russia)*

*The article suggests a method of image compression based on the division of the original image into sections and identification of the model parameters by the brightness characteristics of these areas. In this case, it is proposed to perform the restoration by implementing a model with the parameters obtained.*

**Doubly stochastic images, parameter identification, image compression, image restoration**

**Е. П. Петров, Н. Л. Харина, П. Н. Сухих**  
Вятский государственный университет, г. Киров

## Метод внутрикадрового предсказания

*В работе предлагается метод устранения избыточности цифровых изображений на основе теории условных марковских процессов. Метод может использоваться в качестве внутрикадрового предсказания в кодеках сжатия видео или в качестве самостоятельного алгоритма.*

**Внутрикадровое предсказание, статистическая избыточность, двумерная цепь Маркова, матрица вероятностей перехода**

Одной из важнейших практических задач в области цифрового телевидения является задача сокращения скорости передачи двоичных символов и, соответственно, требуемой полосы частот канала связи. Эта задача может быть решена путем уменьшения избыточности информации, передаваемой в телевизионном сигнале без заметного для глаза ухудшения качества воспроизведенного телевизионного изображения. Уменьшение избыточности обеспечивает также сокращение требуемого объема памяти запоминающих устройств при записи видео или отдельных изображений.

В современных стандартах сжатия видео для устраниния избыточности предусмотрены режимы внутрикадрового (Intra) и межкадрового (Inter) предсказания. Например, стандартом H.264/AVC в режиме Intra для блоков  $4 \times 4$  используются 9 режимов и 4 режима для блоков  $16 \times 16$ . В H.265/HEVC реализовано 35 режимов предсказания [1,2].

В работе предлагается метод устранения избыточности цифровых изображений (ЦИ) на основе теории условных марковских процессов [3,4]. Будем полагать, что каждый кадр является g-разрядным ЦИ и аппроксимируется двумерной цепью Маркова с несколькими состояниями  $N = 2^g$  с вектором вероятностей начальных состояний  $P = \|p_1, p_2, \dots, p_N\|^T$  и матрицами вероятностей переходов (МВП)  ${}^1P$  и  ${}^2P$  из состояния  $M_i$  в соседнее состояние  $M_j$  ( $i, j \in N$ ) по горизонтали и вертикали соответственно [3]:

$${}^1\Pi = \begin{vmatrix} {}^1p_{11} & {}^1p_{12} & \dots & {}^1p_{1N} \\ {}^1p_{21} & {}^1p_{22} & \dots & {}^1p_{2N} \\ \dots & \dots & \dots & \dots \\ {}^1p_{N1} & {}^1p_{N2} & \dots & {}^1p_{NN} \end{vmatrix}, \quad {}^2\Pi = \begin{vmatrix} {}^2p_{11} & {}^2p_{12} & \dots & {}^2p_{1N} \\ {}^2p_{21} & {}^2p_{22} & \dots & {}^2p_{2N} \\ \dots & \dots & \dots & \dots \\ {}^2p_{N1} & {}^2p_{N2} & \dots & {}^2p_{NN} \end{vmatrix}. \quad (1)$$

Разработка метода сжатия ЦИ в таком представлении требует больших вычислительных ресурсов, особенно для ЦИ с большим разрешением. Поэтому предлагается разделить g-разрядное ЦИ на g разрядных двоичных изображений (РДИ) (рис.1).

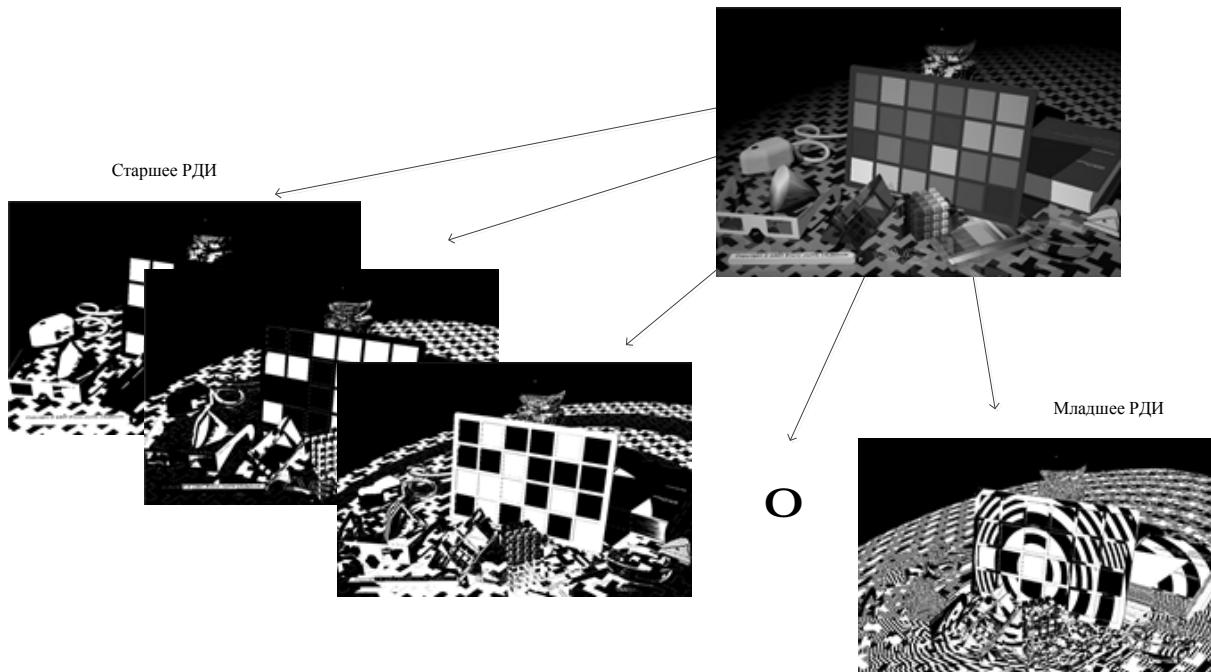


Рис. 1. Пример разделения тестового ЦИ на РДИ

Каждое РДИ представляет собой двумерную цепь Маркова с двумя ( $N = 2$ ) равновероятными состояниями ( $p_1 = p_2$ ) и МВП  ${}^1\mathbf{P}$  и  ${}^2\mathbf{P}$  из состояния  $M_i$  в соседнее состояние  $M_j$  ( $i, j \in N$ ) по горизонтали и вертикали соответственно [3]:

$${}^1\mathbf{P} = \begin{bmatrix} {}^1p_{11} & {}^1p_{12} \\ {}^1p_{21} & {}^1p_{22} \end{bmatrix}, \quad {}^2\mathbf{P} = \begin{bmatrix} {}^2p_{11} & {}^2p_{12} \\ {}^2p_{21} & {}^2p_{22} \end{bmatrix}. \quad (2)$$

РДИ представляет марковское случайное поле размером  $m' n$  на несимметричной полуплоскости, полученной классической разверткой с левого верхнего угла слева направо вниз (рис.2).

Состояние элемента  $M_{i,j}$  ( $i \in m, j \in n$ ) зависит только от известных элементов некоторого подмножества  $L_{i,j}$ , называемого окрестностью элемента  $M_{i,j}$ . Лучшим образом удовлетворяющее условию каузальности является конфигурация окрестности (рис. 3) [3]:

$$L_{i,j} = \{M_{i,j-1}, M_{i-1,j}, M_{i-1,j-1}\}. \quad (3)$$

Вероятности перехода от состояний элементов  $L_{i,j}$  к состоянию  $M_{i,j}$  образуют МВП вида:

$$\mathbf{P} = \begin{bmatrix} p_{111} & p_{121} & p_{211} & p_{221} \\ p_{112} & p_{122} & p_{212} & p_{222} \end{bmatrix}^T. \quad (4)$$

Элементы матрицы  $\mathbf{P}$  (4) связаны с элементами МВП (2) следующими соотношениями:

$$\begin{aligned} p_{111} &= \frac{{}^1p_{11} \times {}^2p_{11}}{3p_{11}}, \quad p_{112} = \frac{{}^1p_{12} \times {}^2p_{12}}{3p_{11}}, \quad p_{121} = \frac{{}^1p_{11} \times {}^2p_{21}}{3p_{12}}, \quad p_{122} = \frac{{}^1p_{12} \times {}^2p_{22}}{3p_{12}}, \\ p_{211} &= \frac{{}^1p_{21} \times {}^2p_{11}}{3p_{21}}, \quad p_{212} = \frac{{}^1p_{21} \times {}^2p_{12}}{3p_{21}}, \quad p_{221} = \frac{{}^1p_{22} \times {}^2p_{21}}{3p_{22}}, \quad p_{222} = \frac{{}^1p_{22} \times {}^2p_{22}}{3p_{22}}. \end{aligned}$$

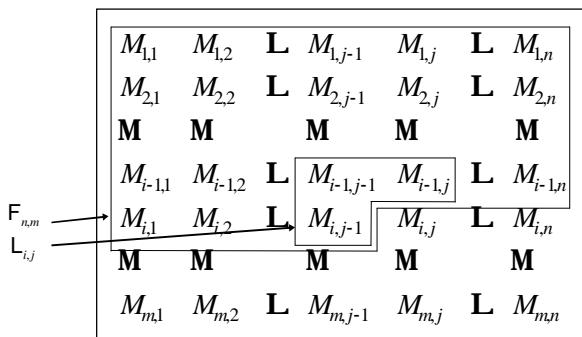


Рис. 2. Модель РДИ

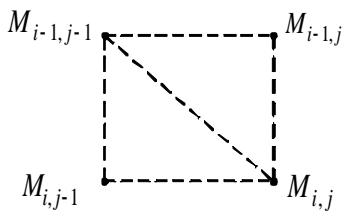


Рис. 3. Окрестность предсказываемого элемента РДИ

где  ${}^3P_{ii}$  - элементы дополнительной МВП  ${}^3P = {}^1P \cup {}^2P \cup \emptyset$ .

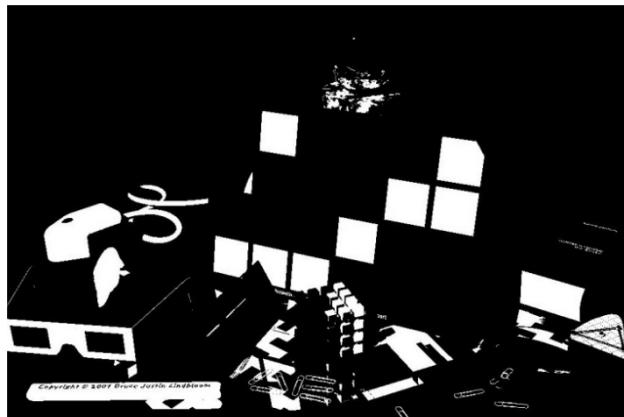
Основой предложенного метода является процедура предсказания состояния каждого пикселя РДИ на основе известной окрестности  $L_{i,j}$  и МВП (4) [4,5]. В результате предсказания формируется РДИ по следующему правилу – если состояние пикселя предсказано верно, то этот пиксель в РДИ принимает состояние 0, иначе 1. При восстановлении РДИ нулевые пиксели заменяются состояниями по результатам предсказания, а единичные – неправильно предсказанные – инверсией результата предсказания.

#### Алгоритм прогнозирования элементов

1. ЦИ разбивается на g РДИ;
2. Для каждого РДИ из классификатора выбираются МВП вида (2) и (4);
3. Прогнозируется элемент  $\hat{M}_{ij}$  на основе МВП (4) и окрестности  $L_{i,j}$ .
4. Если  $M_{ij} = \hat{M}_{ij}$ , то элемент предсказан верно и  $M_{ij} = \hat{M}_{ij}$ ;
5. Если  $M_{ij} \neq \hat{M}_{ij}$ , элемент  $M_{ij}$  предсказан неверно;
6. Формирование РДИ после предсказания;

При восстановлении изображения хранящиеся в файле пиксели используются как опорные для предсказания остальных, точность предсказания составляет 100%.

На рис. 4 приведены примеры предсказания для двух старших РДИ тестового ЦИ. Белым цветом выделены неверно предсказанные биты.



а. РДИ 8



с. РДИ после предсказания, коэффициент сжатия РДИ = 25

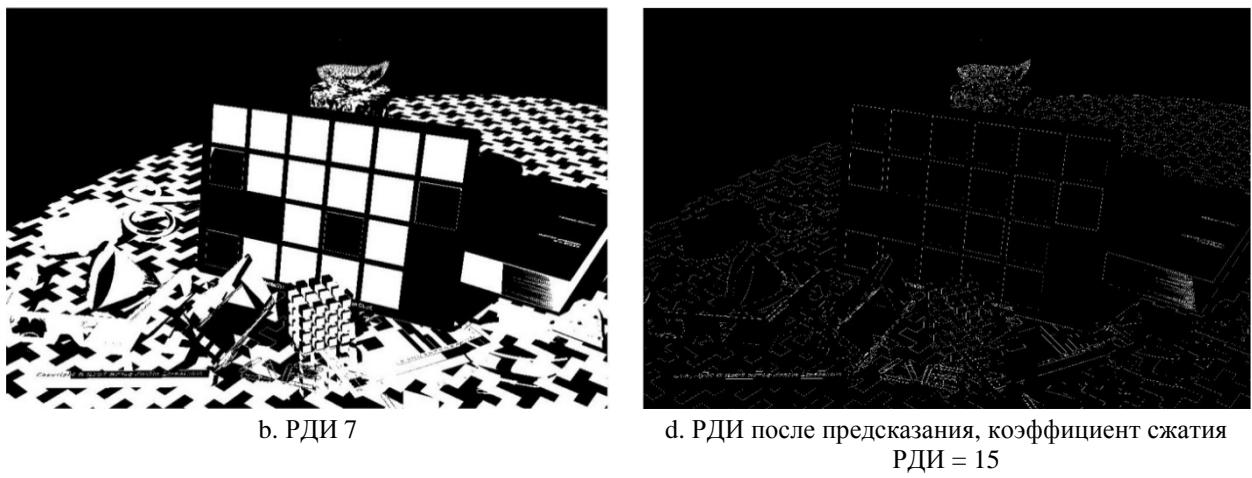


Рис. 4. Результаты предсказания для двух старших РДИ тестового ЦИ

Для исследования эффективности предложенного метода построен алгоритм сжатия ЦИ без потерь, для этого после процедуры предсказания добавлены шаги [4,5]:

7. РДИ сжимается кодом RLE в одномерный поток;
8. Одномерный поток сжимается кодом Хаффмана;
9. Сжатые данные записываются в файл.

Для оценки эффективности предложенного алгоритма (MRK) произведено сравнение с наиболее известными алгоритмами сжатия без потерь PNG, JPEG-LS, JPEG 2000 на статистике 100 ЦИ. Результаты сжатия снимков предложенным методом и аналогами представлены в таблице 1. При исследовании предложенный метод применялся в режиме последовательной обработки РДИ из-за трудности реализации режима параллельной обработки.

Результаты исследований алгоритмов сжатия

Таблица 1.

Алгоритм сжатия	Коэффициент сжатия реальных снимков	Скорость обработки, Мб/с
PNG	1,86	3,72
JPEG-LS	1,95	15,67
JPEG 2000	1,78	3,88
MRK	1,75	44,8

Анализ результатов показывает, что данный алгоритм эффективно «сжимает» старшие РДИ, а младшие РДИ практически не поддаются сжатию, так как они представляют собой случайное поле, близкое по структуре к белому гауссовскому шуму с вкраплением мелких деталей исходного изображения, которыми не всегда можно пренебречь. Поэтому для младших РДИ разработан модифицированный алгоритм, позволяющий выделить и удалить области, близкие к БГШ, с последующим заполнением выборками БГШ при восстановлении изображения [5], который в данной работе не рассматривается.

Достоинствами алгоритма является возможность обработки ЦИ неограниченной разрядности, отсутствие вычислительных операций и возможность применения параллельного вычисления, так как обработка РДИ производится независимо друг от друга. Метод может использоваться в качестве внутrikадрового предсказания в кодеках сжатия видео или в качестве самостоятельного алгоритма.

1. Дворкович В.П., Дворкович А.П. Цифровые видеоинформационные системы (теория и практика). М.: Техносфера, 2012. - 1009с.
2. Дворкович В.П., Дворкович А.П. Метрологическое обеспечение видеоинформационных систем. М.: Техносфера, 2015. - 784с.
3. Петров Е.П., Медведева Е.В., Харина Н.Л. Математическая модель цифровых полутоновых изображений Земли из космоса // Актуальные проблемы ракетно-космической техники. 2011. С. 179-185.
4. Петров Е.П., Харина Н.Л., Сухих П.Н. Метод сжатия многоразрядных спутниковых снимков без потерь // Современные проблемы дистанционного зондирования Земли из космоса, № 2, 2016. – с. 203-210
5. Петров Е.П., Харина Н.Л. Способ сжатия изображений //Патент на изобретение № 2573770

## METHOD OF AN INTRAFRAME PREDICTION

**E.P.Petrov, N.L.Kharina, P.N. Sukhikh**

Vyatka State University, Kirov

*In this paper the method of redundancy reduction of digital images on the basis of the theory of the conditional Markov processes is offered. The method can be used as an intraframe prediction in codecs of video compression or as an independent algorithm.*

**Intraframe prediction, statistical redundancy, two-dimensional Markov chain, matrix of transition probabilities**

**Нгуен Ван Чыонг, Доан Тиен Бан**

Санкт-Петербургский национальный исследовательский университет  
информационных технологий, механики и оптики (Университет ИТМО)

# Исследование блочных алгоритмов кодирования преобразованием в стандартах сжатия мультимедийных данных

*Работа посвящена исследованию распространенных блочных алгоритмов кодирования преобразованием остаточных изображения (после межкадрового и/или внутрикадрового кодирования). При этом предложен лучший алгоритм для применения в дальнейшем исследовании.*

**Дискретное косинусное преобразование, дискретное преобразование Хартли, преобразование Уолша-Адамара**

## Введение

Как известно, существуют преобразования на основе всего изображения (дискретное wavelet-преобразование DWT (Discrete Wavelet Transform)) и блочные преобразования.

Wavelet – преобразование применяется по строкам и по столбцам для каждого уровня разложения. Достоинствами метода являются: отсутствие блокинг-эффекта (без каких-либо блоков, так как преобразование применяется к полному изображению, в декодированном изображении, естественно, отсутствует блокинг-эффект); более высокая степень сжатия (преобразование всего входного сигнала позволяет использовать корреляцию между всеми соседними выборками, а не только между выборками одного и того же блока данных. Это обычно позволяет достичь более высоких степеней сжатия); масштабируемое разрешение (при использовании двоичных декомпозиций, используемых в DWT, можно увеличить или уменьшить пространственное разрешение восстановленных данных, просто увеличив или уменьшив количество декодированных коэффициентов. Эта функция масштабирования качества и пространственного разрешения очень полезна для сжатия изображений (и видео)).

Однако, данному решению присущ серьезный недостаток – это высокая сложность. Выполнение преобразования всего входного сигнала вместо его разделения на меньшие блоки имеет более высокую стоимость с точки зрения сложности. При большем количестве входных выборок число операций, необходимых для выполнения преобразования, также увеличивается.

По этой причине, в нашей работе рекомендовано применение алгоритмов преобразования на основе блоков (блочные преобразования) в замену DWT. Алгоритмов этого типа много, в данной статье рассмотрены только широко распространенные алгоритмы, применяемые в современных системах кодирования мультимедийных данных.

## Блочное преобразование

Блочные преобразования работают с квадратными блоками изображения, элементами которых служат сэмплы (обычно это сэмплы самого изображения или величины разностей из кадра – прогноза), и после ряда операций порождает равносторонний блок коэффициентов. Любой блок изображения можно восстановить с помощью линейной комбинации  $N \times N$  базисных шаблонов, где базисные шаблоны

умножаются на соответствующие весовые множители (коэффициенты преобразования). Рассмотрим некоторые важные блочные алгоритмы кодирования:

### **Дискретное косинусное преобразование DCT (Discrete Cosine Transform).**

Дискретное двухмерное косинусное преобразование определяется соотношением [1]:

$$Y(u, v) = \frac{2}{N} C(u) C(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X(i, j) \cos \frac{(2i+1)up}{2N} \cos \frac{(2j+1)vq}{2N},$$

где  $N \times N$  размеры матрицы преобразования,  $C_w = \sqrt{\frac{1}{2}}$  ( $w=0$ ),  $C_w = 1$  ( $w>0$ ),  $X(i, j)$  - значение яркости кадра в точке с координатами  $(i, j)$ , а  $Y(u, v)$  - коэффициенты разложения.

Достоинства:

Быстрая реализация с использованием только реальных вычислений. DCT может быть реализован с использованием быстрых алгоритмов, которые могут значительно уменьшить количество операций и, следовательно, вычислительную сложность. В дополнение к этому, DCT требует только реальных вычислений, избегая манипулирования комплексными числами, как в дискретное преобразование Фурье.

Уменьшение блогкинг-эффекта. Свойства DCT с точки зрения периодичности помогают избежать граничных разрывов; это может значительно уменьшить появление блогкинг-эффекта.

Недостаток:

Существует один существенный недостаток DCT. Хотя входные данные из предварительно обработанных блоков  $8 \times 8$  являются целочисленными, выходные значения обычно являются вещественными. Таким образом, нам нужен шаг квантования, чтобы принимать некоторые решения относительно значений в каждом блоке DCT и выдавать результат, который является ценным для интергебра.

### **Дискретное преобразование Хартли DHarT (Discrete Hartley Transform).**

Дискретное двухмерное преобразование Хартли [2] определяется соотношением:

$$H(u, v) = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) cas\left(\frac{2pu i}{N} + \frac{2pv j}{N}\right),$$

где  $H(u, v)$  - коэффициенты разложения, а  $casx = \sin x + \cos x$ .

### **Дискретное преобразование Хаара DHaaT (Discrete Haar transformation).**

Матрица ядра преобразования Хаара для  $N=8$  имеет вид:

$$H_8 = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{vmatrix}$$

При этом анализируется не весь спектр, а его отдельные фрагменты. Данная идея подобна идее wavelet-преобразований.

### Преобразование Уолша-Адамара WHT(Walsh-Hadamard Transform).

Матрица ядра преобразования Уолша-Адамара для  $N = 2^m$  может быть описана как результат прямого умножения т матриц  $E_2$  размера 2x2 [3]:

$$A_N = \prod_{i=1}^w E_{2(i)},$$

где  $E_2 = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}$  - матрица порядка 2.

Основным преимуществом WHT является быстрая и простая реализация. Матрицы преобразования Адамара являются чисто вещественными, они содержат значения +1 или -1. Таким образом, WHT должен выполнять очень простые реальные операции, значительно уменьшая сложность преобразования. Более того, с использованием алгоритма WHT считается лучшим преобразованием с точки зрения сложности.

Основным недостатком WHT является скромное сжатие энергии. С точки зрения уплотнения энергии WHT не так эффективен, как альтернативные преобразования, такие как DCT; фактически, по сравнению со всеми другими преобразованиями, представленными в этой главе, WHT имеет худшую производительность уплотнения.

### Экспериментальные результаты

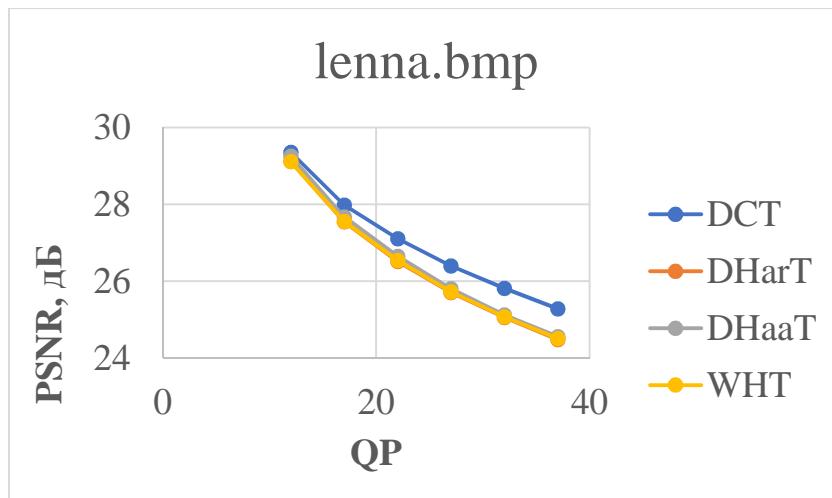


Рис. 1. Кривые зависимости PSNR от параметра квантования QP

На рис. 1 представлены результаты сравнения существующих блочных преобразований для тестового изображения “lenna.bmp” размером 512x512 пикселей. Для дальнейших исследований предлагается дискретное косинусное преобразование, как алгоритма с наибольшим коэффициентом сжатия.

В общем, методы преобразований на основе блоков обладают следующими преимуществами:

- значительная декорреляция сигнала ведет к сокращению избыточности;
- перераспределение энергии сигнала, что значительно повышает эффективность последующих этапов кодирования;
- возможность работы с многомерным сигналом как с комбинацией одномерных.

### Библиографический список

1. Konrad, J. Bayesian estimation of motion vector field / Konrad J., Dubois E. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1992. – 14. – 9. – P. 910-927.
2. Брейсуэлл, Р. Преобразование Хартли. Теория и приложения / Р. Брейсуэлл. – М.: Мир, 1990. – 225 с.

3. Глущик, Р. В. Процедуры распознавания и локализации объектов на изображении / Р. В. Глущик // Современные технологии. Труды молодых ученых ИТМО. – 2001. – С.106-109.

## **INVESTIGATION OF BLOCK TRANSFORM ALGORITHMS FOR MULTIMEDIA DATA COMPRESSION STANDARDS**

**Nguyen Van Truong, Doan Tien Ban**  
ITMO University

*The work deals with the study of common block coding algorithms by transforming the residual images (after interframe and/or intraframe coding). At the same time, the best algorithm is proposed for use in further research.*

**Discrete cosine transform, discrete Hartley transform, Walsh-Hadamard transform, Walsh-Hadamard transform**

**Ю. В. Кузнецов<sup>1</sup>, А. А. Щаденко<sup>2</sup>**

<sup>1</sup>СПбГИКиТ

<sup>2</sup>СПбГУПТД

## Оптимальное кодирование изображений в печати

*В допечатной однокадровой ТВ системе, в отличие от экранного отображения, исходный числовой массив многоуровневых отсчетов изображения преобразуют в двухуровневую битовую карту, чтобы управлять печатью в режиме да-нет (есть краска – нет краски). Частота выборки в этом повторном кодировании, задана т.н. линиатурой автотипного растра, а шкала квантования – числом знаков алфавита печатных элементов, присущим данной полиграфической технологии. Рассматриваются вопросы оптимальной организации этого процесса путем согласования локальных свойств исходного изображения с каналом передачи информации (печатью) и параметров результата передачи (оттиска) с характеристиками зрительного анализатора получателя.*

**Локально-адаптивное кодирование, детальность, четкость, ЧКХ зрения, качество изображения**

### Введение

Один из критериев оптимального кодирования — передача полезной получателю информации, но не той части исходного сообщения, которую он не в состоянии воспринять. Поэтому свойства источника согласуют со свойствами канала передачи, а параметры ее результата с характеристиками восприятия получателя. Примером служит ограничение частотного спектра цветоразностных сигналов в вещательном ТВ и в ряде способов сжатия цифровых изображений, где, учитывая снижение способности зрения различать тон и цвет с уменьшением размеров деталей, лишь яркостной компоненте предоставляют полную полосу частот или шкалу квантования.

Другой пример — диагональная выборка, сама по себе вдвое уменьшающая цифровой поток при сохранении качества изображения. Шахматная геометрия решетки пространственной дискретизации обоснована угловыми анизотропиями, присущими и источнику, и каналу передачи, и зрению [1-3]. Несмотря на то, что, в силу, наверное, «ортогональности» мышления, цифровые файлы сегодня представляют горизонтальными строками и вертикальными столбцами, шахматное расположение присуще, тем не менее, размещению зеленочувствительных сенсоров цифровых камер, печатных элементов оттисков, а также таких устройств низкого разрешения, как матричные принтеры.

Однако традиционная и цифровая печать далеко не оптимально организованы в том отношении, что даже линии толщиной в десятичные доли миллиметра здесь формально приписывают все 256 уровней квантования или количества каждой из четырех CMYK (голубой, пурпурной, желтой, черной) красок тогда как наблюдатель с трудом различит две-три градации яркости и затруднится сказать голубая она или зеленая. Эта объективно присутствующая на оттиске, но остающаяся втуне для получателя информация обозначена индексом 1 на рис. 1, схематически иллюстрирующем поток данных в системе: изобразительный оригинал (слайд или файл) — печатный канал — зрение.



Рис. 1. Результаты неоптимального кодирования изображения на допечатной стадии: 1 — присутствующие на оттиске, но невоспринимаемые различия тона и цвета мелких деталей; 2 — их геометрия и резкость, утраченные в результате растрового преобразования

Отмеченные там же разрушения (2) мелких деталей печатными элементами (растровыми точками) как результат автотипного преобразования, сменившего трудоемкое ручное гравирование в конце 19-го века, иллюстрирует правая часть рис. 2.

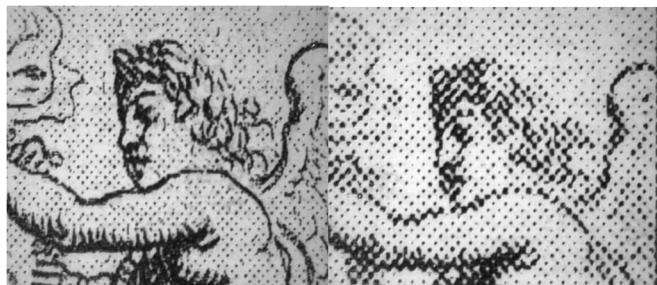


Рис. 2. Увеличенные фрагменты гравированного (слева) и автотипного (справа) оттиска

Коэффициент отражения точки на светочувствительном слое фотографии может изменяться в пределах всего своего интервала. По всему диапазону яркостей изменяется и интенсивность излучения элемента на ТВ экране или компьютерном мониторе (Рис. 3а). Однако принцип автотипии предполагает разнесение этих уровней по площади ячейки оттиска так, чтобы каждый ее субэлемент или микроточка отвечали своим почернением лишь на изменение между какими-либо соседними исходными уровнями квантования. Мера тона может быть распределена в таком периоде периодически с разной геометрией и законом монотонного возрастания/убывания или в случайном порядке с помощью весовых функций, показанных на рис. 3. (б, в).

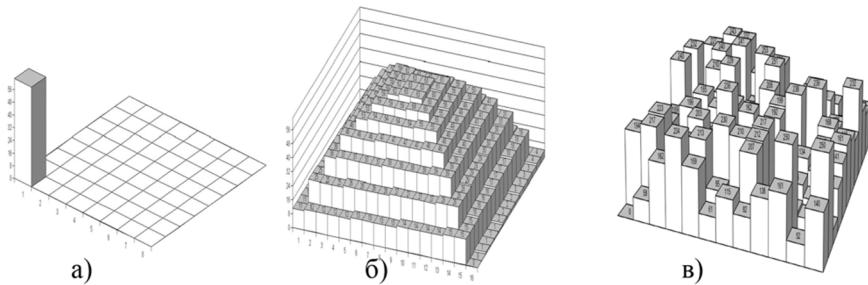


Рис. 3. Многоуровневую меру тона, целиком приписываемую в исходном файле одному элементу изображения (а), подвергают монотонной (б) или случайной (в) дисперсии в некотором пространственном периоде растрового оттиска

Набор (алфавит) растровых точек представляют срезами этих функций — битовыми картами, количественное соотношение и относительное расположение нулей и единиц в которых определяет площадь и форму печатных элементов и пробелов (Рис. 4).

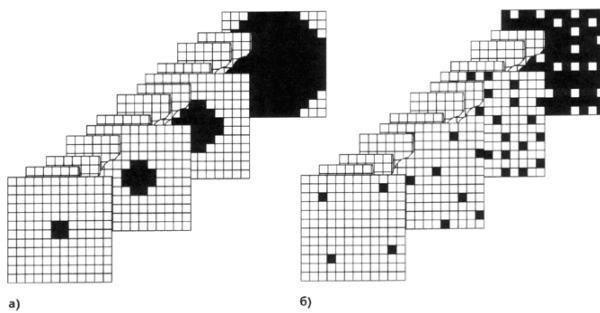


Рис. 4. Представление алфавита печатных элементов битовыми картами с упорядоченным (а) и случайным (б) распределением субэлементов в периоде раstra

С увеличением периода дисперсии растет плавность тонопередачи и контраст оттиска, но снижается его четкость. В этом заключается специфический, присущий автотипии конфликт в обеспечении этих важных показателей качества изображения. Подтвержденный печатной практикой приоритет контраста в выборе частоты раstra обусловливает десятикратное недоиспользование разрешающей способности в отношении четкости тоновой иллюстрации [4]. Даже при своей относительно высокой частоте, например, 70 лин/см растр передает на оттиске всего лишь 3,5 лин/мм. Однако та же печать, обладая разрешающей способностью 40 лин/мм, надежно воспроизводит линии толщиной 25 мкм в ее штриховом режиме, не нуждающемся в рассмотренной дисперсии. Зрение, в свою очередь, способно различать порядка 10 лин/мм. Еще на гравюрах - иллюстрациях книг середины 19 века размеры информационно значимых штрихов составляли менее 0,1 мм.

Существенно отличаясь друг от друга по своим свойствам, штриховые, полутонаовые и микроштриховые (растрированные печатные) оригиналы принадлежат к принципиально различным по своим статистическим свойствам категориям изображений. Поэтому в технологии «оцифровки» прошлых печатных изданий содержание каждой страницы сегментируют, как минимум, на текст и иллюстрацию, чтобы кодировать их по-разному, с лучшим качеством, сокращением времени и цифрового потока.

Весьма актуальной задачей в указанной связи остается устранение упомянутого выше разрыва между разрешением печати и автотипного метода передачи полутонов путем динамического изменения алгоритма кодирования по полю тонового изображения.

#### **Локально-адаптивная технология тоновой печати высокой четкости**

Участки полутонового оригинала могут в разной степени представлять собою собственно тоновое или штриховое изображение. Можно лишь формально отделить т.н. фоновую часть от рисунка, чтобы применить для них принципиально различные алгоритмы растирования, один из которых, соответствующа классическому автотипному методу, обеспечил бы надлежащую тонопередачу, а другой, штриховой — высокую геометрическую точность контуров и мелких деталей. Детали могут иметь разный контраст, их контуры могут быть как резкими, так и в разной степени размытыми. Присутствуют и вырождающиеся тонкие штрихи и контуры с монотонно изменяющимся по их протяженности контрастом и размытостью. Резкое изменение геометрии несущей изображение растровой структуры воспринимается здесь зрением как помеха.

В указанной связи отечественная технология адаптивного растирования предусматривают плавное взаимное замещение таких режимов под управлением

дифференциального оператора, характеризующего т. н. *детальность* (*busyness*) изображения в его некоторой локальной области [5; 6]. Наиболее общим критерием детальности служит мощность высокочастотных составляющих в спектре сигнала, представленного значениями данного и окрестных отсчетов. Вполне приемлемо его представляет и максимальный модуль разности  $q$  пары значений — их градиент в «окне» таких отсчетов.

С учетом этого параметра исходное значение тона  $S$  разделяют на две составляющие:  $S_1 = (1 - q) S$  и  $S_2 = q S$ .

Печатные элементы первой подлежащей запечатыванию площади (количества краски)  $S_1$  формируют по наперед заданной для всего изображения растровой функции, например, «горке» весовых значений (Рис. 3б и 5а), случайному распределению (Рис. 3 в) или по другому способу, гарантирующему плавную тонопередачу.

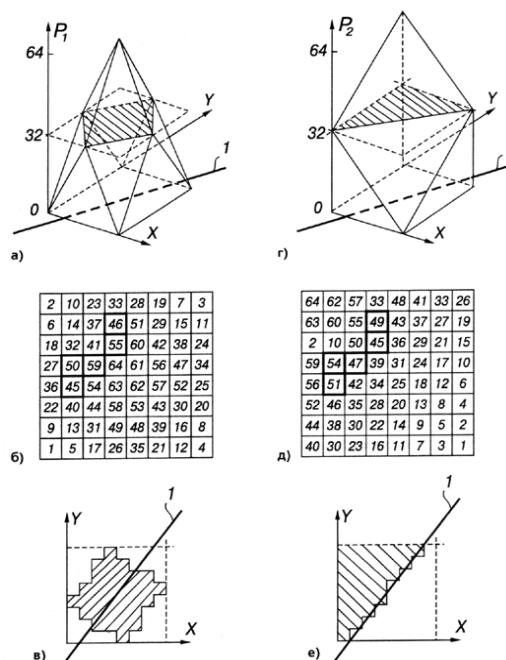


Рис. 5. Воспроизведение контура 1 полного контраста: в) — с использованием наперед заданных весовых значений (а, б); е) — на основе весовой функции (г, д), удовлетворяющей геометрии контура

В последней версии технологии Растворной печати высокой четкости (High Definition Halftone Printing — HDHP) для сигнала  $S_2$  используют растровые функции вдвое меньшего периода, но различной геометрии [7]. Благодаря этому, одновременно с повышением пространственной частоты отклика печати на изменения тона оригинала, более точна конфигурация деталей в пределах зон отсчетов исходного сигнала. Ансамбль из 32 таких дополнительных весовых функций представлен на рис. 6а. Одна из них и три из 36 знаков реализуемого с ее помощью алфавита печатных элементов показаны на рис. 6 (б и в). Подходящую по геометрии дополнительную функцию выбирают на основе сопоставления значений исходных отсчетов по их малой окрестности.

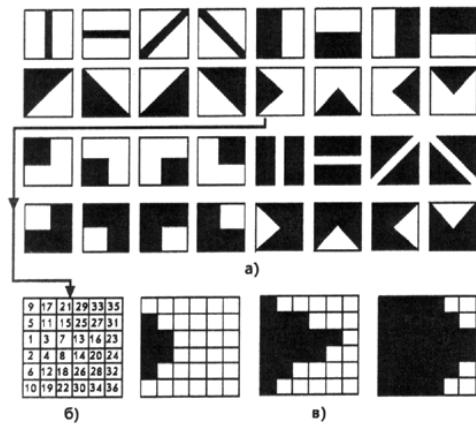


Рис. 6. Геометрии дополнительных растровых алфавитов (а), весовая функция (б) одного из них и примеры формируемых по ней печатных элементов (в) для трех значений тона

На стационарных, фоновых участках значение детальности  $q = 0$ , а  $S_1 = S$  и растровые точки формируют целиком по базовой весовой функции (Рис. 5а). Для контура или тонкой линии полного контраста, где  $q = 1$ , а  $S_2 = S$ , целиком используется только дополнительная весовая функция надлежащей геометрии. В результате контур не разрушается растровой точкой, а передается с полным использованием разрешения печати. «Мозаика» из более 1000 знаков, формируемых по ансамблю функций на рис. 6, позволяет воспроизводить на оттиске мелкие детали достаточно сложной конфигурации, практически полностью разрушающиеся в обычной печати, как показано на рис. 7.

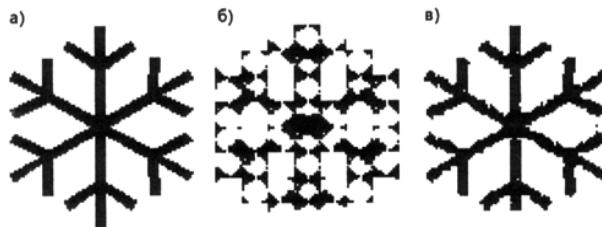


Рис. 7. Деталь сложной конфигурации в штриховом режиме (а), в тривиальном (б) и адаптивном (в) растировании

Для контура промежуточного контраста, с детальностью, например,  $q = 0,5$ ,  $S_1 = S_2$  и по базовой функции формируют только половину подлежащей запечатыванию площади, используя для другой знаки дополнительных алфавитов. Этим не только сохраняется средняя яркость в области контура, но и более достоверно воспроизводится его контраст. На моделях рис. 8 (а, б) элементы, формируемые в результате сравнения текущего значения сигнала с базовыми весами, показаны более темными и образуют привычные растровые точки, а те, что построены по дополнительным функциям - более светлыми. Резкость в этом примере оказывается средней между обычным растированием и штриховой записью, что адекватно отражает неполную «силу» контура на оригинале. Исключается также заметное глазу резкое изменение геометрии раstra вдоль такой границы, когда контраст плавно возрастает или убывает по ее протяженности. Имея вдвое меньший период, каждая из дополнительных весовых функций обеспечивает вчетверо меньший алфавит печатных элементов. Все это обеспечивает плавный по полу изображения динамический обмен между шкалой квантования и частотой пространственной дискретизации, благодаря чему способность плавной тонопередачи печатной системы обменивается на ее возможности в отношении геометрической

точности мелких деталей и контуров с использованием элементов искусственного интеллекта, имитирующих работу гравера.

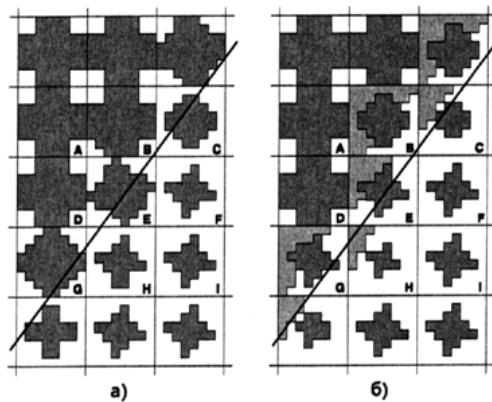


Рис. 8. Растирание контура промежуточного контраста ( $q_1 = 0,5$ ) тривиальным (а) и адаптивным (б) методами

Метод использует относительно малое число операций, а возможность их параллельного выполнения существенно упростила его программную и аппаратную реализацию.

Количественная оценка степени растровых искажений методом, предложенным в Л[8], объективно подтверждает преимущество данного варианта растиривания над традиционными.

Встречаемость контуров и мелких деталей той или иной конфигурации по ансамблю, представленному на рис. 6, можно наглядно отображать гистограммами востребованности тех или иных дополнительных весовых функций в данном способе допечатной подготовки изобразительного файла (Рис. 9), что дает некоторый побочный эффект в части семантического анализа изображений, помогая судить о технике исполнения, характере и общей «детальности» изображения.

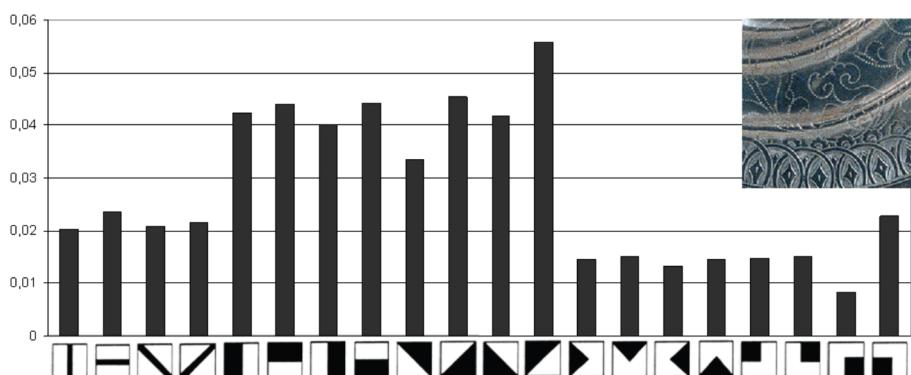


Рис. 9. Гистограммы востребованности программой HDHP дополнительных растровых функций различной геометрии для фрагмента изображения

### Практические результаты

Эффект общего увеличения четкости и резкости зависит в данном методе как от характера оригинала, так и от особенностей печатной технологии. Он особенно выражен, например, при репродуцировании гравюр (Рис. 10), изображений с высоким содержанием важных мелких деталей и, более всего, при относительно низких частотах раstra, например, в трафаретной печати. Повышению четкости и резкости сопутствует известный рост общего визуально воспринимаемого контраста изображения.

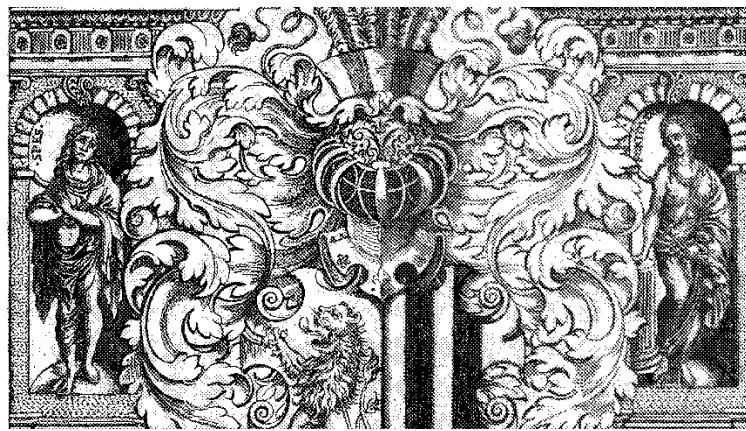


Рис. 10. Репродукция гравюры, левая часть файла которой закодирована для печати по описываемому методу, а правая – общепринятым способом

В сопоставлении различных технологий по тому или иному показателю качества важно сохранить равными прочие условия, например, путем печати изображений на одном листе разными способами растрового преобразования (Рис. 11). Наиболее эффективна визуальная оценка, когда различным образом обработаны бесшовно разделенные правая и левая части сюжета (Рис. 10, 12, 13).

В сравнении с работой ряда растровых программ эффективность метода иллюстрируют фрагменты, взятые с одного и того же оттиска на рис. 11. По ним можно отметить относительно слабый, по части той же четкости, эффект гибридного растирования Agfa Sablina за счет повышения линиатуры со 175 (ABS) до 210 и 240 Lpi.

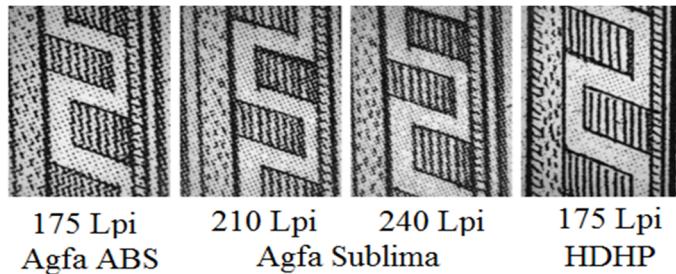


Рис. 11. Увеличенные фрагменты офсетных репродукций гравюры с растровыми преобразованиями фирмы Агфа и по отечественному HDHP методу

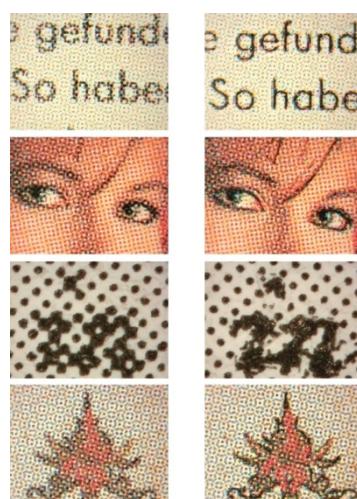


Рис. 12. Увеличенные фрагменты офсетных многокрасочных оттисков с линиатурой 70 лин/см (175 Lpi) по технологии Scitex Class Screening (слева) и данным аддитивным методом (справа)

Апробация метода в «цифровой» печати показала значительное повышение четкости и резкости изображений в большинстве электрографических и струйных устройств (Рис. 13).

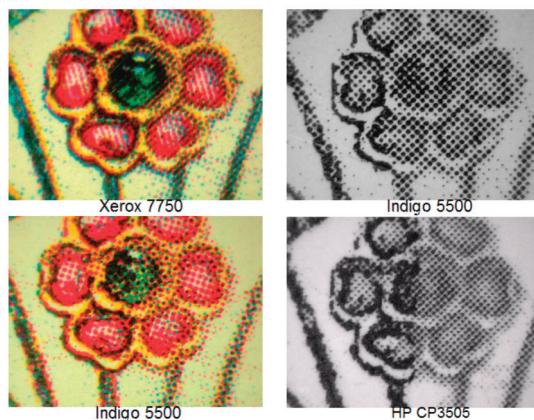


Рис.13. Увеличенные фрагменты оттисков, левая часть каждого из которых закодирована описанным методом, а правая — в драйвере или РИПе цифровой печатной машины

Можно отметить в заключение, что, если, например, т.н. телевидение высокой четкости (HDTV) с переходом от 625 строк стандартного разложения к 1200 и выше строкам требует использования других, более широкополосных (оптоволоконных, цифровых) каналов связи, то технология HDHP практически незатратна. Аналогичный зрительный эффект обеспечивается здесь в широком спектре печатных технологий без увеличения объема исходных файлов, применения каких-либо новых, специальных материалов и оборудования.

### Заключение

Апробация рассмотренного метода в тестовом и тиражном режимах применительно к различным видам традиционной и «цифровой» печати наглядно выявила следующие основные преимущества:

- вдвое большую четкость при любой частоте автотипного раstra стационарной области изображения;
- воспроизведение тонких линий сплошным красочным слоем или пробелом без их разрушений печатными элементами (растровыми точками);
- повышение воспринимаемого контраста, сопутствующее росту четкости и резкости;
- обеспечение этих улучшений качества при стандартных объемах исходных файлов и без каких-либо дополнительных требований к краскам, формам, бумагам и их взаимодействию в печати.

### Библиографический список

1. Грудзинский М. А. и др. Дискретизация ТВ изображений при цифровом кодировании // Техника кино и телевидения.-1978.- № 11.- с. 19-23.
2. Цуккерман И. И. Цифровое кодирование ТВ изображений.- М.: Радио и связь, 1981.
3. Кузнецов Ю. В. Подготовка иллюстраций к печати. Растирование.- М.: Мир книги, 1998.
4. Kouznetsov Y.V. Does some philosophy still exist for the halftone frequency selection? / Proc. of IS&T NIP15: Int. Conf. on Digital Printing Technologies, Oct 17-22, 1999, Orlando.-p. 362-365.- (англ.).
5. Кузнецов Ю.В., Коган А.А., Нишианидзе Г.А.. Устройство для растрового воспроизведения полиграфических оригиналов. Авт. свид. 1246408, заявл. 10.12.1982.
6. Патенты RU 2126598; DE 4498946; GB 2300328; US 5822086.
7. Патенты RU 2335094; US 8004720.

8. Желудев Д.Е. Исследование и разработка метода оценки искажений деталей полутоновых оригиналов в автотипном процессе. Канд. дисс., СПГУТД, СПб, 2009.

## OPTIMAL ENCODING THE IMAGES IN PRINTING

**Y.V. Kuznetsov, A.A. Schadenko**

St. Petersburg State Institute of Cinema and Television

St. Petersburg State University of Industrial Technology and Design

*Unlike to soft copy production the source file of multilevel image samples is used to be transformed into bi-level bitmap to control the printing in yes-no mode (ink – no ink). At this anew image file encoding the sampling frequency is defined by the halftone screen ruling value while the quantization scale by the print symbols amount inherent in given printing technology. Such encoding optimal organization is considered in the light of providing the conformity between the source image local properties and the facilities of a data transmitting channel (printing) as well as between the result of transfer (print) parameters and the specific of user visual system.*

**Locally adaptive encoding; business; definition; MTF of vision; image quality.**

**А. И. Бобровский<sup>2</sup>, М. М. Альмахрук<sup>1</sup>, Ш. С. Фахми<sup>1</sup>, А. Салем<sup>1</sup>**

<sup>1</sup>Санкт-Петербургский государственный электротехнический  
университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

<sup>2</sup>Научно-исследовательский институт прикладных проблем

## Саккадические движения глаз и концепция кодирования изображений по опорным точкам

*Рассматриваются основные принципы формирования образа в задачах технического зрения, которые эффективно используются в учебном процессе по дисциплине «Автоматизация проектирования больших интегральных схем», проводится сравнительная характеристика между биологической системой «глаз- мозг» и электронной видеосистемой на кристалле «фотоприёмник-вычислитель».*

### **Саккадические движения глаз, опорные точки объектов, триангуляция, опорные точки**

Перспективным направлением применения вычислительной техники для семантического анализа и кодирования выделенной полезной информации от фона изображений является слияние двух важных принципов формирования образа в задачах искусственного зрения. Первый принцип - это движение глаз при рассматривании объектов на основе регулярных саккадических движений. Саккадические движения – это очень быстрые крошечные автоматические скачкообразные смещения (вibrations) глаз по опорным точкам объектов изображений. Второй принцип - это формирование двумерных изображений по опорным точкам с минимальной ошибкой, что обеспечивает триангуляцию Делоне на этапе восстановления визуальных данных.

Саккады делают зрение идеально четким. Саккадические движения как бы завлекают свет в глаза, что в значительной степени улучшает использование глазом света, цветовосприятие и четкость зрения. Саккады легко увидеть, когда вы смотрите в глаза человека, имеющего четкое зрение и не носящего очки.

Саккадические движения очень быстрые, короткие, легкие и расслабленные. Глаз в результате этих постоянных движений от одной точки до другой точки рассматриваемого объекта формирует сжатый образ для хранения и распознавания [1,2]. Глаза движутся свободно в разных направлениях и могут пробегать взглядом многие части объекта в доли секунды. Саккадические движения являются причиной искрения и мерцания в глазах. В каждый момент времени человек смотрит на одну крошечную часть объекта, которая видна наиболее четко в центре поля зрения.

Выполнение неспешных передвижений глазами и глубокое их расслабление восстанавливают саккадические и другие типы глазных движений до нормы, улучшая тем самым остроту зрения. Практикуя выполнение медленных продолжительных перемещений, например, большие повороты, человек стимулирует быстрые перемещения (саккады, микроскопические перемещения), которые происходят сами по себе как автоматическая функция глаз. При этом формируется набор опорных точек для передачи в мозг [2].

В каждый момент времени смотрите на одну крошечную часть объекта, которая видна наиболее четко в центре поля зрения. Практикуйте перемещение взгляда на волокнах нити в кусочке ткани.

Существует много видов глазных движений: микроскопические, высокочастотные движения; оптический дрейф и другие крохотные, подсознательные автоматические движения; большие, средние и маленькие движения, которые совершаются, когда глаз перемещает взгляд с одного объекта на другой, с одной части объекта на другую, а также, когда оба глаза сводятся на одном объекте (конвергенция) или когда происходит аккомодация для лучшего зрения вблизи.

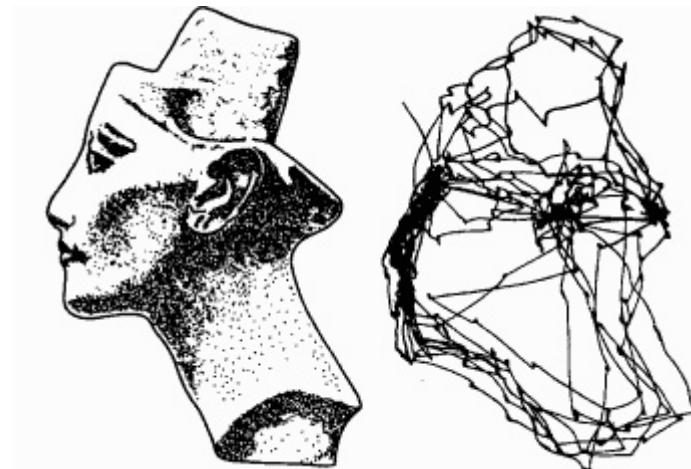


Рис. 1. Саккадические движения глаз при рассматривании головы Нефертити

Если эти движения становятся медленными, менее частыми, то зрение становится нечетким. Если же эти движения очень медленны, нечасты или вообще не совершаются, то это соответствует случаю наихудшего зрения и способствует развитию других глазных болезней. Когда зрение плохое, особенно, когда человек носит очки, глаза смещаются более медленно, нечасто, с меньшей амплитудой и с усилиями. Это состояние, усугубленное уменьшением частоты и амплитуды саккад, является причиной плохого зрения. Глазные движения жизненно необходимы для четкого, ясного зрения и здоровья глаз.

Выполнение передвижения глазами и глубокое расслабление восстанавливают саккадические и другие типы глазных движений до нормы, улучшая тем самым остроту зрения. Практикуя выполнение медленных продолжительных перемещений (например, большие повороты или медвежьи покачивания), быстрые перемещения (саккады, микроскопические перемещения) происходят сами по себе автоматически как автоматическая функция глаз. Вот почему перемещения являются правильной зрительной привычкой.

Ученые определили, что при восприятии лица другого человека, саккадные движения совершаются преимущественно по треугольнику между двумя глазами и ртом собеседника. Именно эти точки выстраивают в нашем сознании индивидуальное восприятие другого человека (рис. 2).

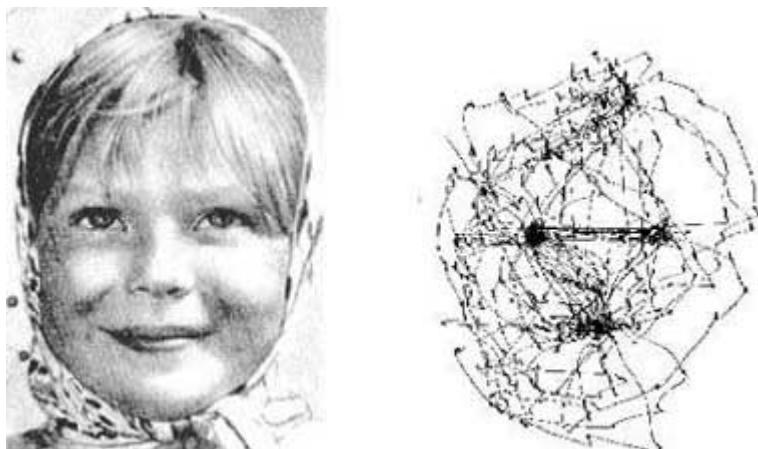


Рис. 2. Саккадные движения совершаются преимущественно по треугольнику между двумя глазами и ртом

В основу синтеза рассматриваемых устройств технического зрения положен тезис о необходимости при высокоскоростном параллельном способе передачи изображения выделять в объекте опорные точки и структурировать видеоданные с целью распознавания образов.

Врачи впервые имплантировали человеку искусственную сетчатку на основе искусственной кремниевой сетчатки (ASR, от artificial silicon retina) компании Optobionics. Сетчатка представляет собой микросхему диаметром 2 мм и толщиной меньше человеческого волоса. На кремниевой пластине размещается порядка 3500 микроскопических солнечных элементов, которые преобразуют свет в электрические импульсы.

С учётом выше изложенного, перспективным направлением становятся научные труды и результаты, полученные в работах [3,4] по кодированию и декодированию на основе полигонально-рекурсивного метода поиска опорных точек изображений.

Кодирование и передача по опорным точкам (ОТ) в процессе анализа изображений включает в себя два этапа (рис. 3). Первый этап - это выделение опорных точек объектов изображения на основе рекурсивного разбиения изображения на области различной формы и размера. Второй этап - это представление опорных точек изображений в удобном для передачи виде (например, одномерный динамический массив для передачи по каналу связи) [3,4].

Вопрос, связанный с поиском ОТ, является наиболее сложным при сжатии изображений по двум причинам. Во-первых, это – необходимость применения рекурсии при поиске ОТ для создания параллельных алгоритмов, во-вторых, – сохранение достаточно высокого коэффициента сжатия. Противоречивость этих обстоятельств приводит к поиску компромиссных решений, что отличает данный метод от известных [A1-A2].

Суть предложенного метода заключается в следующем: исходное квадратное изображение разбивается на равные полигоны, т.е. накладывается равномерная полигональная сетка; далее, в каждом полигоне производится поиск ОТ; эти ОТ соединяются и в результате образуется новая деформированная сетка, после закраски которой получается результирующее изображение, удовлетворяющее требованию субъективного качества [A2].

Для нахождения ОТ предложены эмпирические варианты, основная идея которых заключается в следующем:

1. Строится уравнение плоскости, проходящей через точки полигона с минимальным (по яркости) среднеквадратическим отклонением (СКО) всех точек по отношению к ней.

2. Если СКО меньше, или равно заданному, то выбирается точка, минимально отклоняющаяся от этой плоскости, иначе данный полигон разбивается на четыре подполигона и выполняется п. 1.

Процесс продолжается до тех пор, пока размер полигона не будет равен одному пикслю.

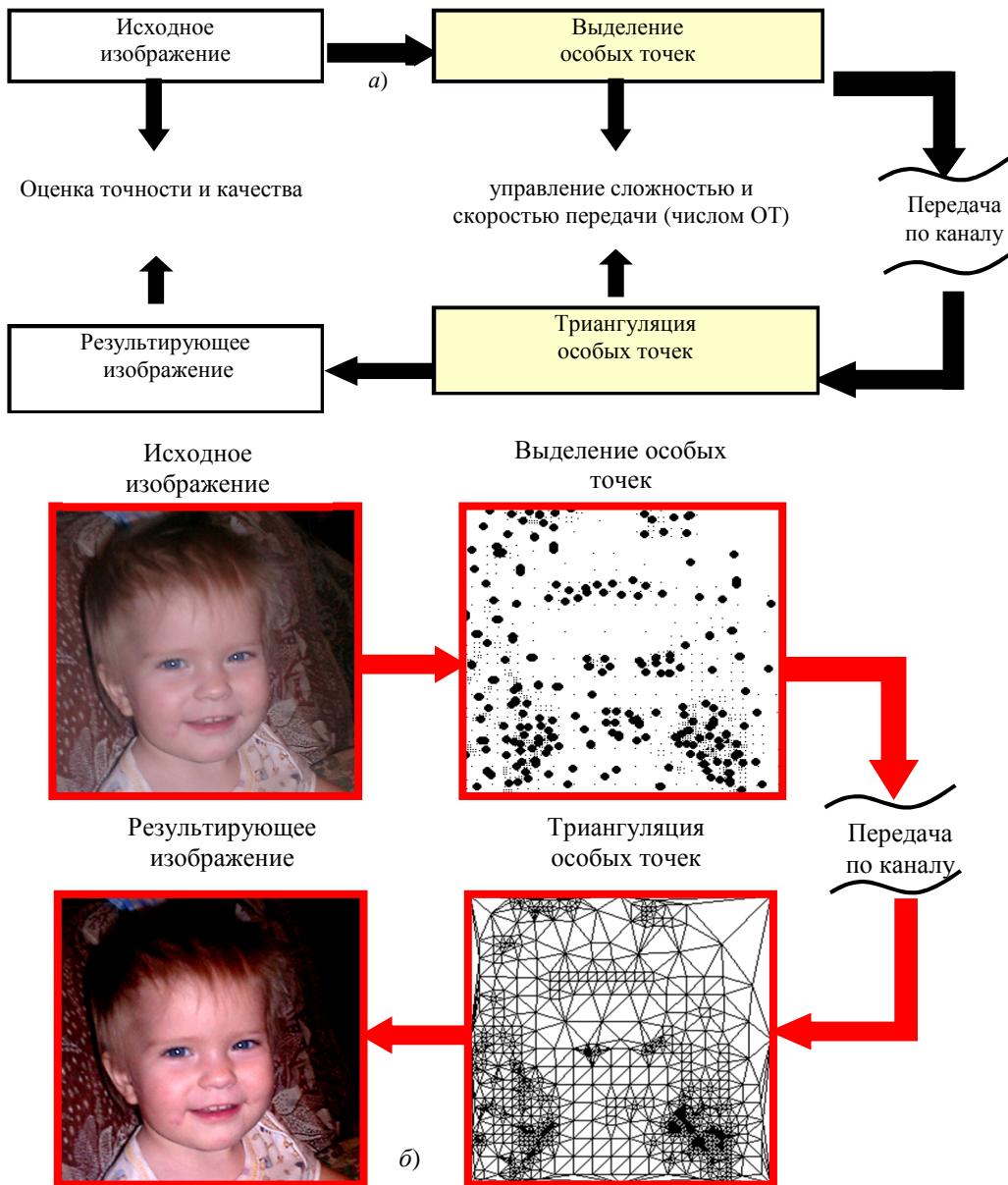


Рис. 3. Пространственно-рекурсивный метод кодирования изображений:  
а) модель кодирования и декодирования по опорным точкам  
б) поиск и триангуляция опорных точек

С точки зрения авторов, представление изображения при анализе в сжатой форме (т.е. выделение особых точек) и восстановление результирующего изображения могут быть реализованы, во-первых, как единый процесс, и, во-вторых, раздельно. В первом

случае сжатие базируется на полигонально-рекурсивном поиске особых точек и формировании сжатого массива (структура видеоданных – сетка, рис. 3. а), а восстановление представляет собой обратный процесс сжатию – соединение этих точек с помощью разработанной структуры по определенному правилу (рис. 3.б).

Во втором случае процесс сжатия аналогичен предыдущему, а процесс восстановления отличается тем, что выполняется классическая триангуляция Делоне над особыми точками без использования структуры данных.

Моделирование алгоритмов второго случая показало, что применение классического алгоритма триангуляции Делоне особого эффекта не дает с точки зрения субъективного качества результирующих изображений. При этом вычислительные затраты заметно увеличиваются. Поэтому в работе не будут рассматриваться те методы и алгоритмы, где процессы сжатия и восстановления не связаны.

Таким образом, процесс «видения» можно сравнить с работой фотокамеры, а систему «глаз- мозг» - с видеосистемой на кристалле «фотоприёмник-вычислитель». В фотокамере световые лучи проходят через набор линз, фокусирующих изображение на пленке. В здоровом глазу лучи света проходят через роговицу и хрусталик, который фокусирует изображение на сетчатке, представляющей собой слой светочувствительных элементов, выстилающих заднюю поверхность глаза.

#### Библиографический список

1. Гиппенрейтер Ю. Б. Движения человеческого глаза/ Ю. Б. Гиппенрейтер. — М.: Изд-во Моск. ун-та, 1978. — 256 с.
2. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов / Д. Марр; Пер. с англ. Н. Г. Гуревич. — М.: «Радио и связь», 1987.- 400 с.
3. Фахми Ш. С., Бобровский А. И., Еид М.М., Альмахрук М., Салем А., Березин В.В. Метод распознавания лиц на основе пространственно-рекурсивных триангуляционных сеток // Вопросы радиоэлектроники, сер. Техника телевидения, 2016, вып. 3, с. 86–91.
4. Фахми Ш. С., Альмахрук М. М., Соколов Ю. М., Бобровский А. И., Еид М. М., Салем А. Точность, скорость и сложность устройств кодирования изображений по опорным точкам// Научно-технический вестник информационных технологий, механики и оптики. 2016. Том 16. № 4. С. 678–689.

#### SACCADIC EYE MOVEMENTS AND THE CONCEPT OF ENCODING IMAGES WITH CONTROL POINTS

**M.M.Almahrouq<sup>1</sup>, S.Sh.Fahmi<sup>1</sup>, A.Salem<sup>1</sup>, V.V.Berezin<sup>2</sup>**

<sup>1</sup> Saccadic eye movements and the concept of selection control points Saint Petersburg Electrotechnical University "LETI"

<sup>2</sup> research Institute of applied problems

*Examines the basic principles of image formation in the problems of technical vision, which are effectively used in educational process on discipline "computer-aided design of large integrated circuits", comparative characteristics of the biological system "eye-brain" and an electronic video system-on-chip "sensor-computer" ..*

**Saccadic eye movements, the reference point of the object, triangulation, control point**

**Н. А. Обухова, И. Г. Зубов**  
СПбГЭТУ "ЛЭТИ"

## Автоматический метод формирования области интереса в прикладных телевизионных системах медицинского назначения

*В статье предложена процедура формирования области интереса на ТВ медицинских изображениях, основанная на алгоритме нечеткой кластеризации с-средних. Эффективность процедуры оценивалась для сиаскопических изображений (СИАграфов) и кольпоскопических мультиспектральных изображений. Проведенные эксперименты показали, что область интереса выделена корректно для 95 % изображений.*

**Обработка медицинских изображений, выделение области интереса, кластеризация, нечеткая кластеризация, алгоритм с-средних**

Эффективность анализа медицинских изображений в значительной мере определяется их предварительной обработкой. Одним из наиболее важных ее шагов является автоматическая сегментация зоны интереса (*region of interest – ROI*) - участка на изображении, для которого целесообразно проводить дальнейший анализ.

Наиболее часто сегментация области интереса реализуют на основе алгоритма *k*-средних. Например, в мультиспектральной телевизионной системе диагностики онкологических изменений шейки матки изображение объекта интереса занимает 70% снимка. Остальная часть может содержать другие ткани, медицинские инструменты, различные артефакты. Предложенная для формирования области интереса процедура на основе алгоритма кластеризации *k* - средних и признака цвета (значение цветовой координаты *a* в системе *Lab*) обеспечивает корректное выделение области интереса с вероятностью 0.95[1].

Процедура сегментации *ROI* на основе алгоритма *k* - средних теряет эффективность, если граница области интереса не является однозначно выраженной. Например, на сиаскопических медицинских изображениях области тканей с изменениями, которые должны быть проанализированы, не имеют резких границ. Соответственно, краевые элементы зоны поражения соответствуют как области интереса, так и области окружающих кожных покровов.

Проведенные эксперименты показали, что применение процедуры основанной на *k*-средних для сиаскопических изображений позволяют выделить корректно *ROI* с вероятностью 0.88.

Сиаскопические изображения (СИАграфы) – это изображения участка кожи 12 мм х 12 мм, полученные в разных спектральных диапазонах, включая видимый и инфракрасный. СИАграфы включают в себя:

Дermatoskopическое изображение кожного образования, с помощью которого оценивают цвет и наружную структуру образования.

Изображения, отражающие уровень меланина (Области более темного цвета отображают места с большой концентрацией пигмента, светлого цвета - с меньшей концентрацией пигмента)

Изображения, отражающие уровень дермального меланина (Белые области указывают на отсутствие меланина в папиллярном слое дермы. Зеленые, синие, красные и черные области показывают увеличивающуюся концентрацию меланина в папиллярном слое дермы).

Изображения, отражающие уровень гемоглобина (Области более темного цвета отображают места с большой концентрацией гемоглобина, светлого цвета - с меньшей концентрацией гемоглобина).

Изображения, отражающие уровень коллагена (Области светлого цвета отображают места с большей концентрацией коллагена, темного цвета - с меньшей концентрацией)

СИАграфы предназначены для диагностики меланомы.

Характерными особенностями сиаскопических изображений затрудняющими корректную работу процедуры сегментации являются: низкий контраст между окружающей кожей и областью поражения, нечеткая и нерегулярная граница поражения, внутренние артефакты, такие как пузырьки воздуха, кровеносные сосуды, волосы и др.

Для повышения эффективности сегментации *ROI* предложено применять математический аппарат нечеткой логики, а именно, алгоритм нечеткой кластеризации *c*-средних.

Алгоритм *c*-средних позволяет сформировать нечеткие кластеры, описываемые матрицей нечеткого разбиения:

$$\mathbf{F} = [\mathbf{m}_{ki}], \mathbf{m}_{ki} \in [0,1], k=1, \dots, M, i=1, \dots, c$$

в которой  $k$ -ая строка содержит степени принадлежности элемента  $k$  к кластерам  $\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_c$ ;  $M$  – число элементов кластеризации,  $c$  – число кластеров. При нечетком разбиении степень принадлежности элемента к кластеру  $\mathbf{m}_{ki}$  принимает значения из интервала  $[0, 1]$ .

Матрица должна обладать следующими свойствами:

$$\begin{aligned} \sum_{i=1}^c \mathbf{m}_{ki} &= 1, \quad k = 1, \dots, M, \\ 0 < \sum_{k=1}^M \mathbf{m}_{ki} &< M, \quad i = 1, \dots, c. \end{aligned} \quad (1)$$

Основные шаги алгоритма нечетких *c*-средних [2]

Шаг 1. Установить параметры алгоритма:  $c$  – число кластеров;  $m$  – экспоненциальный вес;  $\epsilon$  – параметр останова алгоритма.

Шаг 2. Случайным образом сгенерировать матрицу нечеткого разбиения  $\mathbf{F}$ , удовлетворяющую условия (1).

Шаг 3. Рассчитать вектора признаков центров кластеров.  $\mathbf{V}_i$  – вектор признаков центра  $i$ -го нечеткого кластера:

$$\mathbf{V}_i = \frac{\sum_{k=1}^N (\mathbf{m}_{ki})^m \mathbf{q}_k}{\sum_{k=1}^N (\mathbf{m}_{ki})^m},$$

где  $\mathbf{q}_k$  – вектор признаков  $k$ -го элемента кластеризации;  $m \in [1, \infty[$  – экспоненциальный вес, определяющий нечеткость кластеров.

Шаг 4. Рассчитать расстояния между элементами кластеризации и центрами кластеров:

$$D_{ki} = \|\mathbf{q}_k - \mathbf{V}_i\|_{L2}, \quad k = 1, \dots, M, i = 1, \dots, c.$$

Шаг 5. Пересчитать элементы матрицы нечеткого разбиения

Если  $D_{ki} > 0$ , то  $m_{ki} = \frac{1}{\sum_{j=1}^c \frac{1}{D_{jk}^2}}$ .

Если  $D_{ki} = 0$ , то  $m_{kj} = \begin{cases} 1, & j = i \\ 0, & j \neq i \quad j = 1, \dots, c \end{cases}$ .

Шаг 6. Проверить условие  $\|\mathbf{F} - \mathbf{F}^*\| < \epsilon$ , где  $\mathbf{F}^*$  – матрица нечеткого разбиения на предыдущей итерации алгоритма. Если "да", то завершить формирование кластеров, иначе – к шагу 3.

Основные шаги процедуры формирования ROI на основе рассмотренного алгоритма с-средних:

Шаг 1. Изображение обрабатывают гауссовым фильтром  $G_\sigma$  с небольшим параметром размытия  $\sigma$  для подавления шума.

Шаг 2. Выполняют конвертацию изображения в цветовое пространство Lab и выделяют слои с координатой  $a$ .

Шаг 3. Осуществляют кластеризацию методом с-средних.

Шаг 4. Выбирают область, расположенную вблизи центра изображения.

На рисунке 1 показаны результаты основных шагов предложенной процедуры, на рисунке 2 сформированные ROI.

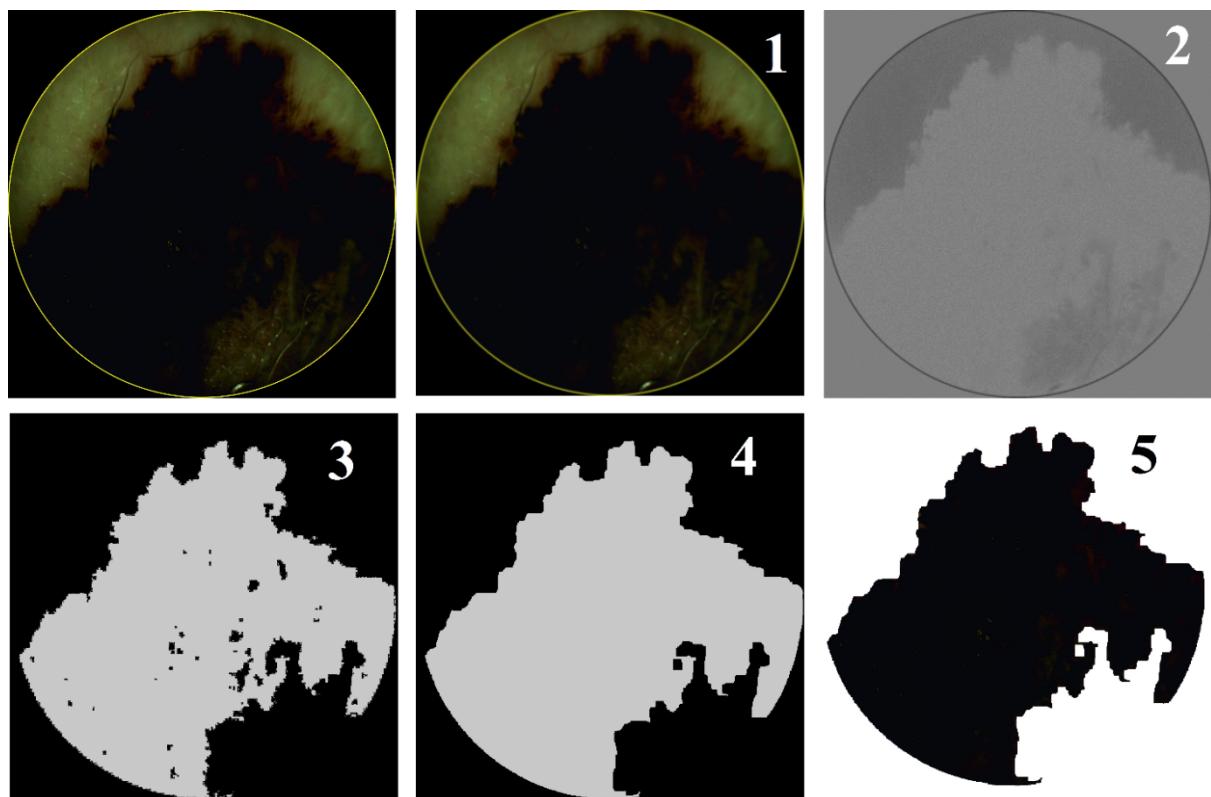


Рис. 1. Результаты основных шагов работы алгоритма

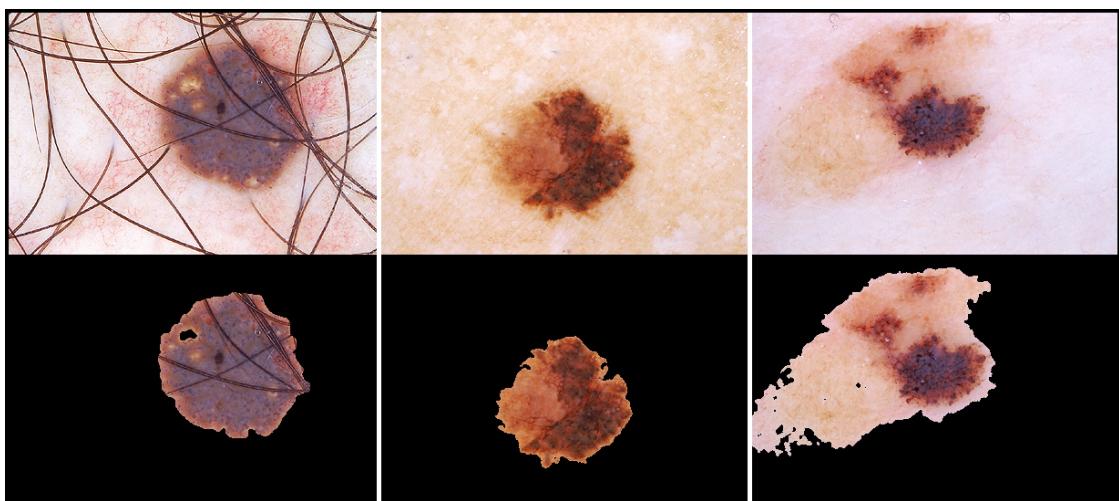


Рис. 2. Результат работы алгоритма выделения зоны интереса

На рисунке 3 представлены результаты выделения зоны интереса при низком контрасте между окружающей кожей и поражением, методами на основе  $k$ -средних и  $c$ -средних.



Рис. 3. Результаты выделения зоны интереса при низком контрасте между окружающей кожей и поражением. 1) исходное изображение; 2) результат работы метода на основе  $k$ -средних; 3) результат работы метода на основе  $c$ -средних

Проведенные эксперименты показали, что  $ROI$  сиаскопических изображений предложенной процедурой выделена корректно для 95 % изображений. Применение процедуры для кольпоскопических изображений также показало ее эффективность. Задача выделения  $ROI$  решена для 95 % изображений.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 17-07-00045.

#### Библиографический список

1. Обухова Н. А., Мотыко А. А. Автоматический метод анализа мультиспектральных кольпоскопических изображений для телевизионной системы диагностики рака шейки матки. // Известия высших учебных заведений России. Радиоэлектроника. СПбГЭТУ, 6(24-33), 2015.2.
2. Zimmerman H.-J. Fuzzy Set Theory and its Applications. Dordrecht: Kluwer Academic Publishers. 1996. 315p.

#### AUTOMATIC METHOD OF REGION OF INTEREST SELECTION IN TELEVISION SYSTEMS FOR MEDICAL PURPOSES

**Natalia A. Obukhova, Ilya G. Zubov**  
Saint Petersburg Electrotechnical University “LETI”

*The article proposes the procedure of region of interest selection for TV medical images, based on the  $c$ -means algorithm. The effectiveness of the procedure was estimated for SIAscopic images (SIAGraphs) and colposcopic multispectral images. The experiments showed that the region of interest was correctly identified for 95% of the images.*

**Image preprocessing in medical systems, the region of interest selection, clustering, fuzzy clustering,  $c$ -means**

**С. В. Гросс**  
СПбГЭТУ «ЛЭТИ» кафедра ТВ

## Встраиваемая система коррекции курса самолета на базе частичного декодирования сигнала H.264/AVC

*В статье представлен подход к использованию визуальной информации для коррекции курса беспилотного самолета (БПЛА), реальная траектория которого изменяется в результате воздействия ветра. Подход основывается на анализе векторов движения видеопотока с телевизионной камеры, устанавливаемой на БПЛА. Данные об изменениях глобального вектора движения позволяют осуществить коррекцию сноса. Предлагаемый метод позволяет реализовать целевую функцию на гомогенных встраиваемых системах с потреблением менее 2 Ватта.*

**Частичное декодирование, вектора движения, H.264/AVC, система на кристалле, SoC, TI DM368**

### Постановка задачи

При воздействии ветра, его скорость и направление изменяет курс любого воздушного судна. Оператор/электроника должны компенсировать скорость и направление ветра, для того, чтобы самолёт выдерживал нужный курс. Контроль пути по указанному направлению главным образом состоит в определении путевого угла и так называемых фактических показателей, а также характерного бокового уклонения от линии заданного пути (ЛЗП). В рамках данной статьи рассматривается противодействие только горизонтальной составляющей сноса, без учета вертикальной динамической или термической составляющей.

Автономные системы навигации современных БПЛА относятся к классу инерциальных навигационных систем. Для таких систем характерна нарастающая во времени ошибка, что решается путем применения глобальных спутниковых систем навигации (ГССН). Такое решение приводит к существенным потерям в автономности и помехоустойчивости. Обеспечить точность навигации возможно, наряду с использованием данных инерциальных датчиков пространственной ориентации (ИДПИ), на основании данных визуального контроля. При этом подходе изображение местности анализируется бортовым компьютером, определяются координаты и ориентация летательного аппарата. В настоящее время в основе навигационных систем БПЛА лежит подход к использованию комплексированного решения на базе показаний приёмников ГССН и ИДПИ. При отсутствии внешних факторов подавления деятельности каждой из систем возможно добиться высокой точности определения положения, до 15 м СКО [1].

Ориентирами для полета, в зависимости от возможности бортового оборудования, могут быть:

1. Наземные линейные ориентиры, которые идут строго параллельно ЛЗП;
2. Периодические измерения и их компенсация (угол сноса, изменения воздушной ситуации по данных наземных служб данного региона) (рис. 1);
3. Специальные отметки.

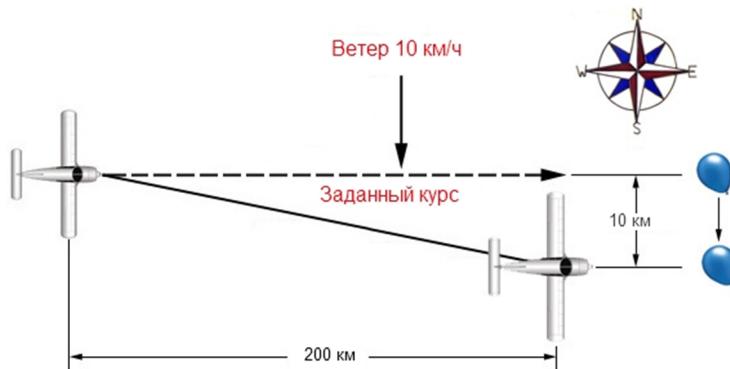


Рис. 1. Смещение курса самолета в результате внешнего воздействия

Рассмотрим технику использования высокочувствительных данных потока H.264/AVC для решения данной задачи в соответствии с пунктом 2 в рамках маломощных систем на кристалле (<2 Ватт).

#### Алгоритм на базе векторов движения

Для реализации алгоритма был выбран декодер H.264 из библиотеки FFmpeg [2]. Для каждого слайса из потока извлекается массив векторов движения. В текущей версии программы не поддерживается обработка В-кадров, поэтому они пропускаются при анализе (предполагается использование baseline-профиля).

Для того чтобы осуществить сегментацию по полному числу сегментов предлагается следующий подход:

- 1) Определение порога движения как 15% от максимальной длины вектора данного кадра (уровень отсечки).
- 2) Выделение в движущиеся области обособленных групп макроблоков с коллинеарными векторами выше уровня отсечки, не менее 15 макроблоков, но не превышающих в сумме 25% от общего числа макроблоков.
- 3) Обнуление векторов движения выбранных макроблоков, уменьшение количества анализируемых макроблоков в кадре
- 4) Вычисление глобального вектора движения как среднего вектора по оставшимся макроблокам.

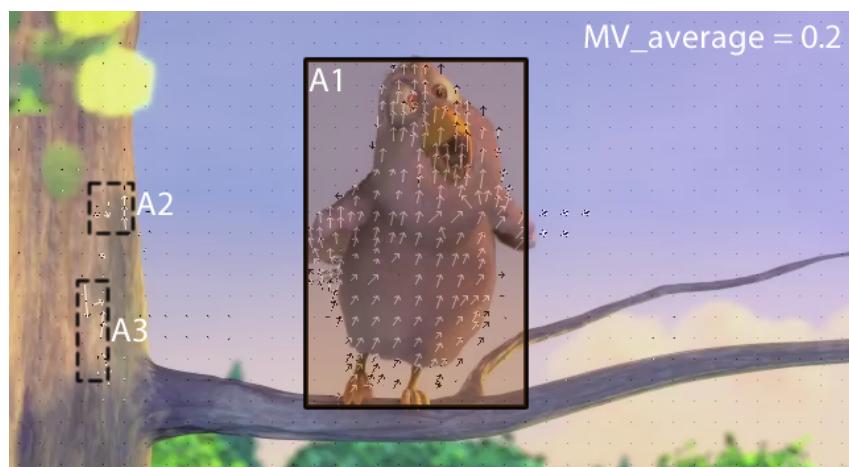


Рис. 2. Визуализация векторов движения на базе библиотеки FFmpeg

На рис. 2 изображен тестовый режим работы программы с наложением отладочной информации. Зона A1 отмечена как зона движения с высокой степенью подобия векторов составляющих ее макроблоков и была исключена из списка анализируемых макроблоков.

В то же время зоны A2 и A3 не являются полноценными объектами из-за малого размера и не исключаются из списка.

### Программная реализация

XY-данные глобального вектора движения являются расчётными параметрами и отправляются на контроллер двигателя через COM-порт в формате:

0xA0 0xA3 0xA5 <смещение по X> <смещение по Y>.

Значения смещения находятся в диапазоне от -128 до 127. Нормировка компенсирующего усилия на реальном оборудовании производится исходя из калибровки для различных скоростных режимов (крейсерская скорость, форсаж).

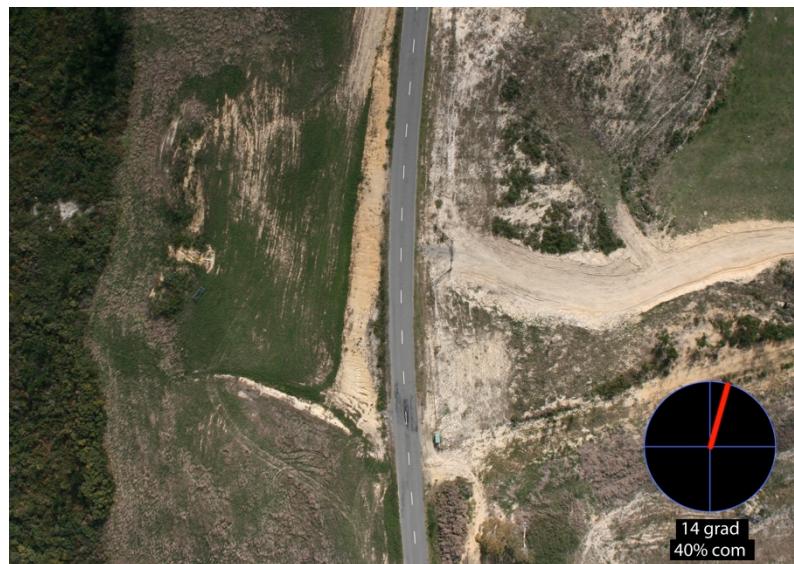


Рис. 3. Общий вид программы коррекции сноса, смещение 14 градусов вправо

Аргументы запуска программы:

[1] Источник сигнала. Возможные режимы: 0 (камера) / путь к файлу.

[2] Порт. В Linux для запуска возможно, например, использовать порт /dev/ttyS2. Для встраиваемых систем UART-порт, зачастую зависит от версии ядра. Для ядер до 2.6.32 часто используется адрес /dev/ttyS2, для более новых версий используется /dev/ttyO2.

[3] Режим записи. Запись ведётся для исходных изображений, с удалением каналов цветности, для каждого проанализированного кадра.

[4] Визуализация (рис.3). Только для целей отладки и только на настольных системах. Выводит окно с выбранным видеопотоком, поверх которого графикой отображаются значительные изменения глобального вектора.

Для ведения курсовой корректировки при помощи программы выбирается баланс между размером кадра и производительностью устройства. Слишком маленький кадр сведёт на нет точность (даже с учётом скорости), слишком большой – приведёт к сильному падению производительности и, как следствие, повышению инерционности показаний программы.

### Реализация и тестирование на встраиваемой системе.

Были проведены эксперименты над видеофрагментами размером 640x480. Для выбранного тестового изображения размера 10000x10000 были проведены ручные смещения на фиксированные размеры пикселов по схеме:

{ 0, 0, +1, +1, +2, +4, ..., +118, +120, +118, .... +2, +1, +1, 0, 0 }

Результирующие изображения были соединены в видеопоток и переданы в тестовую программу. В таблицах 1 и 2 приведены результаты усреднения показаний по отношению к эталонным сдвигам по всему диапазону.

Для сравнения с пиксельными методами был выбран метод RANSAC [3] (стабильный метод оценки параметров модели на основе случайных выборок), устойчивый к зашумлённости исходных данных. В результате применения RANSAC строится наилучшая матрица гомографии.

Сравнение точности методов

Таблица 1.

	RANSAC		H264_MV	
	x	y	x	y
Без искажений	100%	100%	99%	100%
Геометрические искажения	95%	97%	67%	94%
Геометрические искажения + цветной гаусс. шум	96%	98%	83%	87%
Изменение яркости	99%	100%	96%	100%
Гаусс.размытие (3 px)	99%	100%	99%	99%

Одним из самых серьезных факторов, затрудняющих использования RANSAC является очень высокая сложность[4]. Требования к производительности устройства обработки данных серьезно ограничивают применение таких систем в реальных условиях в свете присутствия в рамках бортового оборудования ряда других систем с высоким потреблением энергии – приемопередающий комплекс, системы GPS/GLONASS и др.

Сравнение производительности методов на базе СнК DM368

Таблица 2.

	RANSAC	H264_MV
Без искажений	3.4 к/с	65 к/с
Геометрические искажения	3.1 к/с	62 к/с
Изменение яркости	3.4 к/с	66 к/с
Гауссовское размытие (3 px)	4.6 к/с	63 к/с



Рис. 4. Общий вид типичного решения на базе СнК DM368

Использование нового метода позволяет использовать оборудование энергоэффективного класса ARM9 (рис.4), что существенно улучшает как массогабаритные, так и энергетические параметры решения.

### Выводы

Решения на базе частичного декодирования сигнала H.264/AVC обладают существенными скоростными преимуществами, позволяющими использовать ресурсы видеопотока даже на устройствах, которым потенциально недоступно даже воспроизведение сигнала в режиме реального времени без применения дополнительных аппаратных средств. В случае с коррекцией курса решение на базе векторов движения позволяет получить корректный результат при более чем десятикратном преимуществе по

производительности при решении проблем с геометрическими искажениями в видеопотоке. Речь идет об использовании камер с глобальным затвором или алгоритмических средств коррекции последствий применения сенсоров с построчным переносом заряда.

### Библиографический список

1. Костюк А.С. Использование малых летательных аппаратов и цифровой съемки для обновления плановых материалов и мониторинга объектов нефтегазового комплекса // Тезисы доклада. 10-я Всероссийская научно-практическая конференция «Геоинформатика в нефтегазовой и горной отраслях», Сургут, 2009
2. FFmpeg [Электронный ресурс]. – Режим доступа: <https://ffmpeg.org/>. (07.05.2017)
3. RANSAC [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/RANSAC>. (05.05.2017)
4. C. J. Wang, Y. Zhang, J. Lu, and W. Xu, “A framework for moving target detection, recognition and tracking in UAV videos,” in Affective Computing and Intelligent Interaction, vol. 137 of Advances in Intelligent and Soft Computing, pp. 69–76, Springer, Berlin, Germany, 2012

## SYSTEM-ON-CHIP UAV TRAJECTORY CORRECTION BASED ON H.264/AVC PARTIAL DECODING

**Gross S.V.**

Saint Petersburg ETU, the Faculty of Radio Engineering and Telecommunications/Television and videotronics

*In this work Unmanned aerial vehicle (UAV) course correction by usage of graphical information approach has been proposed. In real coditions trajectory of UAV is changing influenced by the wind. Approach is based on UAV TV camera H.264/AVC video motion vectors. Global vector components allows weighted correction of drift. The proposed method makes possible to form the objective function on homogeneous embedded systems with a consumption of less than 2 Watt*

Partial decoding, moving vectors, H.264/AVC, system-on-chip, TI DM368

**В. А. Сухарев, А. В. Богословский**  
ВУНЦ ВВС «ВВА»

## Моделирование турбулентных искажений

*В работе рассмотрены некоторые методы восстановления изображений, смазанных атмосферной турбулентностью. Предложен подход, позволяющий усовершенствовать известные методы восстановления изображений.*

**Турбулентность, восстановление изображения, моделирование.**

С 1972 года, когда в космос был запущен искусственный спутник Земли Landsat с установленной на его борту аппаратурой MSS (Multispectral Scanner System), запросы военных ведомств разных стран стимулируют развитие различных методов получения информации об объекте на расстоянии без непосредственного контакта с ним измерительного прибора.

При распространении электромагнитного излучения в вакууме его интенсивность не изменяется, однако между изучаемым объектом и регистрирующей аппаратурой всегда находится слой атмосферы, в состав которой входят разнообразные газы, молекулы водяного пара и атмосферные частицы.

В нижней атмосфере всегда присутствует атмосферная турбулентность, которая приводит к изменению плотности воздуха, а следовательно, и коэффициента преломления. В результате таких изменений фаза электромагнитной волны искажается, что приводит к смазу изображения.

В настоящее время эта проблема решена только для определенного класса изображений, в которых искажение на верхних частотах незначительно, а контуры объекта, после воздействия атмосферной турбулентности, остаются узнаваемы.



Рис. 1. Фотография вышки при атмосферной турбулентности (а), результат цифровой фильтрации изображения (б)

На рисунке 1а представлена фотография сделанная сотрудниками ФГУП «Государственный лазерный Центр «Радуга» им. И. С. Косьминова», (<http://mi.mathnet.ru/qe14274>; <https://www.youtube.com/watch?v=avo3LRTWU3k>) [1] с

расстояния 2400 метров при воздействии атмосферной турбулентности, на рисунке 16 представлен результат цифровой обработки.

Восстановление происходит за счет суммирования одинаково расположенных в произвольно выбранном прямоугольнике пикселей, а сам прямоугольник от кадра к кадру определяется по корреляционным свойствам.

В работах [2, 3] предлагается восстанавливать смазанные изображения путем решения системы алгебраических уравнений, описывающих процесс формирования отсчетов видеосигнала изображения, искаженного большим колебательным линейным смазом:

где  $\{f_l, \mathbf{K}, f_{m+n}\}$  - отсчеты видеосигнала «смазанного» изображения;  $\{b_l, \mathbf{K}, b_m\}$  - отсчеты видеосигнала неискаженного смазом изображения объекта;  $\{a_1, \mathbf{K}, a_n, a_{m+1}, \mathbf{K}, a_{m+n}\}$  - отсчеты видеосигнала изображения неподвижного фона.

Моделирование восстановления изображений по системе алгебраических уравнений (1) позволяет предположить возможность применения данного подхода и для компенсации влияния турбулентной среды на формируемое изображение.

Для построения модели турбулентной среды необходимо иметь набор изображений, искаженных различной по свойствам турбулентностью.

Для решения указанной задачи была создана демонстрационная установка, представляющая собой стеклянный куб заполненный жидкостью. Внутри куба, посредством мотора перекачивающего жидкость, создаются турбулентные потоки. Регулировка данных потоков осуществляется с панели управления и позволяет регулировать мощность потока и изменять угол подачи жидкости в горизонтальной и вертикальной плоскостях.

На рисунке 2 приведены примеры типичных искажений при различной степени турбулентности. Влияние турбулентности увеличивается от рисунка 2а к рисунку 2г.

Анализ данных изображений показывает, что размывание части изображения, даже при сравнительно слабой турбулентности, происходит по-разному. Особенно это очевидно при совместном рассмотрении рисунков 2а и 2б. При сильной турбулентности формируемое изображение начинает разрушаться (рисунок 2в, 2г). Разрушение изображения в разных местах также может происходить по-разному.

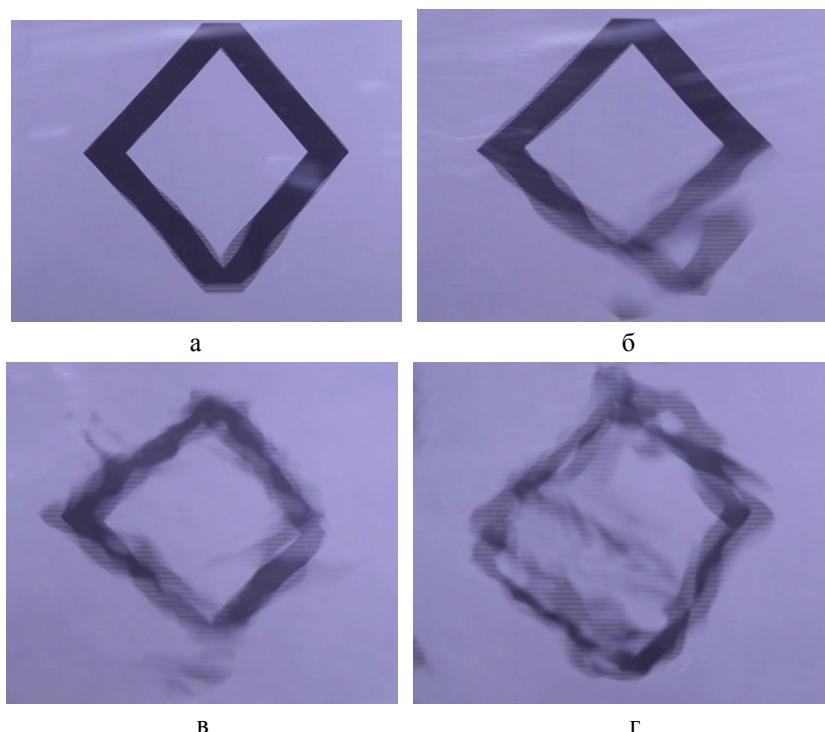


Рис. 2. Изображения, искаженные различными турбулентными потоками

Всё это свидетельствует о том, что при восстановлении изображений сильно искаженных турбулентностью, необходимо учитывать локальные характеристики исследуемого изображения.

#### Библиографический список

1. А. П. Аверин, В. С. Пряничников, В. В. Тяпин, “Компьютерная коррекция изображения, искаженного турбулентной атмосферой”, Квантовая электроника, 40:5 (2010), 418–420 [Quantum Electron., 40:5 (2010), 418–420]
2. Богословский, А.В. Линейный смаз/ А.В. Богословский, Е.А. Богословский, И.В. Жигулина, В.В. Васильев, А.В. Пономарев// Радиотехника. 2015. № 12. С. 147-153.
3. Богословский, А.В. Алгоритм компенсации линейного смаза изображения движущегося объекта/ А.В. Богословский, Е.А. Богословский, И.В. Жигулина, В.В. Васильев, А.В. Пономарев// Цифровая обработка сигналов и ее применение (DSPA-2016): материалы XVIII Международ. науч. конф. Москва: РНТОРЭС им. А.С. Попова, 2016. С.817-822.

#### MODELING OF TURBULENT DISTORTIONS

**Sukharev V.A., Bogoslovsky A.V.**

Air Force Education and Research Center «The Zhukovsky and Gagarin Air Force Academy» (Voronezh)

*Some methods of reconstruction of images blurred by atmospheric turbulence are discussed. An approach is proposed to improve the known methods of image reconstruction.*

Turbulence, image reconstruction, simulation.

**А. Н. Бучатский, С. П. Куликов**  
СПбГУТ им. проф. М.А. Бонч-Бруевича

# Сравнительный анализ характеристик измерительных приборов для контроля параметров сигнала цифрового эфирного телевидения стандарта DVB-T2

*В связи с переходом России на цифровое эфирное телевидение возникает потребность в качественном контроле параметров таких сигналов. При этом не все производители измерительного оборудования на рынке РФ стараются соответствовать государственным стандартам РФ. В данной статье анализируются измерительные приборы по критериям, связанным с их техническими характеристиками и соответствуем ГОСТР 55696-2013.*

## DVB-T2, SFN, R&SETL, R&SEFL, REFEREE 2, IT-15T2, ГОСТ Р 55696-2013

Россия постепенно переходит на новый стандарт цифрового наземного (эфирного) телевизионного вещания DVB-T2, который обеспечивает лучшую помехозащищенность по сравнению со стандартом DVB-T, обладает большей пропускной способностью, а также отличается очень важной для региональных вещателей особенностью – наличием нескольких физических уровней PLP физических каналов передачи.

В связи с этим со стороны специалистов радиосвязи и радиовещания наблюдается повышенный интерес к системам мониторинга, контроля и измерительному оборудованию. Учитывая особенности построения одночастотных сетей вещания (SFN), контроль качества телевизионного сигнала как на входе и выходе передающего оборудования, так и на местах приема сигнала остается первостепенной задачей.

В данной статье рассматривается соответствие ГОСТ Р 55696-2013 четырех измерительных приборов, часто используемых для контроля параметров цифрового эфирного телевидения специалистами РТРС, кабельных служб и службами радиочастотного контроля:

- R&S EFL
- R&S ETL
- Enensys ReFeree 2
- Planar IT-15T2

Также приводятся результаты экспериментального измерения параметров сигнала DVB-T2 на площадке в г.Санкт-Петербург.

По результатам измерений и сравнений измерительных средств сделаны выводы о соответствии данных приборов законодательству РФ и возможности использования их специалистами в области телевидения.

## Библиографический список

Контроль DVB-T2 - это просто/ С. Круглов // Электронный журнал "Broadcasting. Телевидение и радиовещание" #2, 2014г. URL: <http://www.broadcasting.ru/articles2/newproducts/kontrol-dvb-t2-eto-prosto> (Дата обращения: 27.04.2017)

**THE COMPARATIVE ANALYSIS OF CHARACTERISTICS OF MEASURING DEVICES OF MEASURING DEVICES FOR CONTROL OF PARAMETERS OF DIGITAL TELEVISION SIGNALS OF THE DVB-T2**

**Buchatsky A. N. , Kulikov S. P.**

*The Bonch-Bruevich Saint - Petersburg State University of Telecommunications*

*Due to the transition of Russia to digital radio television there is a need for high-quality control of parameters of this television. At the same time not all producers of the measuring equipment try to conform to state standards of the Russian Federation. In this article measuring devices on their technical characteristics and on their compliance are considered by GOST P 55696-2013.*

**DVB-T2, SFN, R&S ETL, R&S EFL, REFEREE 2, IT-15T2, GOST P 55696-2013**

## Методы повышения качества видеоинформации в подводной среде

*Приведены стандартные методы повышения качества видеоинформации.  
Рассмотрены методы повышения качества видеоинформации относительно их применения в подводной среде*

### **Качество видеоинформации, повышение качества видеоинформации, подводное телевидение**

При разработке подводной телевизионной системы важно качество получаемой видеоинформации. В связи со спецификой водной среды получаемая видеоинформация зачастую оказывается недостаточно хорошего качества, что приводит к ошибкам при работе с ней в режиме автоматического принятия решений, либо к невозможности работы оператора ввиду низкого качества видеоизображения. Это связано в первую очередь со спецификой распространения света в воде, а также возможностями телевизионной системы.

Для повышения качества видеоинформации обычно используются следующие методы: повышение контрастности, расширение динамического диапазона и повышение отношения сигнал/шум. Базовая формула для определения отношения сигнал/шум в телевизионной системе подводного видения имеет следующий вид[1]:

$$\psi = \frac{D_{3p} e^{-\alpha_1 L}}{N_c L} K \sqrt{\frac{W_0 R_{ob} \eta_\phi}{2e} F(L, \vartheta_I, d_{ob})},$$

где  $N_c$  – число элементов в строке изображения при квадратном кадре,  $K$  – контраст изображения,  $W_0$  – энергия лазера, излучаемая за время кадра,  $F$  – коэффициент передачи энергии по трассе источник – объект – приемник,  $L$  – расстояние наблюдения,  $d_{ob}$  – размер объекта,  $\vartheta_I$  – угол излучения. Для уменьшения количества шума на изображении существует множество методов фильтрации (усредняющий, пороговый, медианный и т.д.) [2]. На реальных изображениях часто возникает импульсная помеха для борьбы с которой можно применить медианный фильтр. Размер маски фильтра выбирается оператором исходя из качества видеоизображения. Пример работы фильтра изображен на рисунке 1. Если телевизионная система имеет возможность накапливать изображения, то реализуется пространственное усреднение которое позволяет добиться лучших результатов, чем при работе с одиночным изображением.



Исходное изображение



Медианный фильтр с маской 3x3



Медианный фильтр с маской 5x5

Рис. 1. Применение шумоподавляющей фильтрации

После уменьшения шума на изображении необходимо повысить его контрастность, Как видно из формулы одним из ключевых параметров влияющих на исходное соотношение сигнал/шум является контрастность. Слабый контраст подводного изображения возникает из-за помехи обратного рассеяния. Для повышения контрастности рассмотрим методы преобразования гистограммы и линейного контрастирования [3]. Результаты применения контрастно ограниченной адаптивной эквализации гистограммы представлены на рисунке 2. Применение контрастирования позволяет повысить вероятность обнаружения объектов в подводной среде.



Рис. 2. Применение контрастно ограниченной адаптивной эквализации гистограммы

В результате последовательного применения нескольких методов фильтрации и контрастирования изображение наблюдается существенное повышение качества видеинформации. Таким образом применение комплексного подхода к повышению качества оправдывает себя, но возникает проблема быстродействия системы. В системах прямого управления (управление оператором) задержки в поступлении видеопотока при деликатных работах могут создать существенные трудности или привести к невозможности выполнения заданных работ. Необходим более глубокий анализ способов повышения качества видеинформации, либо поиск методов ускорения работы уже существующих.

#### Библиографический список

1. Соломатин В. А. Системы контроля и измерения с многоэлементными приемниками. М: Машиностроение, 1992. 128 с.
2. Красильников Н. Н. Цифровая обработка 2D- и 3D-изображений. СПб.: БХВ-Петербург, 2011. 608 с.
3. Гонсалес, Р. Цифровая обработка изображений/ Р. Гонсалес, Р. Вудс. М.: Техносфера, 2005. 1072 с.

#### METHODS FOR IMPROVING THE QUALITY OF VIDEO INFORMATION IN AN UNDERWATER ENVIRONMENT

**Prokonich A.V.**

Saint Petersburg Electrotechnical University "LETI"

*Standard methods for improving the quality of video information are given. Methods for improving the quality of video information regarding their use in an underwater environment*

**Quality video information, improving the quality of video, underwater television.**

**А. К. Цыцулин<sup>1</sup>, И. А. Зубакин<sup>2</sup>**

<sup>1</sup>АО «НИИ телевидения»

<sup>2</sup>СПбГЭТУ «ЛЭТИ»

## Проблема разделения информации на доминантную и фоновую

*В докладе рассматривается способ оценки доминантной и фоновой информации, из которых, согласно принципу доминантной информации, состоит информация (в том числе синтаксическая).*

**Качество информации, доминантная информация, фоновая информация, шумовая информация**

Согласно принципу доминантной информации [1, 2], информация разделяется по своему качеству на доминантную, фоновую и шумовую. В связи с тем, что концепция качества синтаксической информации [1, 2], обосновывающая оптимальные дискретизацию и квантование, использует функционалы, включающие разность между энтропией доминантного сигнала и пропускной способностью, актуальна оценка энтропий доминантного и фонового сигнала для корректной оценки требуемого сжатия сигнала без потерь доминантной информации. В первом приближении будем считать, что фотонным шумом и шумом канала связи можно пренебречь, поэтому ниже основное внимание будет уделено оценке доминантной и фоновой информации и гипотезе об их аддитивности.

Фоновая информация может быть удалена из изображения путем обработки последнего фильтром высоких частот. В качестве такого фильтра может быть использован фильтр Лапласа. Полученное на выходе фильтра Лапласа изображение может быть использовано для оценки доминантной информации в исходном изображении путем оценки его энтропии. Были вычислены оценки энтропии исходного изображения, его межстроковой разности, межстолбцовой разности, межстроковой и межстолбцовой разности и на выходе фильтра Лапласа (для различных масок фильтра, учитывающих вклад диагональных элементов). Лучший результат с точки зрения приближения к истинной энтропии дала оценка энтропии изображения на выходе фильтра Лапласа.

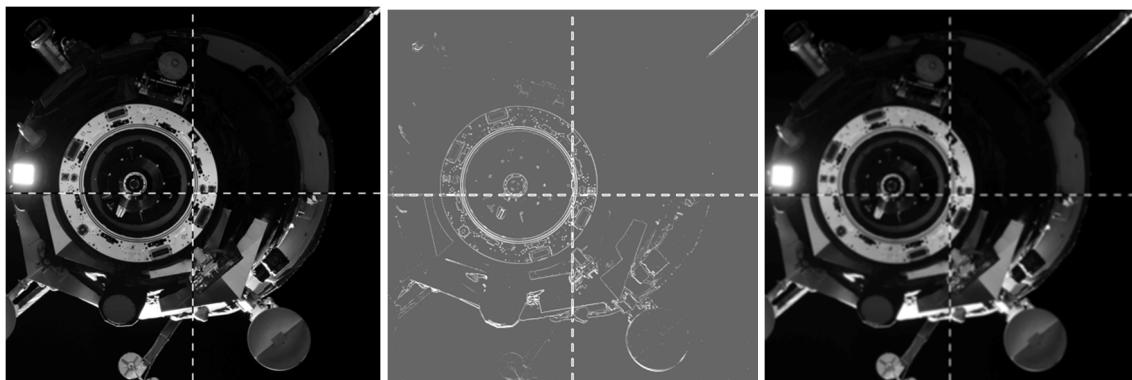


Рис. 1. Исходное изображение (слева), изображение на выходе фильтра Лапласа (в центре, усилено), изображение на выходе усредняющего фильтра (справа)

Рассуждая аналогично, оценка фоновой информации производилась путем обработки изображения фильтром низких частот с последующим вычислением энтропии полученного изображения.

Полученные таким образом оценки доминантной и фоновой информации опровергают гипотезу об их аддитивности. Однако, данные оценки являются оценками сверху, поэтому в сумме они могут быть и не равны исходному количеству информации в изображении.

На рисунке показаны изображения: исходное, на выходе фильтра Лапласа и на выходе усредняющего фильтра с маской  $15 \times 15$ .

Дальнейшие исследования в этой области должны подтвердить или опровергнуть гипотезу об аддитивности этих двух видов информации.

### **Выводы**

Актуален поиск корректных оценок доминантной и фоновой информации, которые позволяют для произвольных изображений находить меру оптимального сжатия без потерь доминантной информации.

### **Библиографический список**

1. Цыцуллин А. К., Адамов Д. Ю., Манцевов А. А., Зубакин И. А. Твердотельные телекамеры: накопление качества информации. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2014, 272 с. (см. <http://www.eltech.ru/ru/fakultety/fakultet-radiotekhniki-i-telekommunikaciy/publikacii>).
2. Цыцуллин А. К., Зубакин И. А. Концепция качества информации в теории связи. Вопросы радиоэлектроники, серия Техника телевидения, 2016, вып. 4, с. 19–25.

### **PROBLEM OF DIVISION OF INFORMATION ON DOMINANT AND BACKGROUND**

**Tsytsulin A. K.<sup>1</sup>, Zubakin I. A.<sup>2</sup>**

<sup>1</sup>Scientific Research Institute of Television,

<sup>2</sup>Saint Petersburg Electrotechnical University “LETI”

*In the report the way of an assessment of dominant and background information of which, according to the principle of dominant information, information consists is considered (including syntactic).*

**Quality of information, dominant information, background information, noise information**

**Г. М. Лабков**

Санкт-Петербургский государственный электротехнический  
университет "ПЭТИ"

## Исследование областей встраивания видеостеганоконтейнера

*Исследованы области встраивания видеостеганоконтейнера на примере алгоритма Куттера - Джордана - Боссена. Рассмотрены преимущества и недостатки сегментации изображения. Показано влияние искажений на распознаваемость видеостеганоконтейнера.*

**Стеганография, видеопоследовательности, скрытая связь, защита информации**

Среди систем, использующих в качестве стеганоконтейнеров изображения, важную роль играют видеопоследовательности, поскольку они обладают высокой пропускной способностью, имеют широкое распространение и не вызывают подозрений при передаче. Такую стеганографическую систему возможно реализовать с использованием алгоритма Куттера-Джордана-Боссена. Алгоритм основан на одном из свойств зрительной системы человека: в модели RGB восприимчивость зрительной системы человека к перепадам яркости синего цвета значительно меньше, чем зеленого и красного [1].

Для того, чтобы встроить один бит сообщения в изображении контейнере псевдослучайным образом выбирается пиксел с координатами  $x$  и  $y$ . Далее, путем модификации яркости

$$\lambda_{x,y} = 0,29890R_{x,y} + 0,58622G_{x,y} + 0,11448B_{x,y}$$

происходит встраивание секретной информации в канал синего цвета по формуле:

$$\hat{B}_{x,y} = \begin{cases} B_{x,y} + u\lambda_{x,y}, & m_i = 1 \\ B_{x,y} - u\lambda_{x,y}, & m_i = 0 \end{cases}$$

где  $u$ -константа, определяющая энергию встраиваемого сигнала. Выбор данного значения следует делать исходя из назначения системы: чем больше  $u$ , тем выше устойчивость системы к искажениям, однако при этом повышается заметность [4].

Извлечение секретного бита происходит путем предсказания не модифицированного пикселя на основе значений соседних пикселов по вертикали и горизонтали, как показано на рисунке 1. В данном случае выбран "крест" 5x5 пикселов.

Оценка интересующего пикселя находится по формуле:

$$\hat{B}^*_{x,y} = \frac{1}{4\sigma} \left[ \sum_{i=-\sigma}^{+\sigma} B^*_{x+i,y} + \sum_{j=-\sigma}^{+\sigma} B^*_{x,y+j} - 2B^*_{x,y} \right],$$

где  $\sigma$  – количество пикселов в одну из сторон "креста", на основании которых производится оценка. На рисунке 1  $\sigma = 2$ .

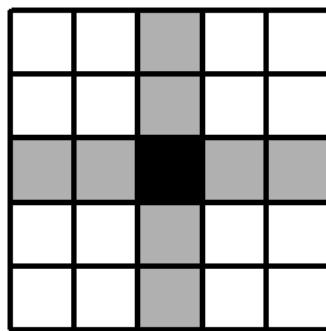


Рис. 1. Извлечение скрытой информации на основе значений соседних пикселов

Чтобы извлечь встроенный бит, необходимо вычислить разность  $\delta$  между текущим и оценочным значением пикселя в синем канале:

$$\delta = B_{x,y}^* - \hat{B}_{x,y}^*$$

Для сегментации областей встраивания стеганоконтейнера могут быть использованы градиентные операторы, посредством которых выделяются контуры и яркостные переходы [2]. Таким образом модернизированный алгоритм предполагает встраивание бит сообщения в области, не имеющие контуров и яркостных переходов, используя "кресты", в которых значения соседних пикселов не имеют резкого изменения значений яркости.

Данный подход уменьшает пропускную способность канала связи, однако повышает вероятность правильного распознавания данных, носящее вероятностный характер. К тому же немаловажным преимуществом данной модификации алгоритма является отсутствие необходимости наличия канала для передачи информации о расположении пикселов с измененной яркостью. Декодер, использующий градиентный оператор с заданным порогом, обнаруживает пиксели, содержащие сообщение. На графиках ниже показано распределение количества ошибок и зависимость количества доступных для встраивания пикселов от величины порога фильтра Собела. В исследовании использовалось изображение с размером 480x360 пикселов, каждый бита сообщения встраивался три раза.

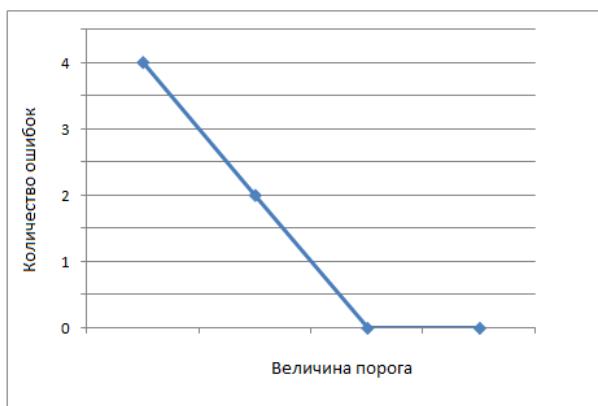


Рис. 2. Зависимость количества ошибок от величины порога оператора Собела

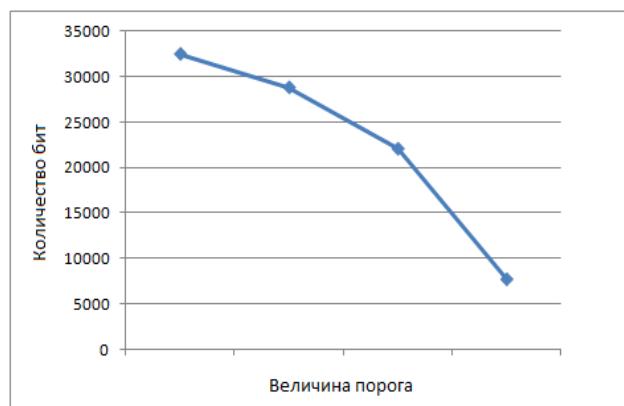


Рис. 3. Зависимость количества доступных для встраивания бит от величины порога оператора Собела

При прохождении по каналу связи стеганоконтейнер может подвергатьсяискажениям, вызванными попытками неавторизованного доступа. В большинстве случаев злоумышленники преследуют цели обнаружения факта присутствия скрытой информации, модификации стеганоконтейнера с целью исказить или полностью

разрушить содержимое передаваемого сообщения, а также извлечения скрытого сообщения. В случае, если по каналу связи передается видеостеганоконтейнер и целью нарушителя является искажение содержимого передаваемого сообщения, неавторизованный пользователь может применить сжатие, геометрические преобразования и использовать фильтрацию изображений. Поскольку данные операции приводят к искажению пикселов, неминуемо происходит частичная или полная потеря информации в сообщении.

В исследовании использовалось изображение размером 480x360 пикселов с глубиной цвета 8 бит. В качестве порога градиентного оператора было выбрано значение 0.1. В таблице 1 показано влияние искажений на распознаваемость видеостеганоконтейнера.

Влияние искажений на распознаваемость видеостеганоконтейнера

Таблица 1.

Вид искажения	Число верно распознанных символов
JPEG сжатие (в диапазоне 5-10%)	2 %
Поворот (в диапазоне 1-5°)	3 %
Масштабирование (с кратностью 0,5-2)	2 %
Линейная фильтрация	1 %

Полученные данные показывают, что модифицированный алгоритм Куттера-Джордана-Боссена крайне чувствителен к JPEG сжатию, геометрическим преобразованиям, а также линейной фильтрации.

#### Библиографический список

1. Kutter M., Jordan F., Bossen F. Digital signature of color images using amplitude modulation // Storage and retrieval for image and video databases. 1997. Proc. SPIE. Vol. 3022. P. 518–526.
2. Красильников Н.Н. Цифровая обработка 2D- и 3D- изображений: учеб. пособие. - СПб.: БХВ-Петербург, 2011. - 608 с.

### EXPLORATION OF VIDEOSTEGANOCONTAINER EMBEDDING AREAS

**G. M. Labkov**

Saint-Petersburg electrotechnical university "LETI"

*Regions for embedding steganocontainer for Kutter-Jordan-Bossen method of steganography is considered. Advantages and disadvantages of image segmentation are explore. The influence of noise on the recognizability of the videosteganocontainer shown.*

**Steganography, video sequences, hidden communication, information protection.**